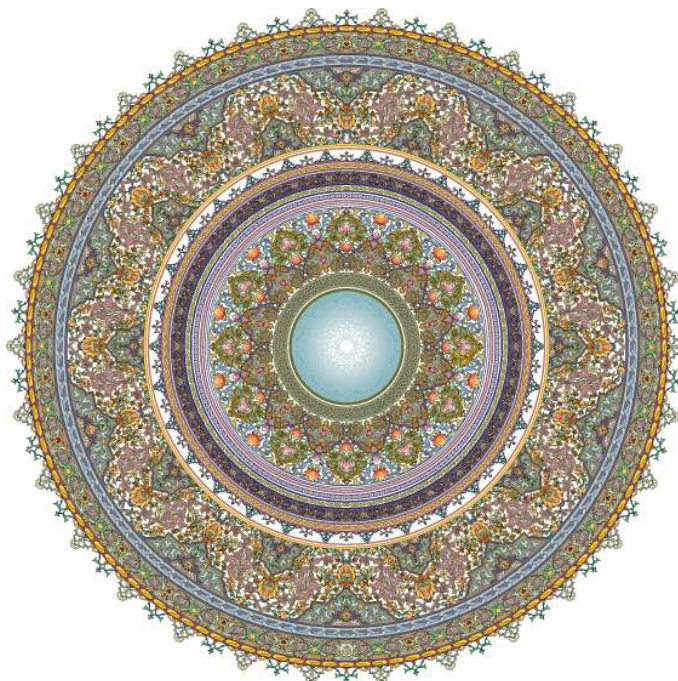

Clases de

Álgebra Lineal

para Ingeniería

Teoría y 300 ejercicios resueltos

Revisión 6.6 - Septiembre 2025



Pedro José Hernando Oter

© Pedro José Hernando Oter, 2025

pedroj@math.uc3m.es
www.clasesde.eu

Departamento de Matemáticas
Escuela Politécnica Superior
UNIVERSIDAD CARLOS III DE MADRID
Avda de la Universidad, 30
28911 Leganés
SPAIN

ISBN: xxx-xx-xxx-xxxx-x

Versión 6.6 - Septiembre 2025
Made by $\text{\LaTeX}$ - $\text{\TeX}$ Live

This is pdfTeX, Version 3.141592653-2.6-1.40.27 (TeX Live 2026/dev/Arch Linux) kpathsea version 6.4.2/dev

Índice

Tema	Página
0 Conceptos Básicos de Rectas y Planos	1
0.1 Sistema de Coordenadas Cartesianas	3
0.2 Ecuaciones de la Recta en el Plano Euclídeo	3
0.2.1 Ecuación de dos Puntos	4
0.2.2 Ecuación Punto-Pendiente	4
0.2.3 Ecuación General o Cartesiana de la Recta	5
0.2.4 Casos Particulares: Rectas Horizontales y Verticales	5
0.2.5 Forma Vectorial y Paramétrica de la Recta	6
0.2.6 Ecuación Normal de la Recta	8
0.2.7 Resumen de Fórmulas: Rectas en el Plano Euclídeo	8
0.2.8 Conversión entre las Formas de una Recta	9
General $\iff$ Punto-Pendiente	9
General $\iff$ Vectorial - Paramétrica	9
General $\iff$ Normal	9
0.3 Ecuaciones del Plano en el Espacio 3D	11
0.3.1 Ecuación General del Plano	11
0.3.2 Forma Vectorial y Paramétrica del Plano	12
0.3.3 Forma Normal del Plano	13
0.3.4 Resumen Fórmulas: Planos en el Espacio 3D	14
0.3.5 Conversión entre Formas de un Plano	14
General $\iff$ Vectorial - Paramétrica	14
General $\iff$ Normal	15
0.4 Ecuaciones de la Recta en el Espacio 3D	15
0.4.1 Forma General de la Recta	15
0.4.2 Forma Vectorial-Paramétrica de la Recta	16
0.4.3 Forma Normal de la Recta	17
0.4.4 Resumen Fórmulas: Rectas en el Espacio 3D	17
0.4.5 Conversión entre Formas	17
General $\iff$ Vectorial - Paramétrica	17
General $\iff$ Normal	18
0.5 Geometría de los Sistemas de Ecuaciones en el Plano y en el Espacio 3D	19
0.5.1 Sistemas de Una Ecuación	19
0.5.2 Sistemas de Dos Ecuaciones	20
0.5.3 Sistemas de Tres Ecuaciones	23
0.6 Métodos Simples de Resolución de Sistemas	25
1 Sistemas de Ecuaciones Lineales	27
1.1 Introducción a los Sistema de Ecuaciones Lineales (SEL)	29
1.1.1 Características de los Sistemas de Ecuaciones	29
1.1.2 Ecuaciones y Sistemas de Ecuaciones	29
Sistemas de Ecuaciones	30
Solución	31
1.1.3 Puntos y Espacios	32
1.1.4 Sistemas de Ecuaciones Lineales (SEL)	34

1.2	Geometría de los SEL en $\mathbb{R}^n$	34
1.2.1	Objetos Geométricos	36
1.2.2	Dimensión de los Objetos Geométricos	36
1.2.3	Fórmula o Ecuación de Equilibrio	37
1.2.4	Relación entre Objetos Geométricos en un SEL	38
1.3	Solución de un SEL en $\mathbb{R}^n$	39
1.3.1	SEL Compatibles e Incompatibles	40
	Conclusiones	40
1.4	Métodos Resolución de SEL en $\mathbb{R}^n$	42
1.4.1	Métodos Simples	42
1.4.2	Ecuaciones Equivalentes	42
1.4.3	Combinación Lineal de Ecuaciones	43
	Propiedades de las Combinaciones Lineales	43
1.4.4	Dependencia e Independencia lineal de Ecuaciones	43
1.4.5	Sistemas Equivalentes	44
1.4.6	Operaciones Elementales de Fila	44
	Notación de las Operaciones Elementales de Fila	45
1.4.7	Sistemas Escalonados o Triangulares	45
	Método de Sustitución hacia Atrás (MSA)	46
	Ventajas de los SEL Triangulares	46
1.4.8	Método de Gauss o Eliminación Gaussiana	46
1.5	Matrices	47
1.5.1	Matrices Asociadas a SEL	48
1.5.2	Operaciones Elementales de Fila en la Matriz Asociada	49
1.5.3	Matriz Equivalente por Filas	49
1.5.4	Matriz Escalonada Equivalente	50
1.5.5	Rango de una Matriz	50
1.5.6	Teorema de Rouché-Fröbenius	51
1.6	Método de Gauss Matricial	51
1.6.1	Identificación de Sistemas Incompatibles	52
1.6.2	Identificación de Sistemas Compatibles Indeterminados	53
1.6.3	Estudio de Soluciones de SEL con Parámetros	55
1.7	Método de Gauss-Jordan	57
1.7.1	Matriz Escalonada Reducida	57
	Unicidad de la Matriz Escalonada Reducida	57
	Proceso del Método Gauss-Jordan	58
1.7.2	Matriz Identidad	58
1.8	Sistemas Homogéneos	59
1.9	SEL en $\mathbb{C}^n$	61
2	Espacios Vectoriales	63
2.1	Vectores	65
2.1.1	Representación Gráfica de Vectores Reales	67
2.2	Operaciones Elementales con Vectores	67
2.2.1	Suma	67
	Regla del Paralelogramo	67
2.2.2	Producto por un Escalar	68
2.2.3	Vector Opuesto y Vectores Estándar	69
2.2.4	Diferencia de Vectores	69
2.2.5	Propiedades Algebraicas de la Suma y Producto por un Escalar	70
2.3	Espacio y Subespacio Vectorial	71
2.4	Sistemas de Vectores	73
2.4.1	Combinación Lineal de Vectores	73
2.4.2	Dependencia e Independencia Lineal	73
2.4.3	Rango de un Sistema de Vectores	76
2.4.4	Cálculo del Rango de un Sistema de Vectores	77
2.4.5	Interpretación Geométrica de la Dependencia e Independencia Vectorial	78

2.5	Conjuntos o Sistemas Generadores	79
2.5.1	Base y Dimensión de un Sistema de Vectores	80
2.5.2	Generación de $\mathbb{R}^n$	81
	Generación del Espacio $\mathbb{R}^2$	81
	Base Canónica o Estándar de $\mathbb{R}^2$	81
	Generación del Espacio $\mathbb{R}^3$	82
	Base Canónica o Estándar de $\mathbb{R}^3$	82
	Generación del Espacio Vectorial $\mathbb{R}^n$	82
2.5.3	Base Canónica o Estándar de $\mathbb{K}^n$	83
3	Matrices	87
3.1	Matrices	89
3.1.1	Tipos de Matrices	89
3.2	Operaciones Básicas con Matrices	91
3.2.1	Suma de Matrices	91
3.2.2	Producto de un Escalar por una Matriz	92
3.2.3	Diferencia o Resta de Matrices	92
3.2.4	Propiedades de la Suma y la Multiplicación por un Escalar	92
3.2.5	Espacio Vectorial de Matrices	92
3.2.6	Combinación Lineal de Matrices	93
3.3	Producto de Matrices	93
3.3.1	Propiedades de la Multiplicación de Matrices	94
3.3.2	Potencia de Matrices Cuadradas	95
	Propiedades de la Potencia de Matrices Cuadradas	95
3.3.3	Producto Matricial Matriz-Vector	95
3.3.4	Producto Matricial Vector-Vector	95
3.3.5	Producto Matriz-Vector Estándar	96
3.4	Expresión Matricial de un SEL	96
3.5	Matriz Transpuesta A^T	97
3.5.1	Matriz Simétrica y Antisimétrica	97
3.5.2	Propiedades de la Matriz Transpuesta	97
3.6	Matriz Transpuesta Conjugada A^H	98
3.6.1	Matriz Hermítica y Antihermítica	99
3.6.2	Propiedades de la Matriz Transpuesta Conjugada	99
3.7	Inversa de una Matriz	99
3.7.1	Cálculo de la Matriz Inversa: Método de Gauss-Jordan	99
	Proceso de Cálculo de la Matriz Inversa (Método Gauss-Jordan)	100
	Propiedades de la Matriz Inversa	100
3.7.2	Inversa y SEL	101
3.7.3	Operaciones con la Matriz Inversa	101
3.8	Matrices por Bloques	101
3.8.1	Propiedades de las Matrices por Bloques	102
3.8.2	Matriz Diagonal por Bloques	103
3.8.3	Propiedades de las Matrices Diagonales por Bloques	103
3.9	Determinante	103
3.9.1	Introducción: Determinante de Matrices 2×2	103
	Propiedades del Determinante de una Matriz 2×2	104
3.9.2	Determinante de Matrices 3×3	104
3.9.3	Regla de Sarrus	104
3.9.4	Determinante de Matrices $n \times n$	105
3.9.5	Cálculo del Determinante mediante Desarrollos de Laplace	106
3.9.6	Determinante e Inversa de una Matriz	107
3.9.7	Propiedades de los Determinantes	107
3.9.8	Método de Gauss para el Cálculo del Determinante	108
3.9.9	Regla de Cramer	109
3.9.10	Matriz Adjunta e Inversa de una Matriz	110
3.10	Subespacios Fundamentales de una Matriz	111

3.10.1	Espacios Columna y Fila de una Matriz	111
3.10.2	Propiedades de las Matrices Equivalentes por Filas	113
3.10.3	Rango, Núcleo y Nulidad de una Matriz	114
3.10.4	Teorema del Rango de las Matrices	116
3.10.5	Teorema Fundamental de las Matrices Invertibles	117
3.10.6	Bases para los Espacios Fila, Columna y Nulo de una Matriz	117
4	Transformaciones Lineales	119
4.1	Concepto de Transformación Lineal	121
4.1.1	Forma Matricial de una Transformación Lineal	122
4.1.2	Operaciones con Transformaciones Lineales	123
4.1.3	Propiedades de las Transformaciones Lineales	123
4.2	Transformaciones Lineales Elementales	125
4.2.1	Transformación Lineal de Vectores Estándar	126
4.2.2	Demostración $\Leftarrow$ de la Sección 4.1.3	126
4.3	Inversa de una TL	127
4.3.1	Función Invertible	127
4.3.2	Inversa de una TL	131
	Propiedades de la TL Inversa	132
	Linealidad de la TL Inversa	132
4.4	Operaciones con TL	132
4.4.1	Suma de Transformaciones Lineales	132
	Propiedades de la Suma de TL	132
4.4.2	Composición de Transformaciones	133
	Dominio e Imagen de la Composición de TL	133
	Propiedades de la Composición de TL	134
	Matriz asociada a la Composición de TL	134
4.5	Imagen y Núcleo de una TL	136
4.5.1	Imagen de una TL	136
	Propiedades de la Imagen de una TL	137
4.5.2	Núcleo o Espacio Nulo de una TL	137
	Propiedades del Núcleo de una TL	139
4.5.3	Nulidad y Rango de una TL	139
	Relación entre los Subespacios Asociados a una TL	141
4.6	Tipos de TL	141
4.6.1	Inyectivas, Sobreyectivas y Biyectivas	141
4.6.2	Homomorfismos, Endomorfismos e Isomorfismos	142
5	Bases	143
5.1	Bases	145
5.1.1	Coordenadas y Matriz asociada a una Base	146
5.2	Cambio de Base	152
5.2.1	Matriz de Cambio de Base	153
5.2.2	Cambio de Base en una TL	158
6	Ortogonalidad	161
6.1	Introducción	163
6.2	Producto Escalar en $\mathbb{R}^n$	163
	Propiedades del Producto Escalar	164
6.2.1	Distancia y Norma en $\mathbb{R}^n$	164
	Propiedades de la Distancia	164
	Distancia de un Punto al Origen	164
	Norma de un Vector en $\mathbb{R}^n$	164
	Propiedades de la Norma	165
	Vectores Unitarios y Normalización	165
6.2.2	Relaciones entre Vectores	166
	Distancia entre dos Vectores	166
	Ángulo entre dos Vectores en $\mathbb{R}^2$	166

	Ángulo entre dos Vectores en $\mathbb{R}^n$	166
	Definición Alternativa del Producto Escalar en $\mathbb{R}^n$	167
6.2.3	Vectores Ortogonales	167
6.3	Ortogonalidad en $\mathbb{R}^n$	168
6.3.1	Conjunto Ortogonal	168
6.3.2	Bases Ortogonales	169
6.3.3	Bases Ortonormales	170
6.3.4	Matrices Ortogonales	171
	Propiedades de las Matrices Ortogonales	172
6.3.5	TL Ortogonales	173
6.4	Complemento Ortogonal	173
6.4.1	Vector Ortogonal a un Subespacio	173
6.4.2	Complemento Ortogonal	174
6.4.3	Características del Complemento Ortogonal	174
6.4.4	Cálculo del Complemento Ortogonal	174
6.4.5	Relaciones de Ortogonalidad entre los Subespacios de una Matriz Real	175
6.4.6	Dimensión del Complemento Ortogonal	177
6.5	Proyección Ortogonal	178
6.5.1	Interpretación Geométrica en $\mathbb{R}^2$ y $\mathbb{R}^3$	178
6.5.2	Vector Perpendicular	180
6.5.3	Proyección Ortogonal sobre un Subespacio	181
	Interpretación Geométrica	182
6.6	Teorema de la Descomposición Ortogonal	183
6.7	Proceso de Gram-Schmidt	184
6.8	Factorización QR	185
6.9	Ortogonalidad en Espacios $\mathbb{C}^n$	187
6.9.1	Producto Interno	187
	Propiedades del Producto Interno	187
6.9.2	Norma en $\mathbb{C}^n$	188
	Propiedades de la Norma	188
6.9.3	Vectores Ortogonales en $\mathbb{C}^n$	188
6.9.4	Complemento Ortogonal en $\mathbb{C}^n$	189
	Subespacios Fundamentales de una Matriz Compleja	189
	Cálculo del Complemento Ortogonal de un Subespacio Vectorial Complejo	191
6.9.5	Proyección Ortogonal en $\mathbb{C}^n$	192
6.9.6	Proceso de Gram-Schmidt en $\mathbb{C}^n$	193
6.9.7	Matrices Unitarias	194
7	Mínimos Cuadrados	195
7.1	Introducción	197
7.2	Mejor Aproximación	197
7.3	Aproximación mediante Mínimos Cuadrados	198
7.3.1	SEL Incompatibles	198
7.3.2	Solución de Mínimos Cuadrados	199
7.3.3	Ecuaciones Normales	200
7.4	Mínimos Cuadrados mediante la Factorización QR	204
7.5	Error en la Aproximación por Mínimos Cuadrados: Residuos	206
7.6	Resumen	206
7.7	Ajuste de Datos mediante Mínimos Cuadrados	207
	Ajuste de datos a una Recta en $\mathbb{R}^2$	208
	Parábola en $\mathbb{R}^2$	209
	Polinomio General en $\mathbb{R}^2$	209
	Plano en $\mathbb{R}^3$	209
	Superficie Cuadrática en $\mathbb{R}^3$	210
	Hiperplano en $\mathbb{R}^m$	210
7.7.1	Ejercicios Resueltos sobre Mínimos Cuadrados	211
7.8	Resumen de Fórmulas de Mínimos Cuadrados	216

8	Autovalores y Autovectores	217
8.1	Introducción: Sistemas Dinámicos Discretos	219
8.2	Autovalores y Autovectores	220
8.2.1	Interpretación Geométrica de los Autovalores y Autovectores	220
8.2.2	Determinación de Autovalores y Autovectores	221
	Procedimiento de Cálculo de Autovalores y Autovectores	222
8.2.3	Autoespacios	223
8.2.4	Multiplicidades Algebraica y Geométrica	225
8.2.5	Autovalores y Matrices Inversas	226
	Teorema Fundamental de las Matrices Invertibles	227
8.2.6	Teoremas sobre Autovalores y Autovectores	227
8.3	Propiedades de los Autovalores de una Matriz	229
8.4	Matrices Semejantes	230
8.4.1	Matrices Triangulares	230
8.4.2	Matrices Semejantes	230
	Propiedades de las Matrices Semejantes	231
8.4.3	Interpretación de la Semejanza de Matrices	232
8.5	Diagonalización	234
8.5.1	Autovalores Reales y Complejos	236
8.6	Diagonalización Ortogonal	237
8.6.1	Interpretación Geométrica de la Diagonalización Ortogonal	238
8.6.2	Teorema Espectral	241
8.6.3	Descomposición Espectral	241
8.7	Teoremas Avanzados Adicionales	242
8.8	Diagonalización Unitaria	243
8.9	Resumen Final del Tema	245

APÉNDICES 247

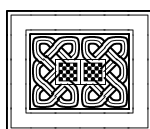
A	Temas de Repaso	249
A.1	Trigonometría	250
A.1.1	Medida de Ángulos: Radianes	250
	Relación Grados-Radianes	250
A.1.2	Funciones Trigonométricas	251
	Definición de Seno, Coseno y Tangente	251
	Valores Trigonométricos: Seno, Coseno y Tangente	252
	Determinación Simplificada	253
	Representación Gráfica de las Funciones Trigonométricas Directas	255
	Inversas de las Funciones Trigonométricas	255
	Representación Gráfica de las Inversas de las Funciones Trigonométricas	256
	Funciones Inversas de las Funciones Trigonométricas	256
	Representación Gráfica de las Funciones Trigonométricas Inversas	257
	Relaciones Trigonométricas Básicas	258
A.2	Funciones Hiperbólicas	260
	Inversar de las Funciones Hiperbólicas	260
	Funciones Hiperbólicas Inversas	260
	Representación Gráfica de las Funciones Hiperbólicas	260
	Relaciones Hiperbólicas Básicas	260
A.3	Productos Notables	262
A.4	Sumatorios	262
A.5	Binomio de Newton	263
A.6	Polinomios	264
	Raíces o Ceros de un Polinomio	264
	División entre Polinomios	265
A.6.1	Regla de Ruffini	266
	Cálculo de las Raíces de un Polinomio	267

Descomposición de un Polinomio en Factores	267
A.7 Distancia Euclídea	269
A.7.1 Valor Absoluto	269
A.7.2 Distancia en $\mathbb{R}$	269
A.7.3 Distancia en $\mathbb{R}^2$	269
A.7.4 Distancia en $\mathbb{R}^3$	270
A.7.5 Distancia en $\mathbb{R}^n$	270
A.8 Vectores en el Plano y en el Espacio	270
Características de los Vectores geométricos	271
A.8.1 Vectores de Posición	271
A.8.2 Vectores que no son de posición: Vectores Generales	271
A.8.3 Vectores Elementales	272
A.9 Vectores en Espacios $\mathbb{R}^n$, $\mathbb{C}^n$ y $\mathbb{K}^n$	272
Nomenclatura	273
A.9.1 Caracterización de Vectores	273
Norma de un Vector en $\mathbb{R}^n$	273
Longitud y Ángulo en $\mathbb{R}^2$	273
A.10 Producto Escalar y Vectorial en $\mathbb{R}^3$	274
A.10.1 Producto Escalar	274
A.10.2 Producto Vectorial	274
A.11 Fórmulas Geométricas sobre Distancias	275
A.11.1 Distancia de un Punto a un Recta en $\mathbb{R}^2$	275
A.11.2 Distancia entre dos Rectas paralelas en $\mathbb{R}^2$	275
A.11.3 Distancia de un Punto a un Plano en $\mathbb{R}^3$	275
A.11.4 Distancia entre dos Planos Paralelos en $\mathbb{R}^3$	275
A.11.5 Distancia de un Punto a un Recta en $\mathbb{R}^3$	276
A.11.6 Distancia entre dos Rectas paralelas en $\mathbb{R}^3$	276
A.12 Introducción a las Funciones	277
A.12.1 Tipos de Funciones	277
Funciones Reales y Complejas	277
Funciones Escalares y Vectoriales	277
A.12.2 Representación Gráfica de Funciones Reales	278
A.12.3 Tipos de Funciones	279
A.12.4 Función Inversa (o Recíproca) $f^{-1}(x)$	280
A.12.5 Composición de Funciones	281
B Aplicaciones Geométricas	283
B.1 Introducción	284
B.2 Aplicaciones Geométricas de la Proyección Ortogonal	284
B.2.1 Distancia de un Punto A a un Recta en $\mathbb{R}^2$ y $\mathbb{R}^3$	284
B.2.2 Distancia entre dos Rectas paralelas en $\mathbb{R}^2$ y $\mathbb{R}^3$	285
B.2.3 Distancia de un Punto A a un Plano en $\mathbb{R}^3$	286
B.2.4 Distancia entre Dos Planos Paralelos en $\mathbb{R}^3$	286
B.3 Transformaciones Geométricas. Movimientos en $\mathbb{R}^2$ y $\mathbb{R}^3$	287
B.3.1 Traslaciones en $\mathbb{R}^2$ y $\mathbb{R}^3$	287
B.3.2 Rotaciones en $\mathbb{R}^2$	288
B.3.3 Rotaciones en $\mathbb{R}^3$	290
B.3.4 Simetría Puntual en $\mathbb{R}^2$	290
B.3.5 Reflexiones sobre Rectas en $\mathbb{R}^2$	290
B.3.6 Reflexiones sobre Rectas en $\mathbb{R}^3$	292
B.3.7 Reflexiones sobre Planos en $\mathbb{R}^3$	293
B.3.8 Homotecias en $\mathbb{R}^2$ y $\mathbb{R}^3$	293
C Números Complejos	295
C.1 Introducción	297
C.2 Formas de los Números Complejos	297
C.2.1 Forma Cartesiana, $z = (a, b)$	297

C.2.2	Forma Binómica, $z = a + bi$	298
C.3	Operaciones Básicas en $\mathbb{C}$	298
C.3.1	Suma de Números Complejos	298
	Propiedades de la Suma	298
C.3.2	Producto de Números Complejos	299
C.3.3	Potencias Enteras de i	299
C.3.4	Elemento Inverso, z^{-1}	300
C.3.5	División en $\mathbb{C}$	300
	Propiedades del Producto	301
C.3.6	El Cuerpo de los Números Complejo	301
C.4	Conjugado de un Número Complejo	301
C.4.1	Propiedades del Conjugado	301
C.4.2	Conjugado y División de Números Complejos	301
C.5	Representación Geométrica de $\mathbb{C}$	302
C.5.1	El Plano Complejo	302
C.5.2	Módulo de un Número Complejo	303
	Propiedades del Módulo, $ z $	303
	Módulo y División de Números Complejos	303
C.5.3	Argumento de un Número Complejo	304
	Determinación General del Argumento	304
C.6	Otras Formas de los Números Complejos	306
C.7	Forma Trigonométrica y Polar	306
C.7.1	Propiedades y Operaciones Básicas de la Forma Polar	307
	Multipliación y División en Forma Polar	307
C.7.2	Potencia n-ésima en Forma Polar	308
C.8	Forma Exponencial	309
C.8.1	Exponencial Compleja. Fórmula de Euler	309
C.8.2	Fórmula de De Moivre	310
C.8.3	Forma Exponencial	311
C.8.4	Operaciones en Forma Exponencial	311
C.9	Logaritmo Complejo	311
C.10	Potencia Compleja	313
C.11	Raíz n-ésima de un Número Complejo	314
	Geometría de las Raíces n-ésimas	315
C.11.1	Raíz Cuadrada Generalizada	316
C.12	Seno y Coseno Complejo	317

Bibliografía

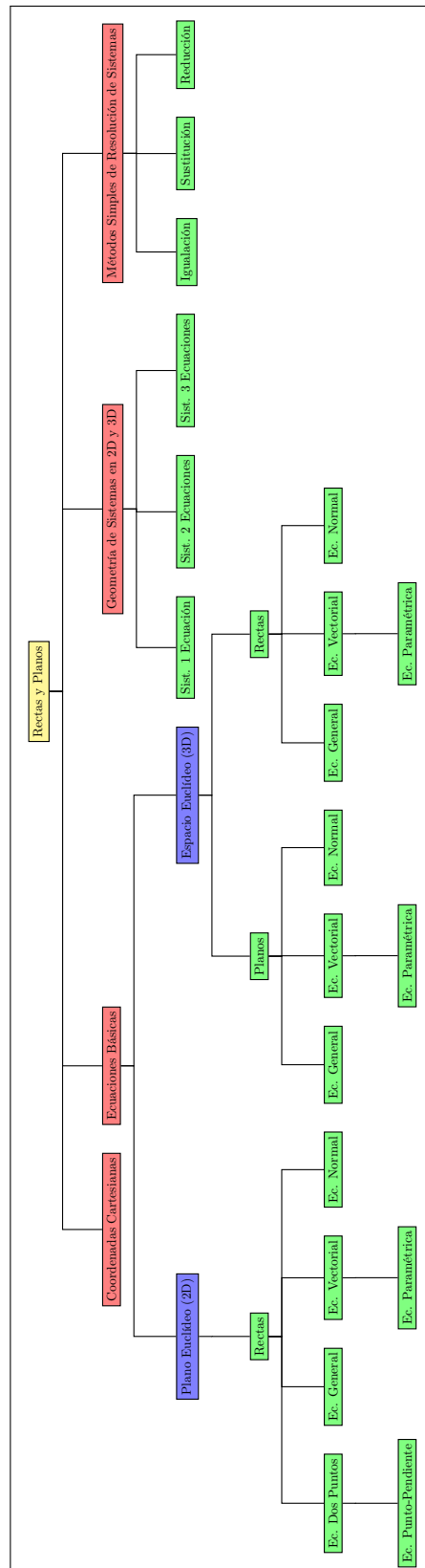
319



Conceptos Básicos de Rectas y Planos

Contenido

0.1	Sistema de Coordenadas Cartesianas	3
0.2	Ecuaciones de la Recta en el Plano Euclídeo	3
0.2.1	Ecuación de dos Puntos	4
0.2.2	Ecuación Punto-Pendiente	4
0.2.3	Ecuación General o Cartesiana de la Recta	5
0.2.4	Casos Particulares: Rectas Horizontales y Verticales	5
0.2.5	Forma Vectorial y Paramétrica de la Recta	6
0.2.6	Ecuación Normal de la Recta	8
0.2.7	Resumen de Fórmulas: Rectas en el Plano Euclídeo	8
0.2.8	Conversión entre las Formas de una Recta	9
0.3	Ecuaciones del Plano en el Espacio 3D	11
0.3.1	Ecuación General del Plano	11
0.3.2	Forma Vectorial y Paramétrica del Plano	12
0.3.3	Forma Normal del Plano	13
0.3.4	Resumen Fórmulas: Planos en el Espacio 3D	14
0.3.5	Conversión entre Formas de un Plano	14
0.4	Ecuaciones de la Recta en el Espacio 3D	15
0.4.1	Forma General de la Recta	15
0.4.2	Forma Vectorial-Paramétrica de la Recta	16
0.4.3	Forma Normal de la Recta	17
0.4.4	Resumen Fórmulas: Rectas en el Espacio 3D	17
0.4.5	Conversión entre Formas	17
0.5	Geometría de los Sistemas de Ecuaciones en el Plano y en el Espacio 3D	19
0.5.1	Sistemas de Una Ecuación	19
0.5.2	Sistemas de Dos Ecuaciones	20
0.5.3	Sistemas de Tres Ecuaciones	23
0.6	Métodos Simples de Resolución de Sistemas	25



0.1. Sistema de Coordenadas Cartesianas

Las *coordenadas cartesianas* o *rectangulares* se usan para definir un sistema de referencia respecto a un solo eje (*recta real*), respecto a dos ejes (*plano euclídeo*) o respecto a tres ejes (*espacio tridimensional*¹).

En el *plano euclídeo* y en el *espacio 3D*, los ejes son perpendiculares entre sí y se cortan en un punto llamado *origen de coordenadas* representado por O .

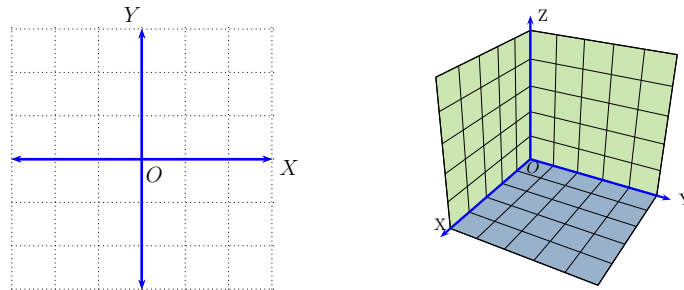


Figura 1: Sistemas de coordenadas cartesianas en el *plano euclídeo* y en el *espacio 3D*.

En el *plano euclídeo*, los ejes se denominan:

- **Eje X o de abscisa:** eje horizontal.
- **Eje Y o de ordenada:** eje vertical.

En el *espacio 3D*, las coordenadas cartesianas se suelen denominar² eje X, eje Y y eje Z.

La localización de un *punto* en el *plano euclídeo* y en el *espacio 3D* se realiza mediante la utilización de las *coordenadas cartesianas*.

- Las *coordenadas cartesianas* de un punto P en el *plano euclídeo* son el par de números reales (x, y) en los cuales rectas perpendiculares a los ejes que pasan por el punto P , cortan a los ejes.
- Las *coordenadas cartesianas* de un punto P en el *espacio 3D* son la terna de números (x, y, z) en los cuales planos perpendiculares a los ejes que pasan por el punto P , cortan a los ejes.

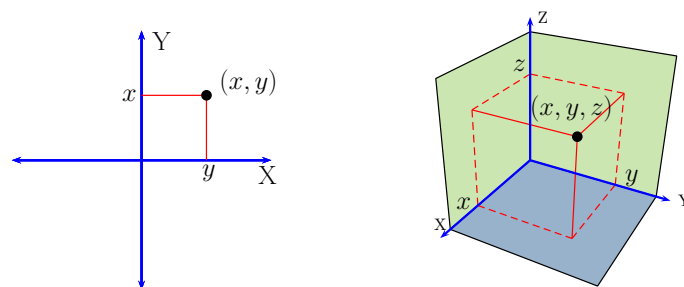


Figura 2: Representación gráfica de puntos en el *plano euclídeo* y en el *espacio 3D*.

0.2. Ecuaciones de la Recta en el Plano Euclídeo

Una **recta** es un conjunto infinito de puntos dispuestos en una misma dirección.

Vamos a ver las diferentes formas básicas que hay de representar matemáticamente una recta en el *plano euclídeo*.

¹El **espacio euclídeo** es un tipo de espacio geométrico donde se satisfacen los axiomas de la geometría clásica de Euclides. La *recta real*, el *plano euclídeo* y el *espacio tridimensional (3D)* son casos especiales de espacios euclídeos de dimensiones 1, 2 y 3 respectivamente. Estos conceptos se pueden extender también al campo de los números complejos.

²La disposición espacial de estos ejes puede variar según las diferentes aplicaciones.

0.2.1. Ecuación de dos Puntos

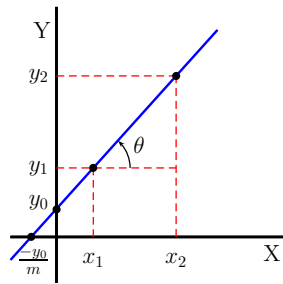


Figura 3: Representación de una recta en el *plano euclídeo* y sus elementos característicos.

Por dos puntos distintos del *plano euclídeo* de coordenadas (x_1, y_1) y (x_2, y_2) pasa una única recta, cuyos infinitos puntos (x, y) cumplen la ecuación:

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{y - y_1}{x - x_1} \quad (1)$$

Despejando la *variable*³ *dependiente* y en función de la *variable independiente* x , se obtiene la siguiente fórmula práctica:

$$y = y_1 + \frac{y_2 - y_1}{x_2 - x_1} (x - x_1) \quad (2)$$

Ejercicio 0.1. Encontrar la ecuación de la recta que pasa por los puntos $P_1 = (1, 2)$ y $P_2 = (-1, 3)$ del *plano euclídeo*.

Vamos a aplicar la **Ecuación (1)**, tomando por ejemplo⁴ $P_1 = (x_1, y_1)$ y $P_2 = (x_2, y_2)$.

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{y - y_1}{x - x_1} \quad ; \quad \frac{3 - 2}{-1 - 1} = \frac{y - 2}{x - 1} \quad ; \quad \frac{-1}{2} = \frac{y - 2}{x - 1} \quad ; \quad -(x - 1) = 2(y - 2)$$

$$-(x - 1) = 2y - 4 \quad ; \quad 2y = 4 - (x - 1) \quad ; \quad y = 2 - \frac{1}{2}(x - 1)$$

Esta misma ecuación se puede obtener aplicando directamente la **Ecuación (2)**. Podemos comprobar como se cumple para las coordenadas de los puntos P_1 y P_2 utilizados en su cálculo.

0.2.2. Ecuación Punto-Pendiente

Se denomina *pendiente* a la inclinación de una recta. Conociendo dos puntos de la misma, su *pendiente* m se puede determinar como:

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad ; \quad \tan \theta = m \quad (3)$$

El valor de m puede ser cualquier número real (negativo, cero o positivo) o bien *infinito* si la recta es vertical⁵. El ángulo de inclinación de la recta θ se puede obtener calculando la arcotangente⁶ de m .

Sustituyendo en la **Ecuación (2)** se obtiene la ecuación de la recta que pasa por un punto (x_1, y_1) y tiene por pendiente m , conocida como ecuación *punto-pendiente*:

$$y = y_1 + m(x - x_1) \quad (4)$$

³Una *variable* es un símbolo que puede ser remplazado o que toma un valor numérico en una ecuación o expresión matemática. Hay diferentes tipos de variables, siendo los más básicos las *variables dependientes*, que son aquellas que dependen del valor de otras variables y las *variables independientes*, que no dependen de otras variables.

⁴Es indiferente el orden que se tome en los puntos, en este caso podríamos tomar $P_1 = (x_2, y_2)$ y $P_2 = (x_1, y_1)$ obteniéndose el mismo resultado.

⁵Ver la **Sección 0.2.4**, página 5.

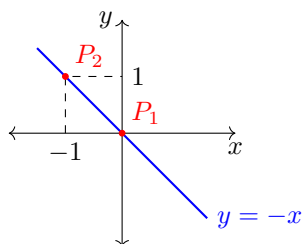
⁶Ver el **Apéndice A.1**, página 250.

Si desarrollamos el paréntesis y denominamos $y_0 = y_1 - mx_1$, se obtiene la ecuación *punto-pendiente* simplificada:

$$y = y_0 + mx \quad (5)$$

El valor y_0 se denomina *ordenada en el origen* y corresponde con la ordenada del punto de corte de la recta con el eje Y , (ver la **Figura 3**).

Ejercicio 0.2. Encontrar la recta en su forma punto-pendiente que pasa por los puntos $P_1 = (0,0)$ y $P_2 = (-1,1)$ y determinar su pendiente y su ordenada en el origen.



Vamos a utilizar la **Ecuación (3)** para determinar la pendiente de la recta:

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1 - 0}{-1 - 0} = -1$$

Por tanto:

$$y = y_1 + m(x - x_1) = 0 - 1(x - 0) = -x \quad ; \quad \boxed{y = -x}$$

La pendiente de esta recta es -1 y su ordenada en el origen 0.

0.2.3. Ecuación General o Cartesiana de la Recta

Una forma muy práctica de representar una recta en el *plano euclídeo* es mediante su *ecuación general* o *cartesiana*⁷:

$$\boxed{ax + by = c} \quad ; \quad a, b, c \in \mathbb{R} \quad (6)$$

donde las constantes a , b y c son número reales ($\mathbb{R}$).

Esta ecuación se puede obtener multiplicando la **Ecuación (5)** por un número b distinto de cero⁸ y llamando $a = bm$ y $c = by_0$.

Ejercicio 0.3. Determinar la ecuación general de la recta que tiene por pendiente $m = 3/2$ y pasa por el punto $P = (-1, 2)$.

Aplicando la **Ecuación (4)** se tiene:

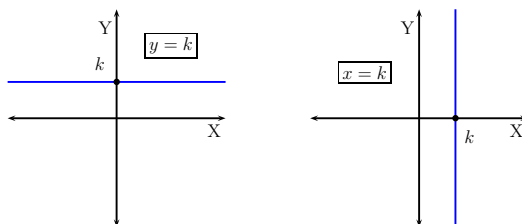
$$y = y_1 + m(x - x_1) = 2 + \frac{3}{2}[x - (-1)] = 2 + \frac{3x}{2} + \frac{3}{2} = \frac{7 + 3x}{2}$$

Reordenando podemos encontrar la ecuación general de esta recta:

$$2y = 7 + 3x \quad = \quad \boxed{3x - 2y = 7}$$

0.2.4. Casos Particulares: Rectas Horizontales y Verticales

Dos casos sencillos de la ecuación general son las *rectas horizontales* ($a = 0$) con ecuación $y = k$, y las *rectas verticales* ($b = 0$) con ecuación $x = k$, siendo k una constante real.

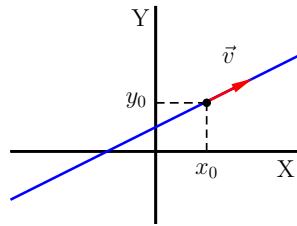


⁷Típicamente se suele denominar *ecuación cartesiana* cuando aparecen las variables a la izquierda del signo igual, y la constante a la derecha. Cuando todos los términos aparecen despejados a la derecha e igualados a cero se suele denominar *ecuación general*.

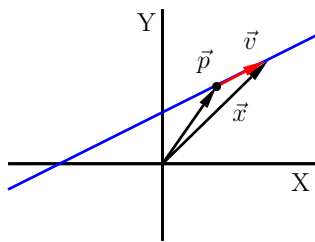
⁸En la ecuación general el valor b puede ser cero dando lugar a una recta vertical con pendiente infinita.

0.2.5. Forma Vectorial y Paramétrica de la Recta

Una forma alternativa y muy útil de describir una recta es representar la dirección de la misma mediante un *vector*⁹, de forma que por un punto del *plano euclídeo* $P = (x_0, y_0)$ y con la dirección dada por el vector $\vec{v} = [v_1, v_2]$ pasa una única recta.



Teniendo esto en cuenta, la *ecuación vectorial de la recta* tiene la siguiente forma:



$$\vec{x} = \vec{p} + t\vec{v} \quad ; \quad t \in \mathbb{R} \quad (7)$$

Los términos de esta ecuación son:

- $\vec{x} = [x, y]$: vector¹⁰ que representa a todos y cada uno de los puntos de la recta.
- $\vec{p} = [x_0, y_0]$: vector que posee su extremo en el punto $P = (x_0, y_0)$.
- $\vec{v}$: vector en la dirección de la recta, conocido con el nombre de *vector director*.
- t : es un **parámetro**¹¹ que puede tomar cualquier valor real. Para cada valor del parámetro t se obtiene un punto de la recta.

El parámetro t modifica el tamaño del vector director $\vec{v}$, alargándolo si $t > 1$, encogiéndolo si $0 < t < 1$ y cambiándolo de dirección si $t < 0$. De esta forma el vector resultante $\vec{x}$ alcanza todos los puntos de la recta.

Teniendo en cuenta que los vectores del *plano euclídeo* tienen dos componentes, la ecuación vectorial de la recta, **Ecuación (7)**, se puede escribir también como¹²:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + t \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad (8)$$

Esta ecuación se puede partir en dos, dando lugar a la conocida como *forma paramétrica de la recta*:

$$\begin{cases} x = x_0 + v_1 t \\ y = y_0 + v_2 t \end{cases} \quad (9)$$

Las **formas vectoriales** y **paramétricas** son elementos esenciales del Álgebra Lineal y serán muy utilizadas en los próximos temas.

⁹Ver una introducción a los vectores en el *plano euclídeo* y en el *espacio 3D* en el **Apéndice A.8, página 270** y un estudio más detallado en el **Tema 2, página 63**.

¹⁰Los vectores se van a representar en este libro encerrando su coordenadas entre corchetes cuadrados, $[\]$.

¹¹Un *parámetro* es un término que puede tomar cualquier valor, pero sin ser una *variable* del sistema.

¹²En este caso los vectores se pueden poner indistintamente como *filas* o como *columnas*.

Ejercicio 0.4. Determinar la ecuación vectorial y paramétrica de la recta que pasa por los puntos $P_1 = (1, -2)$ y $P_2 = (-3, 0)$.

Restando las coordenadas de los puntos dados se obtiene un vector con la dirección de la recta (vector director).

$$\vec{v} = \overrightarrow{P_1 P_2} = [-3 - 1, 0 - (-2)] = [-4, -2]$$

Como vector director $\vec{v}$ podemos tomar o bien éste que acabamos de calcular o cualquiera de sus múltiplos no nulos, como por ejemplo: $[-2, -1]$ ó $[2, 1]$ ya que todos tienen la misma dirección¹³.

Con un vector director y un punto cualquiera de la recta, por ejemplo P_1 ó P_2 , podemos obtener una¹⁴ *ecuación vectorial* de esta recta:

$$\vec{x} = \vec{x}_0 + t\vec{v} \quad ; \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Podemos comprobar como dando los valores $t = 0$ y $t = -2$ se obtienen respectivamente los puntos P_1 y P_2 de la recta. Finalmente una¹⁵ *ecuación paramétrica* de esta recta se obtiene partiendo la *forma vectorial* en dos ecuaciones:

$$\begin{cases} x = 1 + 2t \\ y = -2 + t \end{cases}$$

Ejercicio 0.5. Encontrar la ecuación vectorial de la recta $3x + y = 1$.

La forma más sencilla de pasar una *ecuación general* a su *forma paramétrica* es asignando un parámetro a cualquiera de sus variables. Por ejemplo asignando el parámetro t a la variable x y despejando se obtiene:

$$3x + 2y = 1 \quad = \quad \begin{cases} x = t \\ y = 1 - 3t \end{cases}$$

que es una *ecuación paramétrica* de esta recta. Fácilmente a través de ella podemos obtener una *forma vectorial* reuniendo los coeficientes en vectores:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

Ejercicio 0.6. Obtener la *ecuación general* de la recta cuya *forma vectorial* es:

$$[x, y] = [-1, 2] + t[2, -3]$$

Fácilmente podemos obtener su *forma paramétrica*:

$$\begin{cases} x = -1 + 2t \\ y = 2 - 3t \end{cases}$$

Como la *ecuación general* no tiene parámetros, debemos eliminar t de la *forma paramétrica*. Para ello podemos despejar t en cada una de las ecuaciones e igualarlas:

$$\begin{cases} x = -1 + 2t \\ y = 2 - 3t \end{cases} = \begin{cases} t = \frac{x+1}{2} \\ t = \frac{2-y}{3} \end{cases} = \frac{x+1}{2} = \frac{2-y}{3} \quad ; \quad 3x + 3 = 4 - 2y$$

Simplificando y reordenando se obtiene finalmente:

$$3x + 2y = 1$$

¹³Lo importante es la dirección del vector y no su sentido, por ello también podemos utilizar el vector $\overrightarrow{P_2 P_1}$ en lugar de $\overrightarrow{P_1 P_2}$.

¹⁴Hay infinitas formas distintas de representar una única recta en su forma vectorial, dependiendo del punto y vector director de la recta elegidos, pero todas ellas representan la misma recta.

¹⁵Al igual que la formas vectoriales, también hay infinitas formas paramétricas distintas de representar una única recta.

0.2.6. Ecuación Normal de la Recta

Siguiendo con los conceptos vectoriales, la *ecuación general de la recta*¹⁶ se puede interpretar como un determinado *producto escalar*¹⁷:

$$ax + by = [a, b] \cdot [x, y] = \vec{n} \cdot \vec{x} = c \quad = \quad \boxed{\vec{n} \cdot \vec{x} = c}$$

Donde:

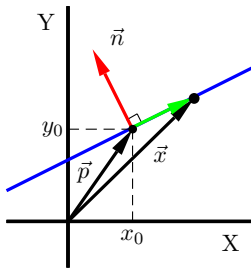
- $\vec{n} = [a, b]$: se denomina *vector normal de la recta* y que tiene por componentes los *coeficientes* de la variables, a y b .
- $\vec{x} = [x, y]$: vector que representa a todos y cada uno los puntos (x, y) de la recta¹⁸.
- c : término independiente (constante) de la ecuación general.

La constante c es un número real y siempre se puede encontrar un determinado punto de la recta $\vec{p} = [x_0, y_0]$ que al multiplicarlo escalarmente por el vector normal $\vec{n}$ se obtenga como resultado c :

$$\vec{n} \cdot \vec{x} = c \quad = \quad \vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{p}$$

Pasando los dos términos a un mismo lado y despejando el vector normal $\vec{n}$ se obtiene la denominada *ecuación normal de la recta*:

$$\vec{n} \cdot \vec{x} - \vec{n} \cdot \vec{p} = 0 \quad = \quad \boxed{\vec{n} \cdot (\vec{x} - \vec{p}) = 0} \quad (10)$$



La interpretación geométrica de esta ecuación vectorial es la siguiente: la diferencia entre un vector $\vec{x} = [x, y]$ con extremo en cualquier punto de la recta y el vector $\vec{p} = [x_0, y_0]$ con extremo en un determinado punto (x_0, y_0) de la recta, da lugar a un nuevo vector $(\vec{x} - \vec{p})$ que tiene la dirección de la recta¹⁹. Según la **Ecuación (10)**, este vector es perpendicular²⁰ al vector $\vec{n}$, luego por tanto $\vec{n}$ es un vector perpendicular (normal) a la recta.

Ejercicio 0.7. Encontrar la ecuación normal de la recta: $x - 2y = 3$.

Podemos reinterpretar esta ecuación en base al *producto escalar* de la siguiente forma:

$$x - y = 1 \quad = \quad [1, -2] \cdot [x, y] = 3$$

Hay infinitos vectores normales a una recta, todos con la misma dirección, uno de ellos es por ejemplo $\vec{n} = [1, -2]$. Tomando un punto cualquiera de la recta, por ejemplo $(3, 0)$ se tiene $\vec{p} = [3, 0]$.

Finalmente una ecuación normal de esta recta es:

$$\boxed{[1, -2] \cdot ([x, y] - [3, 0]) = 0}$$

Desarrollando el producto escalar y simplificando se puede volver a recuperar la ecuación general.

0.2.7. Resumen de Fórmulas: Rectas en el Plano Euclídeo

Ecuación	Fórmula
Punto-pendiente	$y = y_0 + mx$
General	$ax + by = c$
Vectorial - Paramétrica	$\vec{x} = \vec{p} + t\vec{v} \quad ; \quad \begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \end{cases}$
Normal	$\vec{n} \cdot (\vec{x} - \vec{p}) = 0$

¹⁶Ver la **Ecuación 6**, página 5.

¹⁷Ver la **Definición 6.1**, página 163.

¹⁸De la misma forma que se utilizó en la *ecuación vectorial de la recta* en la **Sección 0.2.5**, página 6.

¹⁹Es decir, es un *vector director de la recta*, como se explicó en **Sección 0.2.5**, página 6.

²⁰Ya que el producto escalar de ambos es cero.

0.2.8. Conversión entre las Formas de una Recta

General Punto-Pendiente

$$y = y_0 + mx \quad = \quad \boxed{mx - y = y_0} \quad \begin{cases} a = m \\ b = -1 \\ c = y_0 \end{cases}$$

$$ax + by = c \quad \stackrel{b \neq 0}{=} \quad \boxed{y = \frac{c}{b} - \frac{a}{b}x} \quad \begin{cases} y_0 = c/b \\ m = -a/b \end{cases}$$

Si en la forma general $b = 0$, la recta es vertical y la pendiente m es infinita, por lo que este tipo de rectas no pueden expresarse en la *forma punto-pendiente*.

General Vectorial - Paramétrica

$$ax + by = c \quad = \quad \begin{cases} x = t \\ y = c - at \end{cases} \quad = \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ c \end{bmatrix} + t \begin{bmatrix} 1 \\ -a \end{bmatrix}$$

$$\begin{cases} x = x_0 + v_1 t \\ y = y_0 + v_2 t \end{cases} \quad = \quad \begin{cases} t = \frac{x - x_0}{v_1} \\ t = \frac{y - y_0}{v_2} \end{cases} \quad = \quad \frac{x - x_0}{v_1} = \frac{y - y_0}{v_2}$$

Esta ecuación se denomina **ecuación continua de la recta** en el *plano euclídeo*. Si $\vec{v}_1 = 0$ ²¹ la ecuación general tiene la forma:

$$\frac{v_2}{v_1}x - y = \frac{v_2}{v_1}x_0 - y_0$$

General Normal

$$\vec{n} \cdot (\vec{x} - \vec{p}) = [n_1, n_2] \cdot ([x, y] - [p_1, p_2]) = 0 = [n_1, n_2] \cdot [x - p_1, y - p_2] = 0 = n_1(x - p_1) + n_2(y - p_2) = 0$$

$$= \boxed{n_1x + n_2y = n_1p_1 + n_2p_2} \quad \begin{cases} a = n_1 \\ b = n_2 \\ c = n_1p_1 + n_2p_2 \end{cases}$$

$$ax + by = c \quad = \quad \boxed{[a, b] \cdot ([x, y] - [0, \frac{c}{b}]) = 0} \quad \begin{cases} \vec{n} = [a, b] \\ \vec{p} = [0, c/b] \end{cases}$$

Si en la forma general $b = 0$, la recta es vertical y la ecuación normal se puede expresar como:

$$[a, 0] \cdot ([x, y] - [a/c, 0]) = 0$$

Ejercicio 0.8. Encontrar las distintas formas de expresar matemáticamente la recta que pasa por los puntos $P_1 = (1, -2)$ y $P_2 = (0, 3)$ del *plano euclídeo* y especificar sus valores característicos: pendiente, vector director, vector normal y puntos de corte con los ejes X e Y .

Aplicando la ecuación de dos puntos podemos obtener la ecuación punto-pendiente de la recta.

$$\frac{y - (-2)}{x - 1} = \frac{3 - (-2)}{0 - 1} \quad ; \quad y = -2 - 5(x - 1) \quad ; \quad \boxed{y = 3 - 5x}$$

Con esta ecuación podemos determinar fácilmente la pendiente y los puntos de corte con el eje X (valor de la x para $y = 0$) y con el eje Y (valor de la y para $x = 0$):

$$\boxed{\text{pendiente: } m = -5} \quad ; \quad \boxed{\text{Corte eje } X: \left(\frac{3}{5}, 0\right)} \quad ; \quad \boxed{\text{Corte eje } Y: (0, 3)}$$

²¹Si $v_1 = 0$, la recta es vertical y la ecuación general es $x = x_0$.

La forma general de la recta se puede expresar como²²:

$$5x + y = 3$$

La forma normal de la recta se puede obtener fácilmente de la ecuación general:

$$[5, 1] \cdot [x, y] = 3 \quad = \quad \text{vector normal: } \vec{n} = [5, 1]$$

Ahora tenemos que encontrar un vector $\vec{p} = [x_0, y_0]$ de forma que el producto escalar $\vec{n} \cdot \vec{p} = 3$. Para ello podemos coger P_1 , P_2 o cualquier otro punto de la recta. Tomando por ejemplo P_2 , $\vec{p} = [0, 3]$, tenemos finalmente una ecuación normal²³:

$$[5, 1] \cdot [x, y] = [5, 1] \cdot [0, 3] \quad ; \quad [5, 1] \cdot ([x, y] - [0, 3]) = 0$$

Para obtener la ecuación vectorial/paramétrica necesitamos un vector director de la recta. Éste se obtiene fácilmente restando vectorialmente dos puntos cualquiera de la recta, por ejemplo tomando P_1 y P_2 :

$$\vec{v} = \vec{p}_1 - \vec{p}_2 = [1, -2] - [0, 3] = [1, -5] \quad ; \quad \text{vector director: } \vec{v} = [1, -5]$$

La forma vectorial se puede obtener con un vector director y un punto cualquiera de la recta, por ejemplo P_1 :

$$[x, y] = [1, -2] + t[1, -5]$$

Por último, una forma paramétrica de la recta se puede obtener partiendo la ecuación vectorial en sus distintas componentes:

$$\begin{cases} x = 1 + t \\ y = -2 - 5t \end{cases} \quad ; \quad t \in \mathbb{R}$$

Ejercicio 0.9. Encontrar en cada caso, la ecuación de la recta en el *plano euclídeo* que pasando por el punto $(1, 2)$ cumple:

- Es paralela a la recta $y = -1 - x$.
- Es perpendicular a la recta $y = 1 + 3x$.

- Si es paralela a esta recta, tiene su misma pendiente, en este caso $m = -1$. La recta buscada tendrá entonces por ecuación: $y = a - x$. El valor de a se puede determinar sabiendo que la recta debe pasar por el punto $(1, 2)$. Sustituyendo podemos obtener la forma punto-pendiente de la recta pedida:

$$2 = a - 1 \quad ; \quad a = 3 \quad = \quad y = 3 - x$$

- En este caso, las rectas deben ser perpendiculares. Teniendo en cuenta que las pendientes m y m' de dos rectas perpendiculares cumplen:

$$m = -\frac{1}{m'}$$

y la pendiente de la recta dada es $m' = 3$, la pendiente de la recta pedida es $m = -1/3$. Como además debe pasar por el punto $(1, 2)$, se tiene finalmente:

$$y = a - \frac{1}{3}x \quad ; \quad 2 = a - \frac{1}{3} \quad ; \quad a = \frac{7}{3} \quad = \quad y = \frac{7}{3} - \frac{x}{3}$$

²²Hay infinitas formas de representar la ecuación general, basta con multiplicar ambas partes de la igualdad por un número distinto de cero.

²³Hay infinitas ecuaciones normales asociadas a una recta, porque existen infinitos vectores normales e infinitos puntos de la misma.

0.3. Ecuaciones del Plano en el Espacio 3D

Un *plano* es un ente bidimensional que posee infinitos puntos e infinitas rectas.

Las formas de representar matemáticamente un plano en el *espacio 3D* son similares a las de las rectas en el *plano euclídeo* vistas en la sección anterior, añadiendo una dimensión²⁴ más (variable z).

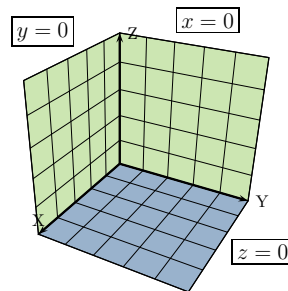
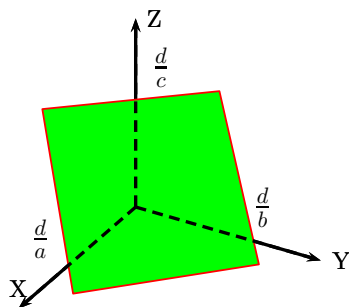
0.3.1. Ecuación General del Plano

La ecuación general de un plano en el *espacio 3D* se puede expresar como:

$$\boxed{ax + by + cz = d} \quad ; \quad a, b, c \in \mathbb{R} \quad \stackrel{c \neq 0}{=} \quad \boxed{z = \alpha x + \beta y + \gamma} \quad \begin{cases} \alpha = -a/c \\ \beta = -b/c \\ \gamma = d/c \end{cases} \quad (11)$$

Características de las constantes a , b , c y d :

- Si $a = 0$ el plano corta al eje X , de otra forma el plano es paralelo a dicho eje.
- Si $b = 0$ el plano corta al eje Y , de otra forma el plano es paralelo a dicho eje.
- Si $c = 0$ el plano corta al eje Z , de otra forma el plano es paralelo a dicho eje.
- Si $d = 0$ el plano pasa por el origen de coordenadas $O(0, 0, 0)$.



Ejercicio 0.10. Encontrar la ecuación general del plano que pasa por los puntos $P_1 = (1, 0, 1)$, $P_2 = (-1, 1, -1)$ y $P_3 = (0, 0, 1)$ del *espacio 3D*.

Vamos a suponer²⁵ que $c = 0$ y determinaremos las constantes α , β y γ de la **Ecuación (11)**.

$$\alpha x + \beta y + \gamma = z$$

Introduciendo los valores de los puntos P_1 , P_2 y P_3 se obtiene el siguiente sistema de ecuaciones:

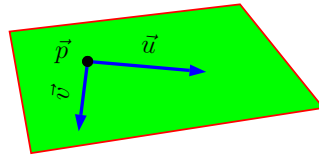
$$\begin{cases} \alpha + \gamma = 1 \\ -\alpha + \beta + \gamma = -1 \\ \gamma = 1 \end{cases} = \begin{cases} \alpha = 0 \\ \beta = -2 \\ \gamma = 1 \end{cases} = -2y - z = -1 = \boxed{2y + z = 1}$$

²⁴El concepto de *dimensión* será estudiado en más detalle en las siguientes temas.

²⁵En caso de que este planteamiento no tenga solución, será debido a que $c = 0$ y tendremos que buscar otra constante a ó b que no sea nula. Tener en cuenta que en todo plano al menos una constante debe ser distinta de cero.

0.3.2. Forma Vectorial y Paramétrica del Plano

Una forma alternativa de describir un plano es a partir de un punto²⁶ $\vec{p} = [x_0, y_0, z_0]$ y dos vectores $\vec{u} = [u_1, u_2, u_3]$ y $\vec{v} = [v_1, v_2, v_3]$, contenidos en el mismo.



El conjunto de todos los puntos del plano (x, y, z) viene dado por las siguientes ecuaciones, donde $t, s \in \mathbb{R}$ son parámetros²⁷.

$$\text{F. Vectorial: } \begin{cases} \vec{x} = \vec{p} + t\vec{u} + s\vec{v} \\ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + t \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + s \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \end{cases}$$

F. Paramétrica:

$$\begin{cases} x = x_0 + tu_1 + sv_1 \\ y = y_0 + tu_2 + sv_2 \\ z = z_0 + tu_3 + sv_3 \end{cases}$$

Notar que la forma vectorial del plano en el *espacio 3D* tiene vectores de tres componentes y posee dos parámetros²⁸.

Ejercicio 0.11. Encontrar las ecuaciones paramétrica y vectorial del plano que pasa por los puntos: $P_1 = (1, 1, 1)$, $P_2 = (0, 0, 0)$ y $P_3 = (2, 0, 1)$ del *espacio 3D*.

Para construir la forma vectorial del plano necesitamos un punto y dos vectores contenidos en el mismo.

$$\vec{x} = \vec{p} + t\vec{u} + s\vec{v}$$

Como punto $\vec{p}$ podemos tomar cualquiera de los dados en el enunciado, (dispuestos como vectores en lugar de puntos), por ejemplo $\vec{p} = [1, 1, 1]$ y para obtener los vectores $\vec{u}$ y $\vec{v}$ podemos restar dos a dos dichos puntos:

$$\vec{v} = \overrightarrow{P_3P_1} = [1, 1, 1] - [0, 0, 0] = [1, 1, 1] \quad ; \quad \vec{u} = \overrightarrow{P_3P_2} = [2, 0, 1] - [0, 0, 0] = [2, 0, 1]$$

Por tanto, unas ecuaciones vectorial y paramétrica de este plano son:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{cases} x = 1 + t + 2s \\ y = 1 + t \\ z = 1 + t + s \end{cases}$$

Ejercicio 0.12. Encontrar la ecuación vectorial del plano $2x - y = 3$ en el *espacio 3D*.

Vamos a asignar dos parámetros t y s a dos variables. Una puede ser la variable $x = t$ y la otra obligatoriamente tiene que ser la variable z que no aparece en la ecuación y por tanto puede tomar cualquier valor $z = s$. Entonces unas ecuaciones paramétrica y vectorial del plano son:

$$2x - y = 3 = \begin{cases} x = t \\ y = -3 + 2t \\ z = s \end{cases} ; \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

²⁶El punto está expresado en forma vectorial, de igual forma que se realizó en la forma vectorial de una recta en el plano, (ver la Sección 0.2.5, página 6).

²⁷Ver la Sección 0.2.5, página 6.

²⁸A diferencia de la forma vectorial de la recta en el *plano euclídeo* donde los vectores tienen dos componentes y solo posee un parámetro, (ver la Sección 0.2.5, página 6).

Ejercicio 0.13. Encontrar la ecuación general del plano cuya forma paramétrica en el *espacio 3D* es:

$$\begin{cases} x = 1 + t - s \\ y = -1 - t - 2s \\ z = -t - s \end{cases}$$

Tenemos que eliminar los parámetros t y s , por ejemplo despejando de la tercera ecuación y sustituyendo en las dos primeras se tiene:

$$t = -z - s \quad = \quad \begin{cases} x = 1 + (-z - s) - s = 1 - z - 2s \\ y = -1 - (-z - s) - 2s = -1 + z - s \end{cases} \quad s = -1 + z - y$$

Sustituyendo el valor de s en la primera ecuación de la x despejada se tiene:

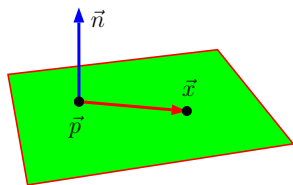
$$x = 1 - z - 2(-1 + z - y) = 3 - 3z + 2y \quad = \quad \boxed{x - 2y + 3z = 3}$$

0.3.3. Forma Normal del Plano

La *forma normal* de un plano se puede obtener de la *forma general* del mismo modo que se realizó con la recta en el *plano euclídeo*²⁹.

$$ax + by + cz = d \quad = \quad \begin{cases} \vec{n} = [a, b, c] \\ \vec{x} = [x, y, z] \end{cases} \quad = \quad \vec{n} \cdot \vec{x} = d \quad = \quad \vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{p} \quad = \quad \boxed{\vec{n} \cdot (\vec{x} - \vec{p}) = 0}$$

Esta ecuación recibe el nombre de *ecuación normal del plano* y $\vec{n}$ es un *vector normal* al plano, que como se deduce de la ecuación normal, es *perpendicular* al mismo.



$$\begin{cases} \vec{x} = [x, y, z] & \text{Todos los puntos del plano (variable)} \\ \vec{n} = [a, b, c] & \text{Un vector normal al plano (constante)} \\ \vec{p} = [x_0, y_0, z_0] & \text{Punto cualquiera del plano (constante)} \end{cases}$$

Notar que la *ecuación normal de un plano* en el *espacio 3D* es exactamente igual que la ecuación normal de una recta en el *plano euclídeo*, aunque los vectores tienen tres componentes en lugar de dos.

Ejercicio 0.14. Encontrar la ecuación normal del plano $x + y - z = 2$ en el *espacio 3D*.

Podemos reinterpretar esta ecuación en base al producto escalar de la siguiente forma:

$$x + y - z = 2 \quad = \quad [1, 1, -1] \cdot [x, y, z] = 2$$

Un vector normal del plano es por tanto $\vec{n} = [1, 1, -1]$. Tomando un punto cualquiera del plano, por ejemplo $(1, 1, 0)$ se tiene $\vec{p} = [1, 1, 0]$.

Finalmente una ecuación normal de este plano es:

$$\boxed{[1, 1, -1] \cdot ([x, y, z] - [1, 1, 0]) = 0}$$

Desarrollando el producto escalar se vuelve a recuperar la ecuación general.

²⁹Ver la [Sección 0.2.6](#), página 8.

Ejercicio 0.15. Encontrar la ecuación general del plano que es perpendicular al plano $x - y + 2z = 1$ y pasa por el punto $P = (1, 1, 1)$ del *espacio 3D*.

El vector normal $\vec{n} = [n_x, n_y, n_z]$ del plano buscado tiene que ser perpendicular al vector normal $\vec{n}_2 = [1, -1, 2]$ del plano dado, por tanto:

$$\vec{n} \cdot \vec{n}_2 = 0 \quad = \quad n_x - n_y + 2n_z = 0$$

Fácilmente encontramos valores que cumplen esta ecuación, como por ejemplo: $\vec{n} = [-1, 1, 1]$. Teniendo el vector normal y un punto del plano podemos encontrar su ecuación normal:

$$[-1, 1, 1] \cdot ([x, y, z] - [1, 1, 1]) = 0 \quad ; \quad [-1, 1, 1] \cdot [x - 1, y - 1, z - 1] = 0 \quad ; \quad -x + 1 + y - 1 + z - 1 = 0$$

$$\boxed{-x + y + z = 1}$$

0.3.4. Resumen Fórmulas: Planos en el Espacio 3D

Ecuación	Fórmula
General	$ax + by + cz = d$
Vectorial - Paramétrica	$\vec{x} = \vec{p} + t\vec{u} + s\vec{v} \quad ; \quad \begin{cases} x = x_0 + tu_x + sv_x \\ y = y_0 + tu_y + sv_y \\ z = z_0 + tu_z + sv_z \end{cases}$
Normal	$\vec{n} \cdot (\vec{x} - \vec{p}) = 0$

0.3.5. Conversión entre Formas de un Plano

General Vectorial - Paramétrica

- Para pasar de la forma general de un plano a su forma vectorial/paramétrica, se necesitan dos vectores contenidos en el plano y un punto del mismo. Éstos se podrían obtener con tres puntos del plano que no estén alineados³⁰ y restándolos dos a dos para obtener los vectores. Sin embargo es mucho más sencillo transformar dos variables en parámetros y expresar la tercera variable en función de los parámetros³¹.

$$ax + by + cz = d \quad = \quad \begin{cases} x = t \\ y = s \\ z = \frac{1}{c}(d - at - bs) \end{cases} \quad = \quad \boxed{[x, y, z] = \left[0, 0, \frac{d}{c}\right] + t \left[1, 0, \frac{-a}{c}\right] + s \left[0, 1, \frac{-b}{c}\right]}$$

- Para pasar de la forma vectorial/paramétrica a la forma general hay que eliminar los parámetros t y s dejando una única ecuación. Esto se puede hacer mediante diferentes métodos³²: por reducción, por sustitución o igualación o bien de forma matricial calculando el siguiente determinante³³:

$$\begin{cases} x = x_0 + tv_x + su_x \\ y = y_0 + tv_y + su_y \\ z = z_0 + tv_z + su_z \end{cases} \quad = \quad \begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ v_x & v_y & v_z \\ u_x & u_y & u_z \end{vmatrix} = 0 \quad = \quad \boxed{ax + by + cz = d}$$

³⁰Para obtener un punto del plano a través de la ecuación general, basta con fijar dos variables, (por ejemplo $x = 0$ e $y = 0$), y despejar el valor de la tercera variable de la ecuación general.

³¹La tercera variable debe tener un coeficiente distinto de cero, para que al despejar no quede una división por cero.

³²Estos métodos son prácticos cuando la ecuación vectorial/paramétrica es simple, es decir, con muchos ceros.

³³Los tres vectores que se utilizan en el determinante son *linealmente dependientes* por estar contenidos en el plano y por tanto su determinante es cero. Esto se estudiará con más detalle en los próximos temas.

General Normal

- Para pasar de la forma general a la forma normal se necesita el vector normal del plano, formado por los coeficientes de las variables, y un punto cualquiera del plano $P = [x_0, y_0, z_0]$.

$$\begin{aligned} ax + by + cz = d &= [a, b, c] \cdot [x, y, z] = d = [a, b, c] \cdot [x_0, y_0, z_0] \\ &= [a, b, c] \cdot ([x, y, z] - [x_0, y_0, z_0]) = 0 \end{aligned}$$

- Para pasar de la forma normal a la forma general basta con desarrollar el producto escalar.

$$\begin{aligned} [a, b, c] \cdot ([x, y, z] - [x_0, y_0, z_0]) = 0 &= [a, b, c] \cdot [x - x_0, y - y_0, z - z_0] = 0 \\ &= a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 = ax + by + cz = ax_0 + by_0 + cz_0 = d \end{aligned}$$

Ejercicio 0.16. Encontrar el plano que pasa por los puntos $P_1 = (1, 1, -1)$, $P_2 = (0, 1, 0)$ y $P_3 = (1, -1, 0)$ del *espacio 3D*, en las formas: general, vectorial, paramétrica y normal.

Dados tres puntos no alineados se pueden obtener fácilmente dos vectores contenidos en el plano que forman:

$$\vec{v}_1 = \vec{p}_2 - \vec{p}_1 = [0, 1, 0] - [1, 1, -1] = [-1, 0, 1] \quad ; \quad \vec{v}_2 = \vec{p}_3 - \vec{p}_1 = [1, -1, 0] - [1, 1, -1] = [0, -2, 1]$$

La forma vectorial y paramétrica se pueden expresar entonces como³⁴:

$$[x, y, z] = [0, 1, 0] + t[-1, 0, 1] + s[0, -2, 1] \quad ; \quad \begin{cases} x = -t \\ y = 1 - 2s \\ z = t + s \end{cases}$$

La forma general se obtiene eliminando los parámetros t y s de la forma paramétrica. En este caso resulta muy sencillo combinando adecuadamente las tres ecuaciones. Por ejemplo, se puede sumar la primera y tercera ecuación para eliminar t y despejando s de la segunda ecuación se sustituye su valor.

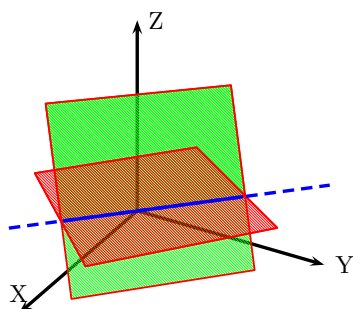
$$\begin{cases} x = -t \\ y = 1 - 2s \\ z = t + s \end{cases} = \begin{cases} x + z = s \\ s = \frac{1}{2}(1 - y) \end{cases} = x + z = \frac{1}{2}(1 - y) = 2x + y + 2z = 1$$

La ecuación normal se puede obtener fácilmente de la ecuación general, obteniendo el vector normal de los coeficientes de las variables y tomando un punto cualquiera del plano, como por ejemplo $P = (0, 1, 0)$.

$$2x + y + 2z = 1 = [2, 1, 2] \cdot [x, y, z] = 1 = [2, 1, 2] \cdot [0, 1, 0] = [2, 1, 2] \cdot ([x, y, z] - [0, 1, 0]) = 0$$

0.4. Ecuaciones de la Recta en el Espacio 3D

0.4.1. Forma General de la Recta



La forma *general* de una recta en el *espacio 3D* se representa mediante las formas *generales* de dos planos no paralelos, quedando definida la recta mediante la intersección de ambos.

$$\begin{cases} a_1x + b_1y + c_1z = d_1 & - & \text{Plano 1} \\ a_2x + b_2y + c_2z = d_2 & - & \text{Plano 2} \end{cases}$$

Esto forma un sistema de dos ecuaciones con tres incógnitas, cuyas infinitas soluciones son los puntos de la recta.

³⁴Hay infinitas formas vectoriales y paramétricas de un plano, en función de los vectores y puntos que se tomen.

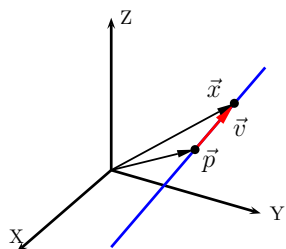
0.4.2. Forma Vectorial-Paramétrica de la Recta

Una recta en el *espacio 3D* queda perfectamente definida con un punto y un vector director³⁵.

F. Vectorial:
$$\begin{cases} \vec{x} = \vec{p} + t\vec{v} \\ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + t \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \end{cases}$$

F. Paramétrica:

$$\begin{cases} x = x_0 + tv_1 \\ y = y_0 + tv_2 \\ z = z_0 + tv_3 \end{cases}$$



$$\begin{cases} \vec{x} = [x, y, z] & \text{Todos los puntos de la recta (variable)} \\ \vec{v} = [v_1, v_2, v_3] & \text{Un vector director de la recta (constante)} \\ \vec{p} = [x_0, y_0, z_0] & \text{Un punto cualquiera de la recta (constante)} \end{cases}$$

Ejercicio 0.17. Obtener la ecuación de la recta en el espacio 3D incluida en el plano $z = 0$, con dirección perpendicular a $\vec{u} = [1, 1, 1]$ y que pasa por el punto $P(1, 1, 0)$.

Como no nos indica nada sobre la forma de la ecuación, la vamos a expresar por ejemplo en forma vectorial:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$$

debiendo encontrar los vectores $[x_0, y_0, z_0]$ y $[v_x, v_y, v_z]$. Como debe pasar por el punto $P(1, 1, 0)$, se tiene:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$$

Como debe ser perpendicular, (ver la [Definición A.10.1, página 274](#)), a $\vec{u} = [1, 1, 1]$:

$$[v_x, v_y, v_z] \cdot [1, 1, 1] = 0 \quad = \quad v_x + v_y + v_z = 0$$

Como debe estar contenida en el plano $z = 0$, el vector director de la recta $[v_x, v_y, v_z]$ debe estar contenido en dicho plano y por tanto $v_z = 0$. Entonces $v_x = -v_y$ y tomando por ejemplo $v_x = 1$ (vale cualquier valor no nulo) se tiene finalmente:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

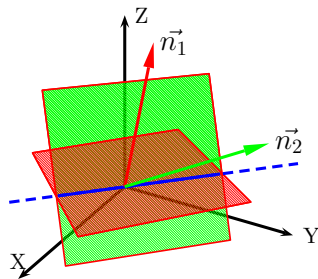
³⁵De igual forma que una recta en el *plano euclídeo* (ver [Sección 0.2.5, página 6](#)), aunque con vectores de tres componentes en lugar de dos.

0.4.3. Forma Normal de la Recta

Vendrá dada por la intersección de dos planos expresados en su forma *normal*.

$$\left\{ \begin{array}{l} \vec{n}_1 \cdot (\vec{x} - \vec{p}_1) = 0 \\ \vec{n}_2 \cdot (\vec{x} - \vec{p}_2) = 0 \end{array} \right.$$

$$\left\{ \begin{array}{ll} \vec{x} = [x, y, z] & \text{Todos los puntos de la recta (variable)} \\ \vec{n}_1 = [n_1, n_2, n_3] & \text{Un vector normal al plano 1 (constante)} \\ \vec{n}_2 = [n'_1, n'_2, n'_3] & \text{Un vector normal al plano 2 (constante)} \\ \vec{p}_1 = [x_0, y_0, z_0] & \text{Un punto cualquiera del plano 1 (constante)} \\ \vec{p}_2 = [x'_0, y'_0, z'_0] & \text{Un punto cualquiera del plano 2 (constante)} \end{array} \right.$$



0.4.4. Resumen Fórmulas: Rectas en el Espacio 3D

Ecuación	Fórmula
General	$\left\{ \begin{array}{l} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{array} \right.$
Vectorial - Paramétrica	$\vec{x} = \vec{p} + t\vec{v} \quad ; \quad \left\{ \begin{array}{l} x = x_0 + tv_x \\ y = y_0 + tv_y \\ z = z_0 + tv_z \end{array} \right.$
Normal	$\left\{ \begin{array}{l} \vec{n}_1 \cdot (\vec{x} - \vec{p}_1) = 0 \\ \vec{n}_2 \cdot (\vec{x} - \vec{p}_2) = 0 \end{array} \right.$

0.4.5. Conversión entre Formas

General Vectorial - Paramétrica

- Para pasar de la forma *general* a la forma *vectorial/paramétrica* lo más sencillo es encontrar dos puntos que pertenezcan a la recta y determinar con ellos un vector director.

$$\left\{ \begin{array}{l} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{array} \right. = \left\{ \begin{array}{l} P_1 = (x_1, y_1, z_1) \\ P_2 = (x_2, y_2, z_2) \end{array} \right. = \vec{v} = [x_2, y_2, z_2] - [x_1, y_1, z_1] = [v_x, v_y, v_z]$$

$$[x, y, z] = [x_1, y_1, z_1] + t[v_x, v_y, v_z] \quad ; \quad \left\{ \begin{array}{l} x = x_1 + tv_x \\ y = y_1 + tv_y \\ z = z_1 + tv_z \end{array} \right.$$

- Para pasar de la forma *vectorial/paramétrica* a la forma *general*, basta con eliminar el parámetro t , y obtener entonces un sistema de dos ecuaciones.

$$\left\{ \begin{array}{l} x = x_1 + tv_x \\ y = y_1 + tv_y \\ z = z_1 + tv_z \end{array} \right. = \frac{x - x_1}{v_x} = \frac{y - y_1}{v_y} = \frac{z - z_1}{v_z}$$

Esta última ecuación se la conoce con el nombre de **ecuación continua de la recta**. Agrupando las igualdades de dos en dos se obtiene una forma general de la recta.

$$\begin{cases} v_y x - v_x y = v_y x_1 - v_x y_1 \\ v_z y - v_y z = v_z y_1 - v_y z_1 \end{cases}$$

General Normal

- Para pasar de la forma *general* a la forma *normal*, basta con pasar cada una de las ecuaciones que forman la *general* a la forma *normal*³⁶.

$$\begin{cases} a_1 x + b_1 y + c_1 z = d_1 \\ a_2 x + b_2 y + c_2 z = d_2 \end{cases} = \begin{cases} [a_1, b_1, c_1] \cdot ([x, y, z] - [x_1, y_1, z_1]) = 0 \\ [a_2, b_2, c_2] \cdot ([x, y, z] - [x_2, y_2, z_2]) = 0 \end{cases}$$

- Para pasar de la forma *normal* a la forma *general*, basta con desarrollar cada uno de los productos escalares de las ecuaciones que forman la *normal*.

$$\begin{cases} [v_x, v_y, v_z] \cdot ([x, y, z] - [x_1, y_1, z_1]) = 0 \\ [u_x, u_y, u_z] \cdot ([x, y, z] - [x_2, y_2, z_2]) = 0 \end{cases} = \begin{cases} v_x x + v_y y + v_z z = v_x x_1 + v_y y_1 + v_z z_1 \\ u_x x + u_y y + u_z z = u_x x_2 + u_y y_2 + u_z z_2 \end{cases}$$

Ejercicio 0.18. Encontrar la recta que pasa por los puntos $P_1 = (1, 1, -1)$ y $P_2 = (0, 1, 0)$ en las formas general, vectorial, paramétrica y normal.

Inicialmente lo más sencillo puede resultar calcular un vector director de la recta y obtener las ecuaciones vectorial y paramétrica.

$$\vec{v} = \vec{p}_1 - \vec{p}_2 = [1, 1, -1] - [0, 1, 0] = [1, 0, -1] \quad ; \quad [x, y, z] = [0, 1, 0] + t[1, 0, -1] \quad ; \quad \begin{cases} x = t \\ y = 1 \\ z = -t \end{cases}$$

Con la ecuación paramétrica se puede obtener la forma general eliminando el parámetro t para obtener dos ecuaciones. En este caso particular se puede sumar la primera y tercera ecuación y coger directamente la segunda.

$$\begin{cases} x + z = 0 \\ y = 1 \end{cases}$$

Por último, la ecuación normal se obtiene transformando la ecuación general en un producto escalar.

$$\begin{cases} x + z = 0 \\ y = 1 \end{cases} \quad ; \quad \begin{cases} [1, 0, 1] \cdot [x, y, z] = 0 \\ [0, 1, 0] \cdot [x, y, z] = 1 \end{cases}$$

Se necesitan dos puntos, uno de cada uno de los planos que forman la forma general. Fácilmente se obtiene $P_1 = (1, 0, -1)$ del primero y $P_2 = (0, 1, 0)$ del segundo. Finalmente:

$$\begin{cases} [1, 0, 1] \cdot ([x, y, z] - [1, 0, -1]) = 0 \\ [0, 1, 0] \cdot ([x, y, z] - [0, 1, 0]) = 0 \end{cases}$$

³⁶Ver la Sección 0.3.5, página 14.

0.5. Geometría de los Sistemas de Ecuaciones en el Plano y en el Espacio 3D

En esta sección se van a analizar algunos casos sencillos de disposiciones geométricas de rectas y planos en el *plano euclídeo* y en el *espacio 3D*.

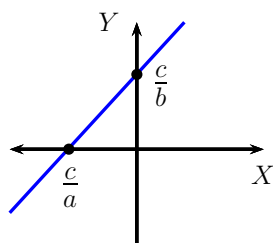
Desde un punto de vista matemático, el conjunto de rectas y planos en un sistema de referencia cartesiano da lugar a un *sistema de ecuaciones lineales*³⁷.

Para simplificar³⁸ vamos a identificar el *plano euclídeo* como $\mathbb{R}^2$ y el *espacio 3D* como $\mathbb{R}^3$.

0.5.1. Sistemas de Una Ecuación

- $ax + by = c$ en $\mathbb{R}^2$

Este es un sistema de ecuaciones lineales formado por una única ecuación de hasta³⁹ dos incógnitas (x e y) en el *plano euclídeo* ($\mathbb{R}^2$) y corresponde a la ecuación general de una *recta*⁴⁰.



Las soluciones de este sistema son todos los (infinitos) puntos de la recta, que son los que verifican la ecuación.

La solución se puede expresar utilizando la forma paramétrica de la recta:

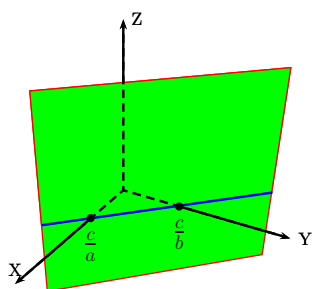
$$\begin{cases} x = t \\ y = \frac{1}{b}(c - at) \end{cases} ; \quad t \in \mathbb{R}$$

Dando valores al parámetro t se obtienen las infinitas soluciones del sistema.

- $ax + by = c$ en $\mathbb{R}^3$

La misma ecuación del caso anterior pero ahora en el *espacio 3D* ($\mathbb{R}^3$) corresponde con un plano. Hay que tener en cuenta que tan importante es la ecuación que aparece, como el espacio ($\mathbb{R}^2$ ó $\mathbb{R}^3$) en el que estamos. En este caso el espacio determina el tipo de elemento que tenemos (en $\mathbb{R}^2$ es una recta y en $\mathbb{R}^3$ es un plano), y la ecuación determina la disposición espacial del correspondiente plano o recta.

En este caso se tiene la ecuación general de un plano en el que falta la variable z , y por tanto corresponderá con un plano paralelo al eje z ⁴¹.



Las soluciones de este sistema corresponden con los infinitos puntos del plano, que son los que verifican la ecuación.

Al igual que antes, la solución se puede expresar utilizando la forma paramétrica del plano⁴²:

$$\begin{cases} x = t \\ y = \frac{1}{b}(c - at) \\ z = s \end{cases} ; \quad t, s \in \mathbb{R}$$

Dando valores a los parámetros t y s se obtienen las infinitas soluciones del sistema.

- $ax + by + cz = d$ en $\mathbb{R}^3$

De forma general, un sistema de ecuaciones lineales en $\mathbb{R}^3$ formado por una única ecuación de hasta 3 incógnitas, corresponde siempre a un plano.

³⁷Ver el Tema 1, página 27.

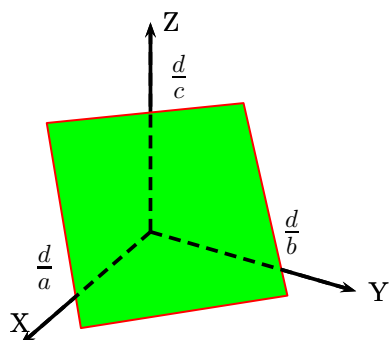
³⁸En la Sección 1.1.3, página 32 se definirá $\mathbb{R}^n$ de forma apropiada.

³⁹En la ecuación podría faltar una de las dos incógnitas, x o y , correspondiente a una recta horizontal o vertical, (ver la Sección 0.2.4, página 5).

⁴⁰Ver la Sección 0.2.3, página 5.

⁴¹Ver la Sección 0.3.1, página 11.

⁴²Ver la Sección 0.3.2, página 12.



La solución se puede expresar utilizando la forma paramétrica del plano, por ejemplo⁴³:

$$\begin{cases} x = t \\ y = s \\ z = \frac{1}{c}(d - at - bs) \end{cases} ; \quad t, s \in \mathbb{R}$$

Dando valores a los parámetros t y s se obtienen las infinitas soluciones del sistema.

Ejercicio 0.19. Encontrar una forma paramétrica y otra vectorial del plano $x + 2y = 2$ en $\mathbb{R}^3$.

Como es un plano, su forma paramétrica tiene que tener dos parámetros. Como en este caso falta la variable z en la ecuación dada, uno de los parámetros será esta variable. El otro parámetro se obtiene despejando una de las dos variables x ó y de la ecuación general y designando la otra variable como un parámetro.

Despejando por ejemplo la variable x se tiene la siguiente forma paramétrica del plano:

$$\begin{cases} x = 2 - 2y = 2 - 2t \\ y = t \\ z = s \end{cases} = \begin{cases} x = 2 - 2t \\ y = t \\ z = s \end{cases}$$

Una forma vectorial se obtiene separando en vectores:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

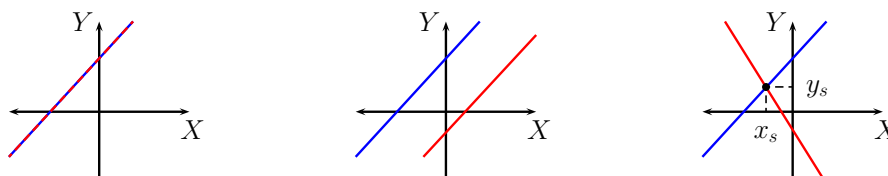
0.5.2. Sistemas de Dos Ecuaciones

$$\bullet \begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases} \quad \text{en } \mathbb{R}^2$$

El sistema de ecuaciones lineales en $\mathbb{R}^2$ formado por dos ecuaciones de hasta 2 incógnitas cada una, corresponde a dos rectas en el *plano euclídeo* que geoméricamente pueden:

- coincidir : si se cortan en infinitos puntos.
- ser paralelas : no se cortan en ningún punto.
- ser secantes: se cortan en un único punto.

Estos tres casos se muestran gráficamente en la siguiente figura:



⁴³Recordar que hay infinitas formas de expresar un plano en su forma vectorial/paramétrica, dependiendo del punto y vectores del plano que se tomen. Por ejemplo otra posible solución a este caso sería:

$$\begin{cases} x = 1/a(d - bt - cs) \\ y = t \\ z = s \end{cases} ; \quad t, s \in \mathbb{R}$$

Por tanto, el sistema puede tener una **única solución** (rectas secantes), **infinitas soluciones** (rectas coincidentes) o **ninguna solución** (rectas distintas paralelas⁴⁴).

La solución se puede obtener, por ejemplo, utilizando uno de los métodos simples de resolución de sistemas⁴⁵.

Ejercicio 0.20. Determinar si los siguientes sistemas de ecuaciones en $\mathbb{R}^2$ corresponden, desde un punto de vista geométrico, con rectas secantes, paralelas o coincidentes.

a) $\begin{cases} 2x + 3y = 4 \\ 4x + 6y = 8 \end{cases}$

b) $\begin{cases} 2x + 3y = 4 \\ 3x + 6y = 5 \end{cases}$

c) $\begin{cases} 2x + 3y = 4 \\ 4x + 6y = 2 \end{cases}$

a) Fijándonos detenidamente podemos comprobar como ambas ecuaciones son *equivalentes*⁴⁶

$$\begin{cases} 2x + 3y = 4 \\ 4x + 6y = 8 \end{cases} \quad - \quad 2(2x + 3y) = 2 * 4 = 8 \quad \Rightarrow \quad 4x + 6y = 8$$

Ya que la primera ecuación se puede obtener multiplicando por 2 la segunda ecuación.

Por tanto se trata de **rectas coincidentes**⁴⁷. Este sistema tiene infinitas soluciones correspondientes a los infinitos puntos de la recta.

b) Fácilmente comprobamos que no se trata de ecuaciones equivalentes⁴⁸, por tanto no son rectas coincidentes. Para determinar si son secantes o paralelas podemos calcular las pendientes de cada una de las rectas:

$$m_1 = \frac{-2}{3} \quad ; \quad m_2 = \frac{-3}{6} = \frac{-1}{2}$$

Como son distintas, se trata de **rectas secantes**. La solución del sistema es un único punto en el *plano euclídeo*, (punto de corte entre las dos rectas).

c) Este caso tampoco corresponde con rectas coincidentes. Calculando las pendientes:

$$m_1 = \frac{-2}{3} \quad ; \quad m_2 = \frac{-4}{6} = \frac{-2}{3}$$

Comprobamos que son iguales y por tanto se trata de **rectas paralelas**. Este sistema no tendrá solución, por no existir ningún punto de corte entre ambas.

$$\bullet \begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases} \quad \text{en } \mathbb{R}^3$$

El mismo sistema de antes, (dos ecuaciones de hasta dos incógnitas) en $\mathbb{R}^3$ es un conjunto de dos **planos**⁴⁹ perpendiculares al eje z , que pueden ser:

- *Coincidentes*⁵⁰ : infinitas soluciones correspondientes a los puntos del único plano que definen.
- *Paralelos*⁵¹ : sin solución.
- *Secantes*⁵² : infinitas soluciones correspondientes a los puntos de la recta de corte.

⁴⁴Las rectas serán **paralelas** si tienen igual pendiente. La pendiente en la forma general se obtiene como $m = -a/b$, de forma que dos rectas serán paralelas si $a_1/b_1 = a_2/b_2$.

⁴⁵Ver la [Sección 0.6, página 25](#).

⁴⁶Ver la [Sección 1.4.2, página 42](#).

⁴⁷Realmente podemos pensar que se trata de un sistema formado por una única ecuación (distinta).

⁴⁸No existe ningún número que al multiplicar una ecuación la transforme en la otra.

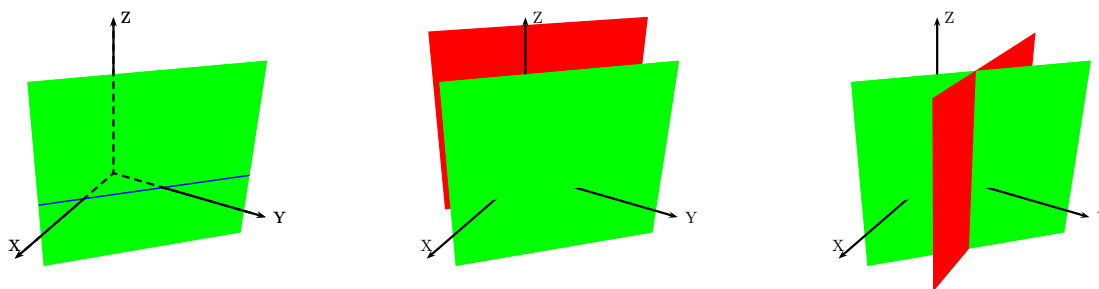
⁴⁹Al estar en $\mathbb{R}^3$ hay tres variables independientes por lo que se trata de un plano y no una recta. Si una o varias variables no aparece en la ecuación, significa que no tienen ninguna restricción y pueden tomar cualquier valor.

⁵⁰Dos planos son coincidentes si aun teniendo aparentemente ecuaciones distintas, representan el mismo conjunto de puntos. En este caso se dice que ambas ecuaciones son *equivalentes* (Ver [Sección 1.4.2, página 42](#)).

⁵¹Dos planos son paralelos si no se intersectan en ningún punto.

⁵²Dos planos son secantes si se cortan en una recta.

En la siguiente figura se muestra gráficamente las tres posibilidades.



Desde el punto de vista de las soluciones, el sistema puede tener:

- Infinitas soluciones: producidas bien por dos planos *coincidentes*⁵³ o bien por dos planos *secantes*⁵⁴.
- Ninguna solución: planos distintos paralelos.

Estos casos se pueden identificar fácilmente calculando los vectores normales de cada uno de los planos, $\vec{n} = [a, b, c]$, y determinando su disposición relativa:

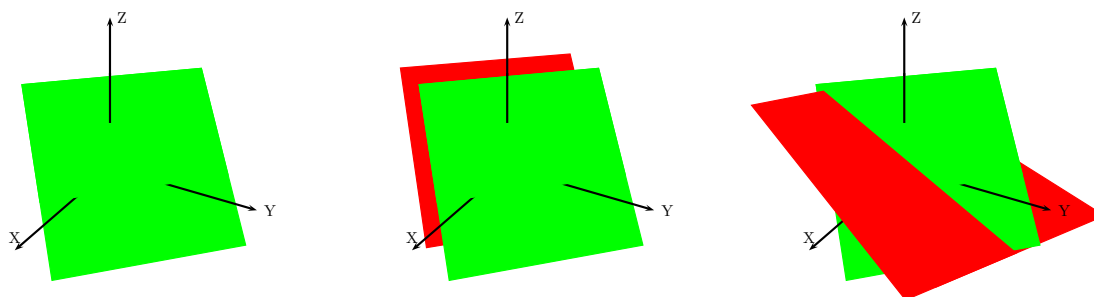
- Si los dos vectores normales $\vec{n}_1$ y $\vec{n}_2$ son paralelos⁵⁵ los planos pueden ser *paralelos* o *coincidentes*.
- Si los vectores normales no son paralelos, los planos serán *secantes*.

Ver en el [Ejemplo 0.21, página 23](#) la resolución de algunos casos prácticos.

$$\bullet \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases} \text{ en } \mathbb{R}^3$$

Un sistema de dos ecuaciones con hasta tres incógnitas en $\mathbb{R}^3$ forma un conjunto de dos planos generales⁵⁶ en el *espacio 3D* que pueden ser:

- *Coincidentes*: infinitas soluciones correspondientes a los puntos del plano.
- *Paralelos*: sin solución.
- *Secantes*: infinitas soluciones correspondientes a los puntos de la recta.



La determinación de estos casos se puede hacer mediante los vectores normales de los planos, de la misma forma que el caso anterior.

⁵³En este caso la solución estará formada por los infinitos puntos del único plano que definen

⁵⁴En este caso los planos se cortan en una recta, cuyos infinitos puntos son la solución del sistema.

⁵⁵Dos vectores $\vec{v}$ y $\vec{u}$ son paralelos si se cumplen $\vec{v} = \alpha\vec{u}$ para cualquier $\alpha \in \mathbb{R}$ no nulo.

⁵⁶No necesariamente paralelos al eje z como el caso anterior.

Ejercicio 0.21. Determinar la disposición relativa de los planos que forman los siguientes sistemas de ecuaciones en $\mathbb{R}^3$:

a) $\begin{cases} x + y - z = 3 \\ 2x + 2y - 2z = 6 \end{cases}$

b) $\begin{cases} x + y - z = 3 \\ 2x - y + z = 5 \end{cases}$

c) $\begin{cases} x + y - z = 3 \\ 2x + 2y - 2z = 5 \end{cases}$

- a) En este caso se puede comprobar claramente como la segunda ecuación es exactamente que la primera multiplicada por 2, y por tanto se trata de dos ecuaciones *equivalentes* que corresponden al caso de planos coincidentes. El sistema en este caso tiene infinitas soluciones que son los puntos del plano resultante.
- b) Los vectores normales de los planos son $\vec{n}_1 = [1, 1, -1]$ y $\vec{n}_2 = [2, -1, 1]$ como claramente no son paralelos, los planos son secantes en una recta y sus puntos son las infinitas soluciones del sistema.
- c) En este caso los vectores normales de los planos son $\vec{n}_1 = [1, 1, -1]$ y $\vec{n}_2 = [2, 2, -2]$ y claramente son paralelos ya que $\vec{n}_2 = 2\vec{n}_1$. Como las ecuaciones no son *equivalentes*, se trata de dos planos paralelos y el sistema por tanto no tiene solución.

0.5.3. Sistemas de Tres Ecuaciones

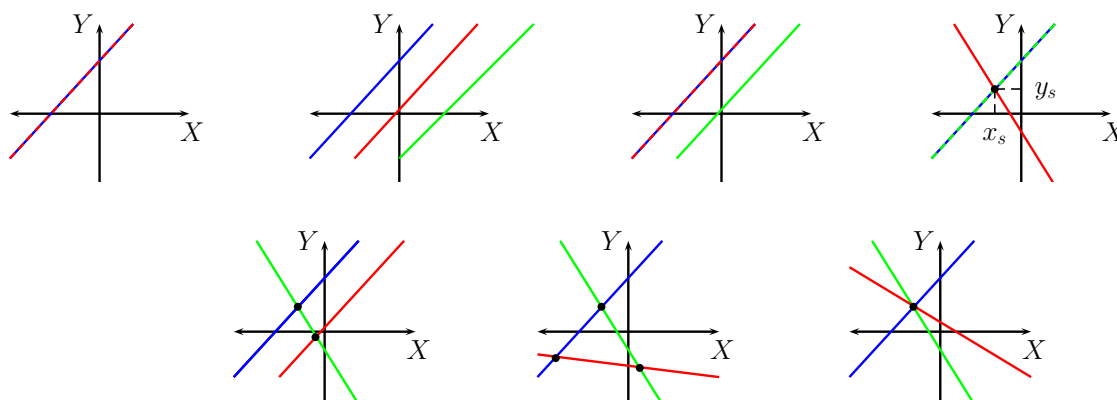
Al aumentar el número de ecuaciones del sistema, el número de posibilidades aumenta grandemente. Aquí analizaremos como ejemplo el caso de sistemas con tres ecuaciones.

$$\bullet \begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \\ a_3x + b_3y = c_3 \end{cases} \text{ en } \mathbb{R}^2$$

Estas tres rectas pueden ser:

- Las tres coincidentes (infinitas soluciones)
- Las tres paralelas (sin solución)
- Dos coincidentes y una paralela (sin solución)
- Dos coincidentes y una secante (solución única)
- Dos paralelas y una secante (sin solución)
- Tres secantes en tres puntos (sin solución)
- Tres secantes en un punto (solución única)

Estos 7 casos se muestran gráficamente en la siguiente figura.

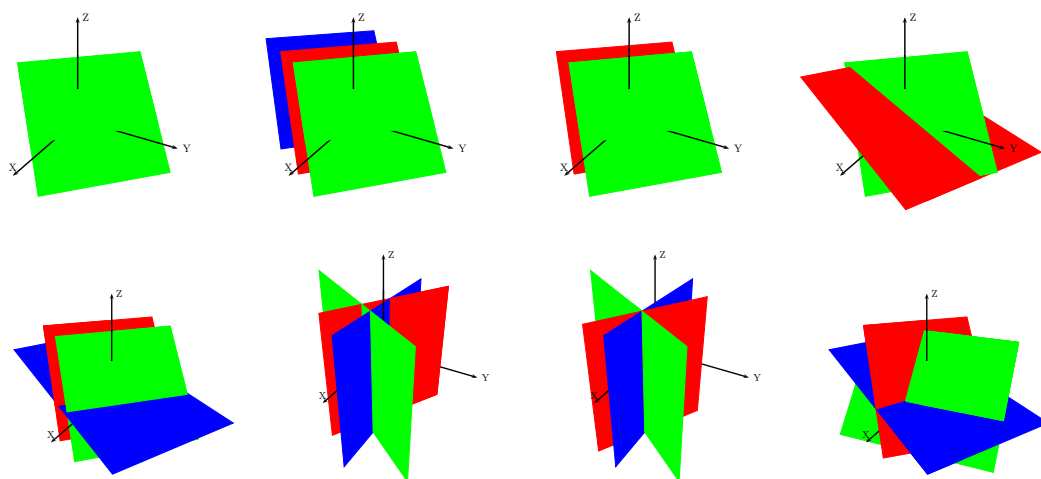


$$\bullet \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases} \quad \text{en } \mathbb{R}^3$$

Estos tres planos pueden ser:

- Los tres coincidentes (infinitas soluciones en el plano)
- Los tres paralelos (sin solución)
- Dos coincidentes y uno paralelo (sin solución)
- Dos coincidentes y uno secante (infinitas soluciones en una recta)
- Dos paralelos y un secante (sin solución)
- Tres secantes en tres rectas distintas (sin solución)
- Tres secantes en una recta (infinitas soluciones en la recta)
- Tres secantes en un punto (solución única)

Estos 8 casos se muestran gráficamente ordenados en la siguiente figura.



El análisis de estos casos resulta pesado realizarlo a través de la pendiente de las rectas y el vector normal de los planos. La situación se complica todavía más cuando se tratan de muchas ecuaciones. La forma eficiente de tratarlo se verá en el [Tema 1, página 27](#).

Ejercicio 0.22. Encontrar los sistemas de ecuaciones lineales correspondientes a los siguientes casos:

- a) En $\mathbb{R}^2$: una recta que pasa por el origen y vector director $\vec{v}_1 = [1, 2]$ junto con otra recta paralela al eje X y que pasa por el punto $(1, -1)$.
- b) En $\mathbb{R}^3$: un plano conteniendo los ejes X y Z , junto con otro plano que pasa por el punto $(0, -1, 0)$ y tiene vector normal $\vec{n}_2 = [0, -1, -2]$.
- c) En $\mathbb{R}^3$: una recta que pasa por el punto $(1, 2, 0)$ y tiene por vector director $\vec{v}_1 = [1, -1, 1]$, junto con el plano paralelo al eje Y , que contiene al vector $\vec{v}_2 = [1, -1, -1]$ y que pasa por el punto $(1, 1, 1)$.

- a) Las ecuaciones vectorial, paramétrica y general de la primera recta son:

$$[x, y] = [0, 0] + t[1, 2] \quad ; \quad \begin{cases} x = t \\ y = 2t \end{cases} = y = 2x$$

La segunda recta, al ser paralela al eje X tiene por vector director $\vec{v}_2 = [1, 0]$ y por tanto sus formas vectorial, paramétrica y general son:

$$[x, y] = [1, -1] + t[1, 0] \quad ; \quad \begin{cases} x = 1 + t \\ y = -1 \end{cases} = y = -1$$

Los SEL que pide el problema deben escribirse utilizando ecuaciones en su forma general, por tanto la solución de este caso se puede expresar como:

$$\begin{cases} 2x - y &= 0 \\ y &= -1 \end{cases} \quad \text{en } \mathbb{R}^2$$

- b) El primer plano que contiene a los ejes X y Z , pasa por el origen $(0, 0, 0)$ y no contiene los coeficientes a y c , (ver la [Sección 0.3.1, página 11](#)), por tanto tiene por ecuación general $y = 0$.

Sabiendo que el vector normal del segundo plano es $\vec{n}_2 = [0, -1, -2]$, su ecuación general tendrá la forma: $-y - 2z = d$. La determinación de d se realiza sabiendo que contiene el punto $(0, -1, 0)$. Sustituyendo obtenemos que $d = 1$. El sistema de ecuaciones es por tanto:

$$\begin{cases} y &= 0 \\ -y - 2z &= 1 \end{cases} \quad \text{en } \mathbb{R}^3$$

- c) La recta en sus formas vectorial, paramétrica y general son:

$$[x, y, z] = [1, 2, 0] + t[1, -1, 1] \quad ; \quad \begin{cases} x = 1 + t \\ y = 2 - t \\ z = t \end{cases} \quad ; \quad \begin{cases} x - z &= 1 \\ y + z &= 2 \end{cases}$$

Con respecto al plano, al ser paralelo al eje Y , contiene al vector $\vec{v}_1 = [0, 1, 0]$ y como además contiene al vector $\vec{v}_2 = [1, -1, -1]$ y pasa por el punto $(1, 1, 1)$, sus ecuaciones vectorial y paramétrica son:

$$[x, y, z] = [1, 1, 1] + t[0, 1, 0] + s[1, -1, -1] \quad ; \quad \begin{cases} x = 1 + s \\ y = 1 + t - s \\ z = 1 - s \end{cases}$$

Para pasar a la forma general del plano tenemos que combinar estas tres ecuaciones para obtener una única ecuación sin parámetros t y s . En este caso basta con sumar la primera ecuación y la tercera obteniéndose la ecuación del plano $x + z = 2$. El sistema pedido es finalmente el conjunto de las dos ecuaciones de la recta junto con esta del plano:

$$\begin{cases} x - z &= 1 \\ y + z &= 2 \\ x + z &= 2 \end{cases} \quad \text{en } \mathbb{R}^3$$

Nota 0.1. En general, hay diferentes formas de resolver un problema y en las soluciones mostradas aquí se presenta solo una de ellas, que suele resultar sencilla o rápida. Naturalmente, cualquier otra forma correcta de resolución del problema es perfectamente válida.

0.6. Métodos Simples de Resolución de Sistemas

Existen tres métodos simples de resolución de sistemas lineales:

• Método de Sustitución

1. Se despeja una de las incógnitas en una de las ecuaciones.
2. Se sustituye la expresión obtenida en el resto de las ecuaciones, obteniéndose un sistema de $n - 1$ ecuaciones.
3. Se repite el proceso de sustitución hasta obtener una única ecuación con una única incógnita, fácil de resolver.
4. La solución del resto de las incógnitas se obtiene resolviendo las ecuaciones obtenidas hacia atrás.

- **Método de Igualación**

1. Se despeja una de las incógnitas en todas las ecuaciones.
2. Se iguala la primera ecuación al resto de las ecuaciones, obteniéndose un sistema de $n - 1$ ecuaciones.
3. Se repite el proceso de igualación hasta obtener una única ecuación con una única incógnita, fácil de resolver.
4. La solución del resto de las incógnitas se obtiene resolviendo las ecuaciones obtenidas hacia atrás.

- **Método de Reducción**

1. Se igualan los coeficientes (salvo el signo) de una de las incógnitas, entre la primera ecuación y el resto de ecuaciones en grupos de dos ecuaciones, mediante multiplicación por números apropiados cada una de las ecuaciones.
2. Se suman las ecuaciones obteniéndose una única ecuación (de cada grupo de dos ecuaciones) con una incógnita menos.
3. Se repite el proceso de reducción hasta obtener una única ecuación con una única incógnita, fácil de resolver.
4. La solución del resto de las incógnitas se obtiene resolviendo las ecuaciones obtenidas hacia atrás.

Ejercicio 0.23. Resolver el siguiente sistema por los métodos de sustitución, igualación y reducción.

$$\begin{cases} 3x - y = 1 \\ x + y = 2 \end{cases}$$

- **Sustitución:** Despejando de la primera ecuación:

$$y = 3x - 1 \quad (12)$$

Sustituyendo en la segunda ecuación el valor de la variable y :

$$x + (3x - 1) = 2 \quad ; \quad 4x - 1 = 2 \quad ; \quad 4x = 3 \quad ; \quad \boxed{x = \frac{3}{4}}$$

Sustituyendo este valor en la **Ecuación (12)**:

$$y = 3\left(\frac{3}{4}\right) - 1 = \frac{9}{4} - 1 \quad ; \quad \boxed{y = \frac{5}{4}}$$

- **Igualación:** Despejando por ejemplo la variable y de ambas ecuaciones se tiene:

$$\begin{cases} y = 3x - 1 \\ y = 2 - x \end{cases}$$

Igualamos ambas ecuaciones y despejamos la variable x :

$$3x - 1 = 2 - x \quad ; \quad 4x = 3 \quad ; \quad \boxed{x = \frac{3}{4}}$$

Sustituyendo este valor en la segunda ecuación del sistema:

$$y = 2 - \frac{3}{4} = \frac{8 - 3}{4} \quad ; \quad \boxed{y = \frac{5}{4}}$$

- **Reducción:** Sumando ambas ecuaciones se obtiene:

$$4x = 3 \quad ; \quad \boxed{x = \frac{3}{4}}$$

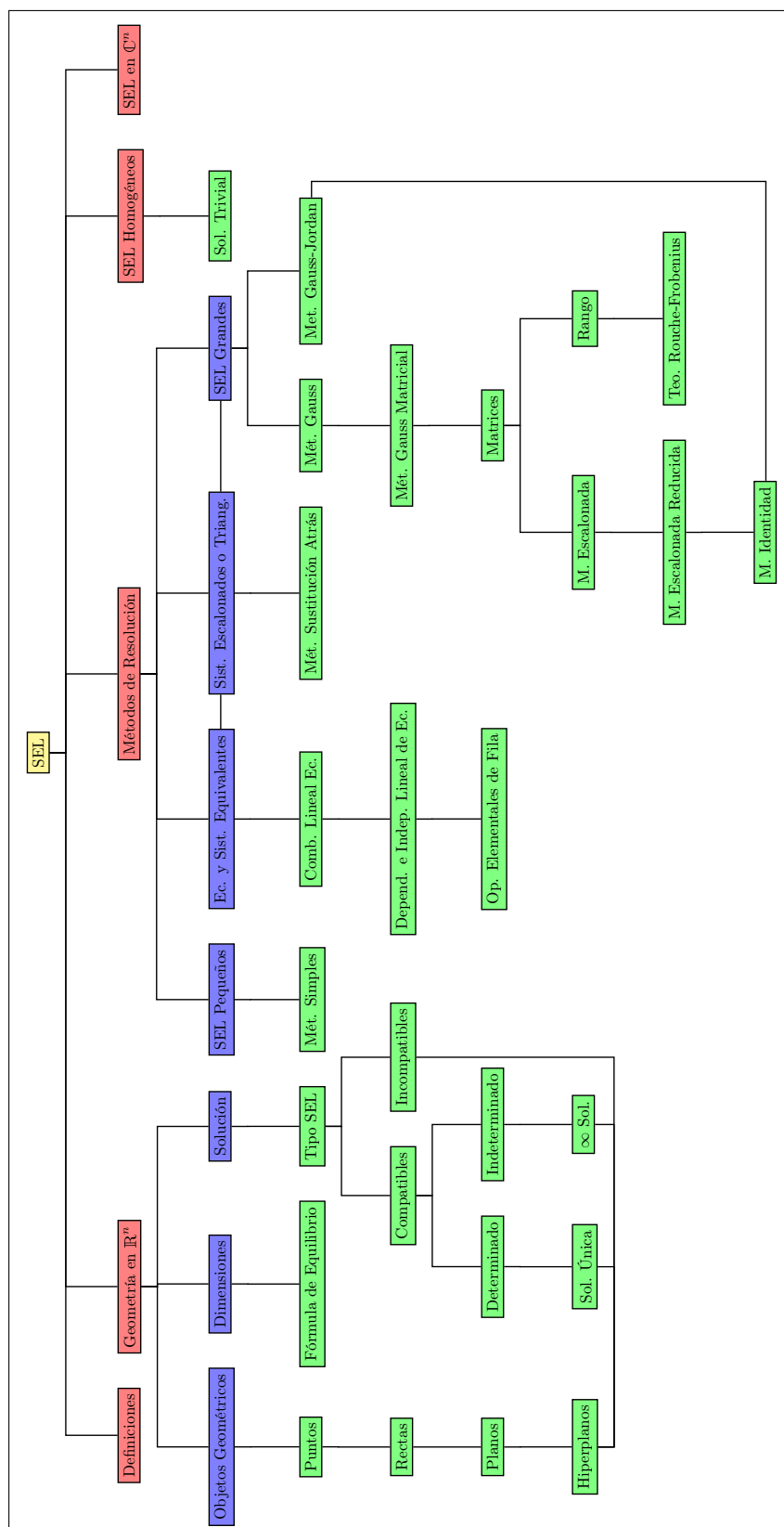
Sustituyendo este valor en la segunda ecuación del sistema:

$$y = 2 - \frac{3}{4} = \frac{8 - 3}{4} \quad ; \quad \boxed{y = \frac{5}{4}}$$

Sistemas de Ecuaciones Lineales

Contenido

1.1	Introducción a los Sistema de Ecuaciones Lineales (SEL)	29
1.1.1	Características de los Sistemas de Ecuaciones	29
1.1.2	Ecuaciones y Sistemas de Ecuaciones	29
1.1.3	Puntos y Espacios	32
1.1.4	Sistemas de Ecuaciones Lineales (SEL)	34
1.2	Geometría de los SEL en $\mathbb{R}^n$	34
1.2.1	Objetos Geométricos	36
1.2.2	Dimensión de los Objetos Geométricos	36
1.2.3	Fórmula o Ecuación de Equilibrio	37
1.2.4	Relación entre Objetos Geométricos en un SEL	38
1.3	Solución de un SEL en $\mathbb{R}^n$	39
1.3.1	SEL Compatibles e Incompatibles	40
1.4	Métodos Resolución de SEL en $\mathbb{R}^n$	42
1.4.1	Métodos Simples	42
1.4.2	Ecuaciones Equivalentes	42
1.4.3	Combinación Lineal de Ecuaciones	43
1.4.4	Dependencia e Independencia lineal de Ecuaciones	43
1.4.5	Sistemas Equivalentes	44
1.4.6	Operaciones Elementales de Fila	44
1.4.7	Sistemas Escalonados o Triangulares	45
1.4.8	Método de Gauss o Eliminación Gaussiana	46
1.5	Matrices	47
1.5.1	Matrices Asociadas a SEL	48
1.5.2	Operaciones Elementales de Fila en la Matriz Asociada	49
1.5.3	Matriz Equivalente por Filas	49
1.5.4	Matriz Escalonada Equivalente	50
1.5.5	Rango de una Matriz	50
1.5.6	Teorema de Rouché-Fröbenius	51
1.6	Método de Gauss Matricial	51
1.6.1	Identificación de Sistemas Incompatibles	52
1.6.2	Identificación de Sistemas Compatibles Indeterminados	53
1.6.3	Estudio de Soluciones de SEL con Parámetros	55
1.7	Método de Gauss-Jordan	57
1.7.1	Matriz Escalonada Reducida	57
1.7.2	Matriz Identidad	58
1.8	Sistemas Homogéneos	59
1.9	SEL en $\mathbb{C}^n$	61



1.1. Introducción a los Sistema de Ecuaciones Lineales (SEL)

Consideremos el siguiente problema sencillo.

Ejercicio 1.1. Determinar la edad de dos hermanos sabiendo que el triple de la diferencia entre el mayor y el menor es dos veces la edad del menor, y que la suma de ambas edades es 24.

Llamando x a la edad del hermano mayor e y a la edad de menor podemos establecer las siguientes ecuaciones del problema:

$$\begin{cases} 3(x - y) = 2y \\ x + y = 24 \end{cases} = \text{Sistema de 2 ecuaciones con 2 incógnitas} \quad \begin{cases} 3x - 5y = 0 \\ x + y = 24 \end{cases}$$

La solución de este *sistema de ecuaciones* dará la solución del problema.

1.1.1. Características de los Sistemas de Ecuaciones

- Un *sistema de ecuaciones* simplifica un problema mediante una estructura de lenguaje (matemático) algebraico, eliminando cualquier complejidad adicional del problema.
- Un mismo *sistema* puede representar a problemas totalmente distintos.
- La solución del problema queda reducida a la resolución del problema *matemático* y la posterior interpretación de los resultados.

En el presente tema estudiaremos los *sistemas de ecuaciones lineales*, cubriendo los conceptos básicos, su interpretación geométrica, teoremas importantes, técnicas de análisis y métodos de resolución, enfocando nuestro interés en los grandes sistemas de ecuaciones.

Primeramente estudiaremos los sistemas de ecuaciones lineales en espacios reales¹ $\mathbb{R}^n$ y posteriormente se generalizarán a los espacios complejos² $\mathbb{C}^n$.

1.1.2. Ecuaciones y Sistemas de Ecuaciones

Vamos a comenzar con algunas definiciones básicas.

Definición 1.1. Ecuación

Una **ecuación** es una *igualdad* que relaciona **incógnitas** (variables) y **constantes**, y que se verifica únicamente para determinados valores de las variables.

Ejemplo 1.1. Ejemplo de ecuación.

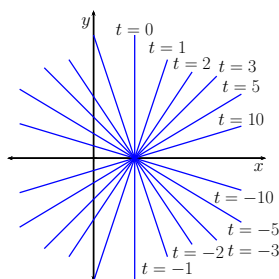
$$3x + 2 \cos(y) = -1 \quad = \quad \begin{cases} \text{Incógnitas} & : x, y. \\ \text{Constantes} & : 3, 2, -1. \end{cases}$$

En algunas ecuaciones además de *incógnitas* y *constantes* aparecen **parámetros**³, que son valores de los cuales dependen las incógnitas.

¹Ver la Sección 1.1.3, página 32.

²Ver la Sección ??, página ??.

³Ver la Sección 0.2.5, página 6.

Ejemplo 1.2. Ejemplo de ecuación con parámetros.

En la siguiente ecuación, x e y representan variables y t es un parámetro:

$$3x - ty = 5$$

Los valores (x, y) que la verifican dependen del valor del parámetro t , por lo que esta expresión representa en realidad una *familia de ecuaciones* que geoméricamente representan un conjunto infinito de rectas.

Definición 1.2. Ecuaciones Lineales y No Lineales

Las ecuaciones se pueden clasificar en dos tipos:

- **Ecuación lineal:** ecuación polinómica de primer grado (potencia 1) en todas las variables.

$$y = 3x + 5 \quad ; \quad -x + 2y - 3z = 5 \quad ; \quad 2x_1 + 3x_2 - 5x_3 = 0$$

- **Ecuación no lineal:** ecuación de tipo no polinómica o polinómica de grado mayor que uno.

$$ax^2 - y = 0 \quad ; \quad yx = \sin(x) + 1 \quad ; \quad \overline{x_1} + x_2 = x_3$$

Sistemas de Ecuaciones**Definición 1.3.** Sistema de Ecuaciones

Un **sistema de ecuaciones** es un conjunto de *ecuaciones* que *comparten incógnitas*.

Definición 1.4. Ecuaciones Acopladas e Independientes

- Se denominan **ecuaciones acopladas** a aquellas que *comparten incógnitas* formando parte de un *sistema de ecuaciones*.
- Un conjunto de ecuaciones no acopladas se denominan **ecuaciones independientes**.

Ejemplo 1.3. Ejemplo de: *ecuaciones acopladas* (a) y *ecuaciones independientes* (b).

$$a) \begin{cases} x^2 + 2y + z = 3 \ln(y) \\ 2x - xy = 0 \end{cases} \quad b) \begin{cases} x - y = -1 \\ z + t^2 = 0 \end{cases}$$

Definición 1.5. Dimensión de un Sistema de Ecuaciones

La **dimensión** de un *sistema de ecuaciones* es el número de ecuaciones y el número de incógnitas del sistema.

Ejemplo 1.4. Dimensión de un sistema.

La dimensión del sistema del **Ejemplo 1.3-a** es: dos ecuaciones y tres incógnitas (x , y y z).

Solución**Definición 1.6. Solución de una Ecuación**

Se denomina **solución** de una ecuación al conjunto de valores $s_1, \dots, s_n$ que al sustituir en las variables, $x_1 = s_1, x_2 = s_2, \dots, x_n = s_n$, se verifica la ecuación.

Definición 1.7. Solución del Sistema de Ecuaciones

Se denomina **solución de un sistema de ecuaciones** al conjunto de valores $s_1, \dots, s_n$ que verifican todas y cada una de las ecuaciones del sistema.

El número de soluciones de una ecuación o sistema de ecuaciones puede ser: una, varias, infinitas o ninguna, dependiendo de las características de las ecuaciones.

Ejemplo 1.5. Diferentes tipos de soluciones en ecuaciones con números reales.

- Ecuación lineal con solución única:

$$x - 5 = 0 \quad = \quad \boxed{x = 5}$$

- Ecuación no lineal con dos soluciones:

$$x^2 - 25 = 0 \quad = \quad \boxed{x_1 = 5 \quad ; \quad x_2 = -5}$$

- Ecuación no lineal con infinitas soluciones:

$$\cos(x) - 1 = 0 \quad = \quad \boxed{x = 2k\pi \quad ; \quad k \in \mathbb{Z}}$$

- Ecuación lineal sin solución.

$$x + 1 = x + 2$$

- Sistema de ecuaciones lineales con solución única:

$$\begin{cases} x + y = 1 \\ x - y = 0 \end{cases} \quad = \quad \boxed{x = 1 \quad ; \quad y = 0}$$

- Sistema de ecuaciones no lineales con dos soluciones:

$$\begin{cases} x^2 + y^2 = 25 \\ x + y = 7 \end{cases} \quad = \quad \boxed{x = 3, y = 4 \quad ; \quad x = 4, y = 3}$$

- Sistema de ecuaciones lineales con infinitas soluciones:

$$\begin{cases} x + y = 1 \\ 2x + 2y = 2 \end{cases} \quad = \quad \boxed{x = t \quad ; \quad y = 1 - t \quad ; \quad t \in \mathbb{R}}$$

- Sistema de ecuaciones lineales sin solución:

$$\begin{cases} x + y = 1 \\ x + y = 0 \end{cases}$$

- Sistema de ecuaciones no lineales sin solución:

$$\begin{cases} y - x^2 = 4 \\ y + 2x^2 = -6 \\ x + y = 0 \end{cases}$$

1.1.3. Puntos y Espacios

A lo largo de este texto vamos a trabajar tanto con números reales $\mathbb{R}$ como complejos $\mathbb{C}$. Si bien el conjunto $\mathbb{C}$ es una generalización de $\mathbb{R}$, cada uno de ellos tiene unas características particulares que les hacen más adecuados en algunas situaciones. Para hacer referencia a los dos casos por separado vamos a utilizar el símbolo $\mathbb{K}$, que representará o bien al conjunto de los números reales $\mathbb{R}$ o bien al conjunto de los números complejos $\mathbb{C}$.

$$\left. \begin{array}{l} \mathbb{R} : \text{Conjunto de Números Reales} \\ \mathbb{C} : \text{Conjunto de Números Complejos} \end{array} \right\} \mathbb{K} = \mathbb{R} \text{ ó } \mathbb{C}$$

Definición 1.8. Punto

Un **punto** es un conjunto de valores ordenados, denominados **coordenadas** o **componentes** del punto. Las *coordenadas* se representan encerradas entre paréntesis:

$$(x_1, x_2, \dots, x_n) \quad ; \quad x_i \in \mathbb{K}$$

El *número de coordenadas* n de un punto se denomina **dimensión** del punto.

Definición 1.9. Espacio $\mathbb{K}^n$

El conjunto de todos los *puntos* de n *componentes* se representa por $\mathbb{K}^n$.

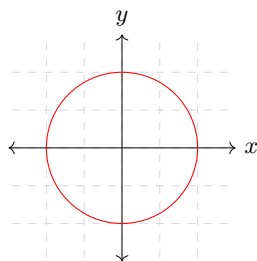
$$\mathbb{K}^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{K}, 1 \leq i \leq n\}$$

El valor $n \geq 1$ se denomina **dimensión** del espacio $\mathbb{K}^n$.

- El espacio $\mathbb{R}$ es el espacio unidimensional representado por todos los puntos (números) de la *recta real*.
- El espacio $\mathbb{C}$ es el espacio unidimensional representado por todos los puntos (números) del *plano complejo*.
- El espacio $\mathbb{K}$ es el espacio unidimensional representado o bien por todos los puntos (números) reales $\mathbb{R}$ o bien todos los puntos (números) complejos $\mathbb{C}$.
- El espacio $\mathbb{R}^2$ es el espacio bidimensional (2D) representado por todos los puntos del *plano euclideo o cartesiano*⁴ con coordenadas (x, y) siendo $x, y \in \mathbb{R}$.
- El espacio $\mathbb{C}^2$ es el espacio bidimensional representado por todos los puntos con coordenadas (x, y) siendo $x, y \in \mathbb{C}$.
- El espacio $\mathbb{R}^3$ es el espacio tridimensional (3D) formado por todos los puntos del espacio euclideo con coordenadas (x, y, z) siendo $x, y, z \in \mathbb{R}$.
- El espacio $\mathbb{R}^n$ es el espacio n -dimensional formado por todos los puntos de n coordenadas reales $(x_1, \dots, x_n)$, $x_i \in \mathbb{R}$.
- El espacio $\mathbb{C}^n$ es el espacio n -dimensional formado por todos los puntos de n coordenadas complejas $(x_1, \dots, x_n)$, $x_i \in \mathbb{C}$.
- El espacio $\mathbb{K}^n$ es el espacio n -dimensional formado por todos los puntos de n coordenadas $(x_1, \dots, x_n)$, $x_i \in \mathbb{K}$ (todas reales o todas complejas).

La solución de una *ecuación* o un *sistema de ecuaciones* se puede representar mediante un *punto* de un espacio $\mathbb{K}^n$, siendo n el número de incógnitas.

⁴Ver la [Sección 0.1](#), página 3.

Ejemplo 1.6. Ecuación no lineal real.

$$x^2 + y^2 = 4$$

Esta es una ecuación no lineal con dos incógnitas x e y reales, que tiene como solución a los infinitos puntos del espacio $\mathbb{R}^2$ (plano euclídeo) que pertenecen a la circunferencia de centro en el origen y radio 2.

$$(x, y) : \begin{cases} x = t \\ y = \pm \sqrt{4 - t^2} \end{cases} ; \quad t \in \mathbb{R}$$

Ejemplo 1.7. Solución de un sistema de ecuaciones lineales reales.

La solución del siguiente sistema de 4 ecuaciones y 3 incógnitas $x, y, z \in \mathbb{R}$:

$$\begin{cases} x - y &= 5 \\ 2x + 3y + z &= 14 \\ -y + 2z &= -3 \\ x - y + z &= 4 \end{cases}$$

es única y podemos representarla mediante el punto $(x, y, z) = (6, 1, -1) \in \mathbb{R}^3$.

Ejemplo 1.8. Solución de una ecuación lineal compleja.

La solución de la siguiente ecuación de dos incógnitas $x, y \in \mathbb{C}$:

$$2x - 5y = 1 + i$$

es un conjunto de infinitos puntos $(x, y) \in \mathbb{C}^2$ dados por:

$$(x, y) : \begin{cases} x = t \\ y = \frac{2t - 1 - i}{5} \end{cases} ; \quad t \in \mathbb{C}$$

Notar que estos puntos no se pueden representar gráficamente ya que cada una de sus coordenadas x e y son números complejos y necesitan dos dimensiones (parte real y parte imaginaria) para su representación en el plano complejo, por lo que necesitaríamos 4 dimensiones en total para representar cada uno de los puntos $(x, y) \in \mathbb{C}^2$.

Ejemplo 1.9. Solución de un sistema de ecuaciones lineales complejo.

La solución del siguiente sistema de 2 ecuaciones y 2 incógnitas $x, y \in \mathbb{C}$:

$$\begin{cases} ix - (2 - i)y &= 2 - 3i \\ x - 2iy &= 2 + i \end{cases}$$

es única y podemos representarla mediante el punto $(x, y) = (-i, -1 + i) \in \mathbb{C}^2$.

1.1.4. Sistemas de Ecuaciones Lineales (SEL)

Definición 1.10. Sistemas de Ecuaciones Lineales

Se denomina **sistema** de m **ecuaciones lineales** con n **incógnitas**, al conjunto *finito*⁵ de ecuaciones lineales *acopladas*⁶ de la forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad \text{en } \mathbb{K}^n$$

donde:

- $a_{ij} \in \mathbb{K}$ se denominan **coeficientes** del sistema (constantes)
- $x_i \in \mathbb{K}$ son las n **incógnitas** del sistema (variables)
- $b_i \in \mathbb{K}$ son los **términos independientes** del sistema (constantes)

En este texto utilizaremos la sigla SEL para referirnos a los *Sistemas de Ecuaciones Lineales*.

Ejemplo 1.10. SEL en $\mathbb{R}^5$

$$\begin{cases} x_1 - 3x_2 + x_3 - x_5 = 9 \\ 2x_1 + x_2 + x_3 - x_4 + 2x_5 = 0 \\ -x_2 + x_4 = -12 \end{cases}$$

La *dimensión*⁷ de este sistema es 3 ecuaciones y 5 incógnitas, donde:

- Las *incógnitas* son: x_1, x_2, x_3, x_4 y x_5 , todas ellas variables reales.
- Los *coeficientes* son: $a_{11} = 1, a_{12} = -3, a_{13} = 3, a_{14} = 0, a_{15} = -1, a_{21} = 2, a_{22} = 1, a_{23} = 1, a_{24} = -1, a_{25} = 2, a_{31} = 0, a_{32} = -1, a_{33} = 0, a_{34} = 1$ y $a_{35} = 0$, todos ellos números reales.
- Los *términos independientes* son: $b_1 = 9, b_2 = 0$ y $b_3 = -12$, valores reales.

1.2. Geometría de los SEL en $\mathbb{R}^n$

En los espacios $\mathbb{R}^n$ se puede hacer una interpretación geométrica de los resultados, mientras que en los espacios $\mathbb{C}^n$ no es posible, consecuencia de la naturaleza de los números complejos⁸ comparado con la simplicidad de los números reales. Esto presenta una gran ventaja que usaremos para entender muchos de los conceptos matemáticos que se desarrollarán a lo largo del curso, de forma que la estrategia general será estudiar primeramente los conceptos en los espacios reales $\mathbb{R}^n$, haciendo uso de la geometría siempre que sea posible, para posteriormente hacer una generalización a los espacios complejos $\mathbb{C}^n$.

⁵Es decir, que no es infinito.

⁶Ver la [Definición 1.4](#), página 30.

⁷Ver la [Definición 1.5](#), página 30.

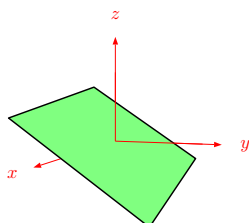
⁸Ver el [Apéndice C.5.1](#), página 302.

⁹Ver una introducción a los planos en el espacio euclídeo tridimensional en el [Sección 0.3](#), página 11.

Definición 1.11. Plano en $\mathbb{R}^3$

Un **plano**⁹ en $\mathbb{R}^3$ es el conjunto de *puntos* (x, y, z) que verifican la ecuación lineal:

$$ax + by + cz = d \quad ; \quad a, b, c, d \in \mathbb{R}$$

Ejemplo 1.11. Representación gráfica.

Plano: $x + 2y + 3z = 0$

El concepto de *plano* se puede generalizar a cualquier espacio $\mathbb{R}^n$ a través del concepto de *hiperplano*.

Definición 1.12. Hiperplano en $\mathbb{R}^n$

Un **hiperplano** de *dimensión*¹⁰ $n - 1$ en $\mathbb{R}^n$ es el conjunto de *puntos* $(x_1, \dots, x_n) \in \mathbb{R}^n$ que verifican la ecuación:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b \quad ; \quad a_i, b \in \mathbb{R}$$

Si el *término independiente* es $b = 0$ el *hiperplano* pasa por el origen $(0, 0, \dots, 0) \in \mathbb{R}^n$.

Nota 1.1. Los hiperplanos de dimensión baja (< 3) en cualquier espacio $\mathbb{R}^n$ reciben nombres particulares:

- Un hiperplano de dimensión 0 se denomina **punto**.

Ejemplos: $3 \in \mathbb{R} \quad ; \quad (1, 0, -1, 2) \in \mathbb{R}^4$.

- Un hiperplano de dimensión 1 se denomina **recta**.

Ejemplos: $x + 2y = -1 \quad \text{en } \mathbb{R}^2 \quad ; \quad \begin{cases} x + y + z = 1 \\ x - y = 0 \end{cases} \quad \text{en } \mathbb{R}^3$.

- Un hiperplano de dimensión 2 se denomina **plano**.

Ejemplos: $2x + y - z = 3 \quad \text{en } \mathbb{R}^3 \quad ; \quad \begin{cases} x_1 + x_3 = -1 \\ x_2 + 2x_3 + x_4 = 0 \end{cases} \quad \text{en } \mathbb{R}^4$.

Ejemplo 1.12. Hiperplanos.

- Ejemplo de *hiperplano de dimensión 3* en $\mathbb{R}^4$: $x - y + 2z - w = -1$

Este hiperplano no pasa por el origen de $\mathbb{R}^4$ ya que el término independiente no es cero. El nombre de las variables, en este caso x, y, z y w , de forma general puede ser cualquiera.

- Ejemplo de *hiperplano de dimensión 5* en $\mathbb{R}^6$: $2x_1 - x_3 + 2x_5 + 3x_6 = 0$

Notar que no es necesario que aparezcan todas las variables del espacio correspondiente, ya que algunos coeficientes pueden ser cero. Este hiperplano sí pasa por el origen de $\mathbb{R}^6$. Además, cuando el espacio tiene una cierta dimensión, se acostumbra a denominar a las variables con una letra y un subíndice como en este caso.

- Ejemplo de *hiperplano de dimensión 9* en $\mathbb{R}^{10}$: $x_1 + 2x_2 = 1$

Notar que no siempre aparecen las variables apropiadas para indicar el espacio al que pertenece, de forma que siempre es necesario indicar el correspondiente $\mathbb{R}^n$ al que pertenece un hiperplano.

¹⁰Ver la **Definición 1.14**, página 36.

1.2.1. Objetos Geométricos

Definición 1.13. Objeto Geométrico

Vamos a denominar **objetos geométricos** a todo conjunto de puntos de $\mathbb{R}^n$ *definido* mediante una o varias ecuaciones *lineales*.

Los **objetos geométricos** son hiperplanos de diferentes dimensiones: **puntos**, **rectas**, **planos** e **hiperplanos** de dimensión mayor de 2.

Ejemplo 1.13. Objetos geométricos en diferentes espacios $\mathbb{R}^n$

a) $2x - 3y = -1$ en $\mathbb{R}^5$

b)
$$\begin{cases} x + y - 2z = 3 \\ x - y - z = -2 \\ x - y - z = 1 \end{cases} \text{ en } \mathbb{R}^3$$

c)
$$\begin{cases} x_1 - x_2 + x_3 = 0 \\ x_1 + 2x_2 - 3x_3 + x_4 = 0 \\ -4x_1 + x_2 - x_3 - 3x_4 = 0 \end{cases} \text{ en } \mathbb{R}^8$$

1.2.2. Dimensión de los Objetos Geométricos

Excepto los puntos individuales, los *objetos geométricos* de cualquier espacio $\mathbb{R}^n$ están formados por infinitos puntos. Para poder cuantificar el tamaño de un objeto geométrico se vuelve a utilizar el término de *dimensión*¹¹.

Definición 1.14. Dimensión de un Objeto Geométrico

La **dimensión** de un *objeto geométrico* de $\mathbb{R}^n$ es un número que expresa el tamaño relativo del mismo.

Los objetos geométricos en $\mathbb{R}^n$ tienen la siguiente dimensión:

- **Punto:** dimensión 0.
- **Recta:** dimensión 1.
- **Plano:** dimensión 2.
- **Hiperplano:** dimensión > 2 .

Estos valores se pueden obtener a través de una fórmula general conocida con el nombre de *Fórmula de Equilibrio*, que se verá posteriormente en la **Definición 1.15, página 37**.

Cuando los objetos geométricos de cualquier espacio $\mathbb{R}^n$ están definidos en su *forma vectorial* o *paramétrica*¹² se puede determinar fácilmente su dimensión contando el número de *parámetros libres* en su ecuación.

¹¹Ver la dimensión de un SEL en la **Definición 1.5, página 30**, dimensión de un punto en la **Definición 1.8, página 32** y la dimensión de un espacio en la **Definición 1.9, página 32**.

¹²Ver en el **Tema 0, página 1** una introducción a las diferentes formas de representar objetos geométricos (rectas y planos) en el plano euclídeo y en el espacio tridimensional.

Dimensión de objetos geométricos en $\mathbb{R}^n$ y su forma paramétrica.

Un punto tiene dim cero	tiene 0 parám	:	$\vec{x} = \vec{p}$	
Una recta tiene dim uno	tiene 1 parám	:	$\vec{x} = \vec{p} + t\vec{v}$	(t parámetro)
Un plano tiene dim dos	tiene 2 parám	:	$\vec{x} = \vec{p} + t\vec{v} + s\vec{u}$	(t, s parámetros)
$\vdots$	$\vdots$		$\vdots$	
Un hiperplano de dim k	tiene k parám	:	$\vec{x} = \vec{p} + t_1\vec{v}_1 + \cdots + t_k\vec{v}_k$	(t_i parámetros)

De forma conceptual, podemos entender la dimensión m de un objeto geométrico en un espacio $\mathbb{R}^n$ como el número de puntos m de $\mathbb{R}^n$ que contiene¹³.

1.2.3. Fórmula o Ecuación de Equilibrio

Como hemos visto, un objeto geométrico en $\mathbb{R}^n$ viene definido por un conjunto de ecuaciones lineales con n incógnitas y puede tener diferentes dimensiones que dependen de su tamaño. Hay una ecuación muy importante, denominada ecuación de equilibrio, que relaciona estos tres conceptos que definen un objeto geométrico:

- Número de ecuaciones lineales que lo definen
- Dimensión del objeto geométrico
- Dimensión n del espacio $\mathbb{R}^n$ donde se encuentra

Definición 1.15. Fórmula o Ecuación de Equilibrio

Relaciona el número de ecuaciones *linealmente independientes*¹⁴ (l.i.) necesarias para definir un determinado objeto geométrico en un determinado espacio.

$$\text{Número de Ecuaciones (l.i.)} = \text{Dimensión Espacio} - \text{Dimensión Objeto}$$

Ejemplo 1.14. Objetos Geométricos

Una recta (dim=1) en $\mathbb{R}^2$	Num. Ec = 2 - 1 = 1	:	$ax + by = c$
Una recta (dim=1) en $\mathbb{R}^3$	Num. Ec = 3 - 1 = 2	:	$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$
Un plano (dim=2) en $\mathbb{R}^3$	Num. Ec = 3 - 2 = 1	:	$ax + by + cz = d$

Ejercicio 1.2. Determinar cuántas ecuaciones (linealmente independientes) se necesitan en $\mathbb{R}^{12}$ para representar un hiperplano de dimensión 7.

Aplicando la ecuación de equilibrio:

$$\text{Núm. de ec. l.i.} = \text{dim. espacio} - \text{dim. objeto} = 12 - 7 = \boxed{5}$$

¹³Un punto tiene dimensión 0 y está formado por un único punto, que interpretaremos como $\infty^0 = 1$. Una recta tiene dimensión 1 y está formada por ∞^1 puntos. Un plano que tiene dimensión 2 está formado por infinitas rectas y posee ∞^2 puntos. Un hiperplano de dimensión 3 está formado por infinitos planos y posee ∞^3 puntos. Y así sucesivamente.

¹⁴Es un tipo especial de ecuaciones que será definido en la **Definición 1.21, página 43**.

Por ejemplo, el siguiente SEL sencillo representa un hiperplano de dimensión 7 en $\mathbb{R}^{12}$:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ x_2 + x_3 + x_4 + x_5 = 1 \\ x_3 + x_4 + x_5 = 2 \\ x_4 + x_5 = 3 \\ x_5 = 4 \end{cases} \quad \text{en } \mathbb{R}^{12}$$

1.2.4. Relación entre Objetos Geométricos en un SEL

Dado un conjunto de m ecuaciones lineales (objetos geométricos) en un espacio $\mathbb{R}^n$ (con n incógnitas), vamos a denotar por E_i a cada una de sus ecuaciones y por X_i cada uno de los términos izquierdos de la igualdad E_i , de la siguiente forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} = \begin{cases} X_1 = b_1 \\ X_2 = b_2 \\ \vdots \\ X_m = b_m \end{cases} = \begin{cases} E_1 \\ E_2 \\ \vdots \\ E_m \end{cases}$$

Cada una de las ecuaciones E_i se puede identificar como un determinado *objeto geométrico* en el espacio $\mathbb{R}^n$.

En un SEL los *objetos geométricos* que representan dos *ecuaciones* cualesquiera E_i y E_j solo pueden ser entre sí uno de los siguientes casos:

- **Coincidentes:** si $E_i = cE_j$, para cualquier $c \in \mathbb{R}$ no nulo.
- **Paralelos:** si $X_i = cX_j$, $b_i = b_j$, para cualquier $c \in \mathbb{R}$ no nulo.
- **Secantes:** si $X_i = cX_j$, para cualquier $c \in \mathbb{R}$ no nulo.

Ejemplo 1.15. Relación entre objetos geométricos.

- Las siguientes dos rectas en $\mathbb{R}^2$ son *coincidentes*, (representan la misma recta).

$$x + 3y = -2 \quad ; \quad 2x + 6y = -4$$

- Los siguientes dos planos en $\mathbb{R}^3$ son *paralelos*, (no tienen puntos en común).

$$x + 2y - z = -2 \quad ; \quad 2x + 4y - 2z = 0$$

- Los siguientes dos hiperplanos de dimensión 3 en $\mathbb{R}^4$ son *secantes*, (tienen algunos puntos en común).

$$x_1 + 2x_2 - x_3 + x_4 = -2 \quad ; \quad x_1 + x_2 - x_3 + x_4 = -2$$

Definición 1.16. Concurrencia

Un conjunto de objetos geométricos es **concurrente** si tiene alguna parte común a todos los objetos.

Ejemplo 1.16. Concurrencia

- Dos *rectas secantes* son concurrentes en un *punto*.
- Dos *planos secantes* son concurrentes en una *recta*.
- Dos *hiperplanos coincidentes* son concurrentes al mismo *hiperplano* que representan.

1.3. Solución de un SEL en $\mathbb{R}^n$

Desde un punto de vista geométrico, la solución de un SEL representa un **objeto geométrico** de una *dimensión*¹⁵ determinada, correspondiente a la parte **concurrente** de **todas** las ecuaciones¹⁶ (**objetos geométricos**).

La *dimensión* de la solución (objeto geométrico) depende de la dimensión del SEL (número de ecuaciones y número de incógnitas¹⁷) y de las relaciones entre ellos (coincidentes, secantes y/o paralelos).

Sea un SEL de $\mathbb{R}^n$ formado por m ecuaciones (objetos geométricos) *linealmente independientes*¹⁸:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Su *solución* es un *objeto geométrico* cuya *dimensión* viene dada por la **fórmula de equilibrio**¹⁹:

$$\text{Dimensión Objeto} = \text{Dimensión Espacio} - \text{Num. Ecuaciones del Sistema} = n - m$$

Ejemplo 1.17. Ejemplos de aplicación de la ecuación de equilibrio.

- Una ecuación en $\mathbb{R}^2$:

$$ax + by = c$$

representa una *recta*, que es un objeto de dimensión $2-1=1$.

- Dos ecuaciones (rectas) secantes en $\mathbb{R}^2$:

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases}$$

tienen por solución un *punto* en $\mathbb{R}^2$, que tiene dimensión $2 - 2 = 0$.

- Una ecuación en $\mathbb{R}^3$:

$$ax + by + cz = d$$

representa un *plano*, que es un conjunto de dimensión $3-1=2$.

- Dos ecuaciones (planos) concurrentes no coincidentes en $\mathbb{R}^3$:

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$$

tienen por solución una *recta* en $\mathbb{R}^3$, que tiene dimensión $3 - 2 = 1$.

- Un SEL formado por cuatro hiperplanos en $\mathbb{R}^4$ *linealmente independientes* :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 = b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 = b_4 \end{cases}$$

tienen como solución un *punto* ($4-4=0$) en $\mathbb{R}^4$.

¹⁵Ver la **Definición 1.14**, página 36.

¹⁶Las ecuaciones de un sistema son la *forma general* de la solución, estando definida en forma *implícita*. Al resolver el sistema, la solución se expresa en su *forma vectorial* o *paramétrica*, siendo ésta su forma *explícita*.

¹⁷El número de incógnitas corresponde con la dimensión del espacio $\mathbb{R}^n$ al que pertenece.

¹⁸Este concepto se verá más adelante, en la **Sección 1.4.4**, página 43, y corresponde con un caso especial de conjunto de objetos geométricos *concurrentes* y *no coincidentes*.

¹⁹Corresponde con la ecuación de equilibrio vista en la **Definición 1.15**, página 37, en la cual se ha despejado la dimensión del objeto en función de la dimensión del espacio y el número de ecuaciones linealmente independientes del sistema.

1.3.1. SEL Compatibles e Incompatibles

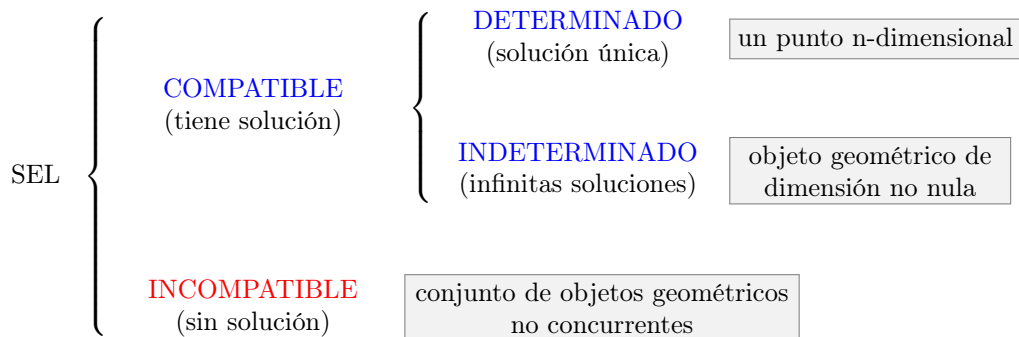
Definición 1.17. Sistema Compatible

Se dice que un SEL es **compatible** si tiene solución. Hay dos tipos:

- **Compatible Determinado**: Si la solución es única (un punto n -dimensional).
- **Compatible Indeterminado**: Si tiene infinitas soluciones (objeto geométrico de dimensión no nula).

Definición 1.18. Sistema Incompatible

Se dice que un SEL es **incompatible** si no tiene solución (conjunto de objetos geométricos no concurrentes).



Para que el SEL en $\mathbb{R}^n$ sea **compatible y determinado** tiene que estar formado por exactamente n ecuaciones *linealmente independientes*.

Conclusiones

Sea un SEL en $\mathbb{R}^n$ de m ecuaciones con n incógnitas.

- Si $m < n$, el sistema puede tener:
 - *ninguna solución*: conjunto de hiperplanos **no concurrentes**, (coincidentes, paralelos y/o secantes en número $k < m$).
 - *infinitas soluciones*: conjunto de hiperplanos **concurrentes**, (coincidentes y/o secantes), en objetos de dimensión no nula (rectas, planos o hiperplanos).
- Si $m \geq n$, el sistema puede tener:
 - *ninguna solución*: conjunto de objetos geométricos **no concurrentes** (coincidentes, paralelos y/o secantes en número $k < n$).
 - *infinitas soluciones*: conjunto de objetos geométricos **concurrentes**, (coincidentes y/o secantes), en un objeto de dimensión no nula (rectas, planos o hiperplanos).
 - *solución única*: conjunto de hiperplanos **concurrentes**, (coincidentes y secantes o solo secantes), en un objeto de dimensión nula (punto n -dimensional).

Ejemplo 1.18. Sistemas compatibles e incompatibles.

- Tres rectas *paralelas*²⁰ y una *secante* en $\mathbb{R}^2$ forman un sistema **incompatible**.
- Dos planos *coincidentes* y uno *secante* en $\mathbb{R}^3$ forman un sistema **compatible indeterminado**.
- Dos rectas *secantes* en $\mathbb{R}^2$ forman un sistema **compatible determinado**.

Ejercicio 1.3. Analizar el siguiente sistema de 2 ecuaciones en $\mathbb{R}^5$:

$$\begin{cases} x_1 + x_2 - 4x_3 = -2 & : E_1 \\ x_2 + 6x_3 - 2x_4 + 9x_5 = 0 & : E_2 \end{cases}$$

Cada una de las ecuaciones representa a un hiperplano de dimensión $5 - 1 = 4$ en $\mathbb{R}^5$. Llamando E_1 a la primera ecuación y E_2 a la segunda, podemos concluir que:

- No son *coincidentes* ya que $E_1 = cE_2$, para cualquier $c \in \mathbb{R}$ no nulo.
- No son *paralelos* ya que $X_1 = cX_2$, para cualquier $c \in \mathbb{R}$ no nulo.
- Por tanto, son *secantes* (*concurrentes*) y la solución del SEL es un objeto de dimensión $5 - 2 = 3$ (hiperplano de dimensión 3 en un espacio de dimensión 5). Su ecuación paramétrica/vectorial tendrá por tanto 3 parámetros libres.

Vamos a encontrar la ecuación que describe la solución. Resolviendo el sistema por ejemplo por el *método de igualación*²¹, se tiene:

$$\left\{ \begin{array}{l} x_2 = -2 - x_1 + 4x_3 \\ x_2 = -6x_3 + 2x_4 - 9x_5 \end{array} \right\} \quad -2 - x_1 + 4x_3 = -6x_3 + 2x_4 - 9x_5$$

Despejando, por ejemplo la variable x_4 :

$$x_4 = -1 - \frac{1}{2}x_1 + 2x_3 + \frac{9}{2}x_5$$

queda en función de las *variables libres*²² x_1 , x_3 y x_5 , que en principio pueden tomar cualquier valor real. Podemos asociar estas *variables libres* a diferentes *parámetros*²³, por ejemplo:

$$\left\{ \begin{array}{l} x_1 = t \\ x_3 = s \\ x_5 = w \end{array} \right\} \quad t, s, w \in \mathbb{R} = \left\{ \begin{array}{l} x_1 = t \\ x_2 = -2 - x_1 + 4x_3 = -2 - t + 4s \\ x_3 = s \\ x_4 = -1 - \frac{1}{2}x_1 + 2x_3 + \frac{9}{2}x_5 = -1 - \frac{1}{2}t + 2s + \frac{9}{2}w \\ x_5 = w \end{array} \right\}$$

siendo las ecuaciones de arriba, la forma *paramétrica*²⁴ del hiperplano de dimensión 3 en el espacio $\mathbb{R}^5$, solución del SEL. La forma vectorial²⁵ correspondiente a este caso se obtiene *agrupando* las constantes asociadas a cada parámetro y las que no tienen parámetro:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 0 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 0 \\ -1/2 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 4 \\ 1 \\ 2 \\ 0 \end{bmatrix} + w \begin{bmatrix} 0 \\ 0 \\ 0 \\ 9/2 \\ 1 \end{bmatrix} = \boxed{\vec{x} = \vec{p} + t\vec{v}_1 + s\vec{v}_2 + w\vec{v}_3}$$

donde hemos denominado $\vec{x}$, $\vec{p}$, $\vec{v}_1$, $\vec{v}_2$ y $\vec{v}_3$ a los vectores de $\mathbb{R}^5$ obtenidos.

²⁰Todo SEL donde hay al menos dos objetos geométricos paralelos es siempre incompatible ya que no puede existir una parte común a todos.

1.4. Métodos Resolución de SEL en $\mathbb{R}^n$

1.4.1. Métodos Simples

Existen tres métodos simples de resolución de sistemas de ecuaciones lineales²⁶:

- Método de Sustitución.
- Método de Igualación.
- Método de Reducción.

Estos métodos son útiles para la resolución “a mano” de SEL de pequeña dimensión. Sin embargo, para la resolución de sistemas más grandes, tanto a mano como por ordenador, estos SEL **no** resultan prácticos, siendo necesarios métodos sistemáticos fáciles de programar.

El objetivo de esta parte es buscar un método de resolución de **grandes SEL**, **eficiente**²⁷ y **sistemático**²⁸, apropiado para su resolución mediante un **ordenador**.

En este tema estudiaremos dos métodos útiles para grandes SEL: el *método de Gauss*²⁹ y el *método de Gauss-Jordan*³⁰. Antes de analizarlos vamos a ver algunos temas necesarios para su conocimiento.

1.4.2. Ecuaciones Equivalentes

Definición 1.19. Ecuaciones Equivalentes

Se dice que dos ecuaciones son **equivalentes** si ambas tienen las mismas soluciones.

Fácilmente se puede obtener una ecuación *equivalente* de otra multiplicando ambas partes de la ecuación por una constante distinta de cero.

Desde un punto de vista geométrico, que dos ecuaciones sean *equivalentes* es lo mismo que sean *coincidentes*³¹, y por tanto representan el mismo objeto geométrico.

$$E_1 = cE_2 \quad ; \quad c \in \mathbb{R}$$

Ejercicio 1.4. Encontrar todas las ecuaciones equivalentes del siguiente plano en $\mathbb{R}^3$.

$$3x + 2y - z = 5$$

Multiplicando la ecuación por un escalar no nulo $a \in \mathbb{R}$:

$$3x + 2y - z = 5 \quad \xrightarrow{\times a} \quad 3ax + 2ay - az = 5a$$

Cualquiera de estas ecuaciones representa exactamente el mismo plano.

²¹Ver la Sección 0.6, página 25.

²²Ver la Sección 1.6.2, página 53.

²³Aparecen exactamente 3 *parámetros libres* ya que la solución tiene dimensión 3.

²⁴Ver el Tema 0, página 1.

²⁵En las formas paramétrica y vectorial, la dimensión de un objeto geométrico coincide con el número de sus parámetros.

²⁶Ver la Sección 0.6, página 25.

²⁷A medida que aumenta el tamaño de un SEL, aumenta siempre el número de operaciones necesario para su resolución. Un método eficiente de resolución será aquel que requiera el menor número posible de operaciones.

²⁸Con un algoritmo sencillo que se pueda programar fácilmente.

²⁹Ver la Sección 1.4.8, página 46.

³⁰Ver la Sección 1.7, página 57.

³¹Ver la Sección 1.2.4, página 38.

1.4.3. Combinación Lineal de Ecuaciones

Definición 1.20. Combinación Lineal de Ecuaciones

Se dice que una **ecuación** $E \in \mathbb{R}^n$ es una **combinación lineal** de un conjunto de m **ecuaciones** $E_1, E_2, \dots, E_m \in \mathbb{R}^n$ si es una suma (término a término) de *ecuaciones equivalentes*³² a ellas.

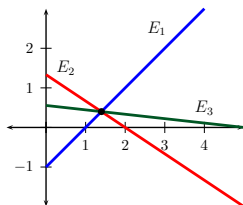
$$E = \sum_{i=1}^m a_i E_i = a_1 E_1 + a_2 E_2 + \dots + a_m E_m \quad ; \quad a_i \in \mathbb{R}$$

Propiedades de las Combinaciones Lineales

1. La **combinación lineal** de ecuaciones **concurrentes** produce una nueva ecuación **concurrente** con las originales.
2. La **combinación lineal** de ecuaciones **paralelas** produce una nueva ecuación **paralela** a las originales.

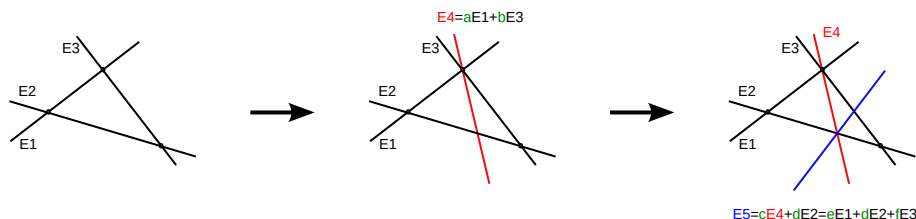
Ejercicio 1.5. Dadas las ecuaciones en $\mathbb{R}^2$: $\begin{cases} E_1 : x - y = 1 \\ E_2 : 2x + 3y = 4 \end{cases}$

Calcular y representar gráficamente la combinación lineal de ecuaciones $E_3 = -3E_1 + 2E_2$.



$$\begin{array}{rcl} (-3) \times & x & -y = 1 \\ (2) \times & 2x & +3y = 4 \\ \hline & x & +9y = 5 \end{array}$$

Ejercicio 1.6. Determinar de forma gráfica una posible *combinación lineal* de las tres rectas E_1 , E_2 y E_3 en $\mathbb{R}^2$ que muestra el siguiente gráfico.



1.4.4. Dependencia e Independencia lineal de Ecuaciones

Definición 1.21. Ecuaciones Linealmente Dependientes e Independientes

- **Ec. Linealmente Dependientes:** aquellas que pueden formarse a partir de alguna *combinación lineal* no nula³³ de otras ecuaciones.
- **Ec. Linealmente Independientes:** aquellas que no pueden formarse por ninguna *combinación lineal* no nula de otras ecuaciones.

- En un espacio $\mathbb{R}^n$ solo puede haber como máximo n ecuaciones *linealmente independientes*.
- La *fórmula de equilibrio*³⁴ solo es aplicable a SEL formados por ecuaciones *linealmente independientes*.

³²Es decir, ecuaciones multiplicadas por una constante, $a_i E_i$. Ver la **Sección 1.4.2, página 42**.

³³Aquellas combinaciones donde no todos los coeficientes que multiplican a las ecuaciones son nulos.

³⁴Ver la **Definición 1.15, página 37**.

Ejercicio 1.7. Dependencia lineal.

Las tres ecuaciones E_1 , E_2 y E_3 del **Ejemplo 1.5** forman un conjunto de objetos geométricos (rectas en $\mathbb{R}^2$) *concurrentes no coincidentes pero linealmente dependientes* y por tanto en este caso no se puede aplicar la *fórmula de equilibrio* para determinar la dimensión de la solución del sistema formado por estas tres ecuaciones. Claramente la solución del mismo es un punto (objeto geométrico de dimensión 0), mientras que la *fórmula de equilibrio* daría un objeto de dimensión $2 - 3 = -1$, lo cual no tiene sentido.

1.4.5. Sistemas Equivalentes**Definición 1.22.** Sistemas Equivalentes

Se dice que dos sistemas de ecuaciones son **equivalentes** si poseen el *mismo conjunto de soluciones*.

Ejemplo 1.19. Sistemas equivalentes.

- Los dos sistemas de $\mathbb{R}^2$:

$$\begin{cases} x - y = 1 \\ x + y = 3 \end{cases} \quad ; \quad \begin{cases} x - y = 1 \\ y = 1 \end{cases}$$

son equivalentes, ya que ambas tienen por única solución $x = 2, y = 1$.

- Los dos sistemas de $\mathbb{R}^3$:

$$\begin{cases} x + y - z = 2 \end{cases} \quad ; \quad \begin{cases} 3x + 3y - 3z = 6 \end{cases}$$

son equivalentes, ya que ambas tienen por solución el plano $z = x + y - 2$.

1.4.6. Operaciones Elementales de Fila**Definición 1.23.** Operaciones Elementales de Fila

Un sistema de ecuaciones lineales puede ser transformado en otro *sistema equivalente* (con las mismas soluciones), aplicando una o varias de las siguientes operaciones, denominadas **operaciones elementales de fila**:

- **Intercambiar el orden de dos ecuaciones entre sí.**

Este cambio no altera el conjunto de ecuaciones y por tanto se obtiene un *sistema equivalente*.

- **Multiplicar una ecuación (ambas partes de la igualdad) por una constante distinta de cero.**

Se obtiene así una *ecuación equivalente* a la original y por tanto el nuevo sistema es *equivalente* al anterior.

- **Sumar un múltiplo de una ecuación a otra ecuación.**

Este proceso se puede realizar múltiples veces³⁵. La nueva ecuación será *concurrente*³⁶ con el conjunto de ecuaciones elegido y por tanto el nuevo sistema es *equivalente* al original.

³⁵Dando lugar a una *combinación lineal* de ecuaciones. Ver la **Sección 1.4.3, página 43**.

³⁶Ver la **Definición 1.16, página 38**.

Las *operaciones elementales de fila* se pueden utilizar para transformar cualquier SEL en un sistema *escalonado equivalente* fácilmente resoluble.

Notación de las Operaciones Elementales de Fila

- $E_i \leftrightarrow E_j$: Intercambiar las ecuaciones E_i y E_j .
- kE_i : Multiplicar la ecuación E_i por el escalar $k \neq 0$.
- $E_i + kE_j$: Sustituir³⁷ E_i por una **comb. lineal** de ec.

Ejemplo 1.20. Ejemplos de *operaciones elementales de fila* sobre un SEL.

$$\left\{ \begin{array}{rcl} 3x + 2y - z & = & 1 \\ x - y + z & = & 2 \\ 2x + y + 3z & = & 3 \end{array} \right\} \xrightarrow{-E_1} \left\{ \begin{array}{rcl} -3x - 2y + z & = & -1 \\ x - y + z & = & 2 \\ 2x + y + 3z & = & 3 \end{array} \right\} \xrightarrow{E_2 \quad E_3} \left\{ \begin{array}{rcl} -3x - 2y + z & = & -1 \\ 2x + y + 3z & = & 3 \\ x - y + z & = & 2 \end{array} \right\}$$

$$\xrightarrow{-E_3 + 2E_1 + E_2} \left\{ \begin{array}{rcl} -3x - 2y + z & = & -1 \\ 2x + y + 3z & = & 3 \\ -5x - 2y + 4z & = & -1 \end{array} \right\}$$

Todos estos sistemas intermedios y cualquier otro SEL resultado de *operaciones elementales de fila* son *equivalentes* al inicial y por tanto poseen el mismo conjunto de soluciones.

1.4.7. Sistemas Escalonados o Triangulares

Existen sistemas de ecuaciones lineales con una estructura especial que permite su resolución sistemática³⁸ de forma fácil, *independientemente* de la dimensión del sistema. Este tipo de sistemas se denominan *escalonados* o *triangulares*.

Definición 1.24. Sistemas Escalonados o Triangulares

Se dice que un sistema está en su **forma escalonada** si permite colocar las ecuaciones con las incógnitas *ordenadas por columnas*, de forma que el primer coeficiente distinto de cero en cada ecuación, denominado **pivote**, se encuentra en alguna columna a la izquierda de cualquier *pivote* debajo de ella.

Ejemplo 1.21. Ejemplo de sistemas en forma escalonada.

$$\left\{ \begin{array}{rcl} x & - & y & - & z & = & 2 \\ & & y & + & 3z & = & 5 \\ & & & & 5z & = & 10 \end{array} \right. ; \quad \left\{ \begin{array}{rcl} x_1 & + & x_2 & - & 3x_3 & + & x_4 & - & x_5 & = & 2 \\ & & x_2 & - & x_3 & - & 2x_4 & + & x_5 & = & 0 \\ & & & & & & 2x_4 & - & 3x_5 & = & -1 \\ & & & & & & & & - & 3x_5 & = & 9 \end{array} \right.$$

Nota 1.2. La forma escalonada que veremos aquí se denomina *triangular superior* y suele ser la más utilizada, sin embargo también se podría utilizar la forma escalonada *triangular inferior*³⁹, siendo el planteamiento del problema y su resolución de forma análoga a la explicada aquí.

³⁷Esta operación se puede repetir varias veces, de forma que mediante esta notación indicamos que se sustituye una ecuación E_i por una combinación lineal de la propia E_i con algunas o todas las ecuaciones restantes, de forma que la nueva ecuación E'_i será:

$$E'_i = aE_i + bE_j + cE_k + \dots ; \quad a \neq 0$$

Notar que el coeficiente a no puede ser nulo ya que de otra forma se perdería la contribución de la ecuación original E_i .

³⁸Es decir, siguiendo siempre un mismo método que consta de una serie de pasos simples bien definidos.

³⁹Las primeras ecuaciones del SEL más cortas que las últimas.

Ejemplo 1.22. Los siguientes sistemas **no** están en *forma escalonada*.

$$\left\{ \begin{array}{rcl} x & - & y & - & z & = & 2 \\ -x & + & y & + & 3z & = & 5 \\ & & & & 5z & = & 10 \end{array} \right. ; \quad \left\{ \begin{array}{rcl} x_1 & + & x_2 & - & 3x_3 & + & x_4 & - & x_5 & = & 2 \\ & & x_2 & - & x_3 & - & 2x_4 & + & x_5 & = & 0 \\ & & & & & & 2x_4 & - & 3x_5 & = & -1 \\ & & & & - & x_3 & + & x_4 & - & 3x_5 & = & 9 \end{array} \right.$$

Método de Sustitución hacia Atrás (MSA)

El método más apropiado para resolver los sistemas escalonados es el *método de sustitución hacia atrás* (MSA).

Definición 1.25. Método de Sustitución Hacia Atrás (MSA)

Consiste en ir resolviendo las ecuaciones del *sistema escalonado* en orden, desde la última ecuación que tiene menos incógnitas hasta la primera que tiene más incógnitas, *sustituyendo* las incógnitas resueltas por los valores encontrados⁴⁰.

Ejercicio 1.8. Resolver los SEL del **Ejemplo 1.22** utilizando el *método de sustitución hacia atrás*.

$$\left\{ \begin{array}{rcl} x & - & y & - & z & = & 2 \\ & & y & + & 3z & = & 5 \\ & & & & 5z & = & 10 \end{array} \right. \rightarrow \begin{array}{l} x = 2 + y + z = 2 - 1 + 2 = 3 \\ y = 5 - 3z = 5 - 3 \cdot (2) = -1 \\ z = 10/5 = 2 \end{array} ; \quad \begin{array}{|l} x = 3 \\ y = -1 \\ z = 2 \end{array} \quad \begin{array}{l} \text{paso 3} \\ \text{paso 2} \\ \text{paso 1} \end{array}$$

$$\left\{ \begin{array}{rcl} x_1 & + & x_2 & - & 3x_3 & + & x_4 & - & x_5 & = & 2 \\ & & x_2 & - & x_3 & - & 2x_4 & + & x_5 & = & 0 \\ & & & & & & 2x_4 & - & 3x_5 & = & -1 \\ & & & & & & & & - & 3x_5 & = & 9 \end{array} \right. \rightarrow \begin{array}{l} x_1 = -x_2 + 3x_3 - x_4 + x_5 \\ x_2 = x_3 + 2x_4 - x_5 \\ x_4 = \frac{1}{2}(-1 + 3x_5) \\ -3x_5 = 9 \end{array} ; \quad \begin{array}{|l} x_1 = 2t + 9 \\ x_3 = t \\ x_4 = -5 \\ x_5 = -3 \end{array}$$

Ventajas de los SEL Triangulares

Los SEL generales de dimensión n necesitan un número de operaciones de orden n^2 y por tanto el tiempo para resolverlo también es proporcional a esa cantidad (en función de la velocidad del ordenador utilizado). Por otra parte, los SEL triangulares de la misma dimensión necesitan un número de operaciones de orden n , consecuencia del *método de sustitución hacia atrás* que se utiliza en su resolución.

Esto es un punto clave para grandes valores de n como puede verse en la **Figura 1.1**, página 47, donde para un caso particular de ordenador, se observa la enorme ventaja de los SEL triangulares a medida que aumenta el tamaño de un SEL.

1.4.8. Método de Gauss o Eliminación Gaussiana

El método de Gauss o eliminación gaussiana⁴¹ es apropiado para la resolución de sistemas de ecuaciones lineales generales *a mano y con un ordenador*, y está basado en las ventajas que muestran los *SEL triangulares o escalonados* frente a los SEL generales.

Definición 1.26. Método de Gauss o Eliminación Gaussiana

Consiste en la transformación de un SEL general en otro SEL *equivalente*⁴² con *forma escalonada*, mediante la utilización de *operaciones elementales por fila*⁴³. El SEL *escalonado* resultante es fácilmente resoluble mediante el *método de sustitución hacia atrás*.

⁴⁰Si el sistema escalonado es *triangular inferior*, consiste en ir resolviendo las ecuaciones en orden inverso.

⁴¹El matemático alemán Carl Friedrich Gauss (1777-1855) desarrolló este método sobre 1794 al resolver un sistema de 17 ecuaciones con 17 incógnitas para determinar la órbita del planeta *Ceres*. Sin embargo, existen evidencias de que este método fue utilizado por los chinos hace 2000 años.

⁴²Ver la **Definición 1.22**, página 44.

⁴³Ver la **Definición 1.23**, página 44.

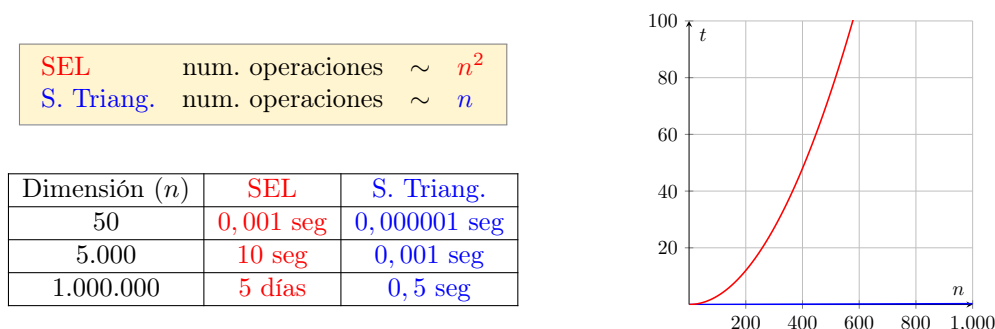


Figura 1.1: Comparación de la velocidad de resolución de un método general de SEL (en rojo) y el método de sustitución hacia atrás para SEL Triangulares (en azul). El tiempo (t) de ejecución es directamente proporcional al número de operaciones necesarios para su resolución. Como se observa en la gráfica, en los SEL generales (en rojo) el tiempo t crece de forma cuadrática con el tamaño n del sistema, mientras que el caso de los SEL triangulares (en azul) el crecimiento es lineal con una pendiente muy pequeña.

Ejercicio 1.9. Resolver mediante el método de *eliminación gaussiana* el siguiente SEL de $\mathbb{R}^3$.

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ 2x_1 + 3x_2 + x_3 = 5 \\ x_1 - x_2 - 2x_3 = -5 \end{cases}$$

Vamos a realizar diferentes operaciones elementales de fila de forma que se eliminen variables en las ecuaciones inferiores hasta obtener un sistema en forma escalonada.

$$\begin{aligned} & \left\{ \begin{array}{ccc} x_1 + x_2 + x_3 = 3 \\ 2x_1 + 3x_2 + x_3 = 5 \\ x_1 - x_2 - 2x_3 = -5 \end{array} \right\} \xrightarrow{-E_2 + 2E_1} \left\{ \begin{array}{ccc} x_1 + x_2 + x_3 = 3 \\ -x_2 + x_3 = 1 \\ x_1 - x_2 - 2x_3 = -5 \end{array} \right\} \\ & \xrightarrow{-E_3 + E_1} \left\{ \begin{array}{ccc} x_1 + x_2 + x_3 = 3 \\ -x_2 + x_3 = 1 \\ 2x_2 + 3x_3 = 8 \end{array} \right\} \xrightarrow{E_3 + 2E_2} \left\{ \begin{array}{ccc} x_1 + x_2 + x_3 = 3 \\ -x_2 + x_3 = 1 \\ 5x_3 = 10 \end{array} \right\} \end{aligned}$$

Aplicando ahora el método de sustitución hacia atrás:

$$\boxed{x_3 = 2} \quad ; \quad \boxed{x_2 = x_3 - 1 = 1} \quad ; \quad \boxed{x_1 = 3 - x_2 - x_3 = 0}$$

Comprobamos el resultado:

$$\begin{cases} x_1 + x_2 + x_3 = 0 + 1 + 2 = 3 \\ 2x_1 + 3x_2 + x_3 = 2 \cdot 0 + 3 \cdot 1 + 2 = 5 \\ x_1 - x_2 - 2x_3 = 0 - 1 - 2 \cdot 2 = -5 \end{cases}$$

1.5. Matrices

En el cálculo de las *operaciones elementales* del *método de Gauss*, es mucho más práctico operar únicamente con los *coeficientes* y *términos independientes* de cada ecuación, sin incluir las *variables*⁴⁴. Para este fin se utilizan las *matrices*.

Aunque se estudiarán en detalle en el **Tema 3**, vamos a recordar la definición de *matriz* numérica.

Definición 1.27. Matriz

Una **matriz** es una disposición de números⁴⁵ ordenados en *filas* y *columnas*.

⁴⁴Ver la **Definición 1.10**, página 34.

⁴⁵Números reales o complejos.

Una matriz A de m filas y n columnas se representa por $A_{m \times n}$ y utilizaremos *corchetes cuadrados* $[]$ en su notación de la siguiente forma:

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad ; \quad a_{ij} \in \mathbb{K} \quad ; \quad A \in \mathbb{K}^{m \times n}$$

donde $\mathbb{K}^{m \times n}$ es el *espacio de las matrices* (reales o complejas) $A_{m \times n}$.

Ejemplo 1.23. Ejemplos de matrices reales y complejas.

$$A_{2 \times 2} = \begin{bmatrix} 1 & -1 \\ 2 & -3 \end{bmatrix} \in \mathbb{R}^{2 \times 2} \quad ; \quad A_{3 \times 4} = \begin{bmatrix} 0 & -2i & 3+i & i \\ -1 & -1+i & 1 & 0 \\ -2 & 4i & -2 & -3 \end{bmatrix} \in \mathbb{C}^{3 \times 4}$$

1.5.1. Matrices Asociadas a SEL

Los SEL se pueden representar de forma matricial asociando matrices a los coeficientes y términos independientes.

Definición 1.28. Matriz de Coeficientes y Matriz Ampliada

Todo SEL de m ecuaciones y n incógnitas se puede representar mediante dos tipos de matrices:

- **Matriz de Coeficientes:** Matriz $\mathbb{K}^{m \times n}$ conteniendo los coeficientes ordenados del sistema.
- **Matriz Ampliada:** Matriz $\mathbb{K}^{m \times (n+1)}$ conteniendo los coeficientes del sistema (*matriz ampliada*) junto con el *vector columna*⁴⁶ de los términos independientes.

Dado un SEL de m ecuaciones con n incógnitas:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad \text{en } \mathbb{K}^n$$

La *matriz de coeficientes* A y la *matriz ampliada* $A \ B$ asociadas a este sistema son:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{K}^{m \times n} \quad ; \quad A \ B = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right] \in \mathbb{K}^{m \times (n+1)}$$

Ejemplo 1.24. Ejemplo de matriz ampliada asociada a un SEL.

$$\begin{cases} 3x + 2y - z = 1 \\ x - y + z = 2 \\ 2x + y + 3z = 3 \end{cases} \quad \text{en } \mathbb{R}^3 \quad = \quad \left[\begin{array}{ccc|c} 3 & 2 & -1 & 1 \\ 1 & -1 & 1 & 2 \\ 2 & 1 & 3 & 3 \end{array} \right]$$

⁴⁶Ver la [Definición 2.3](#), página 65.

1.5.2. Operaciones Elementales de Fila en la Matriz Asociada

Las *operaciones elementales de fila* realizadas en el método de Gauss sobre un SEL para transformarlo en otro *sistema equivalente escalonado*⁴⁷, se pueden realizar de forma más conveniente en su correspondiente *matriz ampliada asociada*.

Definición 1.29. Operaciones Elementales de Fila en Matrices

- **Intercambiar dos filas entre sí.** $(F_i \quad F_j)$
Esta operación corresponde al cambio de orden de las ecuaciones.
- **Multiplicar una fila por una constante distinta de cero.** (kF_i)
Esta operación corresponde con la sustitución de una ecuación por otra equivalente (coincidente) a la original.
- **La adición de un múltiplo de una fila a otra fila.** $(F_i + kF_j)$
Esta operación corresponde con la sustitución de una determinada ecuación por una combinación lineal de ecuaciones del sistema⁴⁸.

Nota 1.3. La notación es igual a la utilizada en las *operaciones elementales de Fila* aplicadas a las ecuaciones, cambiando ecuaciones (E) , por filas (F) .

1.5.3. Matriz Equivalente por Filas

Definición 1.30. Matriz Equivalente por Filas

Las *operaciones elementales por fila* aplicadas en una matriz la transforman en otra matriz denominada **matriz equivalente por filas**.

Ejemplo 1.25. Operaciones elementales sobre la matriz asociada al SEL del [Ejemplo 1.24](#).

$$\left[\begin{array}{ccc|c} 3 & 2 & -1 & 1 \\ 1 & -1 & 1 & 2 \\ 2 & 1 & 3 & 3 \end{array} \right] \xrightarrow{-F_1} \left[\begin{array}{ccc|c} -3 & -2 & 1 & -1 \\ 1 & -1 & 1 & 2 \\ 2 & 1 & 3 & 3 \end{array} \right] \xrightarrow{F_2 \quad F_3} \left[\begin{array}{ccc|c} -3 & -2 & 1 & -1 \\ 2 & 1 & 3 & 3 \\ 1 & -1 & 1 & 2 \end{array} \right]$$

$$\xrightarrow{-F_3 + 2F_1 + F_2} \left[\begin{array}{ccc|c} -3 & -2 & 1 & -1 \\ 2 & 1 & 3 & 3 \\ -5 & -2 & 4 & -1 \end{array} \right]$$

Nota 1.4. Las propiedades de las *matrices equivalentes por filas* se estudiarán en el [Sección 3.10.2, página 113](#).

⁴⁷Ver la [Sección 1.4.6, página 44](#).

⁴⁸De la misma forma que se indicó en la [Sección 1.4.6, página 45](#), mediante esta notación queremos indicar que se sustituye una fila F_i por una combinación lineal de la propia F_i con algunas o todas las filas restantes, de forma que la nueva fila F'_i será:

$$F'_i = aF_i + bF_j + cF_k + \cdots \quad ; \quad a \neq 0$$

1.5.4. Matriz Escalonada Equivalente

El método de Gauss⁴⁹ puede aplicarse a la matriz ampliada del SEL, transformándola mediante *operaciones elementales de fila* a una matriz *equivalente por filas*, en forma escalonada o triangular, correspondiente al SEL escalonado buscado.

Definición 1.31. Matriz Escalonada o Triangular

Una matriz escalonada (o triangular) es aquella que cumple las siguientes propiedades:

- Cualquier fila de ceros está en la parte inferior de la matriz.
- En cada fila no nula, el primer elemento distinto de cero, (denominado **elemento principal** o **pivote**⁵⁰), se encuentra al menos una columna a la izquierda de cualquier *pivote* de cualquier fila por debajo de ella.

Ejemplo 1.26. Ejemplo de matrices escalonadas.

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 5 \\ 0 & 0 & 4 \end{bmatrix} \quad \begin{bmatrix} 2 & 4 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

Definición 1.32. Matriz Escalonada Equivalente

- Se dice que la matriz escalonada de una dada, obtenida a través de operaciones elementales de fila, es una **matriz escalonada equivalente** a la original.
- El proceso de obtención de la matriz escalonada equivalente se denomina *reducción de una matriz a su forma escalonada equivalente*.

1.5.5. Rango de una Matriz

Definición 1.33. Rango de una Matriz

Se define el **rango**⁵¹ de una matriz como el número de filas *no nulas* que tiene en cualquiera de sus formas escalonadas equivalentes⁵². El rango de una matriz A se identifica a través de:

$$\text{rang}(A)$$

Se dice que una matriz $A \in \mathbb{R}^{m \times n}$ es de **rango completo** si $\text{rang}(A) = \min(m, n)$.

Ejercicio 1.10. Determinar el rango de la siguiente matriz: $A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \xrightarrow{F_2 - 2F_1} \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix} = \boxed{\text{rang}(A) = 2}$$

⁴⁹Ver la Sección 1.4.8, página 46.

⁵⁰De igual forma que se denominan a los primeros elementos no nulos en los SEL *escalonados* o *triangulares*, ver la Definición 1.24, página 45.

⁵¹En la Definición 3.24, página 114 se verá una definición alternativa del *rango* de una matriz.

⁵²Coincide con el número de *pivotes* de dicha matriz. Una matriz no nula tiene infinitas formas escalonadas equivalentes, pero todas ellas tienen el mismo número de filas no nulas, por tanto el rango de la matriz se puede determinar con cualquiera de ellas.

1.5.6. Teorema de Rouché-Fröbenius

Teorema 1.1. Teorema de Rouché-Fröbenius

Sea $A \in \mathbb{R}^{m \times n}$ una matriz de coeficientes correspondiente a un sistema de m ecuaciones con n incógnitas, y sea $A|B$ su matriz ampliada. Entonces, el sistema es **compatible** (tiene solución) si y solo si⁵³:

$$\text{rang}(A) = \text{rang}(A|B)$$

- Si $\text{rang}(A) = n$ = Sistema *compatible determinado* (solución única).
- Si $\text{rang}(A) < n$ = Sistema *compatible indeterminado* (infinitas soluciones).

$$\text{Num. variables libres (parámetros)} = n - \text{rang}(A)$$

Nota 1.5. El número de *variables libres* coincide con la *dimensión* de la solución.

1.6. Método de Gauss Matricial

Dado el sistema general de m ecuaciones con n incógnitas:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad \text{en } \mathbb{K}^n$$

la aplicación del método de Gauss (eliminación gaussiana) utilizando la matriz ampliada (forma matricial) consiste en los siguientes pasos:

1. Obtener la matriz ampliada del SEL.

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right]$$

2. Seleccionamos la primera fila, $i = 1$.

3. Tomamos como *pivote*⁵⁴ el elemento $p = a_{ii} = 0$. Si $p = 0$ se intercambia la fila i con cualquier otra por debajo hasta que $p = a_{ii} \neq 0$.

4. Mediante operaciones elementales por fila, hacer ceros todos los elementos de la columna i del *pivote* p , por debajo de él. Para ello, cada fila F_j por debajo del *pivote* ($j > i$) se substituye por la combinación lineal:

$$F_j = F_j - \frac{a_{ji}}{a_{ii}} F_i \quad ; \quad j > i$$

5. Avanzamos a la siguiente fila $i = i+1$, repitiendo los pasos 3-5 hasta completar todas las filas, obteniéndose finalmente una *matriz escalonada*⁵⁴ *ampliada*.

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & p_{22} & \cdots & a'_{2n} & b'_2 \\ \vdots & \vdots & & \vdots & \vdots \end{array} \right]$$

⁵³Este teorema también es conocido con el nombre de *Teorema del Rango*, sin embargo esta forma es más confusa al existir más de un teorema del rango en diferentes áreas.

⁵⁴Ver la [Definición 1.31](#), página 50.

Ejercicio 1.11. Resolver mediante eliminación gaussiana el siguiente SEL en $\mathbb{R}^3$.

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ 2x_1 + 3x_2 + x_3 = 5 \\ x_1 - x_2 - 2x_3 = -5 \end{cases}$$

Utilizando el método de Gauss matricial se tiene, se tiene,

$$\begin{bmatrix} 1 & 1 & 1 & | & 3 \\ 2 & 3 & 1 & | & 5 \\ 1 & -1 & -2 & | & -5 \end{bmatrix} \xrightarrow{F_2 - 2F_1} \begin{bmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 1 & -1 & | & -1 \\ 1 & -1 & -2 & | & -5 \end{bmatrix} \xrightarrow{F_3 - F_1} \begin{bmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 1 & -1 & | & -1 \\ 0 & -2 & -3 & | & -8 \end{bmatrix}$$

$$\xrightarrow{F_3 + 2F_2} \begin{bmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 1 & -1 & | & -1 \\ 0 & 0 & -5 & | & -10 \end{bmatrix}$$

Utilizando el método de sustitución hacia atrás, se obtiene la solución:

$$\boxed{x_3 = 2} \quad ; \quad \boxed{x_2 = x_3 - 1 = 1} \quad ; \quad \boxed{x_1 = 3 - x_2 - x_3 = 0}$$

1.6.1. Identificación de Sistemas Incompatibles

Si una o más filas de la *matriz escalonada ampliada*⁵⁵ tiene todos sus coeficientes *nulos* menos el último correspondiente al término independiente, el sistema es **incompatible**⁵⁶, es decir, no tiene solución.

Ejercicio 1.12. Resolver mediante eliminación gaussiana el siguiente SEL en $\mathbb{R}^3$.

$$\begin{cases} x_1 - x_2 + 2x_3 = 3 \\ x_1 + 2x_2 - x_3 = -3 \\ 2x_2 - 2x_3 = 1 \end{cases}$$

$$\begin{bmatrix} 1 & -1 & 2 & | & 3 \\ 1 & 2 & -1 & | & -3 \\ 0 & 2 & -2 & | & 1 \end{bmatrix} \xrightarrow{F_2 - F_1} \begin{bmatrix} 1 & -1 & 2 & | & 3 \\ 0 & 3 & -3 & | & -6 \\ 0 & 2 & -2 & | & 1 \end{bmatrix} \xrightarrow{\frac{1}{3}F_2} \begin{bmatrix} 1 & -1 & 2 & | & 3 \\ 0 & 1 & -1 & | & -2 \\ 0 & 2 & -2 & | & 1 \end{bmatrix}$$

$$\xrightarrow{F_3 + 2F_2} \begin{bmatrix} 1 & -1 & 2 & | & 3 \\ 0 & 1 & -1 & | & -2 \\ 0 & 0 & 0 & | & 5 \end{bmatrix}$$

La última fila de esta matriz ampliada corresponde con la ecuación:

$$0x_1 + 0x_2 + 0x_3 = 5$$

corresponde para todos los valores de las variables x_1 , x_2 y x_3 a la ecuación $0 = 5$, que no tiene solución y por tanto el sistema es **incompatible**⁵⁷.

⁵⁵Ver la Sección 1.6, página 51.

⁵⁶Ya que el rango de la matriz de coeficientes es menor que el rango de la matriz ampliada y por el teorema de Rouché-Fröbenius, (ver el Teorema 1.1, página 51), el sistema es incompatible.

⁵⁷Recordar que la solución de un sistema de ecuaciones (ver la Definición 1.7, página 31) es el conjunto de valores que asignado a las variables verifican todas y cada una de las ecuaciones del sistema. Por tanto, con que exista una sola ecuación del sistema que no tenga solución, el sistema será incompatible.

1.6.2. Identificación de Sistemas Compatibles Indeterminados

Si el SEL no es incompatible y su *matriz escalonada ampliada* tiene menos pivotes que incógnitas⁵⁸, el sistema es **compatible indeterminado**, teniendo infinitas soluciones.

Ejercicio 1.13. Analizar si el siguiente SEL es compatible (determinado o indeterminado) o incompatible

$$\begin{cases} x_1 + x_2 + 2x_3 = 1 \\ x_1 + 2x_2 + 2x_3 = 1 \\ x_1 + 2x_3 = 1 \end{cases}$$

Vamos a expresar el SEL en forma matricial, obtener una matriz ampliada escalonada equivalente y analizar el tipo de solución.

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 1 & 0 & 2 & 1 \end{array} \right] \xrightarrow[F_3-F_2]{F_2-F_1} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array} \right] \xrightarrow{F_3+F_2} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

En este caso se obtiene una fila completa de ceros en la matriz ampliada escalonada resultante y por tanto el sistema es **compatible indeterminado**.

- Una fila completa llena de ceros en una matriz ampliada corresponde con una ecuación del SEL que es *linealmente dependiente*⁵⁹ del resto, y por tanto no da información adicional al sistema, por lo que puede eliminarse del mismo.
- Toda fila de la matriz ampliada que en su forma *esalonada equivalente* tenga *pivote*⁶⁰, corresponden con ecuaciones *linealmente independientes*, que son las aportan información relevante al sistema.

Vamos a estudiar las características especiales de los SEL *compatibles indeterminados* para encontrar una forma apropiada de representar sus infinitas soluciones.

Definición 1.34. Variables Principales y Libres

En un SEL *compatible indeterminado* hay dos tipos de variables:

- **Variables Principales:** correspondientes (en número) a las filas no nulas y por tanto al número de ecuaciones linealmente independientes del sistema.
- **Variables Libres:** diferencia entre variables totales (n) y variables principales,

$$\text{Variables Libres} = n - \text{Variables Principales}$$

- Cualquiera de las variables del sistema puede tomarse como *principal* o *libre*.
- Las *variables libres* pueden tener cualquier valor y por tanto normalmente se les asocia un **parámetro libre**.
- Las variables *principales* dependen de los valores tomados por las variables *libres*.

⁵⁸A veces en estos casos aparecen una o más filas completas de ceros.

⁵⁹Ver la [Definición 1.21, página 43](#).

⁶⁰Ver la [Definición 1.24, página 45](#).

Ejercicio 1.14. Resolver mediante eliminación gaussiana el siguiente SEL de $\mathbb{R}^4$.

$$\begin{cases} w - x - y + 2z = 1 \\ 2w - 2x - y + 3z = 3 \\ -w + x - y = -3 \end{cases}$$

Por el teorema de Rouché-Fröbenius ([Teorema 1.1, página 51](#)) este sistema solo puede ser *incompatible* o *compatible indeterminado* por tener más incógnitas (4) que ecuaciones (3). Vamos utilizar el método de Gauss matricial para resolverlo:

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array} \right] \xrightarrow{F_2 - 2F_1} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ -1 & 1 & -1 & 0 & -3 \end{array} \right] \xrightarrow{F_3 + F_1} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & -2 \end{array} \right] \xrightarrow{F_3 + 2F_2} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] = \text{Sistema compatible indeterminado.}$$

La fila llena de ceros no se tiene en cuenta, ya que corresponde con una *ecuación linealmente dependiente* de las demás, (ver la [Sección 1.4.4, página 43](#)), por tanto el SEL asociado a esta última matriz es ahora:

$$\begin{cases} w - x - y + 2z = 1 \\ y - z = 1 \end{cases}$$

El cual es un sistema de 2 ecuaciones con 4 incógnitas y por tanto tiene un número infinito de soluciones. Como tenemos 4 incógnitas y 2 ecuaciones, se pueden tomar 2 *variables libres*:

$$4 \text{ variables totales} - 2 \text{ ecuaciones linealmente independientes} = 2 \text{ variables libres}$$

Estas dos *variables libres* pueden ser cualquiera de las cuatro existentes w, x, y, z y se les asignará a cada una un *parámetro libre*. El resto de variables serán *variables principales*.

Resolvemos el sistema mediante el *método de sustitución hacia atrás*:

- De la segunda ecuación podemos despejar cualquiera de las dos variables y o z . Despejando por ejemplo la variable y (variable principal) en función de la variable z (variable libre) se tiene:

$$y = 1 + z$$

La variable libre z puede tomar cualquier valor y por ello le asignamos un *parámetro libre*, por ejemplo t :

$$z = t ; \quad t \in \mathbb{R} \quad = \quad y = 1 + z = 1 + t$$

- De la primera ecuación podemos despejar w o x en función de las demás. Despejando por ejemplo w (variable principal) y asignando a la *variable libre* x un *parámetro libre* s se tiene:

$$x = s ; \quad s \in \mathbb{R} \quad = \quad w = 1 + x + y - 2z = 1 + s + (1 + t) - 2t = 2 - t + s$$

La solución, en *forma paramétrica*⁶¹, se puede expresar como:

$$\begin{cases} x = s \\ y = 1 + t \\ z = t \\ w = 2 - t + s \end{cases} ; \quad t, s \in \mathbb{R}$$

Dando valores a los parámetros t y s se obtienen las infinitas soluciones del sistema.

Esta solución en forma paramétrica no es única ya que existen infinitas *formas paramétricas* de expresar la solución de un sistema, en función de las variables que asignemos como libres o principales. Sin embargo, todas las diferentes formas expresan exactamente la misma solución⁶².

En muchos casos interesa tener la *forma vectorial* de la solución, que en este caso se puede expresar como:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} s \\ 1+t \\ t \\ 2-t+s \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 2 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Esta solución corresponde geoméricamente con un *plano*⁶³ en un espacio de dimensión 4, (correspondiente al número de incógnitas).

De la misma manera que la *forma paramétrica*, existen infinitas *formas vectoriales* de representar la misma solución de un sistema.

1.6.3. Estudio de Soluciones de SEL con Parámetros

El método matricial de Gauss es muy apropiado para el estudio de las soluciones de un SEL que contiene uno o varios parámetros⁶⁴.

El procedimiento a seguir consiste en obtener la matriz escalonada del sistema en función de los parámetros, y entonces estudiar los valores de estos parámetros que anulan los pivotes y términos independientes, dando lugar a los distintos tipos de sistemas: compatibles determinados e indeterminados, e incompatibles.

Ejercicio 1.15. Estudiar el siguiente sistema de ecuaciones en función del parámetro $a \in \mathbb{R}$.

$$\begin{cases} ax - 2y + (a-1)z = 4 \\ -2x + 3y - 6z = 2 \\ -ax + y - 6z = 6 \end{cases}$$

Formamos la matriz ampliada del sistema y aplicamos el método de Gauss hasta obtener la matriz escalonada dependiente del parámetro a . Tener en cuenta que siempre que sea posible, suele resultar más sencillo que los pivotes no contengan parámetros.

$$\begin{aligned} \left[\begin{array}{ccc|c} a & -2 & a-1 & 4 \\ -2 & 3 & -6 & 2 \\ -a & 1 & -6 & 6 \end{array} \right] & \xrightarrow[F_1 \leftrightarrow F_2]{F_3 + F_1} \left[\begin{array}{ccc|c} -2 & 3 & -6 & 2 \\ a & -2 & a-1 & 4 \\ 0 & -1 & a-7 & 10 \end{array} \right] \xrightarrow[F_2 - aF_1]{-\frac{1}{2}F_1} \left[\begin{array}{ccc|c} 1 & -\frac{3}{2} & 3 & -1 \\ 0 & -2 + \frac{3}{2}a & -1 - 2a & 4 + 4 \\ 0 & -1 & a-7 & 10 \end{array} \right] \\ & \xrightarrow[-F_3]{F_2 + (-2 + \frac{3}{2}a)F_3} \left[\begin{array}{ccc|c} 1 & -\frac{3}{2} & 3 & -1 \\ 0 & 0 & \frac{3}{2}a^2 - \frac{29}{2}a + 13 & 16a - 16 \\ 0 & 1 & 7 - a & -10 \end{array} \right] \end{aligned} \quad (1.1)$$

Estudiamos los valores de a que anulan la segunda fila de esta última matriz escalonada (1.1):

- $\frac{3}{2}a^2 - \frac{29}{2}a + 13 = 0$; $32a^2 - 29a + 26 = 0$; $a = \frac{29 \pm \sqrt{29^2 - 12 \cdot 26}}{6} = \frac{29 \pm 23}{6} = \left\{ 1, \frac{26}{3} \right\}$
- $16a - 16 = 0$; $a = 1$

Por tanto, teniendo en cuenta la matriz escalonada (1.1) podemos deducir:

- Si $a = \frac{26}{3}$ se anula la parte izquierda de la segunda fila pero no el término independiente y por tanto el sistema es **incompatible** (sin solución).

⁶¹Ver el Tema 0, página 1.

⁶²Conjunto de vales $[x, y, z, w]$ que cumplen el sistema de ecuaciones.

⁶³Objeto geométrico de dimensión 2, correspondiente con el número de parámetros libres que aparecen en su definición, ver la Sección 1.2.2, página 36.

⁶⁴Estos sistemas en realidad son un conjunto de infinitos SEL con una estructura determinada.

- Si $a = 1$ la segunda fila se anula totalmente y el sistema es **compatible indeterminado** (infinitas soluciones). Teniendo en cuenta los valores de la matriz escalonada (1.1) para $a = 1$, los valores de estas infinitas soluciones son:

$$\text{Fila 3 : } y + 6z = -10 \quad = \quad y = -10 - 6z$$

$$\text{Fila 1 : } x - \frac{3}{2}y + 3z = -1 \quad = \quad x = -1 - \frac{3}{2}y - 3z = 14 + 6z$$

Haciendo a la variable z el parámetro real t se tiene el siguiente conjunto infinito de soluciones para $a = 1$:

$$\boxed{x = 14 + 6t \ ; \ y = -10 - 6t \ ; \ z = t} \quad , \quad t \in \mathbb{R}$$

- Si $a = \{1, \frac{26}{3}\}$ el sistema es **compatible determinado**, cuyas soluciones únicas, teniendo en cuenta los valores de la matriz escalonada (1.1), son:

$$\text{Fila 2 : } \left(\frac{3}{2}a^2 - \frac{29}{2}a + 13\right)z = 16a - 16 \quad = \quad \boxed{z = \frac{16a - 16}{\frac{3}{2}a^2 - \frac{29}{2}a + 13}}$$

$$\text{Fila 3 : } y + (7 - a)z = -10 \quad = \quad \boxed{y = -10 + (a - 7)z = -10 + \frac{16(a - 1)(a - 7)}{\frac{3}{2}a^2 - \frac{29}{2}a + 13}}$$

$$\text{Fila 1 : } x - \frac{3}{2}y + 3z = -1 \quad = \quad \boxed{x = -1 + \frac{3}{2}y - 3z = 14 - \frac{24(a - 1)(a + 5)}{\frac{3}{2}a^2 - \frac{29}{2}a + 13}}$$

Ejercicio 1.16. Estudiar el siguiente sistema de ecuaciones en función del parámetro $m \in \mathbb{R}$.

$$\begin{cases} x + y + 2t = 3 \\ 3x - y + z - t = 1 \\ 5x - 3y + 2z - 4t = m \\ 2x + y + z + t = 2 \end{cases}$$

Formamos la matriz ampliada del sistema y aplicamos el método de Gauss hasta obtener la matriz escalonada dependiente del parámetro m .

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 3 \\ 3 & -1 & 1 & -1 & 1 \\ 5 & -3 & 2 & -4 & m \\ 2 & 1 & 1 & 1 & 2 \end{array} \right] \xrightarrow{F_j - a_{j1}F_1} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 3 \\ 0 & -4 & 1 & -7 & -8 \\ 0 & -8 & 2 & -14 & m - 15 \\ 0 & -1 & 1 & -3 & -4 \end{array} \right] \xrightarrow{\substack{F_3 - 2F_2 \\ F_2 + 4F_4}} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 3 \\ 0 & 0 & -3 & 5 & 8 \\ 0 & 0 & 0 & 0 & m + 1 \\ 0 & -1 & 1 & -3 & -4 \end{array} \right]$$

Si $m = -1$ el sistema es incompatible (sin solución). Para $m \neq -1$ el sistema es compatible indeterminado con infinitas soluciones, dadas por:

$$\text{Fila 2 : } -3z + 5t = 8 \quad = \quad t = \frac{8 + 3z}{5}$$

$$\text{Fila 4 : } -y + z - 3t = -4 \quad = \quad y = 4 + z - 3t = 4 + z - \frac{3}{5}(8 + 3z) = -\frac{4}{5} - \frac{4}{5}z$$

$$\text{Fila 1 : } x + y + 2t = 3 \quad = \quad x = 3 - y - 2t = 3 + \frac{4}{5} + \frac{4}{5}z - \frac{2}{5}(8 + 3z) = \frac{3}{5} - \frac{2}{5}z$$

Haciendo a la variable z el parámetro real α se tiene el siguiente conjunto infinito de soluciones para $m \neq -1$:

$$\boxed{x = \frac{3}{5} - \frac{2}{5}\alpha \ ; \ y = -\frac{4}{5} - \frac{4}{5}\alpha \ ; \ z = \alpha \ ; \ t = \frac{1}{5}(8 + 3\alpha)} \quad , \quad \alpha \in \mathbb{R}$$

1.7. Método de Gauss-Jordan

El método (de eliminación) de Gauss-Jordan⁶⁵ es una modificación del método de Gauss⁶⁶ que simplifica la fase de *sustitución hacia atrás*⁶⁷ en los casos compatibles determinados.

El método de Gauss es el que menor número de operaciones requiere en la resolución de SEL, sin embargo en algunos casos, las operaciones a realizar (aunque haya una mayor cantidad) resultan más sencillas de operar a mano siguiendo el método alternativo de Gauss-Jordan que veremos a continuación. Desde un punto de vista computacional (con un ordenador) solo tiene interés el método de Gauss, pero desde un punto de vista teórico el método de Gauss-Jordan tiene una gran importancia.

1.7.1. Matriz Escalonada Reducida

Definición 1.35. Matriz Escalonada Reducida

Una matriz *escalonada reducida* es una matriz escalonada con las siguientes dos características:

- Todas los *pivotes*⁶⁸ son la unidad, (denominado uno principal)
- Cada columna con *pivote*, tiene ceros en toda posición distinta del *pivote*.

Ejemplo 1.27. Las siguientes matrices están en forma *escalonada reducida*.

$$\begin{bmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} \end{bmatrix} ; \begin{bmatrix} \mathbf{1} & 2 & 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Definición 1.36. Matriz Escalonada Reducida Equivalente

Se define la **matriz escalonada reducida equivalente** de una determinada matriz, a la *matriz escalonada reducida* obtenida mediante *operaciones elementales de fila*.

La *matriz escalonada reducida equivalente* de una matriz A se puede obtener de dos formas:

- Obtener la *matriz escalonada equivalente*⁶⁹ y posteriormente conseguir los *pivotes* unidad y ceros correspondientes. Este método es computacionalmente preferible para un ordenador.
- Ir sacando los *pivotes* unidad a medida que se va obteniendo la matriz escalonada equivalente. Este método es más útil para cálculos a mano.

Unicidad de la Matriz Escalonada Reducida

Existen infinitas matrices escalonadas correspondientes a una matriz dada, en función de la colección de operaciones elementales de fila que se realicen en su obtención. Sin embargo, la matriz escalonada reducida correspondiente a una matriz dada es **única**, independientemente de las operaciones elementales por fila realizadas en su obtención.

⁶⁵W. Jordan (1842-1899) fue un ingeniero y profesor alemán de topografía geodésica cuya contribución a la solución de los sistemas lineales fue un método sistemático de sustitución hacia atrás directamente relacionado con el método de Gauss-Jordan.

⁶⁶Ver la Sección 1.4.8, página 46 y la Sección 1.6, página 51.

⁶⁷Ver la Definición 1.25, página 46.

⁶⁸Ver la Definición 1.31, página 50.

⁶⁹Ver la Definición 1.32, página 50.

Proceso del Método Gauss-Jordan

El método de Gauss-Jordan consiste en la obtención de la *matriz escalonada reducida* de un SEL. Si el sistema es *compatible determinado*⁷⁰, su solución vendrá dada directamente por la columna de los términos independientes, sin necesidad de aplicar el método de sustitución hacia atrás.

Ejercicio 1.17. Resolver mediante el método de Gauss-Jordan el siguiente SEL en $\mathbb{R}^3$.

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ 2x_1 + 3x_2 + x_3 = 5 \\ x_1 - x_2 - 2x_3 = -5 \end{cases}$$

$$\begin{aligned} & \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array} \right] \xrightarrow{F_2 - 2F_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 1 & -1 & -2 & -5 \end{array} \right] \xrightarrow{F_3 - F_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & -2 & -3 & -8 \end{array} \right] \\ & \xrightarrow{F_3 + 2F_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & -5 & -10 \end{array} \right] \xrightarrow{F_1 - F_2} \left[\begin{array}{ccc|c} 1 & 0 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & -5 & -10 \end{array} \right] \xrightarrow{-\frac{1}{5}F_3} \left[\begin{array}{ccc|c} 1 & 0 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 2 \end{array} \right] \\ & \xrightarrow{F_2 + F_3} \left[\begin{array}{ccc|c} 1 & 0 & 2 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array} \right] \xrightarrow{F_1 - 2F_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array} \right] \end{aligned}$$

Como el rango de la matriz de coeficientes es 3, igual que la dimensión del sistema, éste es *compatible y determinado*.

La solución es directamente la columna de *términos independientes*.

$$x_1 = 0 \quad ; \quad x_2 = 1 \quad ; \quad x_3 = 2$$

1.7.2. Matriz Identidad

Definición 1.37. Matriz Identidad

Se denomina **matriz identidad** de orden n a una matriz *cuadrada*⁷¹ de n filas y n columnas, cuyos elementos diagonales a_{ii} son todos la unidad y el resto cero.

$$I_n = \begin{bmatrix} \mathbf{1} & 0 & \cdots & 0 \\ 0 & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Un SEL de $\mathbb{R}^n$ tiene solución *única*, (compatible y determinado), si y solo si la *matriz escalonada reducida equivalente* a su matriz de coeficientes es una **matriz identidad** de orden n , (I_n).

⁷⁰Los SEL *incompatibles* y *compatibles indeterminados* se pueden identificar de igual forma que se vio en la [Sección 1.6.1, página 52](#) y [Sección 1.6.2, página 53](#).

⁷¹Con el mismo número de filas que de columnas. Ver los tipos de matrices en la [Sección 3.1.1, página 89](#).

1.8. Sistemas Homogéneos

Definición 1.38. Sistemas Homogéneos

Un SEL se denomina **homogéneo** si el término independiente en cada ecuación es cero.

El siguiente es un SEL homogéneo de m ecuaciones con n incógnitas:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

Por el teorema de Rouché-Fröbenius⁷² los SEL *homogéneos* son siempre *compatibles*, ya que⁷³ $\text{rang}(A) = \text{rang}(A \ B)$, pudiendo ser *determinados* o *indeterminados* en función de la matriz de coeficientes $A \in \mathbb{R}^{m \times n}$:

- Si $\text{rang}(A) = n$, el sistema es *compatible determinado*, y su solución única es la solución nula $\vec{x} = \vec{0} = [0, \dots, 0] \in \mathbb{R}^n$ conocida con el nombre de **solución trivial**.
- Si $\text{rang}(A) < n$, el sistema es *compatible indeterminado*, con infinitas soluciones, (una de ellas la solución trivial).

Ejercicio 1.18. Resolver el siguiente sistema de ecuaciones lineales homogéneo.

$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ 2x_1 + 3x_2 + x_3 = 0 \\ x_1 - x_2 - 2x_3 = 0 \end{cases}$$

Vamos a resolverlo aplicando el método de Gauss:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 3 & 1 & 0 \\ 1 & -1 & -2 & 0 \end{array} \right] \xrightarrow{F_2 - 2F_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & -2 & 0 \end{array} \right] \xrightarrow{F_3 - F_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -2 & -3 & 0 \end{array} \right]$$

$$\xrightarrow{F_3 + 2F_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -5 & 0 \end{array} \right]$$

Analizando la *matriz escalonada equivalente* del sistema podemos observar como el rango de la matriz de coeficientes es 3, igual que el número de incógnitas del sistema y por tanto corresponde con un sistema *compatible determinado*. Por el método de sustitución hacia atrás se comprueba fácilmente que la solución única es la *trivial*:

$$x = 0 \quad ; \quad y = 0 \quad ; \quad z = 0$$

Cuando se resuelven SEL homogéneos por el método de Gauss o Gauss-Jordan, la columna nula de términos independientes no se ve modificada por las operaciones elementales de fila y por tanto se puede prescindir de ella en los cálculos matriciales.

⁷²Ver el Teorema 1.1, página 51.

⁷³Debido a que la columna de ceros de los términos independientes no aportan nada al rango.

Ejercicio 1.19. Resolver el siguiente sistema de ecuaciones lineales.

$$\begin{cases} x_1 + x_2 + x_3 - x_4 = 0 \\ -x_1 + x_2 - x_3 = 0 \\ -x_2 - x_3 - x_4 = 0 \end{cases}$$

Observamos que es un SEL *homogéneo* de 3 ecuaciones con 4 incógnitas y por tanto es obligatoriamente un sistema *compatible indeterminado*, ya que el rango de la matriz de coeficientes como mucho es 3, mientras que el número de ecuaciones es 4. Vamos a resolverlo aplicando el método de Gauss, omitiendo la columna de ceros de los términos independientes por ser inalterable por las operaciones.

$$\left[\begin{array}{cccc} 1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & -1 & -1 \end{array} \right] \xrightarrow{F_2 + F_1} \left[\begin{array}{cccc} 1 & 1 & 1 & -1 \\ 0 & 2 & 0 & -1 \\ 0 & -1 & -1 & -1 \end{array} \right] \xrightarrow{F_2 + 2F_3} \left[\begin{array}{cccc} 1 & 1 & 1 & -1 \\ 0 & 0 & -2 & -3 \\ 0 & -1 & -1 & -1 \end{array} \right]$$

Esta matriz ya está en forma escalonada, (aunque no estén ordenadas las filas), y podemos resolver el sistema por el método de *sustitución hacia atrás* (ver la [Definición 1.25, página 46](#)):

- **Fila 2:** $-2x_3 - 3x_4 = 0 \Rightarrow x_3 = -\frac{3}{2}x_4 \quad \underline{x_4 = 2t} \quad x_3 = -3t.$
- **Fila 3:** $-x_2 - x_3 - x_4 = 0 \Rightarrow x_2 = -x_3 - x_4 = -(-3t) - 2t = t.$
- **Fila 1:** $x_1 + x_2 + x_3 - x_4 = 0 \Rightarrow x_1 = -x_2 - x_3 + x_4 = -t - (-3t) + 2t = 4t.$

Quedando finalmente la solución del sistema:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} 4 \\ 1 \\ -3 \\ 2 \end{bmatrix}$$

Teorema 1.2.

Sea $\vec{x}_p$ una solución *particular* del SEL⁷⁴ $A\vec{x} = \vec{b}$, y sea $\vec{x}_h$ una solución del SEL homogéneo⁷⁵ $A\vec{x}_h = \vec{0}$ asociado al anterior⁷⁶, entonces:

$$\vec{x} = \vec{x}_p + \vec{x}_h$$

Demostración 1.2.

$$A\vec{x} = A(\vec{x}_p + \vec{x}_h) = \underbrace{A\vec{x}_p}_{\vec{b}} + \underbrace{A\vec{x}_h}_{\vec{0}} = \vec{b}$$

Dado un SEL $A\vec{x} = \vec{b}$ compatible y $\vec{x}_p$ una solución del mismo. Si $\vec{x}_h$ es la solución general⁷⁷ del SEL homogéneo entonces la solución general del SEL viene dada por:

$$\vec{x} = \vec{x}_p + \vec{x}_h$$

⁷⁴Si el SEL fuera compatible indeterminado, entonces $\vec{x}_p$ sería una cualquiera de sus infinitas soluciones, mientras que si fuera un S.C.D. correspondería a su única solución.

⁷⁵Si éste fuera un S.C.D. correspondería con su solución trivial $\vec{x}_h = \vec{0}$, de otra forma sería cualquier otra solución del mismo.

⁷⁶Es decir, con la misma matriz de coeficientes A .

⁷⁷Conjunto de todas las soluciones del sistema.

1.9. SEL en $\mathbb{C}^n$

En los sistemas de ecuaciones lineales en $\mathbb{C}^n$, tanto los **coeficientes** a_{ij} como los **términos independientes** b_i y las **incógnitas** x_i , son números *complejos*.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} ; \quad a_{i,j}, b_i, x_i \in \mathbb{C}$$

De igual forma, la *matriz de coeficientes* A y la *matriz ampliada* AB asociadas a este sistema son complejas.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{C}^{m \times n} ; \quad AB = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right] \in \mathbb{C}^{m \times (n+1)}$$

Los métodos de análisis y resolución de este tipo de sistemas son los mismos que hemos visto para los SEL en $\mathbb{R}^n$, pero teniendo en cuenta las particularidades en las operaciones con este tipo de números⁷⁸.

Veamos algunos ejemplos prácticos.

Ejercicio 1.20. Resolver el siguiente sistema de ecuaciones lineales complejo.

$$\begin{cases} 2x + iy = 3 + 4i \\ ix + 5y = 9 - 4i \end{cases}$$

Para sistemas pequeños y/o sencillos se puede utilizar los métodos simples de igualación, sustitución y reducción, siguiendo el mismo procedimiento que se realiza con sistemas reales⁷⁹.

- **Método de Sustitución**

Despejando por ejemplo la variable x de la primera ecuación se tiene:

$$x = \frac{3 + 4i - iy}{2} \quad (1.2)$$

y sustituyendo este valor en la segunda ecuación:

$$i \frac{3 + 4i - iy}{2} + 5y = 9 - 4i ; \quad i(3 + 4i - iy) + 10y = 18 - 8i ; \quad 3i - 4 + 11y = 18 - 8i$$

$$y = \frac{22 - 11i}{11} = 2 - i = \boxed{y = 2 - i}$$

Sustituyendo en la ecuación (1.2):

$$x = \frac{3 + 4i - iy}{2} = \frac{3 + 4i - i(2 - i)}{2} = \frac{3 + 4i - 2i - 1}{2} = \frac{2 + 2i}{2} = 1 + i = \boxed{x = 1 + i}$$

- **Método de Igualación**

Despejando por ejemplo la variable y de las dos ecuaciones del problema e igualando se tiene⁸⁰:

$$\left\{ \begin{array}{l} y = \frac{3 + 4i - 2x}{i} = -i(3 + 4i - 2x) = -3i + 4 + 2xi \\ y = \frac{9 - 4i - ix}{5} \end{array} \right\} \quad -3i + 4 + 2xi = \frac{9 - 4i - ix}{5}$$

$$-15i + 20 + 10xi = 9 - 4i - ix ; \quad (10 + 1)xi = 9 - 20 - 4i + 16i ; \quad 11xi = 11(-1 + i)$$

$$x = \frac{-1 + i}{i} = -i(-1 + i) = 1 + i = \boxed{x = 1 + i}$$

Sustituyendo en la primera ecuación donde se despejó y :

$$y = -3i + 4 + 2xi = -3i + 4 + 2i(1 + i) = 2 - i = \boxed{y = 2 - i}$$

⁷⁸Ver el Apéndice C, página 295.

• Método de Reducción

Eliminamos por ejemplo la variable x multiplicando por i la primera ecuación y por -2 la segunda:

$$\begin{array}{rcl} i & \times & (2x + iy) = i \times (3 + 4i) \\ -2 & \times & (ix + 5y) = -2 \times (9 - 4i) \\ \hline 0 & + & -y - 10y = 3i - 4 - 18 + 8i = -22 + 11i \end{array}$$

Por tanto:

$$y = \frac{-22 + 11i}{11} = \boxed{-2 + i} = x = \frac{3 + 4i - iy}{2} = \frac{3 + 4i - i(2 - i)}{2} = \frac{2 + 2i}{2} = \boxed{1 + i}$$

Ejercicio 1.21. Resolver el siguiente sistema de ecuaciones lineales complejo.

$$\begin{cases} (1 + 2i)x + (-2 + i)y = 3 + 6i \\ 3x + 3iy = 8 \end{cases}$$

Aplicamos el método de reducción, multiplicando la primera ecuación por -3 y la segunda por $1 + 2i$.

$$\begin{array}{rcl} -3 & \times & (1 + 2i)x + (-2 + i)y = -3 \times (3 + 6i) \\ (1 + 2i) & \times & (3x + 3iy) = (1 + 2i) \times 8 \\ \hline 0 & + & (6 - 3i)y + (-6 + 3i)y = -9 - 18i + 8 + 16i = -1 - 2i \end{array}$$

Quedando:

$$0 = -1 - 2i$$

que al ser una contradicción, indica que el sistema **no tiene solución**.

Ejercicio 1.22. Resolver el siguiente sistema de ecuaciones lineales.

$$\begin{cases} (1 + i)x + (-2 + i)y + iz = 1 + 6i \\ x + iy - (1 + i)z = -2 - 2i \\ -2x - z = -3 - 4i \end{cases}$$

Construimos la matriz ampliada del sistema y aplicamos el método de Gauss para su resolución.

$$\left[\begin{array}{ccc|c} 1+i & -2+i & i & 1+6i \\ 1 & i & -1-i & -2-2i \\ -2 & 0 & -1 & -3-4i \end{array} \right] \xrightarrow[F_3+2F_2]{F_1-(1+i)F_2} \left[\begin{array}{ccc|c} 0 & -1 & 3i & 1+10i \\ 1 & i & -1-i & -2-2i \\ 0 & 2i & -3-2i & -7-8i \end{array} \right]$$

$$\xrightarrow{F_3+2iF_1} \left[\begin{array}{ccc|c} 0 & -1 & 3i & 1+10i \\ 1 & i & -1-i & -2-2i \\ 0 & 0 & -9-2i & -27-6i \end{array} \right]$$

Una vez que tenemos la matriz escalonada (desordenada), aplicamos el método de sustitución hacia atrás:

$$\text{Fila 3 : } z = \frac{-27 - 6i}{-9 - 2i} = \frac{(-27 - 6i)(-9 + 2i)}{9^2 + 2^2} = \frac{255}{85} = 3 = \boxed{z = 3}$$

$$\text{Fila 1 : } y = 3iz - (1 + 10i) = -1 - i = \boxed{y = -1 - i}$$

$$\text{Fila 2 : } x = -2 - 2i - iy + (1 + i)z = -2 - 2i + i - 1 + 3 + 3i = \boxed{x = 2i}$$

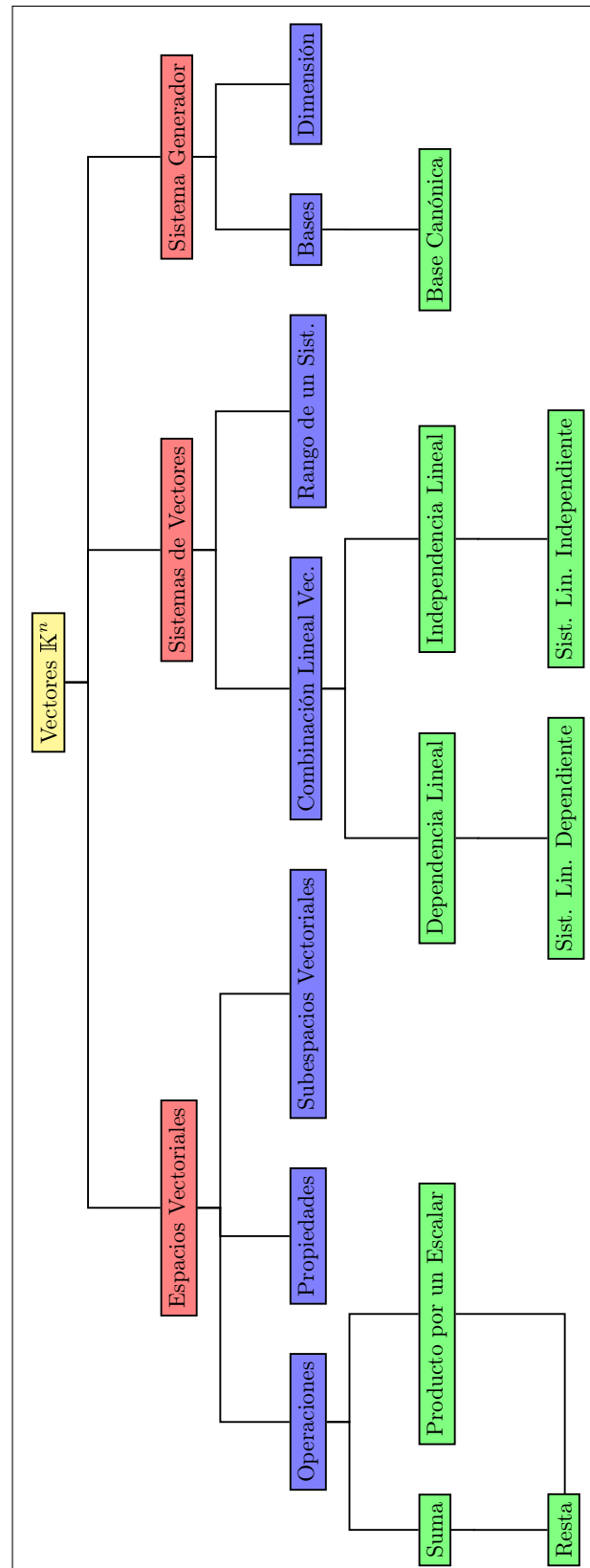
⁷⁹Ver la Sección 0.6, página 25.

⁸⁰Notar que $\frac{1}{i} = \frac{i}{i \cdot i} = -i$.

Espacios Vectoriales

Contenido

2.1	Vectores	65
2.1.1	Representación Gráfica de Vectores Reales	67
2.2	Operaciones Elementales con Vectores	67
2.2.1	Suma	67
2.2.2	Producto por un Escalar	68
2.2.3	Vector Opuesto y Vectores Estándar	69
2.2.4	Diferencia de Vectores	69
2.2.5	Propiedades Algebraicas de la Suma y Producto por un Escalar	70
2.3	Espacio y Subespacio Vectorial	71
2.4	Sistemas de Vectores	73
2.4.1	Combinación Lineal de Vectores	73
2.4.2	Dependencia e Independencia Lineal	73
2.4.3	Rango de un Sistema de Vectores	76
2.4.4	Cálculo del Rango de un Sistema de Vectores	77
2.4.5	Interpretación Geométrica de la Dependencia e Independencia Vectorial	78
2.5	Conjuntos o Sistemas Generadores	79
2.5.1	Base y Dimensión de un Sistema de Vectores	80
2.5.2	Generación de $\mathbb{R}^n$	81
2.5.3	Base Canónica o Estándar de $\mathbb{K}^n$	83



2.1. Vectores

Hay múltiples definiciones del concepto de vector, ya que aparece en diferentes ámbitos. Nosotros nos vamos a centrar al campo de las matemáticas. Empecemos con la definición más general y abstracta.

Definición 2.1. Vector

Un *vector* es un elemento de un *espacio vectorial*¹.

Los *espacios vectoriales* son muy diversos y por tanto sus elementos tienen características muy distintas. En este texto nos vamos a centrar en los *espacios vectoriales euclidianos*, tanto reales $\mathbb{R}^n$ como complejos $\mathbb{C}^n$, con $n > 0$ finito².

La mayoría de los temas los vamos a tratar desde el punto de vista real, por ser operacionalmente más sencillo y por tener una interpretación geométrica para espacios de dimensión baja hasta $n = 3$. Sin embargo, el caso complejo $\mathbb{C}^n$ incluye al caso real y, salvo que se indique lo contrario, se puede sustituir en todos los temas $\mathbb{R}$ por $\mathbb{C}$ verificándose las mismas propiedades.

Para hacer referencia a los dos casos utilizaremos el símbolo $\mathbb{K}^n$, donde $\mathbb{K}$ se definió en la [Sección 1.1.3, página 32](#), y puede ser o bien el conjunto de los números reales $\mathbb{R}$ o bien el conjunto de los números complejos $\mathbb{C}$.

Definición 2.2. Vector en un Espacio Euclídeo

Un *vector*³ $\vec{v}$ es una colección ordenada de números:

$$\vec{v} = [v_1, v_2, \dots, v_n] \quad ; \quad v_i \in \mathbb{K}$$

Cada uno de los números se denomina *componente* o *coordenada* y el número de componentes n se denomina *dimensión* del vector.

Definición 2.3. Vectores Fila y Columna

Hay dos tipos básicos de vectores⁴:

- **Vectores fila:** Las componentes se ordenan en una fila de izquierda a derecha.

$$\vec{v} = [v_1 \quad v_2 \quad \dots \quad v_n]$$

- **Vectores columna:** Las componentes se ordenan en una columna de arriba a abajo.

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Definición 2.4. Vector Transpuesto

Un *vector fila* se puede transformar en un *vector columna* y al contrario mediante la **transposición** del vector. Un *vector transpuesto* se indica mediante una letra T (mayúscula o minúscula) en el superíndice.

¹Ver la [Definición 2.12, página 71](#).

²Los *espacios de Hilbert* son una generalización de los espacios euclídeos incluyendo a los espacios de dimensión infinita.

³Esta definición corresponde a los conocidos como *vectores generales*. Además hay otros tipos de vectores como los *vectores geométricos* y *de posición*. Ver una introducción en el [Apéndice A.8, página 270](#).

⁴Ambos son equivalentes y solo es importante su tipo cuando tienen un carácter *matricial*, ver el [Tema 3, página 87](#). En algunos textos, por defecto un vector es siempre de tipo columna.

Ejemplo 2.1. Vectores Transpuestos.

$$\vec{v} = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad ; \quad \vec{u} = \begin{bmatrix} -2+i \\ -i \\ 3 \end{bmatrix}^t = \begin{bmatrix} -2+i & -i & 3 \end{bmatrix}$$

Definición 2.5. Vector Transpuesto Conjugado

El **vector transpuesto conjugado** de un vector $\vec{v} \in \mathbb{K}^n$ se representa mediante $\vec{v}^H$ (con una H en el superíndice), y corresponde con el vector transpuesto $\vec{v}^T$ en el que a cada elemento del vector se sustituye por su complejo conjugado.

$$\vec{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \vec{v}^H = \begin{bmatrix} \bar{v}_1 \\ \bar{v}_2 \\ \vdots \\ \bar{v}_n \end{bmatrix} \quad ; \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \vec{v}^H = \begin{bmatrix} \bar{v}_1 & \bar{v}_2 & \cdots & \bar{v}_n \end{bmatrix}$$

Si el vector es real, $\vec{v} \in \mathbb{R}^n$, entonces $\vec{v}^T = \vec{v}^H$.

Ejemplo 2.2. Vector Transpuesto y Transpuesto Conjugado.

$$\vec{v} = \begin{bmatrix} -2, & 1+i, & -2i \end{bmatrix} = \vec{v}^T = \begin{bmatrix} -2 \\ 1+i \\ -2i \end{bmatrix} \quad ; \quad \vec{v}^H = \begin{bmatrix} -2 \\ 1-i \\ 2i \end{bmatrix}$$

Definición 2.6. Espacio $\mathbb{K}^n$

El conjunto de todos los *vectores* de n componentes (números $\mathbb{K}$) se representa por $\mathbb{K}^n$.

$$\mathbb{K}^n = \{x_1, x_2, \dots, x_n\} : x_i \in \mathbb{K}, 1 \leq i \leq n$$

Ejemplo 2.3. Espacios $\mathbb{K}^n$

$$\begin{aligned} \mathbb{R}^2 &= \{x, y\} : x, y \in \mathbb{R} & - & \quad [1, 0] \in \mathbb{R}^2 \\ \mathbb{R}^3 &= \{x, y, z\} : x, y, z \in \mathbb{R} & - & \quad [1, -1, 1/3] \in \mathbb{R}^3 \\ \mathbb{C}^3 &= \{z_1, z_2, z_3\} : z_i \in \mathbb{C} & - & \quad [2-3i, -1-i, 1/3+\pi i] \in \mathbb{C}^3 \end{aligned}$$

2.1.1. Representación Gráfica de Vectores Reales

Los vectores reales de baja dimensión⁵, $\mathbb{R}^2$ y $\mathbb{R}^3$, se pueden representar gráficamente como un segmento dirigido desde el origen hasta el punto dado por las coordenadas del vector.

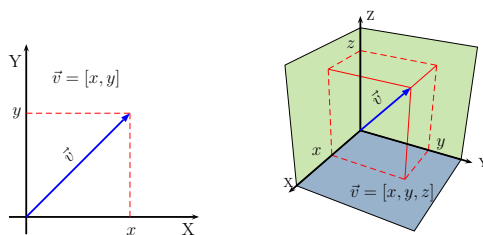


Figura 2.1: Representación gráfica de vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$.

En un espacio real n -dimensional $\mathbb{R}^n$ existe una relación directa entre un *punto* $P = (x_1, \dots, x_n)$ y un *vector* con las mismas componentes $\vec{x} = [x_1, \dots, x_n]$, ya que asociado a cada punto de $\mathbb{R}^n$ podemos considerar que existe un vector con inicio en el origen $(0, \dots, 0)$ en $\mathbb{R}^n$ y final en el punto en cuestión. Por tanto, cuando hablamos de espacios $\mathbb{R}^n$ nos podemos referir indistintamente a espacios de *puntos*⁶ y/o espacios de *vectores*.

2.2. Operaciones Elementales con Vectores

Vamos a definir las operaciones básicas entre vectores de *suma*, *producto por un escalar* y *diferencia* entre vectores.

2.2.1. Suma

Definición 2.7. Suma de Vectores

Se define la *suma* de dos vectores $\vec{u}, \vec{v} \in \mathbb{K}^n$ como el vector $\vec{s} \in \mathbb{K}^n$ dado por:

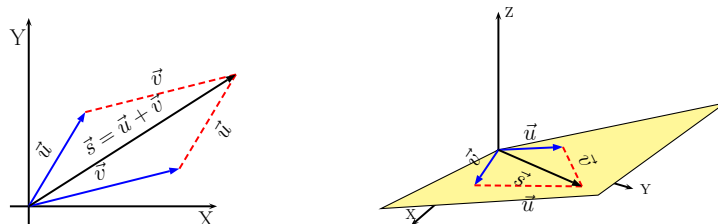
$$\vec{s} = \vec{u} + \vec{v} = [u_1, u_2, \dots, u_n] + [v_1, v_2, \dots, v_n] = [u_1 + v_1, u_2 + v_2, \dots, u_n + v_n]$$

Ejercicio 2.1. Determinar la suma de los vectores de $\mathbb{R}^4$: $\vec{u} = [1, 2, 3, 4]$ y $\vec{v} = [0, -2, 3, -1]$

$$\vec{s} = \vec{u} + \vec{v} = [1, 2, 3, 4] + [0, -2, 3, -1] = [1, 0, 6, 3]$$

Regla del Paralelogramo

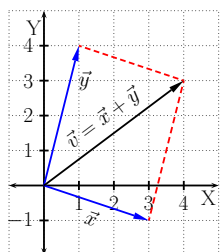
En los casos particulares de $\mathbb{R}^2$ y $\mathbb{R}^3$ donde los vectores se pueden representar gráficamente, la suma de dos $\vec{u} + \vec{v}$ representa el vector diagonal del paralelogramo de lados $\vec{u}$ y $\vec{v}$, en el plano definido por ambos vectores.



⁵Los vectores $\mathbb{R}^n$ para $n > 3$ **no** se pueden representar gráficamente.

⁶Ver la [Definición 1.9](#), página 32.

Ejercicio 2.2. En $\mathbb{R}^2$, determinar $\vec{x} + \vec{y}$ siendo $\vec{x} = [3, -1]$ e $\vec{y} = [1, 4]$.



$$\vec{v} = \vec{x} + \vec{y} = [3, -1] + [1, 4] = [3+1, (-1)+4] = [4, 3]$$

2.2.2. Producto por un Escalar

Definición 2.8. Producto por un Escalar

Se define la multiplicación de un vector $\vec{v} = [v_1, v_2, \dots, v_n] \in \mathbb{K}^n$ por un *escalar*⁷ $\alpha \in \mathbb{K}$ de la siguiente forma:

$$\alpha \vec{v} = \alpha[v_1, v_2, \dots, v_n] = [\alpha v_1, \alpha v_2, \dots, \alpha v_n] \quad ; \quad \vec{v} \in \mathbb{K}^n, \alpha \in \mathbb{K}$$

Ejercicio 2.3. En $\mathbb{R}^2$ determinar el producto del vector $\vec{u} = [1, 2]$ por el escalar $\alpha = -3$.

$$\vec{v} = \alpha \vec{u} = (-3)[1, 2] = [-3, -6]$$

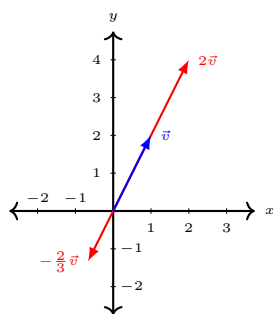
Ejercicio 2.4. Dados los vectores de $\mathbb{C}^3$: $\vec{v} = [2, -2 + i, -i]$, $\vec{u} = [0, 2i, 1]$ y $\vec{w} = [1 + i, 1 - i, -1 + i]$, calcular el siguiente vector:

$$\vec{a} = i\vec{v} - 2\vec{u} + (-1 - i)\vec{w}$$

$$\begin{aligned} \vec{a} &= i[2, -2 + i, -i] - 2[0, 2i, 1] + (-1 - i)[1 + i, 1 - i, -1 + i] = \\ &= [2i, -1 - 2i, 1] + [0, -4i, -2] + [-2i, -2, 2] = [0, -3 - 6i, 1] \end{aligned}$$

Nota 2.1. Desde un punto de vista geométrico, es decir en $\mathbb{R}^n$, el producto de un vector por una constante (escalar) únicamente modifica su *longitud* y/o *sentido*, sin modificar su *dirección*. Su longitud variará si el escalar es distinto de 1 y el sentido cambiará si el escalar es negativo. A esto se le denomina *escalamiento* de un vector.

Ejercicio 2.5. Dado $\vec{v} = [1, 2] \in \mathbb{R}^2$, calcular y representar gráficamente $2\vec{v}$ y $-\frac{2}{3}\vec{v}$.



- $2\vec{v} = 2[1, 2] = [2, 4]$
- $-\frac{2}{3}\vec{v} = -\frac{2}{3}[1, 2] = \left[-\frac{2}{3}, -\frac{4}{3}\right]$

⁷La palabra *escalar* proviene del latín *scala* que significa *escalera*, y de forma general hace referencia a cualquier número real o complejo.

2.2.3. Vector Opuesto y Vectores Estándar

Definición 2.9. Vector Opuesto

El vector opuesto (o negativo) de un vector $\vec{v} \in \mathbb{K}^n$ es el vector $-\vec{v}$ dado por:

$$-\vec{v} = (-1)\vec{v} = (-1)[v_1, v_2, \dots, v_n] = [-v_1, -v_2, \dots, -v_n]$$

Definición 2.10. Vectores Estándar

Se define el vector *estándar* $\vec{e}_i \in \mathbb{K}^n$ como aquel vector que tiene todas sus componentes 0 excepto la componente i -ésima que vale 1.

$$\vec{e}_i = [0, \dots, 0, \underbrace{1}_{\text{pos. } i}, 0, \dots, 0]$$

Nota 2.2. Los vectores estándar de cualquier espacio $\mathbb{K}^n$ son siempre reales.

2.2.4. Diferencia de Vectores

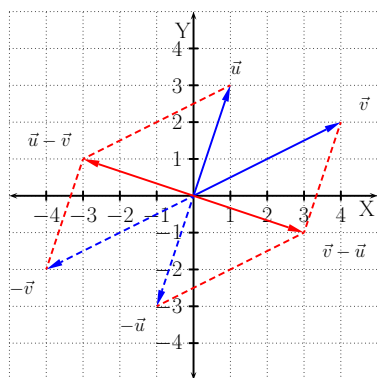
Definición 2.11. Diferencia o Resta de Vectores

La resta $\vec{u} - \vec{v}$ de dos $\vec{u}, \vec{v} \in \mathbb{K}^n$ corresponde a la suma de $\vec{u}$ con el vector opuesto de $\vec{v}$.

$$\vec{u} - \vec{v} = \vec{u} + (-\vec{v}) \quad ; \quad \vec{u}, \vec{v} \in \mathbb{K}^n$$

Por tanto la resta no se considera una operación si no un caso particular de la operación suma.

Ejercicio 2.6. En $\mathbb{R}^2$, dados $\vec{u} = [1, 3]$ y $\vec{v} = [4, 2]$ determinar y representar $\vec{u} - \vec{v}$ y $\vec{v} - \vec{u}$.



- $\vec{u} - \vec{v} = [1, 3] - [4, 2] = [1 - 4, 3 - 2] = [-3, 1]$
- $\vec{v} - \vec{u} = [4, 2] - [1, 3] = [4 - 1, 2 - 3] = [3, -1]$

2.2.5. Propiedades Algebraicas de la Suma y Producto por un Escalar

Dados los vectores $\vec{u}, \vec{v}, \vec{w} \in \mathbb{K}^n$ y los escalares $\alpha, \beta \in \mathbb{K}$, se verifican las siguientes propiedades:

- **Conmutativa** : $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
- **Asociativa de la Suma** : $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
- **Asociativa del Producto** : $\alpha(\beta\vec{u}) = \alpha\beta\vec{u} = (\alpha\beta)\vec{u}$
- **Elemento Neutro 1** : $(\vec{u} + \vec{0}) = \vec{0} + \vec{u} = \vec{u}$
- **Elemento Neutro 2** : $0\vec{u} = \vec{0} \quad ; \quad \alpha\vec{0} = \vec{0}$
- **Elemento Unitario** : $1 * \vec{u} = \vec{u} * 1 = \vec{u}$
- **Elemento Opuesto 1** : $\vec{u} + (-\vec{u}) = (-\vec{u}) + \vec{u} = \vec{0}$
- **Elemento Opuesto 2** : $(-\alpha)\vec{u} = \alpha(-\vec{u}) = -(\alpha\vec{u})$
- **Distributiva 1** : $\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v} = (\vec{u} + \vec{v})\alpha$
- **Distributiva 2** : $(\alpha + \beta)\vec{u} = \alpha\vec{u} + \beta\vec{u} = \vec{u}(\alpha + \beta)$

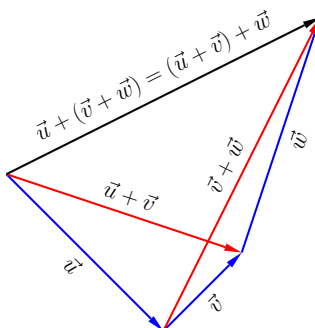
Ejercicio 2.7. Calcular $\vec{u} - (3\vec{v} + 2\vec{w})$ siendo $\vec{u} = [3, 4]$, $\vec{v} = [1, -2]$ y $\vec{w} = [0, 5]$ en $\mathbb{R}^2$.

$$\vec{u} - (3\vec{v} + 2\vec{w}) = [3, 4] - (3[1, -2] + 2[0, 5]) = [3, 4] - ([3, -6] + [0, 10]) = [3, 4] - [3, 4] = [0, 0] = \vec{0}$$

Ejercicio 2.8. Demostrar la propiedad conmutativa de la suma de vectores.

$$\begin{aligned} \vec{u} + \vec{v} &= [u_1, u_2, \dots, u_n] + [v_1, v_2, \dots, v_n] = [u_1 + v_1, u_2 + v_2, \dots, u_n + v_n] = \\ &= [v_1 + u_1, v_2 + u_2, \dots, v_n + u_n] = [v_1, v_2, \dots, v_n] + [u_1, u_2, \dots, u_n] = \vec{v} + \vec{u} \end{aligned}$$

Ejercicio 2.9. Representar gráficamente la propiedad asociativa de la suma de vectores en $\mathbb{R}^3$.



2.3. Espacio y Subespacio Vectorial

Definición 2.12. Espacio Vectorial

Se denomina **espacio vectorial**⁸ a todo conjunto V en el que se definen dos operaciones entre sus elementos, llamadas *suma*⁹ y *multiplicación por un escalar*¹⁰, las cuales cumplen las propiedades algebraicas¹¹ de los vectores en $\mathbb{K}^n$. Un espacio vectorial se representa por la *terna*¹² $(V, +, *)$.

- A los elementos $v \in V$ de un *espacio vectorial* $(V, +, *)$ se les denomina de forma general *vectores*.
- Por simplicidad, vamos a denotar al espacio vectorial $(V, +, *)$ a través del conjunto V , teniendo en cuenta que al referirse al *espacio vectorial* V nos referimos a la terna.

Ejemplo 2.4. Espacios vectoriales.

- El espacio $\mathbb{K}^n$ es un *espacio vectorial* para cualquier $n \geq 1$, con las operaciones habituales entre vectores $\vec{v} \in \mathbb{K}^n$ de suma y multiplicación por escalares ($\mathbb{R}$ ó $\mathbb{C}$).
- El conjunto $\mathbb{P}_n$ de los *polinomios* en $x \in \mathbb{K}$ de grado $\leq n$:

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n ; \quad a_i \in \mathbb{K}$$

junto con las operaciones:

- suma: $(p+q)(x) = p(x) + q(x)$, $p, q \in \mathbb{P}_n$.
- multiplicación por un escalar: $(cp)(x) = cp(x)$, $p \in \mathbb{P}_n$, $c \in \mathbb{K}$

es un **espacio vectorial** de *polinomios*. Notar que los polinomios $\mathbb{P}_n$ pueden representarse como vectores $\mathbb{K}^{n+1}$ a través de los coeficientes de sus monomios, de la siguiente forma:

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \in \mathbb{P}_n \quad = \quad \vec{p} = [a_0, a_1, \dots, a_n] \in \mathbb{K}^{n+1}$$

- El conjunto $\mathbb{F}$ de todas las funciones reales de variable real en el que se definen las operaciones:
 - suma: $(f+g)(x) = f(x) + g(x)$, $f, g \in \mathbb{F}$.
 - multiplicación por un escalar: $(cf)(x) = cf(x)$, $c \in \mathbb{R}$, $f \in \mathbb{F}$.

es un **espacio vectorial** de funciones.

Definición 2.13. Subespacio Vectorial

Un subconjunto¹³ W de un espacio vectorial V , $W \subseteq V$, se denomina **subespacio vectorial** de V si cumple las siguientes tres propiedades:

1. W contiene el elemento nulo de V .
2. W es *cerrado* bajo la suma, es decir, si $w_1, w_2 \in W$ $\Rightarrow w_1 + w_2 \in W$.
3. W es *cerrado* bajo el producto por un escalar, es decir, si $w \in W$ y $k \in \mathbb{K}$ $\Rightarrow kw \in W$.

⁸También denominado a veces *espacio lineal*.

⁹También llamada *adición*.

¹⁰Estas operaciones, dependiendo de la naturaleza de los elementos, pueden ser muy variadas y no necesariamente la *suma* y *producto por un escalar* definidas anteriormente para vectores.

¹¹Ver la [Sección 2.2.5, página 70](#).

¹²Grupo de tres.

¹³La expresión $W \subseteq V$ indica que W *está contenido* (es una parte de él) o *es igual que* V .

Un subespacio vectorial es un subconjunto de un espacio vectorial que cumple las propiedades de espacio vectorial¹⁴.

Ejemplo 2.5. Subespacios vectoriales.

- El conjunto de todos los vectores $\vec{v} = [x, y, z] \in \mathbb{R}^3$ pertenecientes al plano¹⁵ $3x + 2y - z = 0$ es un subespacio vectorial de $\mathbb{R}^3$, ya que cumplen las tres propiedades anteriores.
- El conjunto de todos los vectores $\vec{v} = [x, y] \in \mathbb{R}^2$ pertenecientes a la recta $3x + 2y = 1$ no son un subespacio vectorial de $\mathbb{R}^2$, ya que este conjunto no contiene al vector nulo $[0, 0] \in \mathbb{R}^2$.
- El conjunto de todos los vectores $\vec{v} = [x, y, z, w, t] \in \mathbb{R}^5$ que cumplen las ecuaciones¹⁶:

$$\begin{cases} x + y - z + 5w - t &= 0 \\ 3x - 5t &= 0 \\ y + 2z - w &= 0 \end{cases}$$

forman un subespacio vectorial de $\mathbb{R}^5$.

- Los subespacios vectoriales de $\mathbb{R}^n$ son *objetos geométricos*¹⁷ de *dimensión*¹⁸ no nula que pasan por el origen¹⁹.
- Los subespacios vectoriales se pueden definir mediante un sistema de ecuaciones lineales *homogéneo*²⁰, (para que pasen por el origen).

Definición 2.14. Dimensión de un Subespacio Vectorial en $\mathbb{R}^n$

Se denomina *dimensión* de un subespacio vectorial $W \subseteq \mathbb{R}^n$ a la *dimensión*²¹ del objeto geométrico que representan.

$$\dim(W) = \text{dim. del objeto geométrico (punto, recta, plano, ...) que representa } W.$$

Ejercicio 2.10. Determinar la dimensión del siguiente subespacio vectorial $W \subseteq \mathbb{R}^5$.

$$W : \begin{cases} x_1 - x_4 &= 0 \\ x_2 + x_3 &= 0 \end{cases}$$

Claramente W es un subespacio vectorial al ser un objeto geométrico definido mediante un sistema de ecuaciones homogéneo formado por dos ecuaciones linealmente independientes.

Su dimensión, calculada mediante la fórmula de equilibrio, es:

$$\dim(W) = 5 - 2 = 3$$

Por tanto se trata de un hiperplano de dimensión 3 en $\mathbb{R}^5$.

¹⁴Ver la Definición 2.12, página 71.

¹⁵Es decir, aquellos vectores $\vec{v} = [x, y, z]$ cuyas componentes x, y y z cumplen la ecuación del plano.

¹⁶Es decir, el conjunto de vectores que pertenecen al *hiperplano* de dimensión $5 - 3 = 2$ en el espacio $\mathbb{R}^5$.

¹⁷Ver la Definición 1.13, página 36.

¹⁸Ver la Sección 1.2.2, página 36.

¹⁹Para que contenga el vector nulo.

²⁰Ver la Sección 1.8, página 59.

²¹Ver la Definición 1.14, página 36 y su forma de cálculo en la Definición 1.15, página 37. En la Definición 2.26, página 80 se da una definición más general de dimensión para subespacios vectoriales en $\mathbb{K}^n$.

2.4. Sistemas de Vectores

Definición 2.15. Sistema de Vectores

Se denomina **sistema** de vectores a todo *conjunto* (finito o infinito) de vectores:

$$S = \vec{v}_1, \vec{v}_2, \dots, \vec{v}_p, \dots \subset \mathbb{K}^n$$

2.4.1. Combinación Lineal de Vectores

Definición 2.16. Combinación Lineal de Vectores

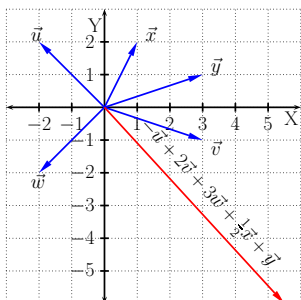
Se dice que un vector $\vec{v} \in \mathbb{K}^n$ es una **combinación lineal** de un *sistema* de m vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{K}^n$ si existen m *escalares* (constantes) $c_1, c_2, \dots, c_m \in \mathbb{K}$ tales que²²:

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_m\vec{v}_m = \sum_{j=1}^m c_j\vec{v}_j$$

A los *escalares* c_i se les denominan **coeficientes de la combinación lineal**.

Ejercicio 2.11. Dados los vectores de $\mathbb{R}^2$: $\vec{u} = [-2, 2]$, $\vec{v} = [3, 1]$, $\vec{w} = [-2, -2]$, $\vec{x} = [1, 2]$, $\vec{y} = [3, -1]$, determinar el vector resultante de la siguiente combinación lineal, representando gráficamente el resultado.

$$-\vec{u} + 2\vec{v} + 3\vec{w} + \frac{1}{2}\vec{x} + \vec{y}$$



$$\begin{aligned} & -\vec{u} + 2\vec{v} + 3\vec{w} + \frac{1}{2}\vec{x} + \vec{y} \\ &= (-1)[-2, 2] + 2[3, 1] + 3[-2, -2] + \frac{1}{2}[1, 2] + [3, -1] \\ &= [2 + 6 - 6 + 0.5 + 3, -2 + 2 - 6 + 1 - 1] \\ &= [5.5, -6] \end{aligned}$$

El vector resultante de una combinación lineal siempre pertenece al subespacio vectorial al que forman los vectores combinados.

2.4.2. Dependencia e Independencia Lineal

Definición 2.17. Dependencia e Independencia Lineal

- Se dice que un vector $\vec{v} \in \mathbb{K}^n$ es **linealmente dependiente** de un *sistema* de m vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{K}^n$, si $\vec{v}$ es una *combinación lineal* de dichos vectores.
- En caso contrario, se dice que $\vec{v}$ es **linealmente independiente** de dichos vectores.

Nota 2.3. La dependencia lineal es *transitiva*: si $\vec{v}$ depende linealmente de unos vectores y cada uno de estos dependen linealmente de otros, entonces $\vec{v}$ depende linealmente de los últimos.

²²Ver un resumen de la definición y propiedades de los *sumatorios* en el [Apéndice A.4, página 262](#).

Ejercicio 2.12. Determinar si el vector $\vec{v} = [1, 0, -1, 0]$ $\mathbb{R}^4$ es linealmente dependiente o independiente de los siguientes sistemas de vectores:

a) $W_1 = \{x_1 + x_2 - x_3 + x_4 = 0\} \subset \mathbb{R}^4$

b) $W_2 : \begin{cases} x_1 + x_3 = 0 \\ x_2 - x_4 = 0 \end{cases} \subset \mathbb{R}^4$

c) $W_3 = \{[1, 0, 1, 1], [0, 0, 1, 0]\} \subset \mathbb{R}^4$

d) $W_4 = \{[1, 0, 1, 1], [0, 0, 1, 0], [0, 0, 0, 1]\} \subset \mathbb{R}^4$

- a) Como $\vec{v}$ no pertenece a W_1 (no cumple su ecuación), no puede obtenerse por combinación lineal de los vectores de W_1 y por tanto es linealmente independiente de él.
- b) En este caso $\vec{v}$ cumple las dos ecuaciones de W_2 , de forma que pertenece a este plano de $\mathbb{R}^4$, y por tanto por combinación lineal de diferentes vectores de W_2 se puede obtener v . De aquí se deduce que v es linealmente dependiente de W_2 .
- c) En este caso hay que determinar si existen escalares $x_1, x_2 \in \mathbb{R}$ que cumplen:

$$x_1[1, 0, 1, 1] + x_2[0, 0, 1, 0] = [1, 0, -1, 0] \quad = \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

Resolviendo el SEL mediante Gauss:

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{array} \right] \xrightarrow{\substack{F_3 - F_1 \\ F_4 - F_1}} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{array} \right]$$

Claramente el sistema es incompatible y por tanto no existe ninguna combinación lineal de los vectores de W_3 que genere $\vec{v}$. Por tanto $\vec{v}$ es linealmente independiente de W_3 .

- d) Al igual que en el apartado anterior vamos a determinar si existen escalares $x_1, x_2, x_3 \in \mathbb{R}$ que cumplen:

$$x_1[1, 0, 1, 1] + x_2[0, 0, 1, 0] + x_3[0, 0, 0, 1] = [1, 0, -1, 0] \quad = \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

Resolviendo el SEL mediante Gauss:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \end{array} \right] \xrightarrow{\substack{F_3 - F_1 \\ F_4 - F_1}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -1 \end{array} \right]$$

Claramente el sistema es compatible determinado, (aunque igual sería si fuera compatible indeterminado), y por tanto sí existe alguna combinación lineal de los vectores de W_4 que genera $\vec{v}$. Por tanto $\vec{v}$ es linealmente dependiente de W_4 .

Definición 2.18. Conjunto Linealmente Dependiente

Se dice que un *sistema* de vectores $S = \vec{v}_1, \vec{v}_2, \dots, \vec{v}_p \subset \mathbb{K}^n$ es **linealmente dependiente** (o **ligado**) si se verifica una de las siguientes condiciones, que son equivalentes entre sí²³:

- Alguno de los vectores de S depende linealmente de los demás.
- Existen escalares $\alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{K}$, no todos nulos²⁴, que cumplen:

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_p \vec{v}_p = \vec{0}$$

Definición 2.19. Conjunto Linealmente Independiente

Se dice que un *sistema* de vectores $S = \vec{v}_1, \vec{v}_2, \dots, \vec{v}_p \subset \mathbb{K}^n$ es **linealmente independiente** (o **libre**) si se verifica una de las siguientes condiciones, que son equivalentes entre sí:

- El sistema S no es linealmente dependiente, es decir, ninguno de los vectores de S depende linealmente de los demás.
- La única combinación lineal de vectores de S posible es la que tiene todos sus coeficientes nulos, es decir:

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_p \vec{v}_p = \vec{0} \quad \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$$

siendo los escalares $\alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{K}$.

Ejercicio 2.13. Determinar si el conjunto de vectores $V = [1, 2, 3], [1, 1, 1], [1, 0, -1]$ es linealmente dependiente o independiente.

Tenemos que encontrar los escalares α_1, α_2 y α_3 que cumplan:

$$\alpha_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2.1)$$

Esta operación vectorial se puede expresar como:

$$\alpha_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ 2\alpha_1 \\ 3\alpha_1 \end{bmatrix} + \begin{bmatrix} \alpha_2 \\ \alpha_2 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \alpha_3 \\ 0 \\ -\alpha_3 \end{bmatrix} = \begin{bmatrix} \alpha_1 + \alpha_2 + \alpha_3 \\ 2\alpha_1 + \alpha_2 + 0 \\ 3\alpha_1 + \alpha_2 - \alpha_3 \end{bmatrix}$$

Por tanto:

$$\begin{bmatrix} \alpha_1 + \alpha_2 + \alpha_3 \\ 2\alpha_1 + \alpha_2 + 0 \\ 3\alpha_1 + \alpha_2 - \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Esto se puede expresar como un SEL homogéneo (Sección 1.8, página 59) que hay que analizar:

$$\begin{bmatrix} \alpha_1 + \alpha_2 + \alpha_3 \\ 2\alpha_1 + \alpha_2 + 0 \\ 3\alpha_1 + \alpha_2 - \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 0 \\ 2\alpha_1 + \alpha_2 = 0 \\ 3\alpha_1 + \alpha_2 - \alpha_3 = 0 \end{cases}$$

Si es compatible determinado, (solución única trivial $\alpha_1 = \alpha_2 = \alpha_3 = 0$), el conjunto será *linealmente independiente*, mientras que si es compatible indeterminado, (infinitas soluciones para α_1, α_2 y α_3), el conjunto de vectores será *linealmente dependiente*.

²³La primera definición es formal o teórica y la segunda operacional o aplicada.

²⁴Al menos dos de ellos deben ser distintos de cero.

Resolviendo por Gauss:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 3 & 1 & -1 \end{bmatrix} \xrightarrow[\substack{F_2-2F_1 \\ F_3-3F_1}]{} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & -2 & -4 \end{bmatrix} \xrightarrow{F_3-2F_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Por tanto el sistema es *compatible indeterminado*, es decir existen infinitos valores α_i no todos nulos que cumplen la **Ecuación (2.1)** y por ello el conjunto de vectores es *linealmente dependiente*.

- En un espacio vectorial $\mathbb{R}^n$ todos sus subespacios vectoriales tienen dimensión $\leq n$, ya que representan objetos geométricos (que pasan por el origen) o es el propio espacio $\mathbb{R}^n$.
- Un sistema de m vectores de $\mathbb{R}^n$ es linealmente dependiente si todos ellos forman un subespacio vectorial de dimensión²⁵ $< m$.
- Un sistema de m vectores de $\mathbb{R}^n$ es linealmente independiente si definen un subespacio vectorial de dimensión $= m$.

En este caso el SEL homogéneo formado con los vectores será compatible y determinado (todos los valores de α_i son nulos).

2.4.3. Rango de un Sistema de Vectores

Definición 2.20. Rango de un Sistema de Vectores

Se denomina *rango*²⁶ de un sistema de vectores S , al *máximo* número de vectores *linealmente independientes* que hay en S . Se representa de la siguiente forma:

$$\text{rang}(S)$$

Sea S un sistema de m vectores:

- Si S es **linealmente independiente**, el *rango* coincide con m , $\text{rang}(S) = m$.
- Si S es **linealmente dependiente**, el *rango* es siempre menor que m , $0 \leq \text{rang}(S) < m$.

Ejemplo 2.6. Rango de un sistema de vectores.

- El rango de $S = [1, 2], [3, 2]$ es $\text{rang}(S) = 2$, ya que los dos son linealmente independientes (no paralelos).
- El rango de $G = [1, 1, 1], [2, 2, 2], [3, 3, 3]$ es $\text{rang}(G) = 1$, ya que solo hay un vector linealmente independiente de los demás.
- El rango de $M = [1, 2, 3], [4, 5, 6], [5, 7, 9]$ es $\text{rang}(M) = 2$, ya que solo hay dos vectores linealmente independientes.
- El rango de $R = [0, 0, 0, 0]$ es $\text{rang}(R) = 0$, ya que el *vector nulo* siempre es *linealmente dependiente*²⁷.

²⁵Este caso es análogo al **Ejercicio 2.13, página 75**, ya que la matriz de vectores forma un SEL homogéneo con alguna fila de ceros en su forma escalonada reducida y por tanto corresponde a un sistema compatible y determinado, (valores de α_i no todos nulos).

²⁶El término *rango* coincide con el utilizado en el *rango de una matriz* (ver **Sección 1.5.5, página 50**) debido a la relación entre ambos conceptos, como se explica en la siguiente sección.

²⁷Ya que según la **Definición 2.18, página 75** se tiene: $\alpha \vec{0} = \vec{0}$, $\forall \alpha \neq 0$. También se puede comprobar usando el método de cálculo del rango de la **Sección 2.4.4, página 77**, ya que la matriz formada por el vector nulo no contiene ningún pivote.

2.4.4. Cálculo del Rango de un Sistema de Vectores

En los casos sencillos, (con pocos vectores y de dimensión baja), el *rango de un conjunto de vectores* se puede determinar fácilmente a simple vista, pero normalmente habrá que determinarlo matemáticamente utilizando la definición de dependencia lineal, [Definición 2.18, página 75](#), e independencia lineal, [Definición 2.19, página 75](#).

Para determinar el rango de un conjunto de k vectores $S = \vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ de $\mathbb{K}^n$ hay que calcular el número de vectores linealmente independientes de S . Para ello vamos a utilizar el SEL homogéneo que aparece en la definición de dependencia lineal²⁸:

$$x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_k \vec{v}_k = \vec{0} \quad = \quad \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_k \end{bmatrix} \vec{x} = \vec{0} \quad = \quad A\vec{x} = \vec{0}$$

en el cual las columnas de la matriz de coeficientes $A \in \mathbb{K}^{n \times k}$ corresponden con los vectores de S .

- El sistema de vectores S será un conjunto *linealmente independiente* si el SEL es compatible determinado con solución única trivial $x_i = 0$.

Al resolver el SEL por Gauss aparece una matriz escalonada con k filas no nulas. En este caso el rango de S será k , (todos los vectores son linealmente independientes).

- El sistema de vectores S será un conjunto *linealmente dependiente* si el SEL es compatible indeterminado con infinitas soluciones para las incógnitas x_i .

Al resolver el SEL por Gauss aparece una matriz escalonada con $r < k$ filas no nulas, lo que indica que solo hay r vectores linealmente independientes²⁹, siendo éste el valor del rango del sistema S .

- El *rango de un conjunto de vectores* V corresponde con el *rango de la matriz*³⁰ A formada con los vectores dispuestos como columnas³¹.
- Para calcular el $\text{rang}(V)$, se obtiene una *matriz escalonada equivalente*³² E de A .
 - El rango del conjunto de vectores corresponde con el número de filas no nulas de E , es decir, con $\text{rang}(A)$.
 - Un conjunto de vectores linealmente independientes de V viene dado por los vectores columna de A correspondientes a la posición de los *pivotes*³³ en E .

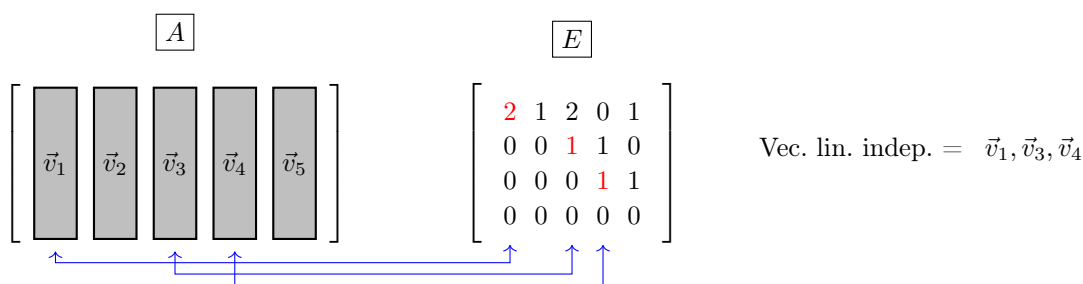


Figura 2.2: Ejemplo de identificación de vectores columna linealmente independientes en una matriz A , mediante la posición de los pivotes en una de sus matrices escalonadas equivalentes E .

²⁸Ver el [Ejercicio 2.13, página 75](#).

²⁹Si eliminamos de S los vectores (columnas de A) correspondientes con las filas nulas de la matriz escalonada equivalente, resulta un sistema compatible determinado, indicando que es un conjunto de vectores linealmente independiente.

³⁰Ver la [Sección 1.5.5, página 50](#).

³¹En el [Teorema 3.5, página 113](#) se demuestra que los vectores se pueden disponer en la matriz indistintamente como filas o como columnas, ya que el rango de la matriz no varía.

³²Ver la [Sección 1.5.4, página 50](#).

³³Ver la [Definición 1.24, página 45](#).

Ejercicio 2.14. Calcular el rango del siguiente conjunto de vectores de $\mathbb{R}^4$:

$$V = [1, 1, 0, -1], [1, 1, -1, -1], [0, 0, 1, 0], [1, 1, 1, -1], [0, 1, 0, 1]$$

Hay que determinar el número de *vectores linealmente independientes* que hay en V . En este caso como mínimo hay uno, ya que no todos los vectores son nulos, y como máximo cinco. Para ello construimos una matriz $A \in \mathbb{R}^{4 \times 5}$ utilizando los vectores de V como columnas (el orden es indiferente aunque debe ser conocido para posteriormente hacer referencia a ellos):

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ -1 & -1 & 0 & -1 & 1 \end{bmatrix}$$

Calculamos el rango de A , el cual coincidirá con el rango de V , encontrando una matriz escalonada equivalente:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ -1 & -1 & 0 & -1 & 1 \end{bmatrix} \xrightarrow{\substack{F_2 - F_1 \\ F_4 + F_1}} \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Como esta matriz tiene 3 filas no nulas, el rango de A y por tanto el rango de V también es 3.

$$\text{rang}(V) = 3$$

Esto indica que en V solo tres vectores del total de cinco son linealmente independientes. La matriz escalonada informa que son el primero $\vec{v}_1 = [1, 1, 0, -1]$, el segundo $\vec{v}_2 = [1, 1, -1, -1]$ y el quinto $\vec{v}_5 = [0, 1, 0, 1]$, correspondientes a las posiciones de los *pivotes* (primer elemento no nulo de cada fila, ver la [Definición 1.24](#), página 45).

Naturalmente esta no es la única terna de vectores de V que son linealmente independientes, aunque este proceso de cálculo nos asegura que estos tres sí lo son.

Por ejemplo en este caso podemos comprobar a simple vista, como también el conjunto de vectores primero, tercero y quinto son linealmente independientes, pero no el primero, segundo y tercero, ya que $\vec{v}_1 = \vec{v}_2 + \vec{v}_3$.

2.4.5. Interpretación Geométrica de la Dependencia e Independencia Vectorial

Definición 2.21. Vectores Paralelos

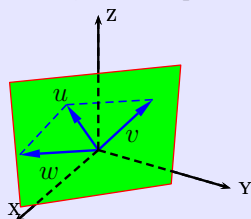
En $\mathbb{R}^n$, se dice que dos vectores son *paralelos* si poseen la misma dirección, es decir:

$$\vec{u} \parallel \vec{v} \iff \vec{u} = \alpha \vec{v} \iff a\vec{u} + b\vec{v} = \vec{0}$$

Un sistema de $N \geq 2$ vectores paralelos es siempre un conjunto *linealmente dependiente*.

Definición 2.22. Vectores Coplanarios

En $\mathbb{R}^3$, se dice que tres vectores son *coplanarios* si están contenidos en el mismo plano, es decir:



$$\vec{u}, \vec{v}, \vec{w} \text{ coplanarios} \iff \vec{u} = \alpha \vec{v} + \beta \vec{w}$$

$$a\vec{u} + b\vec{v} + c\vec{w} = \vec{0} \quad ; \quad a, b, c \in \mathbb{R}$$

Un sistema de $N \geq 3$ vectores coplanarios es siempre un conjunto *linealmente dependiente*.

Definición 2.23. Vectores Co-hiperplanarios

En $\mathbb{R}^n$, se dice que n vectores son *co-hiperplanarios* si están contenidos en el mismo hiperplano, y por tanto forman un conjunto *linealmente dependiente*.

$$\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \text{ co-hiperplanarios} \quad \vec{v}_1 = \alpha_2 \vec{v}_2 + \alpha_3 \vec{v}_3 + \dots + \alpha_n \vec{v}_n$$

$$a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_n \vec{v}_n = \vec{0} \quad ; \quad a_i \in \mathbb{R}$$

Un sistema de $N \geq m$ vectores co-hiperplanarios a un hiperplano de dimensión $m + 1$ es siempre un conjunto *linealmente dependiente*.

Ejercicio 2.15. Determinar si los vectores del siguiente conjunto de vectores V son co-hiperplanarios, y en caso afirmativo calcular la dimensión mínima del hiperplano al que pertenecen y su ecuación vectorial.

$$V = [1, 0, 0, 1], [1, 1, 0, 0], [0, 1, 0, -1], [1, -1, 0, 2] \subset \mathbb{R}^4$$

Los k vectores de un sistema $V \subset \mathbb{R}^n$ son co-hiperplanarios (linealmente dependientes) si el $\text{rang}(V) = r < n$, ya que el rango de cualquier conjunto de vectores de $\mathbb{R}^n$ solo puede ser como máximo n . La dimensión mínima del hiperplano al que pertenecen coincidirá con el número de vectores linealmente independientes:

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 2 \end{bmatrix} \xrightarrow{F_4 - F_1} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix} \xrightarrow{F_4 + F_2} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Como $\text{rang}(V) = 2$, el conjunto V es coplanario a un plano ($\dim = 2$), cuya ecuación vectorial es:

$$[x_1, x_2, x_3, x_4] = t[1, 0, 0, 1] + s[1, 1, 0, 0] \quad ; \quad t, s \in \mathbb{R}$$

Notar que hay infinitos hiperplanos de dimensión 3 que contienen a este plano y por tanto también son co-hiperplanarios a estos hiperplanos pero de dimensión mayor³⁴.

2.5. Conjuntos o Sistemas Generadores

Definición 2.24. Conjunto o Sistema Generador

Dado un sistema de vectores G , subconjunto de un subespacio vectorial³⁵ W , $G \subseteq W$, se dice que G es un **conjunto o sistema generador** de W , si todo vector $\vec{w} \in W$ es linealmente dependiente de G .

$$\vec{w} \in W \quad \vec{w} = c_1 \vec{g}_1 + \dots + c_k \vec{g}_k \quad \vec{g}_i \in G$$

- Un *conjunto generador* de un subespacio vectorial $W \subseteq \mathbb{K}^n$ es cualquier conjunto G de vectores de W que pueden generar por combinación lineal todos los vectores de W .
- Todo subespacio vectorial $W \subseteq \mathbb{K}^n$ no nulo tiene infinitos *conjuntos generadores*.
- Todo *conjunto generador* G de un subespacio vectorial $W \subseteq \mathbb{K}^n$ de $\dim(W) = d \leq n$, debe tener $\text{rang}(G) = d$ y por tanto estar formado por al menos d vectores de W .

³⁴Por ejemplo en $\mathbb{R}^3$, pensar en varios vectores que son paralelos y por tanto pertenecen a una única recta. A su vez, hay infinitos planos que contiene a esta recta y por tanto los vectores son coplanarios a estos planos.

³⁵Típicamente G es un conjunto finito de vectores y W es un conjunto infinito, siempre que $\dim(W) \neq 0$.

Ejercicio 2.16. Encontrar tres conjuntos generadores del plano de $\mathbb{R}^3$:

$$W : \quad x + 2y - z = 0$$

En este caso el subespacio vectorial $W \subset \mathbb{R}^3$ está formado por los infinitos vectores de $\mathbb{R}^3$ que pertenecen al plano.

Como la dimensión del plano es³⁶ $3 - 1 = 2$, cualquier conjunto de vectores de este plano que contenga al menos **dos** vectores linealmente independientes será un conjunto generador del subespacio W .

Tiene por tanto infinitos conjuntos generadores. Por ejemplo tres de ellos son:

$$\begin{aligned} G_1 &= [1, 0, 1], [1, 1, 3] & ; & & G_2 &= [1, 0, 1], [1, 1, 3], [-1, 2, 3] \\ G_3 &= [2, 2, 6], [-7, 4, 1], [1, -1/2, 0], [0, 1/2, 1], [2, -1, 0] \end{aligned}$$

Donde los valores $[x, y, z]$ utilizados cumplen la ecuación del plano $x + 2y - z = 0$.

A través de la notación $\text{gen}()$ se representa el **espacio generado** por *combinación lineal* de los vectores de un conjunto generador $G = \vec{g}_1, \dots, \vec{g}_k$:

$$\text{gen}(\vec{g}_1, \dots, \vec{g}_k) = \vec{w} : \vec{w} = \alpha_1 \vec{g}_1 + \dots + \alpha_m \vec{g}_k, \quad \alpha_1, \dots, \alpha_k \in \mathbb{R}$$

2.5.1. Base y Dimensión de un Sistema de Vectores

Definición 2.25. Base de un Subespacio Vectorial

Una **base** de un subespacio vectorial W es un *conjunto generador* de W formado por **vectores linealmente independientes**.

Todas las *bases* de un subespacio vectorial $W \subseteq \mathbb{K}^n$ de $\dim(W) = d \leq n$ son *conjuntos generadores* de W que contienen exactamente d vectores *linealmente independientes*.

Ejemplo 2.7. Base.

De los tres conjuntos generadores del subespacio vectorial dado en el [Ejemplo 2.16, página 80](#), solo el G_1 es una base de W , ya que su número de vectores coincide con la dimensión del subespacio vectorial (plano de $\mathbb{R}^3$).

Definición 2.26. Dimensión de un Subespacio Vectorial

Al *número* de vectores que contiene cualquier *base* de un subespacio vectorial W se denomina **dimensión**³⁷ de W , y coincide con $\text{rang}(W)$ (número de vectores linealmente independientes).

³⁶Ver la *fórmula de equilibrio* en la [Definición 1.15, página 37](#).

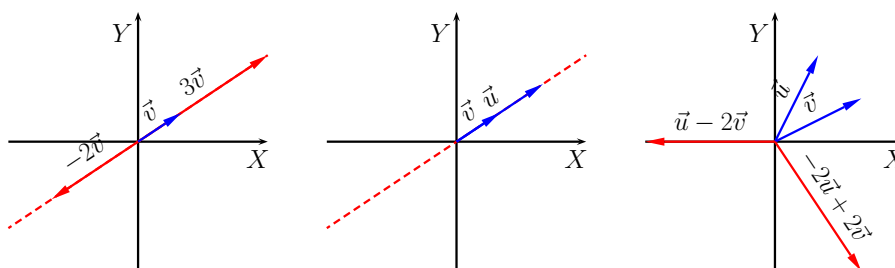
³⁷En la [Definición 2.14, página 72](#) se establece el concepto de dimensión de un subespacio vectorial desde un punto de vista geométrico. Esta nueva definición es más general y sirve para definir la dimensión de subespacios vectoriales tanto reales como complejos, en los cuales no tiene sentido la geometría.

2.5.2. Generación de $\mathbb{R}^n$

Vamos a aplicar los conceptos anteriores en la generación de los espacios $\mathbb{R}^n$.

Generación del Espacio $\mathbb{R}^2$

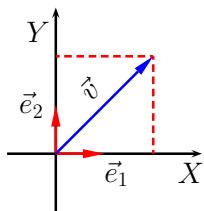
- Con todas las combinaciones lineales de *uno o más vectores paralelos*, (linealmente dependientes), se obtienen todos los vectores contenidos en la recta que tiene la dirección de los vectores. Por tanto, el rango de cualquier conjunto de vectores paralelos en $\mathbb{R}^2$ es 1 y **no** pueden ser una *base* de $\mathbb{R}^2$.
- *Cualquier* vector de $\mathbb{R}^2$ se puede obtener como combinación lineal de *dos vectores no paralelos* (linealmente independientes). Por tanto, cualquier conjunto de vectores que contenga al menos dos vectores no paralelos, tiene *rango* 2, y por tanto es un *conjunto generador* de $\mathbb{R}^2$.
- Una *base* de $\mathbb{R}^2$ será cualquier conjunto de exactamente *dos vectores no paralelos*.



Base Canónica o Estándar de $\mathbb{R}^2$

Tomando los dos vectores *estándar*³⁸ de $\mathbb{R}^2$:

$$\vec{e}_1 = [1, 0] \quad ; \quad \vec{e}_2 = [0, 1]$$



Cualquier vector de $\mathbb{R}^2$ de la forma $\vec{v} = [v_1, v_2]$ se puede expresar como combinación lineal de los dos vectores estándar:

$$\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 = v_1 [1, 0] + v_2 [0, 1] = [v_1, v_2]$$

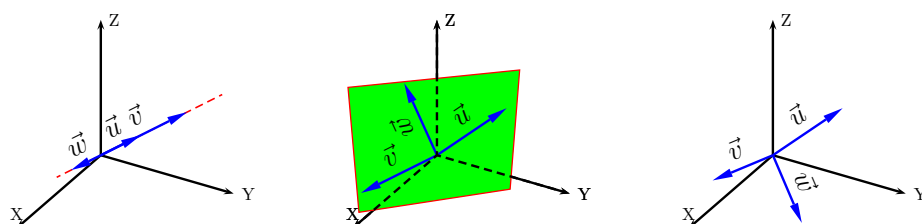
Por tanto, el conjunto $B = \vec{e}_1, \vec{e}_2 = [1, 0], [0, 1]$ es una *base* de $\mathbb{R}^2$.

A los vectores *estándar* se les suele denotar como $\vec{i} = [1, 0]$ y $\vec{j} = [0, 1]$, y a la *base* $\mathcal{B} = \vec{i}, \vec{j}$ se le denomina **base canónica o estándar** de $\mathbb{R}^2$.

³⁸Ver la Definición 2.10, página 69.

Generación del Espacio $\mathbb{R}^3$

- Con todas las combinaciones lineales de *uno o más vectores paralelos*, (linealmente dependientes), se obtienen todos los vectores contenidos en la recta que tiene la dirección de los vectores. Por tanto, el *rango* de cualquier conjunto de vectores paralelos en $\mathbb{R}^3$ es 1 y **no** pueden ser una *base* de $\mathbb{R}^3$.
- Con todas las combinaciones lineales de al menos dos vectores *coplanarios no paralelos* y por tanto linealmente independientes, solo se pueden obtener todos los vectores contenidos en el plano descrito por los vectores. Por tanto, el *rango* de cualquier conjunto de vectores coplanarios en $\mathbb{R}^3$ es 2 y **no** puede ser una *base* de $\mathbb{R}^3$.
- *Cualquier* vector de $\mathbb{R}^3$ se puede obtener como combinación lineal de *tres vectores no coplanarios*. Por tanto, el *rango* de cualquier conjunto de vectores que contenga al menos 3 vectores no coplanarios en $\mathbb{R}^3$ es 3.
- Una *base* de $\mathbb{R}^3$ será cualquier conjunto de exactamente tres vectores *no coplanarios*.



Base Canónica o Estándar de $\mathbb{R}^3$

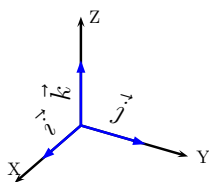
Tomando los vectores estándar de $\mathbb{R}^3$:

$$\vec{e}_1 = [1, 0, 0] \quad ; \quad \vec{e}_2 = [0, 1, 0] \quad ; \quad \vec{e}_3 = [0, 0, 1]$$

Cualquier vector $\vec{v} = [v_1, v_2, v_3]$ en $\mathbb{R}^3$ se puede expresar como *combinación lineal* de estos vectores:

$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 = v_1[1, 0, 0] + v_2[0, 1, 0] + v_3[0, 0, 1] = [v_1, v_2, v_3]$$

Por tanto, el conjunto $B = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\} = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ es una *base* de $\mathbb{R}^3$.



A los vectores estándar de $\mathbb{R}^3$ se les suele denotar como $\vec{i} = [1, 0, 0]$, $\vec{j} = [0, 1, 0]$ y $\vec{k} = [0, 0, 1]$, y a la base $B = \{\vec{i}, \vec{j}, \vec{k}\}$ se le denomina **base canónica** o **estándar** de $\mathbb{R}^3$.

Generación del Espacio Vectorial $\mathbb{R}^n$

- Cualquier conjunto de vectores de $\mathbb{R}^n$ co-hiperplanarios de un hiperplano de dimensión $k < n$ es un *conjunto linealmente dependiente* de rango k .
- Todo conjunto $G \subseteq \mathbb{R}^n$ con al menos n vectores linealmente independientes, $\text{rang}(G) = n$, es un *conjunto generador*³⁹ de $\mathbb{R}^n$.
- Una *base* de $\mathbb{R}^n$ será cualquier conjunto de **exactamente** n vectores *linealmente independientes*⁴⁰.

³⁹Y por tanto, todo $\vec{v} \in \mathbb{R}^n$ se puede obtener por combinación lineal de los vectores de V .

⁴⁰Vectores no *co-hiperplanarios*.

2.5.3. Base Canónica o Estándar de $\mathbb{K}^n$ **Definición 2.27.** Base Canónica o Estándar de $\mathbb{K}^n$

La **base canónica** o **estándar** de $\mathbb{K}^n$ es aquella formada por los *vectores estándar*⁴¹ de $\mathbb{K}^n$:

$$\mathcal{B} = \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n = [1, 0, 0, \dots, 0], [0, 1, 0, \dots, 0], \dots, [0, 0, 0, \dots, 1] \subset \mathbb{K}^n$$

Ejercicio 2.17. Completar el conjunto de vectores S para que sea una base de $\mathbb{R}^3$.

$$S = [1, 0, 1], [1, 1, 1]$$

Una base de $\mathbb{R}^3$ es un conjunto de 3 vectores de $\mathbb{R}^3$ linealmente independientes. Claramente los dos vectores de S son linealmente independientes, (es decir, no paralelos). Necesitamos un tercer vector $\vec{x} = [x, y, z] \in \mathbb{R}^3$ que sea linealmente independiente de los dos vectores de S , es decir, que por combinación lineal de ambos no se pueda obtener $\vec{x}$:

$$\alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad = \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Esto es equivalente a encontrar unos valores x , y y z que produzcan un SEL incompatible:

$$\left[\begin{array}{cc|c} 1 & 1 & x \\ 0 & 1 & y \\ 1 & 1 & z \end{array} \right] \xrightarrow{F_3 - F_1} \left[\begin{array}{cc|c} 1 & 1 & x \\ 0 & 1 & y \\ 0 & 0 & z - x \end{array} \right]$$

Por tanto, basta con que $z = x$. Hay infinitos vectores que cumplen esta condición, por ejemplo:

$$\vec{x} = [x, y, z] = [0, 0, 1]$$

En este caso, la base de $\mathbb{R}^3$ podría ser:

$$\mathcal{B} = [1, 0, 1], [1, 1, 1], [0, 0, 1]$$

Ejercicio 2.18. Encontrar un conjunto generador y una base del subespacio vectorial $W \subset \mathbb{R}^3$ definido por:

$$W : \begin{cases} x + y - z = 0 \\ x - y = 0 \end{cases}$$

Primero tenemos que encontrar el conjunto de los infinitos vectores de $\mathbb{R}^3$ que pertenecen al subespacio W y para ello tenemos que resolver el SEL que los define. Como las dos ecuaciones del SEL son claramente *linealmente independientes* (ver la [Sección 1.4.4, página 43](#)), la dimensión de W es $3 - 2 = 1$ y por tanto es una recta en $\mathbb{R}^3$ que vendrá definida en forma paramétrica por un único parámetro. Resolviendo el SEL por Gauss:

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 \end{array} \right] \xrightarrow{F_2 - F_1} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 1 & 0 \end{array} \right] \rightarrow \begin{cases} z = 2y = 2t \\ y = t \\ x = -y + z = -t + 2t = t \end{cases}$$

En forma vectorial, el conjunto de todos los vectores de $\mathbb{R}^3$ que forman parte de W son:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \quad ; \quad t \in \mathbb{R}$$

⁴¹Ver la [Definición 2.10, página 69](#).

Un conjunto generador de W será cualquier conjunto formado por al menos un vector no nulo de W , como por ejemplo cualquiera de los tres siguientes⁴²:

$$G_1 = [1, 1, 2] \quad ; \quad G_2 = [2, 2, 4], [-1, -1, -2] \quad ; \quad G_3 = [3, 3, 6], [2, 2, 4], [0, 0, 0]$$

Una base de W será cualquier conjunto de vectores formado por un solo vector⁴³ no nulo de W , como por ejemplo cualquiera de los dos siguientes:

$$B_1 = [1, 1, 2] \quad ; \quad B_2 = [-2, -2, -4]$$

Ejercicio 2.19. Encontrar un conjunto generador y una base del subespacio vectorial $W \subset \mathbb{R}^5$ donde sus vectores $\vec{v} = [x, y, z, w, t]$ cumplen:

$$W : x + y - z = 0$$

En este caso la $\dim(W) = 5 - 1 = 4$, por tanto habrá 4 parámetros libres en su definición vectorial. Resolviendo el SEL de una ecuación con 5 incógnitas se tiene:

$$\left. \begin{array}{l} x = -y - z = -\alpha + \beta \\ y = \alpha \\ z = \beta \\ w = \gamma \\ t = \delta \end{array} \right\} \begin{bmatrix} x \\ y \\ z \\ w \\ t \end{bmatrix} = \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \delta \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad ; \quad \alpha, \beta, \gamma, \delta \in \mathbb{R}$$

Dando valores a los parámetros α , β , γ y δ se obtienen todos los vectores del subespacio W .

Un conjunto generador de W estará formado por *al menos* 4 vectores⁴⁴ linealmente independientes de W , como por ejemplo, cualquiera de los dos siguientes:

$$G_1 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \quad ; \quad G_2 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 8 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Una base de W está formada por cualquier conjunto de *exactamente* 4 vectores linealmente independientes de W , como los dos ejemplos siguientes:

$$B_1 = G_1 \quad ; \quad B_2 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 8 \end{bmatrix} \right\}$$

⁴²Estos vectores se obtienen de la solución en forma paramétrica, dando valores al parámetro t .

⁴³Ya que la dimensión de W es uno.

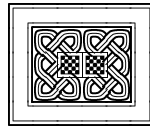
⁴⁴Ya que este valor es la dimensión del subespacio vectorial W .

Ejercicio 2.20. Comprobar que en el espacio vectorial de los polinomios⁴⁵ reales $\mathbb{P}_2$, el conjunto $\mathcal{B} = [-1 - x + x^2, -1 + x^2, -1 + x]$ es una base.

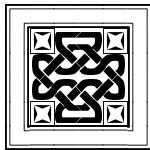
Para simplificar los cálculos, los elementos del espacio $\mathbb{P}_2 = a_0 + a_1x + a_2x^2$ se pueden representar por los vectores $\vec{p} = [a_0, a_1, a_2] \in \mathbb{R}^3$, de forma que podemos comprobar si el conjunto $\mathcal{B} = [-1, -1, 1], [-1, 0, 1], [-1, 1, 0]$ es una base de $\mathbb{R}^3$ calculando el rango de la siguiente matriz:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{rang} = 3$$

Con lo que se comprueba que efectivamente es una base del espacio correspondiente, lo cual indicaría que todo el conjunto de vectores de segundo grado $p(x) = a_0 + a_1x + a_2x^2$ puede obtenerse por combinación lineal de los tres polinomios del conjunto $\mathcal{B}$.



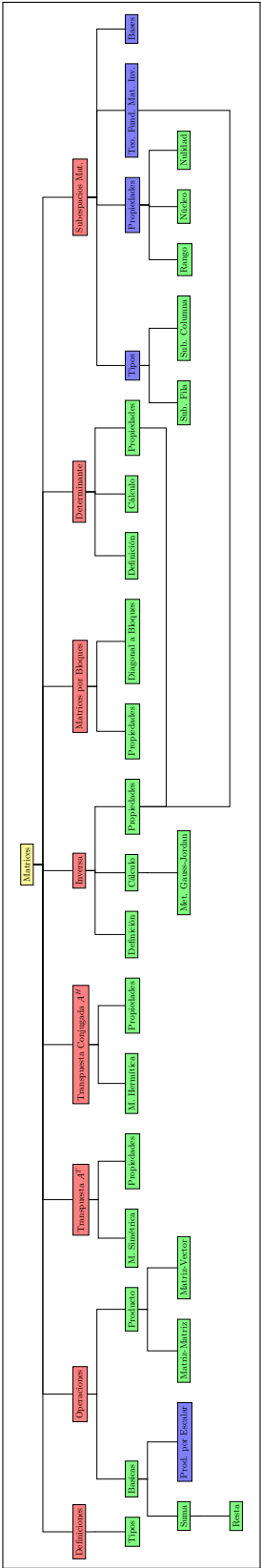
⁴⁵Ver el Ejemplo 2.4, página 71.



Matrices

Contenido

3.1	Matrices	89
3.1.1	Tipos de Matrices	89
3.2	Operaciones Básicas con Matrices	91
3.2.1	Suma de Matrices	91
3.2.2	Producto de un Escalar por una Matriz	92
3.2.3	Diferencia o Resta de Matrices	92
3.2.4	Propiedades de la Suma y la Multiplicación por un Escalar	92
3.2.5	Espacio Vectorial de Matrices	92
3.2.6	Combinación Lineal de Matrices	93
3.3	Producto de Matrices	93
3.3.1	Propiedades de la Multiplicación de Matrices	94
3.3.2	Potencia de Matrices Cuadradas	95
3.3.3	Producto Matricial Matriz-Vector	95
3.3.4	Producto Matricial Vector-Vector	95
3.3.5	Producto Matriz-Vector Estándar	96
3.4	Expresión Matricial de un SEL	96
3.5	Matriz Transpuesta A^T	97
3.5.1	Matriz Simétrica y Antisimétrica	97
3.5.2	Propiedades de la Matriz Transpuesta	97
3.6	Matriz Transpuesta Conjugada A^H	98
3.6.1	Matriz Hermítica y Antihermítica	99
3.6.2	Propiedades de la Matriz Transpuesta Conjugada	99
3.7	Inversa de una Matriz	99
3.7.1	Cálculo de la Matriz Inversa: Método de Gauss-Jordan	99
3.7.2	Inversa y SEL	101
3.7.3	Operaciones con la Matriz Inversa	101
3.8	Matrices por Bloques	101
3.8.1	Propiedades de las Matrices por Bloques	102
3.8.2	Matriz Diagonal por Bloques	103
3.8.3	Propiedades de las Matrices Diagonales por Bloques	103
3.9	Determinante	103
3.9.1	Introducción: Determinante de Matrices 2×2	103
3.9.2	Determinante de Matrices 3×3	104
3.9.3	Regla de Sarrus	104
3.9.4	Determinante de Matrices $n \times n$	105
3.9.5	Cálculo del Determinante mediante Desarrollos de Laplace	106
3.9.6	Determinante e Inversa de una Matriz	107
3.9.7	Propiedades de los Determinantes	107
3.9.8	Método de Gauss para el Cálculo del Determinante	108
3.9.9	Regla de Cramer	109
3.9.10	Matriz Adjunta e Inversa de una Matriz	110
3.10	Subespacios Fundamentales de una Matriz	111
3.10.1	Espacios Columna y Fila de una Matriz	111
3.10.2	Propiedades de las Matrices Equivalentes por Filas	113
3.10.3	Rango, Núcleo y Nulidad de una Matriz	114
3.10.4	Teorema del Rango de las Matrices	116
3.10.5	Teorema Fundamental de las Matrices Invertibles	117
3.10.6	Bases para los Espacios Fila, Columna y Nulo de una Matriz	117



3.1. Matrices

Definición 3.1. Matriz

Una **matriz** es una disposición de números ordenados en *filas* y *columnas*.

Una matriz A de m filas y n columnas se denota por $A_{m \times n}$ y las representaremos mediante corchetes cuadrados¹ de la siguiente forma:

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} ; \quad a_{ij} \in \mathbb{K} \quad ; \quad A = [a_{ij}] \in \mathbb{K}^{m \times n}$$

- El **orden**, **dimensión** o **tamaño** de una matriz es la cantidad de filas y columnas que posee, $(m \times n)$.
- Se denominan **elementos diagonales** de la matriz $A_{m \times n} = [a_{ij}]$ a los elementos a_{ii} que poseen igual fila que columna.

$$\begin{bmatrix} * & \cdots & \cdots \\ \cdots & * & \cdots \\ \cdots & \cdots & * \end{bmatrix}$$

- Al conjunto de todos los *elementos diagonales* se denomina **diagonal principal**. Se definen las **subdiagonales** como aquellas *diagonales* por encima y por debajo de la *diagonal principal*.
- Si las n columnas de la matriz $A_{m \times n}$ son los vectores $\vec{c}_1, \vec{c}_2, \dots, \vec{c}_n \in \mathbb{K}^m$, podemos representarla como un *vector de vectores columna*:

$$A = [\vec{c}_1 \quad \vec{c}_2 \quad \cdots \quad \vec{c}_n] \quad ; \quad \vec{c}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}$$

- Si las m filas de la matriz $A_{m \times n}$ son los vectores $\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m \in \mathbb{K}^n$, podemos representarla como un *vector de vectores fila*:

$$A = \begin{bmatrix} \vec{f}_1 \\ \vec{f}_2 \\ \vdots \\ \vec{f}_m \end{bmatrix} \quad ; \quad \vec{f}_i = [\vec{a}_{i1} \quad \vec{a}_{i2} \quad \cdots \quad \vec{a}_{in}]$$

Definición 3.2. Matrices Iguales

Se dice que dos matrices son iguales si poseen las mismas *dimensiones* (número de filas y columnas) y si son iguales todos sus correspondientes elementos.

3.1.1. Tipos de Matrices

- **Matriz Fila:** Matriz formada por una única fila.

$$A_{1 \times n} = [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}]$$

Es equivalente a un *vector fila*.

¹También es muy frecuente representarla por paréntesis.

- **Matriz Columna:** Matriz formada por una única columna.

$$A_{m \times 1} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

Es equivalente a un *vector columna*.

- **Matriz Nula, $0_{m \times n}$:** Matriz con todos sus elementos iguales a cero.

$$\{0_{m \times n} = [a_{ij}] : a_{ij} = 0, \quad 0 \leq i \leq m, \quad 0 \leq j \leq n\} \quad ; \quad 0_{m \times n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

- **Matriz Cuadrada:** Matriz con igual número de filas que de columnas.

$$A_{n \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

A las matrices $A_{m \times n}$ que **no** son cuadradas $m \neq n$ se las denomina **rectangulares**.

Definición 3.3. Traza de una Matriz Cuadrada

Se denomina *traza* de una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ y se representa por $tr(A)$ a la suma de sus elementos diagonales.

$$tr(A) = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

- **Matriz Triangular Superior:** Matriz cuadrada donde sus entradas por debajo de la diagonal principal son cero.

$$\{A_{n \times n} = [a_{ij}] : a_{ij} = 0 \text{ si } i > j\} \quad ; \quad T = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & * \end{bmatrix}$$

- **Matriz Triangular Inferior:** Matriz cuadrada donde sus entradas por encima de la diagonal principal son cero.

$$\{A_{n \times n} = [a_{ij}] : a_{ij} = 0 \text{ si } i < j\} \quad ; \quad T = \begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix}$$

- **Matriz Diagonal:** Matriz cuadrada donde sus entradas no diagonales son cero.

$$\{A_{n \times n} = [a_{ij}] : a_{ij} = 0 \text{ si } i \neq j\} \quad ; \quad D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

- **Matriz Tridiagonal:** Matriz cuadrada con elementos no todos nulos en la *diagonal principal* y en las *subdiagonales* justo por encima y por debajo de la principal, siendo el resto de los elementos cero.

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & \cdots & 0 \\ 0 & a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & & & \vdots & \\ 0 & 0 & \cdots & a_{nn-1} & a_{nn} \end{bmatrix}$$

- **Matriz Escalar:** *Matriz diagonal* con todos sus elementos diagonales iguales.

$$D_n = \begin{bmatrix} d & 0 & \cdots & 0 \\ 0 & d & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d \end{bmatrix}$$

- **Matriz Identidad de orden n, I_n :** *Matriz escalar* con elementos diagonales de valor unidad. También puede considerarse como una matriz cuyos vectores fila y columna son los *vectores estándar*² de $\mathbb{K}^n$.

$$\{I_n = [a_{ij}] : a_{ij} = 1 \text{ si } i = j, \quad a_{ij} = 0 \text{ si } i \neq j\}$$

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [\vec{e}_1 \quad \vec{e}_2 \quad \cdots \quad \vec{e}_n] = \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vdots \\ \vec{e}_n \end{bmatrix}$$

- **Matriz Elemental:** Cualquier matriz que se obtiene al realizar una única operación elemental por fila³ sobre una matriz identidad I_n . Ejemplos de *matrices elementales*:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} ; \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} ; \quad \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

3.2. Operaciones Básicas con Matrices

3.2.1. Suma de Matrices

Definición 3.4. Suma de Matrices

Dadas dos matrices $A, B \in \mathbb{K}^{m \times n}$, se define la **suma** de ambas $A + B$ como la matriz $C \in \mathbb{K}^{m \times n}$ resultado de la operación:

$$\begin{aligned} A + B = [a_{ij}] + [b_{ij}] &= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \\ &= \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} = [a_{ij} + b_{ij}] = [c_{ij}] = C \end{aligned}$$

²Ver la Definición 2.10, página 69.

³Ver la Sección 1.4.6, página 44.

3.2.2. Producto de un Escalar por una Matriz

Definición 3.5. Producto de un Escalar por una Matriz

Se define el **producto de un escalar** $k \in \mathbb{K}$ por una **matriz** $A \in \mathbb{K}^{m \times n}$ a la matriz $B \in \mathbb{K}^{m \times n}$ resultado de la operación:

$$kA = k \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ka_{11} & \cdots & ka_{1n} \\ \vdots & & \vdots \\ ka_{m1} & \cdots & ka_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} k = Ak = B$$

Definición 3.6. Matriz Opuesta

Dada una matriz $A \in \mathbb{K}^{m \times n}$, se define la matriz **opuesta** de A como la matriz $-A = (-1)A$.

3.2.3. Diferencia o Resta de Matrices

Definición 3.7. Diferencia de Matrices

Dadas dos matrices $A, B \in \mathbb{K}^{m \times n}$, se define la **diferencia** $A - B$ como la *matriz suma* de A más la *opuesta* de B , $A - B = A + (-B) \in \mathbb{K}^{m \times n}$.

Al igual que con los vectores⁴, la *resta de matrices* no se considera una operación si no un caso particular de la *suma matricial*.

3.2.4. Propiedades de la Suma y la Multiplicación por un Escalar

Sean $A, B, C \in \mathbb{K}^{m \times n}$ y $\alpha, \beta \in \mathbb{K}$. Entonces,

- $A + B = B + A$ Conmutativa
- $(A + B) + C = A + (B + C)$ Asociativa
- $A + 0_{m \times n} = A$ Elemento Neutro
- $1A = A$ Elemento Unitario
- $A + (-A) = 0_{m \times n}$ Elemento Opuesto
- $\alpha(A + B) = \alpha A + \alpha B$ Distributiva 1
- $(\alpha + \beta)A = \alpha A + \beta A$ Distributiva 2
- $\alpha(\beta A) = (\alpha\beta)A$ Asociativa 2

3.2.5. Espacio Vectorial de Matrices

El conjunto de todas la matrices $A \in \mathbb{K}^{m \times n}$ con las operaciones matriciales *suma* y *producto por un escalar* es un *espacio vectorial*⁵.

⁴Ver la Sección 2.2.4, página 69.

⁵Ver la Definición 2.12, página 71.

3.2.6. Combinación Lineal de Matrices

Definición 3.8. Combinación Lineal de Matrices

Se dice que una matriz $A \in \mathbb{K}^{m \times n}$ es una **combinación lineal** de k matrices $A_1, A_2, \dots, A_k \in \mathbb{K}^{m \times n}$, si existen k constantes $c_1, c_2, \dots, c_k \in \mathbb{K}$ tales que:

$$A = c_1 A_1 + c_2 A_2 + \dots + c_k A_k = \sum_{j=1}^k c_j A_j$$

- Se dice que las matrices $A_1, A_2, \dots, A_k \in \mathbb{K}^{m \times n}$ son **linealmente independientes** si la única solución de la ecuación:

$$c_1 A_1 + c_2 A_2 + \dots + c_k A_k = 0_{m \times n}$$

es la *solución trivial*⁶: $c_1 = c_2 = \dots = c_k = 0$.

- Se dice que las matrices $A_1, A_2, \dots, A_k \in \mathbb{K}^{m \times n}$ son **linealmente dependientes** si existen soluciones no triviales de la ecuación:

$$c_1 A_1 + c_2 A_2 + \dots + c_k A_k = 0_{m \times n}$$

Definición 3.9. Espacios de Matrices

Se define el **espacio** generado por un conjunto de matrices como el conjunto de todas las posibles *combinaciones lineales* de las matrices del conjunto.

3.3. Producto de Matrices

Definición 3.10. Multiplicación de Matrices

Dadas las matrices $A \in \mathbb{K}^{m \times n}$ y $B \in \mathbb{K}^{n \times p}$, se define el **producto** o **multiplicación** de la matriz A por la matriz B , como aquella matriz $C \in \mathbb{K}^{m \times p}$ resultado de la operación:

$$AB = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{bmatrix} = C$$

donde cada elemento c_{ij} de la matriz resultado C se puede expresar como:

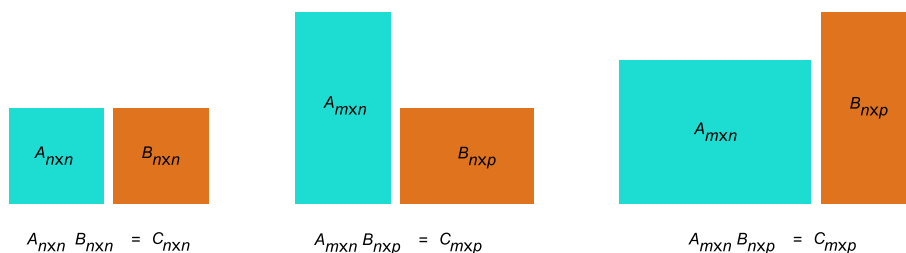
$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj}$$

- El número de **columnas** n de la matriz A debe coincidir con el número de **filas** de la matriz B .
- Los valores m (**filas** de A) y p (**columnas** de B) pueden ser cualquiera.
- Los elementos de la matriz resultante C se obtienen multiplicando los elementos de cada una de las **filas** de A por los correspondientes elementos de cada una de las **columnas** de B .
- La matriz resultante C tiene el mismo número de **filas** que A y el mismo número de **columnas** de B .

$$A_{m \times n} B_{n \times p} = C_{m \times p}$$

⁶Ver la Sección 1.8, página 59.

Nota 3.1. El número de columnas de la matriz A debe coincidir con el número de filas de la matriz B .



Ejercicio 3.1. Calcular AB siendo:

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 4} \quad ; \quad B = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 3 & 0 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 3}$$

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ -1 & 3 & 0 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 1(1) + 2(-1) + (-1)1 + 0(0) & 1(1) + 2(3) + (-1)1 + 0(0) & 1(2) + 2(0) + (-1)(-2) + 0(1) \\ 2(1) + 1(-1) + 1(1) + 1(0) & 2(1) + 1(3) + 1(1) + 1(0) & 2(2) + 1(0) + 1(-2) + 1(1) \end{bmatrix} = \\ &= \begin{bmatrix} -2 & 6 & 4 \\ 2 & 6 & 3 \end{bmatrix} \in \mathbb{R}^{2 \times 3} \end{aligned}$$

Ejercicio 3.2. Calcular AB siendo:

$$A = \begin{bmatrix} i & 2 & -1-i \\ 0 & 1+i & 2i \end{bmatrix} \in \mathbb{C}^{2 \times 3} \quad ; \quad B = \begin{bmatrix} 3i & 1-i \\ -1 & 0 \\ -i & -1 \end{bmatrix} \in \mathbb{C}^{3 \times 2}$$

$$\begin{aligned} AB &= \begin{bmatrix} i & 2 & -1-i \\ 0 & 1+i & 2i \end{bmatrix} \begin{bmatrix} 3i & 1-i \\ -1 & 0 \\ -i & -1 \end{bmatrix} = \begin{bmatrix} i(3i) + 2(-1) + (-1-i)(-i) & i(1-i) + 2(0) + (-1-i)(-1) \\ 0(1-i) + (1+i)(0) + 2i(-1) & 0(1-i) + (1+i)0 + 2i(-1) \end{bmatrix} = \\ &= \begin{bmatrix} -6+i & 2+2i \\ 1-i & -2i \end{bmatrix} \in \mathbb{C}^{2 \times 2} \end{aligned}$$

3.3.1. Propiedades de la Multiplicación de Matrices

- **Elemento Neutro:** $AI_n = I_m A = A$; $A \in \mathbb{K}^{m \times n}$
- **Asociativa 1:** $A(BC) = (AB)C$; $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{n \times p}$, $C \in \mathbb{K}^{p \times s}$
- **Asociativa 2:** $(kA)B = A(kB) = k(AB)$; $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{n \times p}$
- **Distributiva 1:** $(A+B)C = AC + BC$; $A, B \in \mathbb{K}^{m \times n}$, $C \in \mathbb{K}^{n \times p}$
- **Distributiva 2:** $A(B+C) = AB + AC$; $A \in \mathbb{K}^{m \times n}$, $B, C \in \mathbb{K}^{n \times p}$

El producto de matrices no es conmutativo y por tanto de forma general: $AB \neq BA$

3.3.2. Potencia de Matrices Cuadradas

Definición 3.11. Potencia de una Matriz Cuadrada

Dada una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ se define la operación A **elevado a la potencia** $n \in \mathbb{N}$, como el resultado del producto matricial:

$$A^n = \underbrace{AA \cdots A}_{n \text{ veces}}$$

Propiedades de la Potencia de Matrices Cuadradas

Dada $A \in \mathbb{K}^{n \times n}$ y $r, s \in \mathbb{N}$, se verifica:

- $A^r A^s = A^{r+s}$
- $(A^r)^s = A^{rs}$
- Por definición $A^0 = I_n$

3.3.3. Producto Matricial Matriz-Vector

El producto matricial de dos matrices da como resultado una nueva matriz.

- El *producto matricial* de una matriz $A \in \mathbb{K}^{m \times n}$ por una matriz (o vector) columna $V \in \mathbb{K}^{n \times 1}$ da como resultado una matriz (o vector) columna $C \in \mathbb{K}^{m \times 1}$:

$$A\vec{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix} = C_{m \times 1} = \vec{v}_c$$

- El *producto matricial* de una matriz (o vector) fila $V \in \mathbb{K}^{1 \times m}$ por una matriz $A \in \mathbb{K}^{m \times n}$ da como resultado una matriz (o vector) fila $C \in \mathbb{K}^{1 \times n}$:

$$\vec{v}A = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} v_1a_{11} + v_2a_{21} + \cdots + v_ma_{m1} & \cdots & v_1a_{1n} + v_2a_{2n} + \cdots + v_ma_{mn} \end{bmatrix} = C_{1 \times n} = \vec{v}_f$$

3.3.4. Producto Matricial Vector-Vector

- El *producto matricial* de una matriz (o vector) columna $\vec{w} \in \mathbb{K}^{n \times 1}$ por una matriz (o vector) fila $\vec{v} \in \mathbb{K}^{1 \times n}$ es igual a una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ de la forma:

$$\vec{w}\vec{v} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} w_1v_1 & w_1v_2 & \cdots & w_1v_n \\ w_2v_1 & w_2v_2 & \cdots & w_2v_n \\ \vdots & \vdots & & \vdots \\ w_nv_1 & w_nv_2 & \cdots & w_nv_n \end{bmatrix} = A_{n \times n}$$

- El *producto matricial* de una matriz (o vector) fila $\vec{v} \in \mathbb{K}^{1 \times n}$ por una matriz (o vector) columna $\vec{w} \in \mathbb{K}^{n \times 1}$ es una matriz (o vector) de un solo elemento $A \in \mathbb{K}^{1 \times 1}$, siendo por tanto un *escalar*.

$$\vec{v}\vec{w} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1w_1 + v_2w_2 + \cdots + v_nw_n \end{bmatrix} = A_{1 \times 1} \in \mathbb{K}$$

El *producto matricial* de un vector fila por un vector columna es una operación equivalente al *producto escalar en sentido real*⁷ de dos vectores.

3.3.5. Producto Matriz-Vector Estándar

Sea $A \in \mathbb{K}^{m \times n}$ y los *vectores estándar*⁸ $\vec{e}_i \in \mathbb{K}^{m \times 1}$ y $\vec{e}_j \in \mathbb{K}^{1 \times n}$. Entonces

- $A\vec{e}_i$ es la i -ésima columna de A : $A\vec{e}_i = \vec{c}_i$.

$$A\vec{e}_i = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{bmatrix} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} = \vec{c}_i$$

- $\vec{e}_j A$ es la j -ésima fila de A : $\vec{e}_j A = \vec{f}_j$.

$$\vec{e}_j A = \begin{bmatrix} 0 & \cdots & 1 & \cdots \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_{j1} & a_{j2} & \cdots & a_{jn} \end{bmatrix} = \vec{f}_j$$

3.4. Expresión Matricial de un SEL

Como se ha visto anteriormente⁹, a todo SEL de m ecuaciones con n incógnitas se le puede asociar una matriz de coeficientes $A \in \mathbb{K}^{m \times n}$ y un vector columna de términos independientes¹⁰ $\vec{b} \in \mathbb{K}^{m \times 1}$:

$$\left. \begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \right\} A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}; \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Si agrupamos las incógnitas en un vector columna $\vec{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{K}^{n \times 1}$, el SEL anterior se puede representar matricialmente como:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \implies A\vec{x} = \vec{b}$$

Todo SEL se puede definir mediante una *ecuación matricial*¹¹ en la que el *producto matricial*¹² entre la matriz de coeficientes A y el vector de incógnitas $\vec{x}$ debe ser igual al vector de términos independientes $\vec{b}$, de la forma:

$$A\vec{x} = \vec{b}$$

⁷Ver la [Definición 6.1](#), página 163.

⁸Ver la [Definición 2.10](#), página 69.

⁹Ver la [Sección 1.5.1](#), página 48.

¹⁰La matriz ampliada $A|B$ está formada por la matriz de coeficientes A y el vector columna de términos independientes $\vec{b}$.

¹¹En este caso los datos conocidos son la matriz de coeficientes A y el vector de términos independientes $\vec{b}$, y la incógnita corresponde con el vector de incógnitas $\vec{x}$.

¹²Ver la [Sección 3.3.3](#), página 95.

3.5. Matriz Transpuesta A^T

Definición 3.12. Matriz Transpuesta

Dada $A \in \mathbb{K}^{m \times n}$, se define su matriz **transpuesta** $A^T \in \mathbb{K}^{n \times m}$ aquella que se obtiene intercambiando las filas y columnas¹³ de A .

$$A = [a_{ij}] \quad \Rightarrow \quad A^T = [a_{ji}] \quad \forall i, j$$

$$A = \begin{bmatrix} \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \end{matrix} \\ \begin{matrix} a_{21} & a_{22} & \cdots & a_{2n} \end{matrix} \\ \vdots \\ \begin{matrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix} \end{bmatrix} \quad \Rightarrow \quad A^T = \begin{bmatrix} \begin{matrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{matrix} & \begin{matrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{matrix} & \cdots & \begin{matrix} a_{m1} \\ a_{m2} \\ \vdots \\ a_{mn} \end{matrix} \end{bmatrix}$$

3.5.1. Matriz Simétrica y Antisimétrica

Definición 3.13. Matriz Simétrica y Antisimétrica

- Una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ se dice que es **simétrica** si $A^T = A$.
- Una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ se dice que es **antisimétrica** si $A^T = -A$.

3.5.2. Propiedades de la Matriz Transpuesta

- $(A^T)^T = A$; $A \in \mathbb{K}^{m \times n}$
- $(A \pm B)^T = A^T \pm B^T$; $A, B \in \mathbb{K}^{m \times n}$
- $(kA)^T = k(A)^T$; $A \in \mathbb{K}^{m \times n}$, $k \in \mathbb{K}$
- $(AB)^T = B^T A^T$; $A, B \in \mathbb{K}^{n \times n}$
- $(A^r)^T = (A^T)^r$; $A \in \mathbb{K}^{n \times n}$, $r \geq 0$

Teorema 3.1.

- AA^T y $A^T A$ son siempre matrices *simétricas*.
- Si A es una matriz cuadrada, entonces $A + A^T$ es una matriz *simétrica*¹⁴.
- Si A es una matriz cuadrada, entonces $A - A^T$ es una matriz *antisimétrica*.

Demostración 3.1.

Por las propiedades de la matriz transpuesta, (ver la [Sección 3.5.2, página 97](#)), se tiene:

- $(AA^T)^T = (A^T)^T A^T = AA^T$
- $(A^T A)^T = A^T (A^T)^T = A^T A$
- $(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$
- $-(A - A^T)^T = -[A^T - (A^T)^T] = -(A^T - A) = A - A^T$

¹³Es decir, la i -ésima fila de A es la i -ésima columna de A^T y la j -ésima columna de A pasa a ser la j -ésima fila de A^T .

¹⁴Ver la [Definición 3.13, página 97](#).

Ejercicio 3.3. Dada la siguiente matriz A , calcular: $A + A^T$ y $A - A^T$.

$$A = \begin{bmatrix} 1 & -1+i \\ 0 & i \end{bmatrix} \in \mathbb{C}^{2 \times 2}$$

$$\text{a) } A + A^T = \begin{bmatrix} 1 & -1+i \\ 0 & i \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1+i & i \end{bmatrix} = \begin{bmatrix} 2 & -1+i \\ -1+i & 2i \end{bmatrix} \quad (\text{Matriz simétrica})$$

$$\text{b) } A - A^T = \begin{bmatrix} 1 & -1+i \\ 0 & i \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ -1+i & i \end{bmatrix} = \begin{bmatrix} 0 & -1+i \\ 1-i & 0 \end{bmatrix} \quad (\text{Matriz antisimétrica})$$

Ejercicio 3.4. Dada la siguiente matriz A , calcular: AA^T y $A^T A$.

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \in \mathbb{R}^{3 \times 2}$$

$$\text{a) } AA^T = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -2 \\ -1 & 1 & -1 \\ -2 & 1 & 2 \end{bmatrix} \quad (\text{Matriz simétrica})$$

$$\text{b) } A^T A = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix} \quad (\text{Matriz simétrica})$$

3.6. Matriz Transpuesta Conjugada A^H

Definición 3.14. Matriz Transpuesta Conjugada

Dada $A \in \mathbb{K}^{m \times n}$, se define su matriz **transpuesta conjugada**¹⁵ $A^H \in \mathbb{K}^{n \times m}$ a la matriz transpuesta de A con sus elementos conjugados.

$$A = [a_{ij}] \implies A^H = [\overline{a_{ji}}] \quad \forall i, j \quad ; \quad A^H = \overline{A^T} = \overline{A}^T$$

Si A es una matriz *real*, entonces $A^H = A^T$.

Ejercicio 3.5. Dada la siguiente matriz compleja A , calcular: A^T y A^H .

$$A = \begin{bmatrix} 1+i & -1 \\ 0 & i \\ -i & -1-i \end{bmatrix} \in \mathbb{C}^{3 \times 2}$$

$$A^T = \begin{bmatrix} 1+i & 0 & -i \\ -1 & i & -1-i \end{bmatrix} \quad ; \quad A^H = \begin{bmatrix} 1-i & 0 & i \\ -1 & -i & -1+i \end{bmatrix}$$

¹⁵También es típico denominarla como A^* .

3.6.1. Matriz Hermítica y Antihermítica

Definición 3.15. Matriz Hermítica y Antihermítica

- Una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ se dice que es **hermítica** si $A^H = A$.
- Una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ se dice que es **antihermítica** si $A^H = -A$.

3.6.2. Propiedades de la Matriz Transpuesta Conjugada

- $(A^H)^H = A$; $A \in \mathbb{K}^{m \times n}$
- $(A \pm B)^H = A^H \pm B^H$; $A, B \in \mathbb{K}^{m \times n}$
- $(kA)^H = \bar{k}(A)^H$; $A \in \mathbb{K}^{m \times n}$, $k \in \mathbb{K}$
- $(AB)^H = B^H A^H$; $A, B \in \mathbb{K}^{n \times n}$
- $(A^r)^H = (A^H)^r$; $A \in \mathbb{K}^{n \times n}$, $r \geq 0$
- $A^H A$ y AA^H son siempre matrices **hermíticas** para toda matriz A .

3.7. Inversa de una Matriz

Definición 3.16. Matriz Invertible y Matriz Inversa

Se dice que una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ es **invertible** o **no singular**, si existe una matriz $A^{-1} \in \mathbb{K}^{n \times n}$ tal que:

$$AA^{-1} = A^{-1}A = I_n$$

donde I_n es la *matriz identidad* de orden n . En caso de existir, la matriz A^{-1} se le denomina **matriz inversa** de A .

Una matriz no invertible se dice que es **singular**.

Nota 3.2. Dada una matriz (no necesariamente cuadrada) $A \in \mathbb{K}^{m \times n}$, se denomina :

- Inversa por la derecha: $R \in \mathbb{K}^{n \times m}$: $AR = I_m$
- Inversa por la izquierda: $L \in \mathbb{K}^{n \times m}$: $LA = I_n$

Si A es invertible entonces $R = L = A^{-1}$.

3.7.1. Cálculo de la Matriz Inversa: Método de Gauss-Jordan

Sea $A \in \mathbb{K}^{n \times n}$ una matriz *invertible*, entonces llamando $X = A^{-1}$ se tiene que cumplir:

$$AX = I_n \implies \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{nn2} & \cdots & a_{nnn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Denotando por $\vec{x}_i$ a los *vectores columna* de X y $\vec{e}_i$ a los *vectores columna* de I_n , el problema del cálculo de la matriz inversa se puede transformar en el cálculo de n SEL de la forma¹⁶:

$$A[\vec{x}_1 \ \vec{x}_2 \ \cdots \ \vec{x}_n] = [\vec{e}_1 \ \vec{e}_2 \ \cdots \ \vec{e}_n] \implies \begin{cases} A\vec{x}_1 = \vec{e}_1 \\ A\vec{x}_2 = \vec{e}_2 \\ \vdots \\ A\vec{x}_n = \vec{e}_n \end{cases}$$

¹⁶Ver la Sección 3.4, página 96.

Notar que puede utilizarse el método de Gauss-Jordan¹⁷ para resolver los n SEL a la vez:

$$\left[\underbrace{\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}}_A \quad \underbrace{\begin{array}{cccc} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{array}}_{I_n} \right] \xrightarrow{\text{Op. Elem. de Fila}} \left[\underbrace{\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & x_{11} & x_{12} & \cdots & x_{1n} \\ 0 & 1 & \cdots & 0 & x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & x_{n1} & x_{n2} & \cdots & x_{nn} \end{array}}_{I_n} \quad \underbrace{\begin{array}{cccc} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{array}}_{A^{-1}} \right]$$

Proceso de Cálculo de la Matriz Inversa (Método Gauss-Jordan)

Para encontrar la inversa de una matriz $A \in \mathbb{K}^{n \times n}$ se puede seguir los siguientes pasos:

1. Formar la matriz $n \times (2n)$ de la forma: $[A|I_n]$.
2. Calcular la matriz escalonada reducida equivalente de la anterior.
 - si el resultado es de la forma $[I_n|X]$, entonces A es invertible y $A^{-1} = X$.
 - si es de otra forma (en la parte izquierda no aparece I_n), entonces A es no invertible (*singular*).

Ejercicio 3.6. Calcular, en caso de que exista, la matriz inversa de la matriz:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 8 & 2 & 0 & 0 & 1 \end{array} \right] \xrightarrow[\substack{F_2-2F_1 \\ F_3-3F_1}]{\substack{F_2-2F_1 \\ F_3-3F_1}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 5 & -1 & -3 & 0 & 1 \end{array} \right] \xrightarrow[\substack{F_1-F_2 \\ F_3-5F_2}]{\substack{F_1-F_2 \\ F_3-5F_2}} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -1 & 7 & -5 & 1 \end{array} \right]$$

$$\xrightarrow[\substack{F_1+F_3 \\ -F_3}]{\substack{F_1+F_3 \\ -F_3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 10 & -6 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & -1 \end{array} \right] \implies A^{-1} = \begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{bmatrix}$$

Propiedades de la Matriz Inversa

Una matriz $A \in \mathbb{K}^{m \times n}$ es invertible si y solo si¹⁸:

- A es una matriz cuadrada, $m = n$.
- $\text{rang}(A) = m = n$. (Ver la definición de *rango de una matriz* en la [Sección 1.5.5, página 50](#)).

Ejercicio 3.7. Determinar si la siguiente matriz es invertible.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

La matriz es cuadrada y por tanto puede ser invertible. Vamos a calcular su *rango*, obteniendo una matriz escalonada equivalente de A .

$$\left[\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{array} \right] \xrightarrow[\substack{F_2-4F_1 \\ F_3-7F_1}]{\substack{F_2-4F_1 \\ F_3-7F_1}} \left[\begin{array}{ccc} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{array} \right] \xrightarrow{F_3-2F_2} \left[\begin{array}{ccc} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{array} \right]$$

Como $\text{rang}(A) = 2 < n = 3$, por tanto A es **no** invertible.

¹⁷Ver la [Sección 1.7, página 57](#).

¹⁸Estas propiedades están basadas en la resolución de SEL. Ver el [Teorema 1.1, página 51](#).

3.7.2. Inversa y SEL

Teorema 3.2.

Dada una matriz cuadrada invertible $A \in \mathbb{K}^{n \times n}$, entonces se verifica:

- A^{-1} es única.
- El SEL $A\vec{x} = \vec{b}$ es *compatible determinado* y su solución *única* viene dada por¹⁹:

$$\vec{x} = A^{-1}\vec{b} \quad ; \quad \forall \vec{b} \in \mathbb{K}^n$$

Dada una matriz $A \in \mathbb{K}^{m \times n}$:

- $\forall \vec{y} \in \mathbb{K}^m$:
 - si A es invertible, el SEL $A\vec{x} = \vec{y}$ tiene una única solución $\vec{x} = A^{-1}\vec{y}$.
 - si A no es invertible, el SEL $A\vec{x} = \vec{y}$ tiene infinitas soluciones o ninguna.
- En el caso especial $\vec{y} = \vec{0}$, el SEL $A\vec{x} = \vec{0}$ siempre tiene como solución $\vec{x} = \vec{0}$.
 - si A es invertible, esta será la única solución.
 - si A no es invertible, tendrá además otras infinitas soluciones más.

3.7.3. Operaciones con la Matriz Inversa

Sean $A, B \in \mathbb{K}^{n \times n}$ matrices invertibles, entonces:

- A^{-1} es invertible y $(A^{-1})^{-1} = A$
- $(cA)^{-1} = \frac{1}{c}A^{-1}$, siendo $c \in \mathbb{K}$, con $c \neq 0$.
- $(AB)^{-1} = B^{-1}A^{-1}$.
- A^T es invertible y $(A^T)^{-1} = (A^{-1})^T$.
- A^H es invertible y $(A^H)^{-1} = (A^{-1})^H$.
- A^r es invertible y $(A^r)^{-1} = (A^{-1})^r = A^{-r}$, $\forall r \geq 0$.
- Toda *matriz elemental*²⁰ es invertible y su inversa es una *matriz elemental* del mismo tipo.

3.8. Matrices por Bloques

Definición 3.17. Matriz por Bloques

Una **matriz por bloques** o **matriz particionada** es aquella que está dividida en matrices más pequeñas llamadas **bloques** o **submatrices**.

En determinados casos prácticos, particionar una matriz en bloques presenta ventajas a la hora de realizar cálculos con ella, como consecuencia de las propiedades que veremos a continuación.

¹⁹De forma general, este no es el método apropiado para resolver un SEL, debido al alto coste computacional del cálculo de la inversa de una matriz, prefiriéndose el método de Gauss en casi todas las ocasiones.

²⁰Ver la [Sección 3.1.1](#), página 89.

Ejemplo 3.1. Matriz por bloques.

$$A = \begin{bmatrix} 1 & 0 & -1 & 2 & 3 & -2 \\ 3 & 1 & 1 & -2 & 0 & 2 \\ -3 & 1 & 2 & 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 2 & 3 & -2 \\ 3 & 1 & 1 & -2 & 0 & 2 \\ -3 & 1 & 2 & 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

Donde:

$$\begin{cases} A_{11} = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & 1 \end{bmatrix} & A_{12} = \begin{bmatrix} 2 & 3 \\ -2 & 0 \end{bmatrix} & A_{13} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \\ A_{21} = \begin{bmatrix} -3 & 1 & 2 \end{bmatrix} & A_{22} = \begin{bmatrix} 2 & 0 \end{bmatrix} & A_{23} = \begin{bmatrix} 0 \end{bmatrix} \end{cases}$$

3.8.1. Propiedades de las Matrices por Bloques

Sean A y B dos matrices de igual tamaño e idéntica partición.

- **Suma:** $A + B = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} & A_{13} + B_{13} \\ A_{21} + B_{21} & A_{22} + B_{22} & A_{23} + B_{23} \end{bmatrix}$

- **Producto por un escalar:** $kA = k \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} = \begin{bmatrix} kA_{11} & kA_{12} & kA_{13} \\ kA_{21} & kA_{22} & kA_{23} \end{bmatrix}$

- **Transpuesta:** $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} A_{11}^T & A_{12}^T \\ A_{21}^T & A_{22}^T \end{bmatrix} \neq \begin{bmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \\ A_{13} & A_{23} \end{bmatrix}$

- **Transpuesta conjugada:** $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \Rightarrow A^H = \begin{bmatrix} A_{11}^H & A_{21}^H \\ A_{12}^H & A_{22}^H \\ A_{13}^H & A_{23}^H \end{bmatrix}$

- **Producto matricial:** Para poder realizar la multiplicación de dos matrices por bloques, es necesario que tanto las dimensiones de las matrices a multiplicar como las dimensiones de cada uno de los bloques, sean los apropiados para que se pueda realizar todas las multiplicaciones matriciales involucradas.

$$AB = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} =$$

$$= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} & A_{11}B_{13} + A_{12}B_{23} + A_{13}B_{33} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} & A_{21}B_{13} + A_{22}B_{23} + A_{23}B_{33} \end{bmatrix}$$

- **Matriz Inversa:** Existen algunas fórmulas para matrices por bloques invertibles, que en determinados casos permiten una reducción considerable en el cálculo computacional de la inversa. Estas fórmulas no se verán aquí.

3.8.2. Matriz Diagonal por Bloques

Definición 3.18. Matriz Diagonal por Bloques

Una **matriz diagonal por bloques** es una *matriz por bloques cuadrada*, tal que los bloques de la diagonal principal son matrices cuadradas y todos los bloques fuera de la diagonal son matrices nulas. Su forma característica sigue el siguiente patrón:

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_n \end{bmatrix} \in \mathbb{K}^{n \times n}$$

donde las matrices $A_i \in \mathbb{K}^{m_i \times m_i}$ son cuadradas y no necesariamente de la misma dimensión $m_i \geq 1$.

3.8.3. Propiedades de las Matrices Diagonales por Bloques

- **Determinante:** $|A| = |A_1| |A_2| \cdots |A_n|$
- **Traza**²¹: $\text{tr}(A) = \text{tr}(A_1) + \text{tr}(A_2) + \cdots + \text{tr}(A_n)$
- **Inversa:**

$$A^{-1} = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_n \end{bmatrix}^{-1} = \begin{bmatrix} A_1^{-1} & 0 & \cdots & 0 \\ 0 & A_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_n^{-1} \end{bmatrix}$$

3.9. Determinante

3.9.1. Introducción: Determinante de Matrices 2×2

El cálculo de la *matriz inversa* de una matriz invertible $A \in \mathbb{K}^{2 \times 2}$ de la forma:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

es sencillo siguiendo el método de Gauss-Jordan. Sin pérdida de generalidad, podemos suponer que $a \neq 0$, entonces:

$$\begin{aligned} \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right] &\xrightarrow{\frac{1}{a}F_1} \left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ c & d & 0 & 1 \end{array} \right] \xrightarrow{F_2 - cF_1} \left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & \frac{ad-cb}{a} & -\frac{c}{a} & 1 \end{array} \right] \\ &\xrightarrow{\frac{a}{ad-cb}F_2} \left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & \frac{-c}{ad-cb} & \frac{a}{ad-cb} \end{array} \right] \xrightarrow{F_1 - \frac{b}{a}F_2} \left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-cb} & \frac{-b}{ad-cb} \\ 0 & 1 & \frac{-c}{ad-cb} & \frac{a}{ad-cb} \end{array} \right] \end{aligned}$$

La matriz inversa de A será:

$$A^{-1} = \begin{bmatrix} \frac{d}{ad-cb} & \frac{-b}{ad-cb} \\ \frac{-c}{ad-cb} & \frac{a}{ad-cb} \end{bmatrix} = \frac{1}{ad-cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

De aquí podemos deducir que la matriz A será invertible, si y solo si $ad - bc \neq 0$.

²¹Ver la Definición 3.3, página 90.

Definición 3.19. Determinante de una Matriz 2×2

Dada una matriz $A \in \mathbb{K}^{2 \times 2}$:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

se denomina **determinante** de A y se denota por $\det(A)$ o bien $|A|$, al escalar:

$$\det(A) = |A| = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Ejemplo 3.2. Determinante de una Matriz.

$$A = \begin{bmatrix} 1+i & 2-i \\ -1-2i & 3i \end{bmatrix} \implies |A| = (1+i)3i - (2-i)(-1-2i) = -3+3i - (-4-3i) = \boxed{1+6i}$$

Propiedades del Determinante de una Matriz 2×2

Dada una matriz $A \in \mathbb{K}^{2 \times 2}$.

- A es invertible si y solo si $\det(A) \neq 0$.
- En caso de ser invertible, su inversa vale: $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

3.9.2. Determinante de Matrices 3×3

Dada una matriz general $A_{3 \times 3}$:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Calculando su inversa por el método de Gauss-Jordan se obtiene:

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{bmatrix} \quad \text{siendo} \quad |A| = aei + bfg + cdh - ceg - afh - bdi$$

De igual forma que en el caso 2×2 , la matriz A^{-1} existirá si y solo si $|A| \neq 0$.

3.9.3. Regla de Sarrus

La regla de Sarrus es un método sistemático para encontrar el determinante de una matriz A de hasta orden 3×3 . Consiste en escribir la matriz A y a su derecha las dos primeras columnas de A . Entonces se multiplica las entradas de las seis diagonales, sumándose o restándose en función de la dirección.

$$|A| = aei + bfg + cdh - ceg - afh - bdi$$

$$\begin{array}{ccccc} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} & \begin{bmatrix} a & b \\ d & e \\ g & h \end{bmatrix} & \begin{bmatrix} b & c \\ e & f \\ h & i \end{bmatrix} & \begin{bmatrix} c & a \\ f & d \\ i & g \end{bmatrix} & \begin{bmatrix} a & b \\ d & e \\ g & h \end{bmatrix} \\ - & - & - & + & + & + \end{array}$$

La regla de Sarrus **no** es válida²² para matrices de orden superior a 3×3 .

²²El valor obtenido de esta forma para matrices de mayor orden no tiene ninguna relación con la inversa de la matriz.

Ejercicio 3.8. Calcular el determinante de la siguiente matriz, utilizando la regla de Sarrus.

$$A = \begin{bmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{bmatrix}$$

$$|A| = 5(0)(3) + 1(-1)(2) + 2(-3)(2) - 2(0)(2) - 2(-1)(5) - 3(-3)(1) = 0 - 2 - 12 - 0 + 10 + 9 = \boxed{5}$$

3.9.4. Determinante de Matrices $n \times n$

Vamos a deducir un método general para calcular el determinante de una matriz $n \times n$. Comenzamos con el determinante de una matriz $A_{3 \times 3}$:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Como hemos visto anteriormente este determinante se puede expresar fácilmente mediante la *regla de Sarrus*. El resultado obtenido se puede reorganizar de la siguiente forma:

$$\begin{aligned} |A| &= aei + bfg + cdh - ceg - afh - bdi = a(ei - fh) - b(di - fg) + c(dh - eg) = \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \end{aligned}$$

que puede interpretarse como un *desarrollo* de determinantes 2×2 a través de los elementos de la primera fila de la matriz. También se puede obtener un resultado análogo *desarrollando* por la primera columna de la matriz A como sigue:

$$\begin{aligned} |A| &= aei + bfg + cdh - ceg - afh - bdi = a(ei - fh) - d(bi - ch) + g(bf - ce) = \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} a & g \\ c & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix} \end{aligned}$$

Se puede comprobar fácilmente que el resultado es el mismo independientemente de la fila o columna que se elija para el *desarrollo*.

Este procedimiento se puede generalizar para el cálculo del determinante de cualquier matriz $n \times n$ de forma recursiva, como función de los determinantes de submatrices $(n-1) \times (n-1)$ de A . A su vez, cada uno de los determinantes obtenidos, se pueden definir mediante los determinantes de submatrices $(n-2) \times (n-2)$ de A , y así sucesivamente.

Vamos a clarificar el cálculo definiendo previamente los conceptos de *menor complementario* y *adjunto*.

Definición 3.20. Menores y Adjuntos de una Matriz Cuadrada

Sea una matriz cuadrada $A \in \mathbb{K}^{n \times n}$ y la correspondiente matriz $A_{ij} \in \mathbb{K}^{(n-1) \times (n-1)}$ formada al eliminar la fila i -ésima y la columna j -ésima de A . Se define entonces:

- **Menor Complementario** (o simplemente **Menor**) ij de A , representado por M_{ij} , al determinante de la matriz A_{ij} .

$$M_{ij} = \det(A_{ij})$$

- **Adjunto** (o **Cofactor**) ij de A , es el menor ij de A al que se le añade un signo $(+/-)$ dado por la expresión:

$$C_{ij} = (-1)^{i+j} \det(A_{ij}) = (-1)^{i+j} M_{ij}$$

La regla de los signos utilizada en el cálculo del *adjunto*, $(-1)^{i+j}$, se puede expresar *intuitivamente* en forma de *tablero de ajedrez* como:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Ejercicio 3.9. Dada la siguiente matriz A , calcular el *menor complementario* M_{23} y el *adjunto* C_{32} .

$$A = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -2 & 1 & 2 & 3 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & 2 & 1 \end{bmatrix}$$

$$M_{23} = |A_{23}| = \begin{vmatrix} 1 & -1 & 0 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{vmatrix} = 1 - 0 - 1 + 0 - 1 + 1 = \boxed{0}$$

$$C_{32} = (-1)^5 |A_{32}| = \begin{vmatrix} 1 & 1 & 0 \\ -2 & 2 & 3 \\ -1 & 2 & 1 \end{vmatrix} = -(2 - 0 - 3 + 0 - 6 + 2) = \boxed{5}$$

3.9.5. Cálculo del Determinante mediante Desarrollos de Laplace

1. Se define el determinante de una matriz $A_{1 \times 1} = [a]$ como el escalar $\det(A) = |A| = a$.
2. Dada una matriz cuadrada $A = [a_{ij}] \in \mathbb{K}^{n \times n}$ se puede definir su **determinante** como el *escalar* obtenido por el *desarrollo* por una fila o por una columna de la forma siguiente:

- Desarrollo a lo largo de la fila i -ésima :

$$\begin{aligned} \det(A) = |A| &= (-1)^{i+1} a_{i1} |A_{i1}| + (-1)^{i+2} a_{i2} |A_{i2}| + \cdots + (-1)^{i+n} a_{in} |A_{in}| = \\ &= \sum_{j=1}^n (-1)^{i+j} a_{ij} |A_{ij}| = \sum_{j=1}^n (-1)^{i+j} a_{ij} M_{ij} = \sum_{j=1}^n a_{ij} C_{ij} \end{aligned}$$

- Desarrollo a lo largo de la columna j -ésima :

$$\begin{aligned} \det(A) = |A| &= (-1)^{1+j} a_{1j} |A_{1j}| + (-1)^{2+j} a_{2j} |A_{2j}| + \cdots + (-1)^{n+j} a_{nj} |A_{nj}| = \\ &= \sum_{i=1}^n (-1)^{i+j} a_{ij} |A_{ij}| = \sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij} = \sum_{i=1}^n a_{ij} C_{ij} \end{aligned}$$

Ejercicio 3.10. Calcular el determinante de la siguiente matriz utilizando el desarrollo de Laplace.

$$A = \begin{bmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{bmatrix}$$

Desarrollando por ejemplo a través de la primera fila:

$$|A| = 5 \begin{vmatrix} 0 & 2 \\ -1 & 3 \end{vmatrix} - (-3) \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + 2 \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} = 10 - 3 - 2 = 5$$

Desarrollando por ejemplo a través de la segunda columna se obtiene el mismo resultado:

$$|A| = -(-3) \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + 0 \begin{vmatrix} 5 & 2 \\ 2 & 3 \end{vmatrix} - (-1) \begin{vmatrix} 5 & 2 \\ 1 & 2 \end{vmatrix} = 3 * (-1) + 0 + 8 = 5$$

- El cálculo de un determinante por el método de los *desarrollos de Laplace* se simplifica al desarrollar por la fila o columna que contenga un mayor número de ceros.
- El cálculo de un determinante mediante el desarrollo de Laplace no es eficiente computacionalmente, siendo el número de operaciones totales a realizar superior a $n!$.

3.9.6. Determinante e Inversa de una Matriz

Dada una matriz *invertible*²³ A , se tiene:

$$AA^{-1} = I_n$$

Tomando determinantes a ambos lados de la igualdad:

$$|A||A^{-1}| = |I_n|$$

Como el determinante de la *matriz identidad* es la unidad²⁴, se tiene:

$$|I_n| = 1 \quad \implies \quad |A||A^{-1}| = 1$$

Y finalmente

$$|A^{-1}| = |A|^{-1} = \frac{1}{|A|}$$

Como consecuencia de esta fórmula se puede deducir la siguiente propiedad:

Una matriz cuadrada A es *invertible* si y solo si $|A| \neq 0$.

3.9.7. Propiedades de los Determinantes

Sean $A, B \in \mathbb{K}^{n \times n}$ matrices cuadradas.

- $|AB| = |A| |B|$
- En general: $|A + B| \neq |A| + |B|$
- $|A^{-1}| = |A|^{-1} = 1/|A|$
- $|A^T| = |A|$
- $|A^H| = |\bar{A}|$

²³Ver la **Definición 3.16**, página 99.

²⁴Basta ir desarrollando por cada uno de los elementos de la diagonal.

²⁵Esto implica de forma general que la sustitución de una de fila por cualquier combinación lineal entre filas que incluya a dicha fila **una sola vez**, no altera el valor del determinante de la matriz.

$$F_i \rightarrow F_i + aF_1 + bF_2 + \dots$$

Sí F_i aparece multiplicado por cualquier escalar distinto de uno, el determinante queda multiplicado por ese valor.

Sean A , B y C matrices cuadradas de orden n , entonces:

- Si A tiene una fila o columna de ceros, entonces $|A| = 0$.
- Si A tiene dos filas o columnas iguales, entonces $|A| = 0$.
- Si B se obtiene de intercambiar dos filas o columnas de A , entonces $|B| = -|A|$.
- Si A es una matriz triangular (superior, inferior o diagonal), su determinante es igual al producto de los elementos de su diagonal principal.
- Si B se obtiene de multiplicar una fila o columna por el escalar k , entonces $|B| = k|A|$. Por tanto $\det(kA) = k^n|A|$.
- Si B se obtiene al sumar un múltiplo de una fila (columna) de A a otra fila (columna) de A , entonces²⁵ $|B| = |A|$.
- Si A , B y C son matrices iguales excepto que la i -ésima fila (o i -ésima columna) de C es la suma de las correspondientes i -ésimas filas (o columnas) de A y B , entonces, $|C| = |A| + |B|$.

3.9.8. Método de Gauss para el Cálculo del Determinante

Como se ha visto en las propiedades del determinante, la sustitución de una fila de la matriz por una combinación lineal del resto de las filas no altera el valor del determinante. Por otra parte, el valor del determinante de una matriz triangular (escalónada) es simplemente el producto de sus elementos diagonales.

Por tanto, un método apropiado para calcular el determinante de una matriz es utilizar el método de Gauss para obtener una matriz escalonada equivalente, y obtener su determinante.

Hay que tener en cuenta que no todas las operaciones elementales de fila dejan inalterable el determinante de la matriz.

- Cuando se multiplica una fila por un escalar, el valor del determinante se multiplica por ese escalar. Por tanto, en el resultado final habrá que dividirlo por él.
- El intercambio de filas, cambia el signo del determinante.

Ejercicio 3.11. Aplicar el método de Gauss para calcular el determinante de la siguiente matriz:

$$A = \begin{bmatrix} 5 & 5 & -5 & 5 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$$

$$\begin{aligned} & \begin{bmatrix} 5 & 5 & -5 & 5 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \xrightarrow[\substack{\frac{1}{5}F_1^* \\ F_2 \leftrightarrow F_3}]{\substack{F_3 + F_1 \\ F_4 - F_1}} \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & -1 \end{bmatrix} \xrightarrow[\substack{F_3 + F_1 \\ F_4 - F_1}]{\substack{F_3 - F_2 \\ F_4 + 2F_3}} \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & -2 \end{bmatrix} \end{aligned}$$

El determinante de la matriz resultante es: $1 * 1 * (-1) * (-2) = 2$. El determinante de la matriz pedida no es este, ya que hemos hecho dos operaciones elementales de fila (las dos primeras marcadas con un asterisco), que han modificado el valor del determinante original. El producto por $\frac{1}{5}$ la primera fila y el intercambio de las filas 2 y 3. Para deshacer estos cambios, habrá que multiplicar por 5 y cambiar de signo el valor del determinante calculado, por tanto el resultado final es:

$$|A| = -10$$

3.9.9. Regla de Cramer

La *regla de Cramer* permite obtener la solución de un SEL compatible determinado mediante la utilización del determinante de la matriz de coeficientes.

Sea el sistema de n ecuaciones lineales con n incógnitas:

$$A\vec{x} = \vec{b} \quad \begin{cases} A \in \mathbb{K}^{n \times n} \\ \vec{b} \in \mathbb{K}^n \end{cases}$$

Denotemos por $A_i(\vec{b})$ a la matriz obtenida al reemplazar la i -ésima columna de A por el vector $\vec{b}$,

$$A_i(\vec{b}) = [\vec{c}_1 \quad \vec{c}_2 \quad \cdots \quad \underset{\substack{\uparrow \\ \text{Columna } i}}{\vec{b}} \quad \cdots \quad \vec{c}_n]$$

A través de esta matriz se puede expresar la solución del anterior sistema de ecuaciones lineales de la forma dada por el siguiente teorema.

Teorema 3.3. Regla de Cramer

Sea una matriz cuadrada e invertible $A \in \mathbb{K}^{n \times n}$ y sea $\vec{b}$ un vector de $\mathbb{K}^n$. Entonces la solución del sistema $A\vec{x} = \vec{b}$ está dada por:

$$x_i = \frac{\det(A_i(\vec{b}))}{|A|} \quad ; \quad \forall i = 1, \dots, n$$

Demostración 3.3.

Como las columnas de la matriz identidad I son los vectores estándar columna $\vec{e}_i$, se tiene que:

$$\begin{aligned} AI_i(\vec{x}) &= A[\vec{e}_1 \quad \cdots \quad \vec{x} \quad \cdots \quad \vec{e}_n] = [A\vec{e}_1 \quad \cdots \quad A\vec{x} \quad \cdots \quad A\vec{e}_n] = \\ &= [\vec{c}_1 \quad \cdots \quad \vec{b} \quad \cdots \quad \vec{c}_n] = A_i(\vec{b}) \end{aligned}$$

Tomando determinantes a ambos lados de la igualdad,

$$|A| \det(I_i(\vec{x})) = \det(A_i(\vec{b}))$$

siendo:

$$\det(I_i(\vec{x})) = \begin{vmatrix} 1 & 0 & \cdots & x_1 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & x_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & x_i & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{n-1} & \cdots & 1 & 0 \\ 0 & 0 & \cdots & x_n & \cdots & 0 & 1 \end{vmatrix} = x_i$$

como se puede ver al desarrollar por la fila i -ésima. Por tanto, despejando de la expresión $|A| x_i = \det(A_i(\vec{b}))$ se obtiene el resultado anterior.

Desde un punto de vista computacional la *regla de Cramer* es **ineficiente** para resolver sistemas, salvo los muy pequeños, debido al gran coste de cálculo de los determinantes.

3.9.10. Matriz Adjunta e Inversa de una Matriz

Dada una matriz cuadrada $A \in \mathbb{K}^{n \times n}$, se puede construir una nueva *matriz de adjuntos* sustituyendo cada elemento a_{ij} por el correspondiente *adjunto*²⁶ de cada elemento,

$$C_{ij} = (-1)^{i+j} M_{ij} = (-1)^{i+j} \det(A_{ij})$$

Definición 3.21. Matriz Adjunta

Se denomina **matriz adjunta** de una matriz A y se representa por $\text{adj}(A)$, a la transpuesta de la *matriz de adjuntos*, dada por:

$$\text{adj}(A) = [C_{ij}]^T = [C_{ji}] = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

Teorema 3.4.

Para toda matriz invertible $A \in \mathbb{K}^{n \times n}$ se tiene:

$$A^{-1} = \frac{1}{|A|} \text{adj}(A)$$

Demostración 3.4.

La inversa de una matriz invertible A es una matriz X tal que:

$$AX = I \quad ; \quad X = [\vec{x}_1 \quad \vec{x}_2 \quad \cdots \quad \vec{x}_n]$$

siendo $\vec{x}_j$ la columna j -ésima de la matriz X . Para determinar cada una de estos vectores columna se puede resolver los sistemas $A\vec{x}_j = \vec{e}_j$ y por la regla de Cramer:

$$x_{ij} = \frac{\det(A_i(\vec{e}_j))}{|A|}$$

y como:

$$\det(A_i(\vec{e}_j)) = \begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & 1 & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = (-1)^{j+i} \det(A_{ji}) = C_{ji}$$

se deduce fácilmente la expresión anterior.

²⁶Ver la Definición 3.20, página 105.

3.10. Subespacios Fundamentales de una Matriz

La información contenida en una matriz está almacenada en el valor, cantidad y disposición de sus elementos. Si la matriz es de grandes dimensiones, resulta muy complicado establecer algún tipo de relación entre todos sus elementos.

En esta sección vamos a estudiar una forma de analizar esta información basada en la determinación de los *subespacios vectoriales* asociados a una matriz.

3.10.1. Espacios Columna y Fila de una Matriz

Una matriz general $A \in \mathbb{K}^{m \times n}$ puede ser vista como un vector de vectores fila ($\vec{f}_i$) o como un vector de vectores columna ($\vec{c}_i$).

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = [\vec{c}_1 \quad \cdots \quad \vec{c}_n] = \begin{bmatrix} \vec{f}_1 \\ \vdots \\ \vec{f}_m \end{bmatrix}$$

Definición 3.22. Espacio Columna de una Matriz

Se define el **espacio columna**²⁷ de $A \in \mathbb{K}^{m \times n}$ como el subespacio vectorial de $\mathbb{K}^m$ generado por la *combinación lineal* de todos los vectores columna $\vec{c}_i$ de A . Se representa por $\text{col}(A)$.

$$\text{col}(A) = \{\vec{x} \in \mathbb{K}^m : \vec{x} = \alpha_1 \vec{c}_1 + \cdots + \alpha_n \vec{c}_n \quad ; \quad \alpha_i \in \mathbb{K}\}$$

Definición 3.23. Espacio Fila de una Matriz

Se define el **espacio fila** de $A \in \mathbb{K}^{m \times n}$ como el subespacio vectorial de $\mathbb{K}^n$ generado por la *combinación lineal* de todos los vectores fila $\vec{f}_i$ de A . Se representa por $\text{fil}(A)$.

$$\text{fil}(A) = \{\vec{x} \in \mathbb{K}^n : \vec{x} = \alpha_1 \vec{f}_1 + \cdots + \alpha_m \vec{f}_m \quad ; \quad \alpha_i \in \mathbb{K}\}$$

Ejercicio 3.12. Determinar el *espacio fila* y el *espacio columna* de la siguiente matriz:

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

El espacio fila es el subespacio vectorial generado por la combinación lineal de los tres vectores filas de A . Analizando los vectores fila de este caso, vemos solamente hay dos filas linealmente independientes (la primera es combinación lineal de la segunda menos la primera) y por tanto el espacio fila corresponde con un plano de dimensión 2 en $\mathbb{R}^4$. Una base del subespacio fila de A es por ejemplo:

$$\mathcal{B}_{\text{fil}} = \{[1, 0, 0, 1], [-1, 0, 1, -1]\}$$

El espacio columna de A está formado por la combinación lineal de sus cuatro vectores columna. Claramente en este caso solo hay dos vectores columna linealmente independientes, por tanto el espacio columna corresponde con un plano (dimensión 2) en $\mathbb{R}^3$, siendo por ejemplo una de sus bases:

$$\mathcal{B}_{\text{col}} = \{[1, -1, 0]^T, [0, 1, 1]^T\}$$

²⁷En la [Definición 4.11, página 136](#) se estudia la relación entre *espacio columna* de una matriz y la *imagen* de ésta.

Ejercicio 3.13. Dada la matriz:

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Determinar si el vector $\vec{u} = [1, 0, 0, 1]^T$ pertenece al espacio columna de A y si el vector $\vec{v} = [1, 2, 3]$ pertenece al espacio de fila de A .

Empezamos con el vector $\vec{u}$. Dicho vector pertenecerá a $\text{col}(A)$ si existe una combinación lineal de los vectores columna de A que genere $\vec{u}$, y por tanto que existan escalares x_1, x_2 y x_3 tales que:

$$x_1 \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Al igual que se hizo en el [Ejercicio 2.13, página 75](#), esto se puede transformar en un SEL de la forma:

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Solo si este sistema es compatible, el vector $\vec{u} \in \text{col}(A)$. Resolviendo por Gauss:

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{array} \right] \xrightarrow[\substack{F_3+F_1 \\ F_4-F_1}]{F_2+F_1} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & -1 & 2 & 1 \\ 0 & 2 & 0 & 0 \end{array} \right] \xrightarrow[\substack{F_4+2F_3}]{F_1-F_3} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 4 & 2 \end{array} \right] \xrightarrow{F_4-2F_2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Que claramente corresponde con un SEL compatible y determinado, (la forma escalonada está desordenada ya que basta con identificarla), y por tanto:

$$\vec{u} \in \text{col}(A)$$

Veamos ahora si el vector $\vec{v}$ pertenece al espacio fila de A . En este caso se tiene que cumplir:

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

en el cual por comodidad, se han colocado los vectores fila de A como vectores columna. Al igual que en el caso anterior, esto corresponde con el siguiente SEL:

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Resolviendo por Gauss:

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 3 \end{array} \right] \xrightarrow[\substack{F_3-F_1}]{F_2+F_1} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 1 \\ 0 & 0 & -1 & 2 & 3 \\ 0 & 2 & 2 & 0 & 2 \end{array} \right] \quad (3.1)$$

Como se obtiene una matriz escalonada (desordenada) el sistema es compatible indeterminado y por tanto:

$$\vec{v} \in \text{fil}(A)$$

Nota 3.3. En este problema no ha sido necesario resolver completamente los SEL, basta con analizarlos, ya que solo se pide determinar la pertenencia o no a los subespacios fila y columna de la matriz. Sería necesario resolverlos en el caso de pedir una posible combinación lineal que demuestre la pertenencia.

Ejercicio 3.14. Encontrar una posible combinación lineal entre los vectores fila de la matriz A que genere el vector $\vec{v}$ correspondientes al **Ejemplo 3.13**.

Resolviendo el SEL **Ecuación (3.1)** del **Ejemplo 3.13** se tiene:

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 1 \\ 0 & 0 & -1 & 2 & 3 \\ 0 & 2 & 2 & 0 & 2 \end{array} \right] \Rightarrow \begin{cases} x_1 = 1 + x_2 + x_3 - x_4 = 1 + 4 - 2t - 3 + 2t - t = 2 - t \\ x_2 = 1 - x_3 = 1 + 3 - 2t = 4 - 2t \\ x_3 = -3 + 2x_4 = -3 + 2t \\ x_4 = t \end{cases}$$

Hay por tanto infinitas soluciones posibles. Por ejemplo, una de ellas se obtiene para $t = 0$:

$$t = 0 \quad \Rightarrow \quad [x_1 = 2 \ ; \ x_2 = 4 \ ; \ x_3 = -3 \ ; \ x_4 = 0]$$

Por tanto, una combinación lineal de los vectores fila de A que produce $\vec{v}$ es:

$$\begin{aligned} x_1 \vec{f}_1 + x_2 \vec{f}_2 + x_3 \vec{f}_3 + x_4 \vec{f}_4 &= 2 \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} + 4 \begin{bmatrix} -1 & 1 & 1 \end{bmatrix} - 3 \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} + 0 \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 2 - 4 + 3 + 0 & -2 + 4 - 0 + 0 & 2 + 4 - 3 + 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \vec{v} \end{aligned}$$

3.10.2. Propiedades de las Matrices Equivalentes por Filas

Teorema 3.5. Propiedades de las Matrices Equivalentes por Filas

Sea $A \in \mathbb{K}^{m \times n}$, entonces se verifica:

- Si B es una matriz *equivalente por filas*²⁸ a la matriz A entonces²⁹, $\text{fil}(A) = \text{fil}(B)$.
- $\dim(\text{fil}(A)) = \dim(\text{col}(A)) = \text{rang}(A)$.

Demostración 3.5.

- Las matrices *equivalente por filas* se obtienen mediante *operaciones elementales por fila*, (ver la **Sección 1.5.2, página 49**): intercambio de filas, multiplicación de una fila por un escalar no nulo y sumar a una fila una combinación lineal de otras filas. Mediante estas tres operaciones se generan distintos vectores que siguen perteneciendo al espacio fila de la matriz original, conservándose además el número de vectores linealmente independientes de la matriz.

Por tanto, los vectores fila de la nueva matriz equivalente a la original constituyen un *sistema generador*, (ver la **Sección 2.5, página 79**), del mismo espacio fila.

- Para determinar $\dim(\text{fil}(A))$, (ver la **Definición 2.26, página 80**), hay que encontrar el número de vectores fila de A linealmente independientes, para ello se determina el rango de la matriz, (ver la **Sección 1.5.5, página 50**).

Para determinar $\dim(\text{col}(A))$ hay que encontrar el número de vectores columna de A linealmente independientes, para ello se determina el rango del conjunto de vectores, (ver la **Sección 2.4.3, página 76**), mediante el cálculo del rango de la matriz.

Por tanto $\dim(\text{fil}(A)) = \dim(\text{col}(A)) = \text{rang}(A)$.

²⁸Ver la **Sección 1.5.3, página 49**.

²⁹Aunque no necesariamente $\text{col}(A) = \text{col}(B)$.

3.10.3. Rango, Núcleo y Nulidad de una Matriz

Definición 3.24. Rango de una Matriz

Dada la matriz $A \in \mathbb{K}^{m \times n}$, se puede definir³⁰ el **rango** de A como la *dimensión* de su *espacio fila* o de su *espacio columna*³¹. Se denota por $\text{rang}(A)$.

$$\text{rang}(A) = \dim(\text{fil}(A)) = \dim(\text{col}(A))$$

Recordamos que el *rango de una matriz* coincide con el *rango del conjunto de vectores*³² tanto fila como columna de la matriz.

Ejercicio 3.15. Calcular el rango del conjunto de vectores:

$$V = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 0 \\ 2 \end{bmatrix} \right\}$$

Hay que determinar el número de *vectores linealmente independientes* que hay en V , (ver la [Sección 2.4.2, página 73](#)). En este caso como mínimo hay uno, (ya que no todos los vectores son nulos), y como máximo cuatro, que es el número de vectores de V . La forma más apropiada de calcular el rango es construir una matriz con los vectores de V dispuestos en *filas* o en *columnas* y determinar el *rango de la matriz*, el cual coincide con el *rango* del conjunto de vectores V , (ver la [Definición 2.20, página 76](#)). En nuestro caso vamos a disponerlos como vectores columna, ya que de esta forma además de poder determinar cuantos vectores linealmente independientes hay, también se puede determinar más fácilmente cuáles son a través de la posición de los *pivotes*. Si se disponen como filas, al realizar alguna operación elemental de intercambio de filas, cambia el orden de los vectores, y es necesario llevar un registro de los cambios realizados para localizar los vectores linealmente independientes. Por tanto:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 2 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 2 \end{bmatrix}$$

Para determinar el *rango* de la matriz A se realizan *operaciones elementales de fila* hasta obtener una matriz en *forma escalonada equivalente*, (ver el método de Gauss matricial en la [Sección 1.6, página 51](#)), la cual posee el mismo *espacio fila* que la original, (ver el [Teorema 3.5, página 113](#)), y por tanto el mismo *rango*. Aplicando Gauss:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 2 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 2 \end{bmatrix} \xrightarrow[\begin{smallmatrix} F_3 - F_1 \\ F_4 - F_1 \end{smallmatrix}]{\begin{smallmatrix} F_2 - 2F_1 \\ F_3 + F_1 \end{smallmatrix}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 2 \end{bmatrix} \xrightarrow[\begin{smallmatrix} F_4 - F_2 \end{smallmatrix}]{\begin{smallmatrix} F_3 + 2F_2 \end{smallmatrix}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

El rango de la matriz resultante es 3, correspondiente a los tres *pivotes*: 1, -1 y 3, (ver la [Definición 1.31, página 50](#) y la [Definición 1.24, página 45](#)).

Por tanto el *rango* del sistema de vectores V es 3, siendo los vectores *linealmente independientes* el primero, segundo y tercero, (correspondientes a las columnas de los *pivotes*).

³⁰En la [Sección 1.5.5, página 50](#) se dio una definición alternativa y equivalente del *rango* de una matriz.

³¹Y por tanto, según la [Definición 2.20, página 76](#), el *rango* de una matriz coincide con el número de vectores fila y vectores columna linealmente independientes que tiene la matriz.

³²Ver la [Definición 2.20, página 76](#).

Definición 3.25. Núcleo o Espacio Nulo de una Matriz

Dada la matriz $A \in \mathbb{K}^{m \times n}$, se define el **núcleo** o **espacio nulo** de A al subespacio de $\mathbb{K}^n$ formado por las soluciones del sistema lineal *homogéneo*³³ $A\vec{x} = \vec{0}$. Se denota³⁴ por $\ker(A)$.

$$\ker(A) = \{\vec{x} \in \mathbb{K}^n : A\vec{x} = \vec{0}\}$$

Definición 3.26. Nulidad de una Matriz

Dada la matriz $A \in \mathbb{K}^{m \times n}$, se define la **nulidad** de A al valor de la dimensión del *núcleo* de una matriz.

$$\text{nul}(A) = \dim(\ker(A))$$

Ejercicio 3.16. Encontrar el *núcleo* y la *nulidad* de la siguiente matriz real:

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Para determinar el *núcleo* tenemos que resolver el sistema homogéneo $A\vec{x} = \vec{0}$. Lo vamos a hacer por Gauss:

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \xrightarrow{F_2+F_1} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \xrightarrow{F_3-F_2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{cases} x_1 = -x_4 \\ x_2 = t_1 \\ x_3 = 0 \\ x_4 = t_2 \end{cases}$$

En forma vectorial:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t_1 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow \ker(A) = \text{gen} \left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right) ; \quad \boxed{\text{nul}(A) = 2}$$

El núcleo corresponde con un *plano* en $\mathbb{R}^4$ y por tanto su dimensión³⁵ es 2, (correspondiente a los dos parámetros libres t_1 y t_2) y por tanto la nulidad es 2.

Los cuatro subespacios fundamentales de toda matriz $A \in \mathbb{K}^{m \times n}$ son:

- **Espacio Fila:** $\text{fil}(A)$.
- **Espacio Columna:** $\text{col}(A)$.
- **Espacio Nulo o Núcleo:** $\ker(A)$.
- **Núcleo de la Transpuesta**³⁶: $\ker(A^T)$.

Estos valores determinan perfectamente la información contenida y las propiedades de toda matriz, de ahí la importancia fundamental de su conocimiento.

³³Ver la [Sección 1.8, página 59](#).

³⁴En lengua inglesa la palabra *kernel* se traduce como *núcleo*.

³⁵Ver la [Sección 1.2.2, página 36](#).

³⁶Este subespacio será tratado en más profundidad en la [Sección 6.4.5, página 175](#).

3.10.4. Teorema del Rango de las Matrices

Teorema 3.6. Teorema del Rango de las Matrices

Sea $A \in \mathbb{K}^{m \times n}$, entonces se verifica:

$$\text{rang}(A) + \text{nul}(A) = n.$$

- Este teorema establece la relación entre dos de los cuatro subespacios fundamentales de una matriz demostrando que no son independientes.
- En la [Sección 6.4.6, página 177](#) se generaliza este teorema estableciendo la relación entre los otros dos subespacios fundamentales.

Demostración 3.6.

$$\text{nul}(A) = \text{Num. de parámetros libres de la solución del Sist. Homog.} = n - \text{rang}(A)$$

$$\implies \text{nul}(A) + \text{rang}(A) = (n - \text{rang}(A)) + \text{rang}(A) = n$$

Ejercicio 3.17. Comprobar el teorema del rango en la matriz del [Ejemplo 3.16, página 115](#).

Esta matriz tiene 4 columnas, por tanto $n = 4$ y la nulidad se calculó resultando ser 2. El rango de la matriz se puede determinar fácilmente a través de la matriz resultado del método de Gauss en el [Ejemplo 3.16](#), donde vemos que posee dos pivotes y por tanto el rango también es 2. Por tanto se tiene:

$$\text{rang}(A) + \text{nul}(A) = n \implies 2 + 2 = 4$$

Teorema 3.7.

Sea $A \in \mathbb{K}^{m \times n}$, entonces se verifica³⁷:

- $\text{rang}(A^T) = \text{rang}(A^H) = \text{rang}(A)$.
- $\text{rang}(A^T A) = \text{rang}(A^H A) = \text{rang}(A)$.
- La matriz $A^T A \in \mathbb{K}^{n \times n}$ es invertible si y solo si $\text{rang}(A) = n$ (igual con la matriz $A^H A$).

Demostración 3.7.

- Ya que se cumple, (ver el [Teorema 3.5, página 113](#)):

$$\text{rang}(A) = \dim(\text{col}(A)) = \dim(\text{fil}(A)) = \text{rang}(A^T)$$

- Como $A^T A \in \mathbb{R}^{n \times n}$, tiene el mismo número de columnas que A y por el teorema del rango de las matrices, (ver el [Teorema 3.6, página 116](#)):

$$\text{rang}(A) + \text{nul}(A) = n = \text{rang}(A^T A) + \text{nul}(A^T A)$$

Entonces, basta con demostrar que $\text{nul}(A) = \text{nul}(A^T A)$ para que $\text{rang}(A) = \text{rang}(A^T A)$. Para ello vamos a demostrar que todo $\vec{x}$ que pertenece a $\text{nul}(A)$ también pertenece a $\text{nul}(A^T A)$ y viceversa.

1. Todo $\vec{x} \in \text{nul}(A)$ cumple $A\vec{x} = \vec{0} \implies A^T A\vec{x} = A^T \vec{0} = \vec{0}$, y por tanto $\vec{x} \in \text{nul}(A^T A)$.
2. Para todo $\vec{x} \in \text{nul}(A^T A)$ se tiene $A^T A\vec{x} = \vec{0} \implies \vec{x}^T A^T A\vec{x} = \vec{x}^T \vec{0} = 0$. Entonces:

$$\vec{x}^T A^T A\vec{x} = (A\vec{x})^T (A\vec{x}) = (A\vec{x}) \cdot (A\vec{x}) = 0 \implies A\vec{x} = \vec{0}$$

y por tanto $\vec{x} \in \text{nul}(A)$.

³⁷Las propiedades de la matriz $A^T A$ serán estudiadas más detenidamente en el [Tema 7, página 195](#).

Por lo cual, $\text{nul}(A) = \text{nul}(A^T A) \implies \text{rang}(A) = \text{rang}(A^T A)$.

- Por las propiedades de la inversa de una matriz, (ver la [Sección 3.7.3, página 101](#)), se tiene:

$$(A^T A)^{-1} = A^{-1}(A^T)^{-1} = A^{-1}(A^{-1})^T$$

Por tanto A debe ser invertible y para ello $\text{rang}(A) = n$, (ver la [Sección 3.7.1, página 100](#)).

3.10.5. Teorema Fundamental de las Matrices Invertibles

Los *subespacios asociados* a una matriz determinan todas las características y propiedades de la misma. En el siguiente teorema se reúnen las interrelaciones que hay entre ellos y el concepto de la *inversa* de una matriz³⁸.

Teorema 3.8. Teorema Fundamental de las Matrices Invertibles

Sea $A \in \mathbb{K}^{n \times n}$. Las siguientes afirmaciones son equivalentes:

- A es *invertible*, (ver la [Definición 3.16, página 99](#)).
- Las matrices A^T y A^H son invertibles.
- $A\vec{x} = \vec{b}$ tiene *solución única*³⁹ $\forall \vec{b} \in \mathbb{K}^n$.
- $A\vec{x} = \vec{0}$ tiene solo la solución *trivial*⁴⁰ $\vec{x} = \vec{0} \in \mathbb{K}^n$.
- La *forma escalonada reducida equivalente*⁴¹ de A es I_n .
- $\det(A) \neq 0$, (ver la [Sección 3.9.6, página 107](#)).
- $\text{rang}(A) = n$, (ver la [Sección 1.5.5, página 50](#) y la [Definición 3.24, página 114](#)).
- $\ker(A) = \{\vec{0}\}$, (ver [Definición 3.25, página 115](#)).
- Tanto los vectores fila como los vectores columna de A son *linealmente independientes* y forman *bases*⁴² de $\mathbb{K}^n$.

3.10.6. Bases para los Espacios Fila, Columna y Nulo de una Matriz

Dada una matriz $A \in \mathbb{K}^{m \times n}$, los pasos a seguir para calcular *bases* para el espacio fila, el espacio columna y/o el espacio nulo son los siguientes:

1. Encontrar la *forma escalonada*⁴³ A . La denominaremos R .
2. Los vectores fila **no nulos** (con *pivotes*⁴⁴) de R forman una base⁴⁵ del $\text{fil}(A)$.
3. Los vectores columna de A **correspondientes** a los vectores columna de R con *pivote*, forman una base del $\text{col}(A)$.
4. Resolver el sistema $R\vec{x} = \vec{0}$, expresando la solución como una combinación lineal, (a través de posibles parámetros libres), de un conjunto de vectores. Estos vectores forman una *base* del $\ker(A)$.

³⁸En el [Teorema 8.2, página 227](#) se ampliará este teorema con el concepto de *autovalor* de una matriz.

³⁹Ver la [Sección ??, página ??](#).

⁴⁰Ver la [Sección 1.8, página 59](#).

⁴¹Ver la [Definición 1.32, página 50](#).

⁴²Ver la [Sección 2.5.1, página 80](#).

⁴³Puede ser la escalonada normal ([Sección 1.5.4, página 50](#)) o la reducida ([Sección 1.7.1, página 57](#)). Esta última es más costosa de calcular pero las bases resultantes pueden, en algunos casos, ser más sencillas y útiles para su posterior utilización.

⁴⁴Ver la [Definición 1.31, página 50](#).

⁴⁵También se pueden tomar las correspondientes filas de la matriz original A , teniendo en cuenta los cambios de fila realizados, en lugar de las filas de la matriz escalonada R como base del $\text{fil}(A)$, aunque normalmente serán más sencillas (con más ceros) y útiles las obtenidas de la matriz R .

Ejercicio 3.18. Determinar bases para los espacios fila, columna y núcleo de la siguiente matriz:

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{K}^{4 \times 5}$$

Vamos a seguir los pasos descritos en la [Sección 3.10.6, página 117](#).

1. Forma escalonada R de A .

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{\substack{F_3 - F_1 \\ F_4 + F_1}} \begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{\substack{F_3 + F_2 \\ F_4 - F_2}} \begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} = R$$

2. Como hay pivote en las cuatro filas de R , nos indica que las cuatro filas de A son linealmente independientes y por tanto forman una base del espacio fila de A . Por otra parte, como la matriz escalonada equivalente R tiene el mismo espacio fila⁴⁶, una base más apropiada⁴⁷ del espacio fila de A se puede formar con las filas de R .

$$\mathcal{B}_{\text{fil}(A)} = \{[1, 2, 0, 1, 0], [0, 1, -1, 0, 0], [0, 0, 0, 0, 1], [0, 0, 1, 1, 1]\}$$

3. Una base del espacio columna de A está formado por el conjunto de vectores columna linealmente independientes de A , y son aquellas columnas de A **correspondientes** a las columnas con pivote de la matriz R . En este caso corresponden con las columnas: 1, 2, 3 y 5 de A .

$$\mathcal{B}_{\text{col}(A)} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Notar que en este caso no se puede tomar como base del espacio columna de A , las columnas pivotes de R , ya que no se conserva el espacio columna en las matrices escalonadas equivalentes⁴⁸.

4. Para encontrar una base del núcleo de A tenemos que operar igual que en el [Ejemplo 3.16, página 115](#). Para ello resolvemos el SEL homogéneo $A\vec{x} = \vec{0}$ a partir de la matriz R .

$$\left\{ \begin{array}{lcl} x_1 & = & -2x_2 - x_4 = t \\ x_2 & = & x_3 = -t \\ x_3 & = & -x_4 - x_5 = -t \\ x_4 & = & t \\ x_5 & = & 0 \end{array} \right\}; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = t \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \Rightarrow \ker(A) = \text{gen} \left(\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right)$$

Una base del núcleo de A será en este caso:

$$\mathcal{B}_{\ker(A)} = \{[1, -1, -1, 1, 0]\}$$

Podemos también comprobar como se cumple el teorema del rango en esta matriz, ya que el rango de A (4) más su nulidad (1) es igual al número de columnas de A (5).

⁴⁶Ver el [Teorema 3.5, página 113](#).

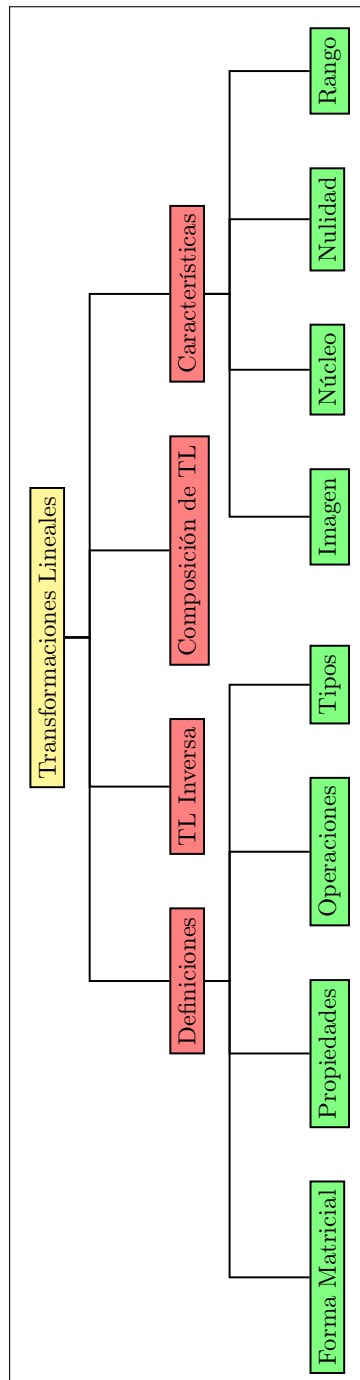
⁴⁷Por ser más sencilla, ya que sus vectores poseen un mayor número de ceros.

⁴⁸En este caso, la posición de las columnas de los pivotes de R sirven para determinar la posición de las columnas de A que son linealmente independientes.

Transformaciones Lineales

Contenido

4.1	Concepto de Transformación Lineal	121
4.1.1	Forma Matricial de una Transformación Lineal	122
4.1.2	Operaciones con Transformaciones Lineales	123
4.1.3	Propiedades de las Transformaciones Lineales	123
4.2	Transformaciones Lineales Elementales	125
4.2.1	Transformación Lineal de Vectores Estándar	126
4.2.2	Demostración $\Leftarrow$ de la Sección 4.1.3	126
4.3	Inversa de una TL	127
4.3.1	Función Invertible	127
4.3.2	Inversa de una TL	131
4.4	Operaciones con TL	132
4.4.1	Suma de Transformaciones Lineales	132
4.4.2	Composición de Transformaciones	133
4.5	Imagen y Núcleo de una TL	136
4.5.1	Imagen de una TL	136
4.5.2	Núcleo o Espacio Nulo de una TL	137
4.5.3	Nulidad y Rango de una TL	139
4.6	Tipos de TL	141
4.6.1	Inyectivas, Sobreyectivas y Biyectivas	141
4.6.2	Homomorfismos, Endomorfismos e Isomorfismos	142



4.1. Concepto de Transformación Lineal

Una expresión del tipo:

$$\begin{cases} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots & \vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases}$$

además de como un SEL¹, puede ser vista como una **función vectorial**² que *transforma* un vector de entrada:

$$\vec{x} = [x_1, x_2, \dots, x_n] \in \mathbb{K}^n$$

en otro vector de salida:

$$\vec{y} = [y_1, y_2, \dots, y_m] \in \mathbb{K}^m$$

La notación utilizada para identificar las funciones vectoriales es la siguiente:

$$f : \mathbb{K}^n \mapsto \mathbb{K}^m \quad ; \quad \vec{x} \xrightarrow{f} \vec{y}$$

Definición 4.1. Transformación Lineal

Se denomina **transformación³ lineal** a toda *función vectorial de varias variables*, **lineal** en todas sus variables, (polinómica de primer grado), entre un *espacio dominio* $\mathbb{K}^n$ a otro *espacio imagen* $\mathbb{K}^m$, de la forma:

$$T : \mathbb{K}^n \mapsto \mathbb{K}^m \quad ; \quad [x_1, x_2, \dots, x_n] \xrightarrow{T} [y_1, y_2, \dots, y_m]$$

$$\begin{cases} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots & \vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases} \quad ; \quad \boxed{\vec{y} = T(\vec{x})}$$

Nota 4.1. Por interés práctico, se denotará mediante las siglas **TL** a las **Transformaciones Lineales**.

Ejemplo 4.1. Transformación lineal.

La transformación lineal:

$$\begin{cases} y_1 &= x_1 + x_2 \\ y_2 &= x_1 - x_2 \end{cases}$$

Convierte un vector de entrada $\vec{x} = [x_1, x_2] \in \mathbb{K}^2$ en otro vector $\vec{y} = [y_1, y_2] \in \mathbb{K}^2$.

Entrada	Salida
$\vec{x} = [1, 0]$	$\vec{y} = [1, 1]$
$\vec{x} = [1, 1]$	$\vec{y} = [2, 0]$
$\vec{x} = [1 + i, 1]$	$\vec{y} = [2 + i, i]$
$\vec{x} = [-2 - 3i, 1 + 2i]$	$\vec{y} = [-1 - i, -3 - 5i]$
$\vdots$	$\vdots$

Las TL pueden utilizarse para *codificar* información como muestra el siguiente ejemplo.

¹Ver la **Definición 1.10**, página 34.

²Ver una introducción a las funciones en **Apéndice A.12**, página 277.

³También se llama **función**, **aplicación**, **operador** o **mapeo**.

4.1.2. Operaciones con Transformaciones Lineales

Definición 4.3. Operaciones con TL

Sean $T, S : \mathbb{K}^n \mapsto \mathbb{K}^m$ dos transformaciones lineales:

$$T(\vec{x}) = A\vec{x} \quad , \quad S(\vec{x}) = B\vec{x} \quad ; \quad A, B \in \mathbb{K}^{m \times n} \quad , \quad \vec{x} \in \mathbb{K}^n$$

Se definen las operaciones **suma** y **multiplicación por una escalar** de TL de la siguiente forma:

- **Suma:** $F(\vec{x}) = T(\vec{x}) + S(\vec{x}) = A\vec{x} + B\vec{x} = (A + B)\vec{x} = C\vec{x}$, $C \in \mathbb{K}^{m \times n}$
- **Producto por un Escalar:** $G(\vec{x}) = kT(\vec{x}) = (kA)\vec{x} = D\vec{x}$; $D \in \mathbb{K}^{m \times n}$, $k \in \mathbb{K}$

Consecuencia de las propiedades de las operaciones con matrices⁷, el conjunto de las transformaciones lineales $T : \mathbb{K}^n \mapsto \mathbb{K}^m$ junto con sus operaciones *suma* y *producto por un escalar*, forman un **espacio vectorial**⁸.

Ejercicio 4.1. Dadas las TL $T, S : \mathbb{R}^2 \mapsto \mathbb{R}^4$ definidas como:

$$T : \begin{cases} y_1 &= x_1 - 2x_2 + x_4 \\ y_2 &= -x_1 - x_3 \end{cases} \quad ; \quad S : \begin{cases} y_1 &= 3x_1 + x_2 + x_3 - x_4 \\ y_2 &= x_2 - x_4 \end{cases}$$

calcular la TL $F(\vec{x}) = 3T(\vec{x}) - 2S(\vec{x})$.

Encontramos primero las matrices asociadas a las TL iniciales $T(\vec{x}) = A\vec{x}$ y $S(\vec{x}) = B\vec{x}$.

$$A = \begin{bmatrix} 1 & -2 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{bmatrix} \quad ; \quad C = \begin{bmatrix} 3 & 1 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Entonces la matriz C asociada a la TL pedida $F(\vec{x}) = C\vec{x}$ es:

$$C = 3A - 2B = 3 \begin{bmatrix} 1 & -2 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{bmatrix} - 2 \begin{bmatrix} 3 & 1 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -4 & -8 & -2 & 5 \\ -3 & -2 & -3 & 2 \end{bmatrix}$$

Por tanto:

$$F : \begin{cases} y_1 &= -4x_1 - 8x_2 - 2x_3 + 5x_4 \\ y_2 &= -3x_1 - 2x_2 - 3x_3 + 2x_4 \end{cases}$$

4.1.3. Propiedades de las Transformaciones Lineales

Teorema 4.1. Propiedades de las TL

Una *transformación* (función) $T : \mathbb{K}^n \mapsto \mathbb{K}^m$ es *lineal* si y solo si, verifica las siguientes dos propiedades:

- $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$; $\forall \vec{u}, \vec{v} \in \mathbb{K}^n$
- $T(k\vec{u}) = kT(\vec{u})$; $\forall \vec{u} \in \mathbb{K}^n, \forall k \in \mathbb{K}$ Nota: Para $k = 0 \implies T(\vec{0}) = \vec{0}$

Las propiedades del **Teorema 4.1** también se utilizan como definición alternativa de *transformación lineal*⁹.

⁷Ver la Sección 3.2.4, página 92.

⁸Ver la Definición 2.12, página 71.

⁹Ver la definición inicial dada en la Definición 4.1, página 121.

Demostración 4.1.**Demostración $\Rightarrow$ (Una TL verifica estas propiedades)**

Vamos a demostrar como una transformación $T: \mathbb{K}^n \mapsto \mathbb{K}^m$ del tipo:

$$\begin{cases} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots & \vdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n \end{cases} \Rightarrow \vec{y} = T(\vec{x}) = A\vec{x} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

cumple las dos propiedades antes citadas.

$$1. T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) \quad ; \quad \forall \vec{u}, \vec{v} \in \mathbb{K}^n$$

La parte izquierda de la igualdad será igual a:

$$T(\vec{u} + \vec{v}) = A(\vec{u} + \vec{v}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} = \begin{bmatrix} a_{11}(u_1 + v_1) + \cdots + a_{1n}(u_n + v_n) \\ \vdots \\ a_{m1}(u_1 + v_1) + \cdots + a_{mn}(u_n + v_n) \end{bmatrix}$$

Por otra parte, el lado derecho de la igualdad será:

$$\begin{aligned} T(\vec{u}) + T(\vec{v}) &= A\vec{u} + A\vec{v} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \\ &= \begin{bmatrix} a_{11}u_1 + \cdots + a_{1n}u_n \\ \vdots \\ a_{m1}u_1 + \cdots + a_{mn}u_n \end{bmatrix} + \begin{bmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{bmatrix} \stackrel{\text{Suma de Vectores}}{=} \\ &\quad \begin{bmatrix} a_{11}(u_1 + v_1) + \cdots + a_{1n}(u_n + v_n) \\ \vdots \\ a_{m1}(u_1 + v_1) + \cdots + a_{mn}(u_n + v_n) \end{bmatrix} \end{aligned}$$

siendo ambos resultados iguales.

$$2. T(k\vec{u}) = kT(\vec{u}) \quad ; \quad \forall \vec{u} \in \mathbb{K}^n, \quad \forall k \in \mathbb{K}$$

La parte izquierda de la igualdad será igual a:

$$T(k\vec{u}) = A(k\vec{u}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} ku_1 \\ \vdots \\ ku_n \end{bmatrix} = \begin{bmatrix} k(a_{11}u_1 + \cdots + a_{1n}u_n) \\ \vdots \\ k(a_{m1}u_1 + \cdots + a_{mn}u_n) \end{bmatrix} = kA\vec{u}$$

Y la parte derecha de la igualdad:

$$kT(\vec{u}) = kA\vec{u} = k \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} ky_1 \\ \vdots \\ ky_n \end{bmatrix} = \begin{bmatrix} k(a_{11}u_1 + \cdots + a_{1n}u_n) \\ \vdots \\ k(a_{m1}u_1 + \cdots + a_{mn}u_n) \end{bmatrix} = kA\vec{u}$$

siendo ambos resultados iguales.

Nota 4.2. La demostración $\Leftarrow$, correspondiente a que una función vectorial (transformación) cumpliendo las dos propiedades anteriores obligatoriamente debe ser *lineal* en todas sus componentes, se verá en la [Sección 4.2.2, página 126](#).

Las dos propiedades del **Teorema 4.1, página 123**, se pueden resumir en una sola:

$$T(\alpha\vec{u} + \beta\vec{v}) = \alpha T(\vec{u}) + \beta T(\vec{v}) \quad ; \quad \forall \vec{u}, \vec{v} \in \mathbb{K}^n, \quad \forall \alpha, \beta \in \mathbb{K}$$

Es decir, la TL de una *combinación lineal* de vectores es igual a la *combinación lineal* de las TL de los vectores.

Ejercicio 4.2. Dada la TL $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ definida por:

$$\begin{cases} y_1 &= 3x_1 + x_2 \\ y_2 &= -x_1 - 2x_2 \end{cases}$$

calcular $T(2\vec{u} - 5\vec{v} + 7\vec{w}) \quad \forall \vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^2$.

Expresándolo en forma matricial se tiene:

$$T(\vec{x}) = \begin{bmatrix} 3 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad ; \quad \vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

Como es una TL, podemos calcularlo de dos formas distintas.

Forma 1:

$$\begin{aligned} T(2\vec{u} - 5\vec{v} + 7\vec{w}) &= \begin{bmatrix} 3 & 1 \\ -1 & -2 \end{bmatrix} \left(2 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - 5 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + 7 \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right) = \begin{bmatrix} 3 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 2u_1 - 5v_1 + 7w_1 \\ 2u_2 - 5v_2 + 7w_2 \end{bmatrix} \\ &= \begin{bmatrix} 3(2u_1 - 5v_1 + 7w_1) + (2u_2 - 5v_2 + 7w_2) \\ -(2u_1 - 5v_1 + 7w_1) - 2(2u_2 - 5v_2 + 7w_2) \end{bmatrix} = \begin{bmatrix} 6u_1 - 15v_1 + 21w_1 + 2u_2 - 5v_2 + 7w_2 \\ -2u_1 + 5v_1 - 7w_1 - 4u_2 + 10v_2 - 14w_2 \end{bmatrix} \end{aligned}$$

Forma 2:

$$\begin{aligned} T(2\vec{u} - 5\vec{v} + 7\vec{w}) &= 2T(\vec{u}) - 5T(\vec{v}) + 7T(\vec{w}) = \\ &= 2 \begin{bmatrix} 3 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - 5 \begin{bmatrix} 3 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + 7 \begin{bmatrix} 3 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 2 \\ -2 & -4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -15 & -5 \\ 5 & 10 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} 21 & 7 \\ -7 & -14 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \\ &= \begin{bmatrix} 6u_1 + 2u_2 \\ -2u_1 - 4u_2 \end{bmatrix} + \begin{bmatrix} -15v_1 - 5v_2 \\ 5v_1 + 10v_2 \end{bmatrix} + \begin{bmatrix} 21w_1 + 7w_2 \\ -7w_1 - 14w_2 \end{bmatrix} = \begin{bmatrix} 6u_1 + 2u_2 - 15v_1 - 5v_2 + 21w_1 + 7w_2 \\ -2u_1 - 4u_2 + 5v_1 + 10v_2 - 7w_1 - 14w_2 \end{bmatrix} \end{aligned}$$

Siendo ambas expresiones iguales.

Estas dos propiedades simplifican en gran medida la utilidad y operaciones de las TL, haciendo de ellas una herramienta matemática de gran interés práctico.

4.2. Transformaciones Lineales Elementales

- **TL Nula** : $T : \mathbb{K}^n \mapsto \mathbb{K}^m$: $T(\vec{x}) = \vec{0} \in \mathbb{K}^m$; $\forall x \in \mathbb{K}^n$

La matriz asociada a la *transformación nula* $T(\vec{x}) = A\vec{x} = \vec{0}$ es la *matriz nula* de orden $m \times n$:

$$0_{m \times n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{K}^{m \times n}$$

- **TL Identidad** : $T : \mathbb{K}^n \mapsto \mathbb{K}^m$: $T(\vec{x}) = \vec{x} \in \mathbb{K}^m$; $\forall x \in \mathbb{K}^n$

En este caso obligatoriamente $n = m$ y la matriz asociada a la *transformación identidad* $T(\vec{x}) = A\vec{x} = \vec{x}$ corresponde con la *matriz identidad* I_n :

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{K}^{n \times n}$$

4.2.1. Transformación Lineal de Vectores Estándar

Dada una transformación lineal cualquiera $T : \mathbb{K}^n \mapsto \mathbb{K}^m$, se tiene¹⁰:

$$T(\vec{e}_i) = A\vec{e}_i = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} ; \quad \vec{e}_i \in \mathbb{K}^n$$

siendo $\vec{e}_i$ el *vector columna estándar*¹¹ i -ésimo. Por tanto, el vector $T(\vec{e}_i)$ corresponde con la columna i -ésima, $\vec{c}_i$, de la matriz asociada a la transformación lineal $T : \mathbb{K}^n \mapsto \mathbb{K}^m$.

$$T(\vec{e}_i) = A\vec{e}_i = \begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_n \end{bmatrix} \vec{e}_i = \vec{c}_i$$

Definición 4.4. Matriz Asociada a una TL a través de vectores estándar

Como $T(\vec{e}_i) = \vec{c}_i$, la *matriz asociada* a cualquier *transformación lineal* $T : \mathbb{K}^n \mapsto \mathbb{K}^m$, se puede definir como:

$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_n) \end{bmatrix}$$

Dado un vector $\vec{x} \in \mathbb{K}^n$ se tiene:

$$T(\vec{x}) = T(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n) = x_1T(\vec{e}_1) + x_2T(\vec{e}_2) + \cdots + x_nT(\vec{e}_n) = x_1\vec{c}_1 + x_2\vec{c}_2 + \cdots + x_n\vec{c}_n$$

siendo $\vec{c}_i$ la columna i -ésima de la matriz asociada a la transformación T .

$$\vec{y} = A\vec{x} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \sum_{i=1}^n x_i \vec{c}_i$$

4.2.2. Demostración $\Leftarrow$ de la Sección 4.1.3

Vamos a utilizar la **Definición 4.4** de una matriz asociada a una TL mediante vectores estándar, para terminar la demostración del **Teorema 4.1**, **página 123**, cuyo resultado era:

Una transformación $T : \mathbb{K}^n \mapsto \mathbb{K}^m$ que verifica las siguientes dos propiedades es una TL.

- (1) $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$, $\forall \vec{u}, \vec{v} \in \mathbb{K}^n$
- (2) $T(k\vec{u}) = kT(\vec{u})$, $\forall \vec{u} \in \mathbb{K}^n$, $\forall K \in \mathbb{K}$

¹⁰Ver la **Sección 3.3.5**, **página 96**.

¹¹Ver la **Sección 2.2.3**, **página 69**.

Demostración

Una transformación es *lineal* si tiene la forma:

$$T(\vec{x}) = A\vec{x} \quad \begin{cases} \vec{x} \in \mathbb{K}^n \\ A \in \mathbb{K}^{m \times n} \end{cases}$$

Ya que el producto matricial cumple todas propiedades, (ver la [Sección 3.3.1, página 94](#)), de la definición de una TL.

Sean $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ los *vectores estándar* de $\mathbb{K}^n$, entonces:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n]$$

Por tanto, aplicando las dos propiedades anteriores (1) y (2) tenemos:

$$\begin{aligned} T(\vec{x}) &= T(x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n) \stackrel{(1)}{=} T(x_1\vec{e}_1) + T(x_2\vec{e}_2) + \dots + T(x_n\vec{e}_n) \stackrel{(2)}{=} \\ &\stackrel{(2)}{=} x_1T(\vec{e}_1) + x_2T(\vec{e}_2) + \dots + x_nT(\vec{e}_n) = [T(\vec{e}_1), T(\vec{e}_2), \dots, T(\vec{e}_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\vec{x} \end{aligned}$$

4.3. Inversa de una TL

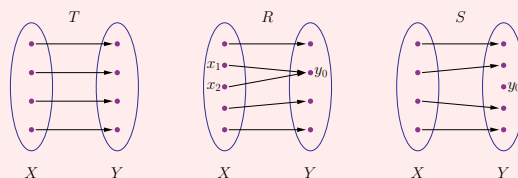
4.3.1. Función Invertible

Definición 4.5. Función Invertible

Se dice que una función $T : X \mapsto Y$ es *invertible*¹² si la ecuación $T(x) = y$ tiene una solución *única* $x \in X$ para cada $y \in Y$. Si una función $T : X \mapsto Y$ es invertible, su inversa $T^{-1} : Y \mapsto X$ se define como:

$$T^{-1}(y) = \{ \text{solución única } x \in X \text{ de } T(x) = y \}$$

Ejercicio 4.3. Determinar si las siguientes TL T , R y S son invertibles, a través de sus correspondientes definiciones dadas mediante conjuntos de Venn.



- T es invertible al ser una función *uno a uno* o biyectiva.
- R no es invertible, ya que $R(x) = y_0$ tiene dos soluciones x_1 y x_2 .
- S no es invertible, ya que $S(x) = y_0$ no tiene solución.

Ejemplo 4.4. Codificación de mensajes.

¹²Ver una definición de *función inversa* en el [Apéndice A.12.4, página 280](#).

Sea el problema de la codificación de mensajes, **Ejemplo 4.2**, dado en este caso por la siguiente TL:

$$\begin{cases} y_1 = x_1 + 3x_2 \\ y_2 = 2x_1 + 5x_2 \end{cases} \implies \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies \vec{y} = A\vec{x}$$

Esta TL codifica una posición de entrada $\vec{x} = [x_1, x_2] \in \mathbb{R}^2$ en otro vector $\vec{y} = [y_1, y_2] \in \mathbb{R}^2$.

Cuando se recibe el mensaje codificado hay que decodificarlo. Por ejemplo si el vector codificado es $\vec{y} = [131, 220]$, conociendo la matriz de codificación A habrá que resolver el SEL:

$$\begin{cases} x_1 + 3x_2 = 131 \\ 2x_1 + 5x_2 = 220 \end{cases}$$

Utilizando Gauss-Jordan:

$$\left[\begin{array}{cc|c} 1 & 3 & 131 \\ 2 & 5 & 220 \end{array} \right] \xrightarrow{F_2 - 2F_1} \left[\begin{array}{cc|c} 1 & 3 & 131 \\ 0 & -1 & -42 \end{array} \right] \xrightarrow[\begin{smallmatrix} F_1 + 3F_2 \\ -F_3 \end{smallmatrix}]{\begin{smallmatrix} F_1 + 3F_2 \\ -F_3 \end{smallmatrix}} \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 42 \end{array} \right]$$

por tanto, la información decodificada es $\vec{x} = [5, 42]$. Lo interesante es encontrar la forma de decodificar cualquier mensaje $\vec{y} = [y_1, y_2]$.

$$\left[\begin{array}{cc|c} 1 & 3 & y_1 \\ 2 & 5 & y_2 \end{array} \right] \xrightarrow{F_2 - 2F_1} \left[\begin{array}{cc|c} 1 & 3 & y_1 \\ 0 & -1 & -2y_1 + y_2 \end{array} \right] \xrightarrow[\begin{smallmatrix} F_1 + 3F_2 \\ -F_3 \end{smallmatrix}]{\begin{smallmatrix} F_1 + 3F_2 \\ -F_3 \end{smallmatrix}} \left[\begin{array}{cc|c} 1 & 0 & -5y_1 + 3y_2 \\ 0 & 1 & 2y_1 - y_2 \end{array} \right]$$

Luego, la función de decodificación (función inversa) es:

$$\begin{cases} x_1 = -5y_1 + 3y_2 \\ x_2 = 2y_1 - y_2 \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -5 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \implies \vec{x} = A^{-1}\vec{y}$$

siendo también una TL. La relación entre las dos matrices A y A^{-1} es la siguiente:

$$\vec{x} \xrightleftharpoons[\text{decodificar con matriz } A^{-1} = \begin{bmatrix} -5 & 3 \\ 2 & -1 \end{bmatrix}]{\text{codificar con la matriz } A = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}} \vec{y}$$

Ejemplo 4.5. TL no invertibles.

No todas las transformaciones lineales son invertibles. Por ejemplo, supongamos que se elige como codificación:

$$\begin{cases} y_1 = x_1 + 2x_2 \\ y_2 = 2x_1 + 4x_2 \end{cases} \implies \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies \vec{y} = A\vec{x}$$

Cuando se recibe la información codificada $\vec{y} = [89, 178]$ y queremos decodificarla, tendremos que resolver el sistema:

$$\begin{cases} x_1 + 2x_2 = 89 \\ 2x_1 + 4x_2 = 178 \end{cases}$$

que tiene infinitas soluciones de la forma:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 89 - t \\ t \end{bmatrix} \quad ; \quad \forall t \in \mathbb{R}$$

Ya que este SEL no tiene solución única, es imposible recuperar la actual posición codificada, es decir, **la transformación de codificación y la matriz de codificación A son no invertibles**. Por tanto este código (TL) no es útil para este fin.

Definición 4.6. TL Invertible

La TL $\vec{y} = A\vec{x}$ es *invertible* si el SEL:

$$A\vec{x} = \vec{y}$$

tiene una única solución $\vec{x} \in \mathbb{K}^n$ para todo $\vec{y} \in \mathbb{K}^m \implies A$ es invertible

Analicemos todos los posibles casos de $A \in \mathbb{K}^{m \times n}$:

- $m < n$

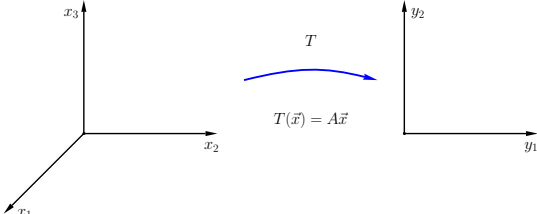
En este caso el sistema tiene menos ecuaciones que incógnitas y por tanto puede ser:

- *incompatible* (sin solución).
- *compatible indeterminado* (infinitas soluciones), para cualquier $\vec{y} \in \mathbb{K}^m$.

Por tanto la TL $\vec{y} = A\vec{x}$ es **no invertible**.

Ejemplo 4.6. TL no invertible.

Una TL real con $n = 3$ y $m = 2$ de la forma $T : \mathbb{R}^3 \mapsto \mathbb{R}^2$, podemos interpretarla gráficamente como:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$


Se puede considerar como una transformación de todos los puntos del espacio $\mathbb{R}^3$ en los puntos del plano $\mathbb{R}^2$. Intuitivamente se aprecia que existen más puntos en el conjunto dominio $\mathbb{R}^3$ que en el conjunto imagen $\mathbb{R}^2$ y por tanto:

- La TL o bien no es *inyectiva* o no es *sobreyectiva* o ambas cosas. $\implies$ no es *biyectiva*¹³.
- El SEL $A\vec{x} = \vec{y}$ puede tener infinitas soluciones o bien no tener solución, para todos o algunos valores $\vec{y} \in \mathbb{R}^2$.

En cualquiera de los dos casos, la TL **no es invertible**.

¹³Ver el Apéndice A.12.3, página 279.

- $m = n$

En este caso el sistema tiene el mismo número de ecuaciones que incógnitas y por tanto puede ser:

- *incompatible* (sin solución).
- *compatible indeterminado* (infinitas soluciones).
- *compatible determinado* (solución única) para cualquier $\vec{y} \in \mathbb{K}^m$.

La TL $\vec{y} = A\vec{x}$ será **invertible** si y solo si¹⁴:

$$\text{rang}(A) = n$$

es decir:

- Cualquier *matriz escalonada equivalente*¹⁵ de A no posee ninguna fila de ceros.
- La *matriz escalonada reducida equivalente*¹⁶ de A es la *matriz identidad* de orden n , I_n .

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{K}^{n \times n}$$

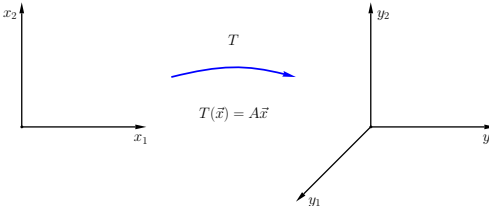
- $m > n$

En este caso el sistema tiene más ecuaciones que incógnitas y si T es inyectiva siempre podemos encontrar algún determinado $\vec{y} \in \mathbb{R}^m$ tal que el sistema $A\vec{x} = \vec{y}$ sea *incompatible* (sin solución).

Por tanto la TL $\vec{y} = A\vec{x}$ es **no invertible**.

Ejemplo 4.7. TL no invertible.

Una TL real con $n = 2$ y $m = 3$ de la forma $T : \mathbb{R}^2 \mapsto \mathbb{R}^3$, se puede interpretar como una aplicación de de todos los puntos del plano $\mathbb{R}^2$ en los puntos del espacio $\mathbb{R}^3$.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$


Claramente existen menos puntos en $\mathbb{R}^2$ que en $\mathbb{R}^3$, de forma que no puede existir una correspondencia *uno a uno* (biyectiva) y por tanto la TL **no** es invertible.

¹⁴Ver la Sección 3.7.1, página 100.

¹⁵Ver la Definición 1.32, página 50.

¹⁶Ver la Definición 1.36, página 57.

4.3.2. Inversa de una TL

Una transformación lineal $T : \mathbb{K}^n \mapsto \mathbb{K}^m$ se puede expresar de las siguientes formas:

$$\left. \begin{array}{lcl} y_1 & = & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 & = & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots & & \vdots \\ y_m & = & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{array} \right\} \Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\boxed{T(\vec{x}) = A\vec{x} = \vec{y}} \quad \left\{ \begin{array}{l} \vec{x} \in \mathbb{K}^n \\ A \in \mathbb{K}^{m \times n} \\ \vec{y} \in \mathbb{K}^m \end{array} \right. ; \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{K}^{m \times n}$$

Definición 4.7. TL Inversa

Si una TL¹⁷ :

$$T : \mathbb{K}^n \rightarrow \mathbb{K}^n \quad ; \quad \vec{y} = T(\vec{x}) = A\vec{x}$$

es invertible¹⁸, se define la TL inversa, $T^{-1}(\vec{y})$, como:

$$T^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n \quad ; \quad \vec{x} = T^{-1}(\vec{y}) = A^{-1}\vec{y}$$

Ejercicio 4.4. Encontrar, en caso de que exista, la TL inversa de la TL, $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$:

$$T : \left\{ \begin{array}{lcl} y_1 & = & x_1 + x_2 + x_3 \\ y_2 & = & 2x_1 + 3x_2 + 2x_3 \\ y_3 & = & 3x_1 + 8x_2 + 2x_3 \end{array} \right\} \Rightarrow \vec{y} = T(\vec{x}) = A\vec{x} \quad ; \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}$$

Aplicando el método de Gauss-Jordan¹⁹:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 8 & 2 & 0 & 0 & 1 \end{array} \right] \xrightarrow[F_3-3F_1]{F_2-2F_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 5 & -1 & -3 & 0 & 1 \end{array} \right] \xrightarrow[F_1-F_2]{F_3-5F_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -1 & 7 & -5 & 1 \end{array} \right]$$

$$\xrightarrow[-F_3]{F_1+F_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 10 & -6 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & -1 \end{array} \right] \Rightarrow A^{-1} = \begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{bmatrix}$$

La TL inversa queda entonces:

$$\left[\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right] = \left[\begin{array}{ccc} 10y_1 & - & 6y_2 + y_3 \\ -2y_1 & + & y_2 \\ -7y_1 & + & 5y_2 - y_3 \end{array} \right] = \left[\begin{array}{ccc} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{array} \right] \left[\begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \right] \Rightarrow \vec{x} = T^{-1}(\vec{y}) = A^{-1}\vec{y}$$

¹⁷Notar que necesariamente la matriz asociada a la TL tiene que ser cuadrada, (con igual número de filas que de columnas).

¹⁸Ver la Definición 4.6, página 129.

¹⁹Ver la Sección 1.7, página 57.

Ejercicio 4.5. Encontrar, en caso de que exista, la TL inversa de la TL, $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies \vec{y} = A\vec{x}$$

La matriz A asociada a la TL es cuadrada pero claramente no tiene vectores fila (ni columna) linealmente independientes, y por tanto la matriz A no es invertible y tampoco lo es la TL definida por ella.

Propiedades de la TL Inversa

Dada una función *biyectiva*²⁰ $T : X \mapsto Y$ se tiene:

- $T^{-1}(T(x)) = T^{-1}(y) = x$, $\forall x \in X$
- $T(T^{-1}(y)) = T(x) = y$, $\forall y \in Y$
- $(T^{-1})^{-1} = T$

Linealidad de la TL Inversa

En caso de que exista, la transformación inversa T^{-1} de una TL:

$$T : \mathbb{K}^n \mapsto \mathbb{K}^m \quad ; \quad T(\vec{x}) = \vec{y} = A\vec{x}$$

también es una TL, ya que²¹:

$$T^{-1}(\alpha\vec{u} + \beta\vec{v}) = A^{-1}(\alpha\vec{u} + \beta\vec{v}) = \alpha A^{-1}\vec{u} + \beta A^{-1}\vec{v} = \alpha T^{-1}(\vec{u}) + \beta T^{-1}(\vec{v})$$

4.4. Operaciones con TL

4.4.1. Suma de Transformaciones Lineales

Definición 4.8. Suma de Transformaciones Lineales

Dadas las transformaciones lineales $F, G : \mathbb{K}^n \mapsto \mathbb{K}^m$, se define la transformación lineal suma $(F + G) : \mathbb{K}^n \mapsto \mathbb{K}^m$ como:

$$(F + G)(\vec{x}) = F(\vec{x}) + G(\vec{x}) = A\vec{x} + B\vec{x} = (A + B)\vec{x} = C\vec{x} \quad ; \quad \vec{x} \in \mathbb{K}^n$$

siendo $C = A + B$ la matriz suma²² de $A \in \mathbb{K}^{m \times n}$ y $B \in \mathbb{K}^{m \times n}$:

$$C = A + B = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

Propiedades de la Suma de TL

- La suma de TL es una TL.
- La suma de TL es *conmutativa*, ya que: $A + B = B + A$.
- La suma de TL es *asociativa*, ya que: $A + (B + C) = (A + B) + C$.
- La operación **diferencia** de TL se define como la suma del opuesto: $A - B = A + (-B)$

²⁰Ver el Apéndice A.12.3, página 279.

²¹Ver el Teorema 4.1, página 123 y la propiedad resumen en la página 125.

²²Ver la Sección 3.2.1, página 91.

Ejercicio 4.6. Dadas las TL:

$$F : \mathbb{R}^2 \mapsto \mathbb{R}^3 \quad \begin{cases} y_1 = x_1 + x_2 \\ y_2 = 2x_1 - x_2 \\ y_3 = x_1 + 2x_2 \end{cases} \quad G : \mathbb{R}^2 \mapsto \mathbb{R}^3 \quad \begin{cases} y_1 = 2x_1 \\ y_2 = -x_1 - x_2 \\ y_3 = 2x_2 \end{cases}$$

calcular las TL: $(F + G)(\vec{x})$ y $(F - G)(\vec{x})$.

$$\begin{aligned} (F + G)(\vec{x}) &= A(\vec{x}) + B(\vec{x}) = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ -1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (A + B)\vec{x} = \\ &= \left(\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ -1 & -1 \\ 0 & 2 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & -2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = C\vec{x} \\ (F - G)(\vec{x}) &= A(\vec{x}) - B(\vec{x}) = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ -1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (A - B)\vec{x} = \\ &= \left(\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ -1 & -1 \\ 0 & 2 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 3 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = D\vec{x} \end{aligned}$$

4.4.2. Composición de Transformaciones

Definición 4.9. Composición de Transformaciones

Dadas las transformaciones:

$$\begin{cases} G : \mathbb{K}^p \mapsto \mathbb{K}^n \\ F : \mathbb{K}^n \mapsto \mathbb{K}^m \end{cases}$$

se define, la *composición de la transformación F y G*, $(F \circ G) : \mathbb{K}^p \mapsto \mathbb{K}^m$, de la siguiente forma:

$$(F \circ G)(\vec{x}) = F[G(\vec{x})] \quad \vec{x} \xrightarrow{G} G(\vec{x}) \xrightarrow{F} F[G(\vec{x})]$$

Dominio e Imagen de la Composición de TL

Dada la composición de TL: $(F \circ G)(\vec{x}) = F[G(\vec{x})]$.

- El *dominio* de F debe de coincidir con la *imagen* de G .
- El *dominio* de la nueva transformación $(F \circ G)(x)$ coincide con el *dominio* de $G(\vec{x})$ y su *imagen* es la *imagen* de $F(\vec{y})$.

Ejercicio 4.7. Dadas las TL:

$$F : \mathbb{R}^2 \mapsto \mathbb{R}^3 \quad \begin{cases} y_1 = x_1 + x_2 \\ y_2 = 2x_1 - x_2 \\ y_3 = x_1 + 2x_2 \end{cases} \quad G : \mathbb{R}^4 \mapsto \mathbb{R}^2 \quad \begin{cases} y_1 = x_1 + x_2 + x_3 + x_4 \\ y_2 = 2x_1 - x_2 - x_4 \end{cases}$$

calcular la TL composición F con G , $(F \circ G)(\vec{x})$.

$$F(\vec{x}) = A\vec{x} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad ; \quad G(\vec{x}) = B\vec{x} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$\begin{aligned}
(F \circ G)(\vec{x}) &= F[G(\vec{x})] = A(B\vec{x}) = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 2 \end{bmatrix} \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \right) = \\
&= \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 + x_2 + x_3 + x_4 \\ 2x_1 - x_2 - x_4 \end{bmatrix} = \begin{bmatrix} (x_1 + x_2 + x_3 + x_4) + (2x_1 - x_2 - x_4) \\ 2(x_1 + x_2 + x_3 + x_4) - (2x_1 - x_2 - x_4) \\ (x_1 + x_2 + x_3 + x_4) + 2(2x_1 - x_2 - x_4) \end{bmatrix} = \\
&= \begin{bmatrix} 3x_1 + x_3 \\ 3x_2 + 2x_3 + 3x_4 \\ 5x_1 - x_2 + x_3 - x_4 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 5 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = C\vec{x}
\end{aligned}$$

La composición de estas dos TL es una nueva TL: $(F \circ G) : \mathbb{R}^4 \mapsto \mathbb{R}^3$.

Propiedades de la Composición de TL

- La composición de TL es una TL: $(F \circ G)(\vec{x}) = A(B\vec{x}) = C\vec{x}$.
- La composición de TL no es conmutativa: $(F \circ G) \neq (G \circ F)$.
- La composición de TL sí es asociativa: $F \circ (G \circ H) = (F \circ G) \circ H$.

Demostración

$$\begin{cases} F : \mathbb{K}^n \mapsto \mathbb{K}^m & ; & F(\vec{x}) = A\vec{x} \\ G : \mathbb{K}^p \mapsto \mathbb{K}^n & ; & G(\vec{x}) = B\vec{x} \end{cases}$$

- $(F \circ G)(\alpha\vec{u} + \beta\vec{v}) = F[G(\alpha\vec{u} + \beta\vec{v})] = A(B(\alpha\vec{u} + \beta\vec{v})) = A(\alpha B\vec{u} + \beta B\vec{v}) = A(\alpha B\vec{u}) + A(\beta B\vec{v}) = \alpha A(B\vec{u}) + \beta A(B\vec{v}) = \alpha(F \circ G)(\vec{u}) + \beta(F \circ G)(\vec{v})$
- Por una parte $(F \circ G) : \mathbb{K}^p \mapsto \mathbb{K}^m$ y por otra $(G \circ F)$ estará definida si y solo si $m = p$. Por tanto de forma general $(F \circ G) \neq (G \circ F)$.
- Ver [Apéndice A.12.5, página 281](#).

Matriz asociada a la Composición de TL

La matriz asociada a la composición de TL corresponde con el *producto matricial*²³ de las matrices asociadas a cada una de las TL.

$$(F \circ G)(\vec{x}) = A(B\vec{x}) = (AB)\vec{x}$$

Demostración

$$\begin{cases} F : \mathbb{K}^n \mapsto \mathbb{K}^m & ; & F(\vec{x}) = A\vec{x} \\ G : \mathbb{K}^p \mapsto \mathbb{K}^n & ; & G(\vec{x}) = B\vec{x} \end{cases}$$

$$(F \circ G)(\vec{x}) = F[G(\vec{x})] = A(B\vec{x}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \left(\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \right) =$$

²³Ver la [Sección 3.3, página 93](#).

$$\begin{aligned}
&= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1p}x_p \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2p}x_p \\ \vdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{np}x_p \end{bmatrix} = \\
&= \begin{bmatrix} a_{11}(b_{11}x_1 + \cdots + b_{1p}x_p) + a_{12}(b_{21}x_1 + \cdots + b_{2p}x_p) + \cdots + a_{1n}(b_{n1}x_1 + \cdots + b_{np}x_p) \\ a_{21}(b_{11}x_1 + \cdots + b_{1p}x_p) + a_{22}(b_{21}x_1 + \cdots + b_{2p}x_p) + \cdots + a_{2n}(b_{n1}x_1 + \cdots + b_{np}x_p) \\ \vdots \\ a_{m1}(b_{11}x_1 + \cdots + b_{1p}x_p) + a_{m2}(b_{21}x_1 + \cdots + b_{2p}x_p) + \cdots + a_{mn}(b_{n1}x_1 + \cdots + b_{np}x_p) \end{bmatrix} = \\
&= \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1})x_1 + \cdots + (a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np})x_p \\ (a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1})x_1 + \cdots + (a_{21}b_{1p} + a_{22}b_{2p} + \cdots + a_{2n}b_{np})x_p \\ \vdots \\ (a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1})x_1 + \cdots + (a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np})x_p \end{bmatrix} = \\
&= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & \cdots & a_{21}b_{1p} + a_{22}b_{2p} + \cdots + a_{2n}b_{np} \\ \vdots & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \\
&= \begin{bmatrix} \vec{f}_{1A} \\ \vec{f}_{2A} \\ \vdots \\ \vec{f}_{mA} \end{bmatrix} [\vec{c}_{1B} \ \vec{c}_{2B} \ \cdots \ \vec{c}_{pB}] = \begin{bmatrix} \vec{f}_{1A} \cdot \vec{c}_{1B} & \vec{f}_{1A} \cdot \vec{c}_{2B} & \cdots & \vec{f}_{1A} \cdot \vec{c}_{nB} \\ \vec{f}_{2A} \cdot \vec{c}_{1B} & \vec{f}_{2A} \cdot \vec{c}_{2B} & \cdots & \vec{f}_{2A} \cdot \vec{c}_{nB} \\ \vdots & \vdots & & \vdots \\ \vec{f}_{mA} \cdot \vec{c}_{1B} & \vec{f}_{mA} \cdot \vec{c}_{2B} & \cdots & \vec{f}_{mA} \cdot \vec{c}_{nB} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = C\vec{x} \\
&C = AB = \begin{bmatrix} \vec{f}_{1A} \cdot \vec{c}_{1B} & \vec{f}_{1A} \cdot \vec{c}_{2B} & \cdots & \vec{f}_{1A} \cdot \vec{c}_{nB} \\ \vec{f}_{2A} \cdot \vec{c}_{1B} & \vec{f}_{2A} \cdot \vec{c}_{2B} & \cdots & \vec{f}_{2A} \cdot \vec{c}_{nB} \\ \vdots & \vdots & & \vdots \\ \vec{f}_{mA} \cdot \vec{c}_{1B} & \vec{f}_{mA} \cdot \vec{c}_{2B} & \cdots & \vec{f}_{mA} \cdot \vec{c}_{nB} \end{bmatrix} \begin{cases} A & \in \mathbb{K}^{m \times n} \\ B & \in \mathbb{K}^{n \times p} \\ AB & \in \mathbb{K}^{m \times p} \end{cases}
\end{aligned}$$

Ejercicio 4.8. Siguiendo con en el problema de la codificación de mensajes. Un espía 1 envía una posición $\vec{x}$ codificada según una TL $\vec{y} = F(\vec{x})$ al espía 2 y éste vuelve a codificar el mensaje según otra TL distinta $\vec{z} = G(\vec{y})$ y lo envía al espía 3. Conociendo las TL de codificación y sabiendo que el mensaje final doblemente codificado es $\vec{z} = [6, 6]$, encontrar el mensaje original $\vec{x}$.

$$\vec{y} = F(\vec{x}) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad ; \quad \vec{z} = G(\vec{y}) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\vec{z} = (G \circ F)(\vec{x}) = G[F(\vec{x})] = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = C\vec{x}$$

Notar que la composición es $(G \circ F)(\vec{x})$ y **no** $(F \circ G)(\vec{x})$, ya que el orden de aplicación de las codificaciones es primero F y después G . Por tanto:

$$\vec{x} = C^{-1}\vec{z}$$

$$C^{-1} = \begin{bmatrix} 1/6 & 1/6 \\ -1/2 & 1/2 \end{bmatrix} \implies \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/6 & 1/6 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

correspondiente al mensaje original $\vec{x}$.

4.5. Imagen y Núcleo de una TL

4.5.1. Imagen de una TL

Definición 4.10. Imagen de una TL

Sea una TL entre un **espacio dominio** X a otro **espacio imagen** Y .

$$T : X \mapsto Y \quad ; \quad \vec{x} = [x_1, x_2, \dots, x_n] \xrightarrow{T} [y_1, y_2, \dots, y_m] = \vec{y}$$

Se denomina **imagen** de la TL al conjunto de vectores $\vec{y}$ del *espacio imagen* Y de T que son *imagen* de algún vector $\vec{x}$ del *espacio dominio* X .

$$\text{im}(T) = \{ \vec{y} \in Y : \vec{y} = T(\vec{x}) \text{ para algún } \vec{x} \in X \}$$

Nota 4.3. Dada una TL, $T : X \mapsto Y$, no confundir la *imagen* de la transformación, $\text{im}(T)$, con el *espacio imagen* de la misma, Y . La *imagen* es un determinado subconjunto del *espacio imagen*, $\text{im}(T) \subseteq Y$, ver la [Sección 4.5.1, página 137](#).

Definición 4.11. Imagen de una Matriz

Se define la *imagen de una matriz* $A \in \mathbb{K}^{m \times n}$ y se denota por $\text{im}(A)$, a la *imagen* de la TL que define dicha matriz.

$$\vec{y} = T(\vec{x}) = A\vec{x} \quad \implies \quad \text{im}(T) = \text{im}(A)$$

El *espacio columna*²⁴ de una matriz coincide con la *imagen* de la matriz²⁵:

$$\text{im}(A) = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{n1} \\ \vdots \\ a_{mn} \end{bmatrix} = \text{col}(A)$$

Ejercicio 4.9. Describir la imagen de la TL $T : \mathbb{R}^2 \mapsto \mathbb{R}^3$ definida por la matriz:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

La *imagen* de T es el conjunto de todos los posibles valores dados por la transformación. Aplicando la relación anterior se tiene:

$$T(\vec{x}) = T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

De forma que la $\text{im}(T)$ se puede obtener con todas las posibles combinaciones lineales de los vectores columna de la matriz A . En este caso en concreto, como los dos vectores son linealmente independientes, corresponde a un *plano* en $\mathbb{R}^3$ que pasa por el origen^a $(0, 0, 0)$.

$$\text{im}(T) = \text{im}(A) = \left\{ s \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} ; \forall s, t \in \mathbb{R} \right\}$$

²⁴Ver la [Definición 3.22, página 111](#).

²⁵Ver demostración en la [página 126](#).

^aSi no pasara por el origen no sería un *subespacio vectorial* (ver [Sección 4.5.1, página 137](#)).

Ejercicio 4.10. Calcular la imagen de la TL de $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ definida por la matriz:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

$$T(\vec{x}) = T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 6 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = (x_1 + 3x_2) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Como $x_1, x_2 \in \mathbb{R}$, la $\text{im}(T)$ está formada por todos los vectores $\vec{y} \in \mathbb{R}^2$ que tienen la dirección del vector $[1, 2]$. Esto corresponde a una *recta* en $\mathbb{R}^2$ que pasa por el origen $(0, 0)$.

$$\text{im}(T) = \text{im}(A) = \left\{ t \begin{bmatrix} 1 \\ 2 \end{bmatrix} ; \forall t \in \mathbb{R} \right\}$$

Propiedades de la Imagen de una TL

Dada una TL entre un *espacio dominio* X a otro *espacio imagen* Y , $T : X \mapsto Y$

- El vector nulo del *espacio imagen*, $\vec{0} \in Y$, siempre está contenido en la *imagen* de T , $\text{im}(T)$.
- La *imagen* de T es *cerrada* con respecto a la suma:

$$\forall \vec{v}_1, \vec{v}_2 \in \text{im}(T) \implies \vec{v}_1 + \vec{v}_2 \in \text{im}(T)$$

- La *imagen* de T es *cerrada* con respecto a la multiplicación por un escalar:

$$\forall \vec{v} \in \text{im}(T) , \forall k \in \mathbb{K} \implies k\vec{v} \in \text{im}(T)$$

Demostración

- $T(\vec{0}) = A\vec{0} = \vec{0}$
- Sean $\vec{v}_1, \vec{v}_2 \in \text{im}(T) \implies \vec{v}_1 + \vec{v}_2 = T(\vec{x}_1) + T(\vec{x}_2) = T(\vec{x}_1 + \vec{x}_2) \implies \vec{v}_1 + \vec{v}_2 \in \text{im}(T)$
- Sea $\vec{v} \in \text{im}(T)$ y $k \in \mathbb{K} \implies k\vec{v} = kT(\vec{x}_1) = T(k\vec{x}_1) = T(\vec{x}_2) \implies k\vec{v} \in \text{im}(T)$

La *imagen* de una TL es un *subespacio vectorial*²⁶ de su *espacio imagen*.

4.5.2. Núcleo o Espacio Nulo de una TL

Definición 4.12. Núcleo de una TL

Dada una TL entre un **espacio dominio** X a otro **espacio imagen** Y .

$$T : X \mapsto Y ; \quad \vec{x} = [x_1, x_2, \dots, x_n] \xrightarrow{T} [y_1, y_2, \dots, y_m] = \vec{y}$$

Se denomina **núcleo** o **espacio nulo** de la TL al conjunto de $\vec{x}$ perteneciente al *espacio dominio* de T , cuya *imagen* es el vector nulo $\vec{0} \in Y$ del *espacio imagen*. Por tanto, el *núcleo* de T serán las soluciones del SEL homogéneo $A\vec{x} = \vec{0}$.

$$\ker(T) = (\vec{x} \in X : \vec{y} = T(\vec{x}) = A\vec{x} = \vec{0})$$

²⁶Ver la **Definición 2.13**, página 71.

Nota 4.4. Dada una TL, $T : X \mapsto Y$, el *núcleo* es un determinado subconjunto del *espacio dominio*, $\ker(T) \subseteq X$, ver la Sección 4.5.2, página 139.

El *núcleo*²⁷ de una matriz $A \in \mathbb{K}^{m \times n}$ coincide con el *núcleo* de la TL que define dicha matriz.

$$\vec{y} = T(\vec{x}) = A\vec{x} \quad \implies \quad \ker(T) = \ker(A)$$

Ejercicio 4.11. Encontrar el núcleo de la TL de $T : \mathbb{R}^3 \mapsto \mathbb{R}^2$ definida por:

$$T(\vec{x}) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

El *núcleo* de T son las soluciones del sistema $A\vec{x} = \vec{0}$. Resolviendo dicho sistema mediante el método de Gauss-Jordan²⁸:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{array} \right] \xrightarrow{F_2 - F_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{array} \right] \xrightarrow{F_1 - F_2} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array} \right]$$

$$\text{Solución}^{29} \text{ del SEL: } x_3 = t, \quad x_2 = -2t, \quad x_1 = t \quad \implies \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

En este caso el *núcleo* de T es la *recta* en $\mathbb{R}^3$ que pasando por el origen, tiene por vector director $[1, -2, 1]$.

$$\ker(T) = \left\{ t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} ; \quad \forall t \in \mathbb{R} \right\}$$

Ejercicio 4.12. Encontrar el núcleo de la TL $T : \mathbb{R}^5 \mapsto \mathbb{R}^4$ definida por la matriz:

$$A = \begin{bmatrix} 1 & 5 & 4 & 3 & 2 \\ 1 & 6 & 6 & 6 & 6 \\ 1 & 7 & 8 & 10 & 12 \\ 1 & 6 & 6 & 7 & 8 \end{bmatrix}$$

Aplicando Gauss-Jordan³⁰ se obtiene la matriz *escalonada reducida equivalente*:

$$A = \begin{bmatrix} 1 & 5 & 4 & 3 & 2 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 7 & 10 \\ 0 & 1 & 2 & 4 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -6 & -12 & -18 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -6 & 0 & 6 \\ 0 & 1 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{Solución del SEL: } \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 6s - 6t \\ -2s + 2t \\ s \\ -2t \\ t \end{bmatrix}$$

²⁷Ver la Sección 3.10, página 111.

²⁸Ver la Sección 1.7, página 57.

²⁹Como se puede apreciar de los cálculos del problema, al aplicar el método de Gauss-Jordan en la determinación del *núcleo* no es necesario incluir la *matriz ampliada* completa, ya que el vector de ceros independiente permanece inalterable con las operaciones.

El *núcleo* de esta TL corresponde con un *plano* en $\mathbb{R}^5$ dado por:

$$\ker(T) = \left\{ s \begin{bmatrix} 6 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -6 \\ 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} ; \quad \forall s, t \in \mathbb{R} \right\}$$

Propiedades del Núcleo de una TL

Dada una TL entre un **espacio dominio** X a otro **espacio imagen** Y : $T : X \mapsto Y$

- El vector nulo del *espacio dominio*, $\vec{0} \in X$, está contenido en el $\ker(T)$.
- El *núcleo* de T es *cerrado* respecto a la suma:

$$\vec{v}_1 + \vec{v}_2 \in \ker(T) \quad ; \quad \forall \vec{v}_1, \vec{v}_2 \in \ker(T)$$

- El *núcleo* de T es *cerrado* con respecto a la multiplicación por un escalar:

$$k\vec{v} \in \ker(T) \quad ; \quad \forall \vec{v} \in \ker(T), k \in \mathbb{K}$$

Demostración

- $A\vec{0} = \vec{0} \implies \vec{0} \in \ker(T)$
- Sean $\vec{v}_1, \vec{v}_2 \in \ker(T) \implies \begin{Bmatrix} A\vec{v}_1 = \vec{0} \\ A\vec{v}_2 = \vec{0} \end{Bmatrix} \implies A(\vec{v}_1 + \vec{v}_2) = A\vec{v}_1 + A\vec{v}_2 = \vec{0} + \vec{0} = \vec{0}$
- Sea $\vec{v} \in \ker(T)$ y $k \in \mathbb{K} \implies A(k\vec{v}) = k(A\vec{v}) = k\vec{0} = \vec{0}$

El *núcleo* de una TL es un *subespacio vectorial* de su *espacio dominio*.

4.5.3. Nulidad y Rango de una TL

Definición 4.13. Nulidad de una TL

La *dimensión*³¹ del *núcleo* de una TL, $T : \mathbb{K}^n \mapsto \mathbb{K}^m$, definida por $\vec{y} = T(\vec{x}) = A\vec{x}$, se denomina **nulidad** de la transformación lineal (o de la matriz A asociada a dicha transformación). Se representa por $\text{nul}(T)$, (o $\text{nul}(A)$):

$$\text{nul}(T) = \dim(\ker T) = \dim(\ker A) = \text{nul}(A)$$

La *nulidad* es la *dimensión* de la solución del *sistema homogéneo*³² $A\vec{x} = \vec{0}$, y por tanto coincide con el número de parámetros de su solución, (que a su vez, coincide con el número de filas nulas que contiene una *matriz escalonada equivalente*³³ de A).

³⁰En este caso y para simplificar como se ha visto en el ejemplo anterior, no incluimos el vector de términos independientes (ceros). El símbolo $\sim$ indica *matrices equivalentes por filas*.

³¹Ver la [Definición 2.26, página 80](#).

³²Ver la [Sección 1.8, página 59](#).

³³Ver la [Definición 1.32, página 50](#).

Definición 4.14. Rango de una TL

La *dimensión* de la *imagen* de una TL, $T : \mathbb{K}^n \mapsto \mathbb{K}^m$, definida por $\vec{y} = T(\vec{x}) = A\vec{x}$, se denomina **rango** de T .

$$\text{rang}(T) = \dim(\text{im } T)$$

El concepto de *rango* ya fue definido en el **Tema 3** sobre matrices³⁴ y como se verá en el siguiente teorema, está íntimamente relacionado con el *rango* de una TL.

Teorema 4.2.

El *rango* de una TL coincide con el *rango* de su matriz asociada A :

$$\text{rang}(T) = \dim(\text{im } T) = \dim(\text{im } A) = \text{rang}(A)$$

Demostración 4.2.

Dada una TL: $\vec{y} = A\vec{x}$

1. La *imagen* de T es el subespacio del *espacio imagen* que posee solución en el SEL: $A\vec{x} = \vec{y}$.
2. La $\dim(T)$ coincide con el número de columnas *linealmente independientes* de A .
3. El número de columnas *linealmente independientes* coincide con el número de filas no nulas en la *matriz escalonada equivalente* y por tanto, coincide con el $\text{rang}(A)$.

Por tanto, si $T : \mathbb{K}^n \mapsto \mathbb{K}^m$, $\vec{y} = T(\vec{x}) = A\vec{x}$:

$$\dim(\text{im}(T)) = m - \text{Num. parámetros libres de la sol.} = m - \text{Num. filas nulas de } A = \text{rang}(A)$$

Teorema 4.3. Teorema del Rango de las TL

Dada una TL, $T : \mathbb{K}^n \mapsto \mathbb{K}^m$, con matriz asociada $A \in \mathbb{K}^{m \times n}$, entonces se verifica:

$$\text{rang}(T) + \text{nul}(T) = n$$

Nota 4.5. Tener en cuenta que n representa la dimensión del *espacio dominio* de la TL y también el *número de columnas* de la matriz asociada A .

Demostración 4.3.

Ver el **Teorema 3.6, página 116**.

Las siguientes ecuaciones son todas equivalentes al teorema del rango:

$$\dim(\ker T) + \dim(\text{im } T) = n$$

$$\dim(\ker A) + \dim(\text{im } A) = n$$

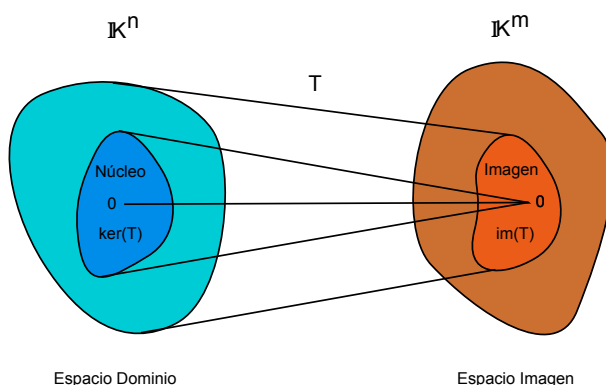
$$\text{nul}(A) + \text{rang}(A) = n$$

³⁴Ver la **Definición 3.24, página 114**.

Relación entre los Subespacios Asociados a una TL

El teorema del rango es muy importante ya que establece la relación que existe entre los subespacios vectoriales *imagen* y *núcleo* de una TL, a través de sus dimensiones, *rango* y *nulidad*.

De forma esquemática, podemos representar el “funcionamiento” de una TL y la relación entre los elementos que la definen: *espacio dominio*, *espacio imagen*³⁵, *núcleo* e *imagen*³⁶ de la transformación, de la siguiente forma:



El teorema del rango establece que el “tamaño”³⁷ de los subespacios *núcleo* e *imagen* juntos coincide con el “tamaño” del espacio *dominio*.

4.6. Tipos de TL

4.6.1. Inyectivas, Sobreyectivas y Biyectivas

Dada una transformación lineal, $T : \mathbb{K}^n \mapsto \mathbb{K}^m$, definida por $\vec{y} = T(\vec{x}) = A\vec{x}$. Entonces³⁸:

- Si $\ker(T) = \{\vec{0}\} \iff T$ es **inyectiva**.
- Si $\dim(\text{im}(T)) = m \iff T$ es **sobreyectiva**.
- T es *invertible* si es **biyectiva**.

Demostración

- Si $\ker(T) = \{\vec{0}\} \iff T$ es *inyectiva*.

Si $\ker(T) = \{\vec{0}\} \implies$ la única solución de $A\vec{x} = \vec{0}$ es $\vec{x} = \vec{0}$.

Dados $\vec{u}, \vec{v} \in \mathbb{K}^n$, dos vectores del *espacio dominio* de T , vamos a ver bajo que condiciones $T(\vec{u}) = T(\vec{v})$.

$$T(\vec{u}) = T(\vec{v}) \ ; \ T(\vec{u}) - T(\vec{v}) = \vec{0} \ ; \ T(\vec{u} - \vec{v}) = \vec{0} \implies \text{Sol. única } \vec{u} - \vec{v} = \vec{0} \implies \vec{u} = \vec{v}$$

Que es la definición de una función *inyectiva*.

- Si $\dim(\text{im}(T)) = m \iff T$ es *sobreyectiva*.

Si $\dim(\text{im}(T)) = m$, todo elemento del *espacio imagen* ($\mathbb{K}^m$) es *imagen* de algún elemento del *espacio dominio* ($\mathbb{K}^n$), y por tanto T es *sobreyectiva*.

- T es *invertible* si es *biyectiva*.

Ver la **Teorema 3.8**, página 117.

³⁵Ver la **Definición 4.1**, página 121.

³⁶Ver la **Sección 4.5**, página 136.

³⁷Representado a través de la *dimensión* del mismo, ver la **Definición 2.26**, página 80.

³⁸Ver el **Apéndice A.12.3**, página 279, las definiciones de funciones *inyectiva*, *sobreyectiva* y *biyectiva*.

4.6.2. Homomorfismos, Endomorfismos e Isomorfismos

- A las *transformaciones lineales* también se las denominan **morfismos**³⁹ (u **homomorfismos**).
- Un **endomorfismo** es una transformación lineal donde los espacios dominio e imagen son iguales. Se dice que es una aplicación lineal de un espacio vectorial en sí mismo.

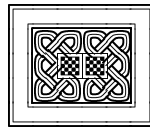
$$T : V \rightarrow V$$

- A las TL *biyectivas* se las denomina **isomorfismos**. Se dice entonces que los *espacios dominio e imagen* son **isomorfos**.

$$T : U \rightarrow V$$

En este caso los espacios dominio e imagen no necesitan ser iguales, pero sí hay una relación *uno a uno* entre sus elementos. Todo *endomorfismo* es un *isomorfismo*, pero no al contrario.

Nota 4.6. En el [Apéndice B.3, página 287](#) se pueden ver algunas aplicaciones geométricas de las TL.

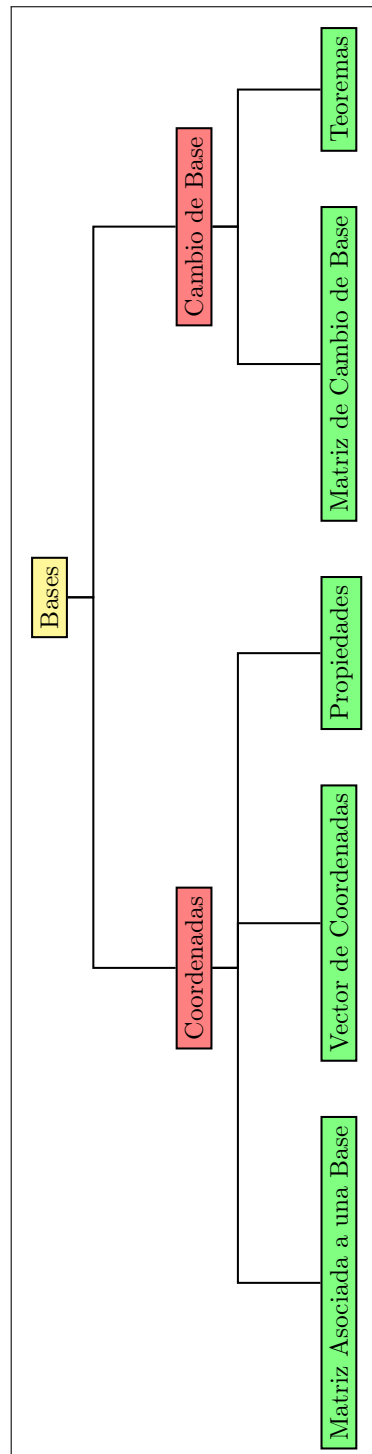


³⁹Aplicaciones entre *estructuras matemáticas* que preservan su estructura interna. Por ejemplo, en álgebra lineal, los morfismos son las transformaciones lineales, y en topología, las funciones continuas.

Bases

Contenido

5.1	Bases	145
5.1.1	Coordenadas y Matriz asociada a una Base	146
5.2	Cambio de Base	152
5.2.1	Matriz de Cambio de Base	153
5.2.2	Cambio de Base en una TL	158



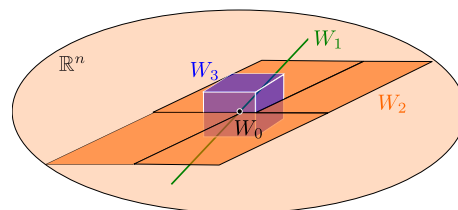
5.1. Bases

Recordamos los conceptos de *conjunto generador* y *base* de un subespacio vectorial.

Sea un subespacio vectorial $W \subseteq \mathbb{K}^n$ de dimensión $m \leq n$. Se dice que un subconjunto $G \subset W$, es:

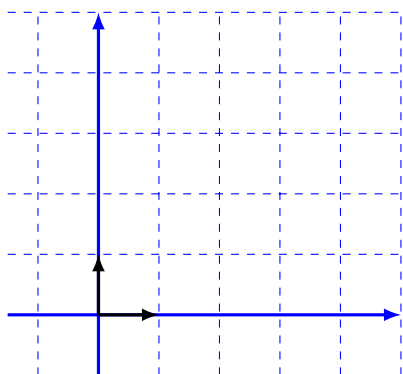
- **Conj. Generador** de W : si $\text{rang}(G) = \dim(W)$
- **Base** de W : si $N^{\circ} \text{ Vec. de } G = \text{rang}(G) = \dim(W)$

Podemos pensar que un espacio vectorial¹ $\mathbb{R}^n$ está estructurado internamente mediante subespacios vectoriales W_m de dimensión $0 \leq m \leq n$. Cada subespacio vectorial puede definirse a través de una base $\mathcal{B}$ formada por un conjunto de m vectores de W_m linealmente independientes, existiendo infinitas bases para cada subespacio vectorial de dimensión no nula.

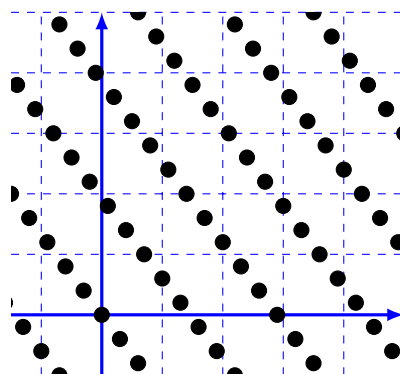


Las bases son las herramientas matemáticas que se usan para trabajar con los subespacios vectoriales a los que representan, es por ello que la elección de una buena base es esencial.

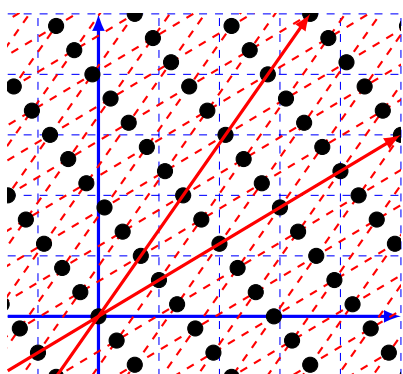
En la siguiente figura vamos a poner un ejemplo gráfico en $\mathbb{R}^2$. La base más sencilla en $\mathbb{R}^2$ es la base canónica $\{[1, 0], [0, 1]\}$ formada por dos vectores unitarios perpendiculares (ver vectores en negro en la figura a). Supongamos que tenemos que trabajar con una serie de elementos² dispuestos en una determinada configuración espacial (ver figura b). Las coordenadas de estos elementos en la base canónica son valores no sencillos ya que no siguen la malla o disposición espacial que genera la base canónica (rayas discontinuas azules).



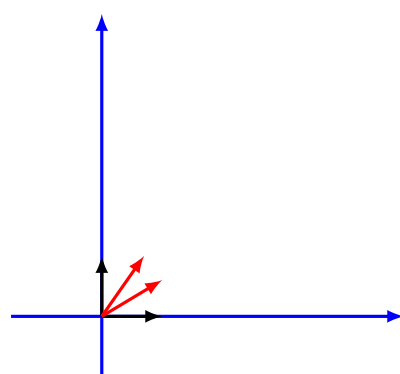
a)



b)



c)



d)

¹Puede extenderse a cualquier espacio vectorial complejo $\mathbb{C}^n$ pero sin el sentido geométrico exclusivo para $\mathbb{R}^n$. En la gráfica, el subespacio vectorial W_0 representa el origen (punto de dimensión 0), W_1 representa una recta (dimensión 1), W_2 un plano (dimensión 2) y W_3 un hiperplano de dimensión 3.

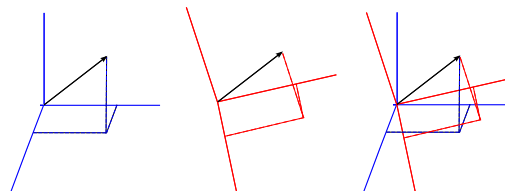
²Conjunto de puntos o vectores correspondientes a diferentes estados (valores de las variables) de un sistema.

Una forma óptima de hacer referencia a los elementos del sistema es buscar una buena base que permita identificarlos de una forma sencilla (coordenadas formadas por números sencillos). En el ejemplo presentado se pueden elegir dos ejes (en rojo) siguiendo dos direcciones particulares de la ordenación (figura c). De esta forma se puede generar una nueva malla (línea discontinua en rojo) que contenga a todos los puntos del sistema y hacer referencia a ellos mediante coordenadas simples.

Por tanto se tienen dos bases (figura d), la base canónica (en negro) útil para la entrada de datos y salida de resultados, y la nueva base que sigue la malla de los puntos (en rojo) útil para los cálculos internos con las coordenadas de los puntos.

Notar que al expresar un vector en distintas bases solo cambia la forma en que se hace referencia a éstos pero no los vectores en sí.

En la siguiente figura se muestra un único vector (en negro) referenciado por dos bases distintas (en azul y rojo).



En la siguiente sección vamos a ver la forma de referenciar vectores mediante bases distintas de la canónica.

5.1.1. Coordenadas y Matriz asociada a una Base

Definición 5.1. Coordenadas de un Vector respecto una Base

Sea una *base* $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_m\}$ de un subespacio $W \subseteq \mathbb{K}^n$ de $\dim(W) = m$. Cualquier vector $\vec{x} \in W$ se puede escribir de forma *única* como:

$$\vec{x} = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \dots + \alpha_m \vec{b}_m$$

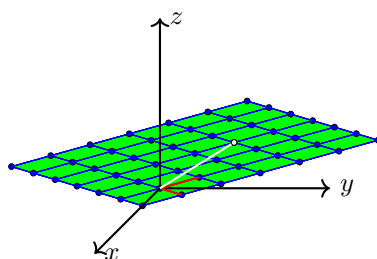
Los escalares $\alpha_1, \dots, \alpha_m$ se denominan **coordenadas de $\vec{x}$ respecto de la base $\mathcal{B}$** y el vector:

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}$$

se denomina **vector de coordenadas de $\vec{x}$ respecto de la base $\mathcal{B}$** .

Ejemplo 5.1. Coordenadas de un Vector en una Base

Sea en $\mathbb{R}^3$ una serie de puntos (en color azul en la figura de abajo) dispuestos en una malla incluida en un plano (en color verde) que corresponde a un subespacio vectorial W de $\dim(W) = 2$.



Cada uno de estos puntos tienen unas coordenadas en $\mathbb{R}^3$ que de forma general no son sencillas. Por ejemplo uno de estos puntos (en color blanco) tiene por coordenadas (expresado como vector) $\vec{u} = [0.5, 1.5, 1] \in W$. Podemos referenciar cada uno de estos puntos de una forma sencilla buscando una base $\mathcal{B}_W$ apropiada (en rojo en la figura) siguiendo alguna de las direcciones características de la malla, por ejemplo la siguiente (en rojo en la figura):

$$\mathcal{B}_W = \{[-1, 0, -0.5], [-0.5, 0.5, 0]\}$$

Todos los puntos de la malla pueden expresarse de forma simple en esta base, por ejemplo el vector $\vec{u}$ anterior puede expresarse mediante la siguiente combinación lineal:

$$\vec{u} = [0.5, 1.5, 1] = -2[-1, 0, -0.5] + 3[-0.5, 0.5, 0]$$

de forma que:

$$\begin{cases} (-2, 3) & : \text{coordenadas de } \vec{u} \text{ en la base } \mathcal{B} \\ [\vec{u}]_{\mathcal{B}} = [-2, 3] & : \text{vector de coordenadas de } \vec{u} \text{ en la base } \mathcal{B} \end{cases}$$

Las bases son los elementos matemáticos utilizados para trabajar con los subespacios vectoriales y para ello se organizan en matrices que permiten cálculos de forma apropiada.

Definición 5.2. Matriz Asociada a una Base

La definición de un vector $\vec{x} \in W$ en una base $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_m\}$ del subespacio W , se puede expresar como una TL³ de la siguiente forma:

$$\vec{x} = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \dots + \alpha_m \vec{b}_m = [\vec{b}_1 \ \dots \ \vec{b}_m] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = P_{\mathcal{B}} [\vec{x}]_{\mathcal{B}}$$

donde $P_{\mathcal{B}} = [\vec{b}_1 \ \dots \ \vec{b}_m] \in \mathbb{K}^{n \times m}$ es la *matriz asociada*⁴ a la TL que define la *base*⁵ y se denomina **matriz asociada a la base $\mathcal{B}$** , y $[\vec{x}]_{\mathcal{B}}$ es el vector de coordenadas⁶ de $\vec{x}$ en la base $\mathcal{B}$.

La matriz $P_{\mathcal{B}}$ asociada a una base $\mathcal{B}$ de un subespacio vectorial $W \subseteq \mathbb{K}^n$ es aquella cuyos vectores **columna** corresponden con los vectores de la base ordenados.

Ejercicio 5.1. Sea la base $\mathcal{B} = \{[3, 1], [-1, 3]\}$ de $\mathbb{R}^2$.

- Encontrar la matriz asociada a la base $\mathcal{B}$, $P_{\mathcal{B}}$.
- Si $\vec{x} = [10, 10]$, encontrar $[\vec{x}]_{\mathcal{B}}$.
- Si $[\vec{u}]_{\mathcal{B}} = [2, -1]$, encontrar el vector $\vec{u}$ en la base canónica.
- Encontrar, en caso de que exista, una base $\mathcal{C}$ de $\mathbb{R}^2$ donde el vector $\vec{w} = [1, -1]$ en la base canónica, tenga por vector de coordenadas $[\vec{w}]_{\mathcal{C}} = [1, 1]$ en la base $\mathcal{C}$.

a)

$$P_{\mathcal{B}} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}$$

b) Tenemos que encontrar el vector de coordenadas del vector $\vec{x}$ en la base $\mathcal{B}$.

$$\vec{x} = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 = P_{\mathcal{B}} [\vec{x}]_{\mathcal{B}} \implies \begin{bmatrix} 10 \\ 10 \end{bmatrix} = \alpha_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

³Ver la Definición 4.1, página 121.

⁴Ver la Definición 4.2, página 122.

⁵Ver la Definición 5.1, página 146.

⁶Los vectores $\vec{b}_i$ se disponen como columnas de la matriz.

Resolviendo el SEL:

$$\left[\begin{array}{cc|c} 3 & -1 & 10 \\ 1 & 3 & 10 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 0 & -10 & -20 \\ 10 & 0 & 40 \end{array} \right] \Rightarrow \begin{cases} \alpha_2 = 2 \\ \alpha_1 = 4 \end{cases} \Rightarrow [\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

- c) Tenemos que encontrar un vector $\vec{u}$ a partir de sus coordenadas dadas en la base $\mathcal{B}$. Para ello, basta con determinar el vector resultado de la combinación lineal:

$$\vec{u} = P_{\mathcal{B}}[\vec{u}]_{\mathcal{B}} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix} \Rightarrow \vec{u} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}$$

- d) Se tiene que cumplir:

$$\vec{w} = P_{\mathcal{C}}[\vec{w}]_{\mathcal{C}} \quad ; \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \begin{cases} a + b = 1 \\ c + d = -1 \end{cases}$$

Este es un sistema compatible indeterminado, con infinitas soluciones:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow \begin{cases} a + b = 1 - t \\ b = t \\ c + d = -1 - s \\ d = s \end{cases} \Rightarrow \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

Una posible solución sería por ejemplo para $t = s = 1$:

$$P_{\mathcal{C}} = \begin{bmatrix} 0 & 1 \\ -2 & 1 \end{bmatrix} \Rightarrow \mathcal{C}_W = \left\{ \begin{bmatrix} 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Ya que los vectores de coordenadas en una base se pueden definir a través de transformaciones lineales, se verifican las propiedades básicas de estas, como muestra el siguiente teorema.

Teorema 5.1. Propiedades del Vector de Coordenadas

Sea $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ una base del subespacio $W \subseteq \mathbb{K}^m$, y sean $\vec{x}, \vec{y} \in W$ y $c \in \mathbb{K}$. Entonces:

1. $[\vec{x} + \vec{y}]_{\mathcal{B}} = [\vec{x}]_{\mathcal{B}} + [\vec{y}]_{\mathcal{B}}$.
2. $[c\vec{x}]_{\mathcal{B}} = c[\vec{x}]_{\mathcal{B}}$.

Demostración 5.1.

Sean:

$$\begin{cases} [\vec{x}]_{\mathcal{B}} = [\alpha_1, \dots, \alpha_n] \Rightarrow \vec{x} = \alpha_1 \vec{b}_1 + \dots + \alpha_n \vec{b}_n \\ [\vec{y}]_{\mathcal{B}} = [\beta_1, \dots, \beta_n] \Rightarrow \vec{y} = \beta_1 \vec{b}_1 + \dots + \beta_n \vec{b}_n \end{cases}$$

Aplicando propiedades de los espacios vectoriales, (ver la [Sección 2.2, página 67](#)), se tiene:

$$\vec{x} + \vec{y} = (\alpha_1 + \beta_1) \vec{b}_1 + \dots + (\alpha_n + \beta_n) \vec{b}_n \Rightarrow [\vec{x} + \vec{y}]_{\mathcal{B}} = [(\alpha_1 + \beta_1), \dots, (\alpha_n + \beta_n)] = [\vec{x}]_{\mathcal{B}} + [\vec{y}]_{\mathcal{B}}$$

$$c\vec{x} = c(\alpha_1 \vec{b}_1 + \dots + \alpha_n \vec{b}_n) \Rightarrow [c\vec{x}]_{\mathcal{B}} = [c\alpha_1, \dots, c\alpha_n] = c[\vec{x}]_{\mathcal{B}}$$

Las dos propiedades del **Teorema 5.1** se pueden resumir en una sola⁷:

$$[c_1 \vec{b}_1 + \cdots + c_n \vec{b}_n]_{\mathcal{B}} = c_1 [\vec{b}_1]_{\mathcal{B}} + \cdots + c_n [\vec{b}_n]_{\mathcal{B}} \quad ; \quad \forall c_i \in \mathbb{K}$$

Esto implica que las *coordenadas* de una *combinación lineal de vectores* en una *base* determinada coincide con la *combinación lineal de las coordenadas* de los vectores en la misma *base*.

Ejercicio 5.2. Comprobar las propiedades del **Teorema 5.1** con los vectores $\vec{a} = [0, 1, -1, 1]$ y $\vec{b} = [1, -2, 1, -1]$ de un subespacio $W \subset \mathbb{R}^4$, siendo $\mathcal{B} = \{[-1, 1, 0, 0], [-1, 0, 1, 0], [0, 0, 0, 1]\}$ una base de W , para un caso en concreto.

Primeramente vamos a formar un vector $\vec{v} \in W$ como una combinación lineal cualquiera de los vectores $\vec{a}$ y $\vec{b}$ de W :

$$\vec{v} = \alpha \vec{a} + \beta \vec{b} = 2[0, 1, -1, 1] + 3[1, -2, 1, -1] = [3, -4, 1, -1]$$

en el cual hemos tomado por ejemplo $\alpha = 2$ y $\beta = 3$. Tenemos que comprobar que el vector de coordenadas del vector $\vec{v}$ en la base $\mathcal{B}$ se puede también obtener realizando la misma combinación lineal pero con los vectores de coordenadas de $\vec{a}$ y $\vec{b}$ en la base $\mathcal{B}$, es decir comprobar que se cumple:

$$[\alpha \vec{a} + \beta \vec{b}]_{\mathcal{B}} = \alpha [\vec{a}]_{\mathcal{B}} + \beta [\vec{b}]_{\mathcal{B}}$$

Si obtenemos el vector de coordenadas del vector $\vec{v}$ se tiene:

$$[\vec{v}]_{\mathcal{B}} = [\alpha \vec{a} + \beta \vec{b}]_{\mathcal{B}} = [-4, 1, -1]$$

Por simplicidad, no se muestran aquí los cálculos realizados para su obtención. Si de igual forma obtenemos los vectores de coordenadas de $\vec{a}$ y $\vec{b}$:

$$\left. \begin{array}{l} [\vec{a}]_{\mathcal{B}} = [1, -1, 1] \\ [\vec{b}]_{\mathcal{B}} = [-2, 1, -1] \end{array} \right\} \implies \alpha [\vec{a}]_{\mathcal{B}} + \beta [\vec{b}]_{\mathcal{B}} = 2[1, -1, 1] + 3[-2, 1, -1] = [-4, 1, -1]$$

Pudiéndose comprobar que se cumple la propiedad en este caso.

En los casos particulares de que los subespacios vectoriales sean el espacio vectorial completo, la obtención de los vectores de coordenadas puede hacerse de una forma particular, como muestra el siguiente teorema.

Teorema 5.2.

En el espacio $\mathbb{K}^n$ las *coordenadas* de cualquier vector $\vec{x} \in \mathbb{K}^n$ en una determinada base $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$, se pueden obtener como:

$$\vec{x} = P_{\mathcal{B}} [\vec{x}]_{\mathcal{B}} = [\vec{b}_1, \dots, \vec{b}_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \implies [\vec{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1} \vec{x}$$

Demostración 5.2.

Como en este caso particular donde $W = \mathbb{K}^n$, la matriz $P_{\mathcal{B}}$ es cuadrada y con vectores columna linealmente independientes (por ser una base), siempre es invertible, pudiéndose por tanto multiplicar por la izquierda las dos partes de la expresión $\vec{x} = P_{\mathcal{B}} [\vec{x}]_{\mathcal{B}}$, obteniéndose el resultado mostrado.

⁷Al igual que se hizo en el **Teorema 4.1**, página 123, con las propiedades de una TL, que fueron resumidas en la página 125.

Ejercicio 5.3. Encontrar una base $\mathcal{B}$ de $\mathbb{R}^4$ distinta de la canónica y dar una determinada transformación lineal que permita obtener el vector de coordenadas de cualquier $\vec{v} = [v_1, v_2, v_3, v_4] \in \mathbb{R}^4$ en la base $\mathcal{B}$.

El ejercicio nos pide que encontremos una matriz asociada a una base $\mathcal{B}$, es decir $P_{\mathcal{B}}$ y a través de ella demos la transformación lineal:

$$[\vec{v}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1} \vec{v}$$

Para encontrar una base de $\mathbb{R}^4$ distinta de la canónica de forma sencilla, podemos por ejemplo dar cuatro vectores simples de $\mathbb{R}^4$, que formen una matriz triangular, (lo cual asegura que los cuatro vectores son linealmente independientes y por tanto forman una base), por ejemplo:

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\} \implies P_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Encontrando la inversa de esta matriz, por ejemplo por Gauss-Jordan:

$$\left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right]$$

Por tanto, para todo $\vec{v} = [v_1, v_2, v_3, v_4]$ la transformación lineal pedida es:

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

Ejercicio 5.4. Dada la base $\mathcal{B} = \{-1-x+x^2, -1+x^2, -1+x\}$ del espacio de polinomios⁸ $\mathbb{P}_2$ encontrar las coordenadas del polinomio $p(x) = 1 + 2x + x^2$ en la base $\mathcal{B}$ y el polinomio $[q(x)]_{\mathcal{B}} = [-1, 2, 1]$ en la base canónica.

Con respecto a la primera pregunta, resolvemos el SEL:

$$\left[\begin{array}{ccc|c} -1 & -1 & -1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} -1 & -1 & -1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -1 & 2 \end{array} \right] \rightarrow \begin{cases} x_1 = -1 - x_2 - x_3 = -1 - 5 + 2 = -4 \\ x_2 = 1 - 2x_3 = 1 + 4 = 5 \\ x_3 = -2 \end{cases}$$

$$\implies [p(x)]_{\mathcal{B}} = [-4, 5, -2]$$

Comprobamos el resultado:

$$p(x) = -4(-1-x+x^2) + 5(-1+x^2) - 2(-1+x) = 4 + 4x - 4x^2 - 5 + 5x^2 + 2 - 2x = 1 + 2x + x^2$$

Con respecto a la segunda pregunta:

$$q(x) = -1(-1-x+x^2) + 2(-1+x^2) + (-1+x) = 1 + x - x^2 - 2 + 2x^2 - 1 + x = -2 + 2x + x^2$$

⁸Ver el Ejercicio 2.20, página 85.

Ejercicio 5.5. Dada la transformación lineal entre polinomios⁹ reales de grado dos y tres, $T : \mathbb{P}_2 \rightarrow \mathbb{P}_3$ que verifica:

$$T(t) = -t + 2 \quad ; \quad T(2t^2 + t) = 2t^2 + t \quad ; \quad T(t^2 + t + 1) = -t^3 - t + 1$$

- Encontrar la matriz asociada a T respecto de las bases canónicas $\mathcal{B}_{\mathbb{P}_2} = [1, t, t^2]$ y $\mathcal{B}_{\mathbb{P}_3} = [1, t, t^2, t^3]$.
- Determinar las dimensiones del núcleo y de la imagen de la transformación, dando (si es posible) una base de cada uno de ellos.
- Demostrar que $\mathcal{B} = \{t, 2t^2 + t, t^2 + t + 1\}$ es una base de $\mathbb{P}_2$.

a) Este problema puede interpretarse en términos de vectores de la siguiente forma:

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}^4 \quad ; \quad T([0, 1, 0]) = [2, 1, 0, 0] \quad ; \quad T([0, 1, 2]) = [0, 0, 2, 4] \quad ; \quad T([1, 1, 1]) = [1, 0, 0, 1]$$

La matriz $A_{4 \times 3}$ asociada a T se puede determinar de la siguiente forma:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \end{bmatrix} \quad \left\{ \begin{array}{l} T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} b \\ e \\ h \\ k \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{array}{l} b = 2 \\ e = -1 \\ h = 0 \\ k = 0 \end{array} \\ \\ T\left(\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} a & 2 & c \\ d & -1 & f \\ g & 0 & i \\ j & 0 & l \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2+2c \\ -1+f \\ 2i \\ 2l \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} \rightarrow \begin{array}{l} c = -1 \\ f = 1 \\ i = 1 \\ l = 0 \end{array} \\ \\ T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} a & 2 & -1 \\ d & -1 & 1 \\ g & 0 & 1 \\ j & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+a \\ d \\ 1+g \\ j \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ -1 \end{bmatrix} \rightarrow \begin{array}{l} a = 0 \\ d = -1 \\ g = -1 \\ j = -1 \end{array} \end{array} \right.$$

$$A = \begin{bmatrix} 0 & 2 & -1 \\ -1 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$$

b) Aplicamos Gauss:

$$\begin{bmatrix} 0 & 2 & -1 \\ -1 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 2 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{cases}$$

Por tanto $\ker(T) = [0, 0, 0]$, siendo $\dim[\ker(T)] = 0$ y por tanto el núcleo de la transformación no tiene base. Por otra parte, la imagen de T corresponde con el espacio columna de A y por tanto $\dim[\text{im}(T)] = 3$, con $\text{im}(T) = \mathbb{P}_2$ y siendo una posible base $\mathcal{B}_{\text{im}(T)} = \{1, t, t^2\}$.

c) Demostrar que $\mathcal{B} = \{t, 2t^2 + t, t^2 + t + 1\}$ es una base de $\mathbb{P}_2$ es equivalente a demostrar que $\mathcal{B} = \{[0, 1, 0], [0, 1, 2], [1, 1, 1]\}$ es una base de $\mathbb{R}^3$ y por tanto comprobamos que son linealmente independientes:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Claramente al ser 3 vectores linealmente independientes forman una base de $\mathbb{R}^3$ y por tanto queda demostrado que $\mathcal{B}$ es una base de $\mathbb{P}_2$.

⁹Ver el Ejemplo 2.4, página 71.

5.2. Cambio de Base

Muchas veces en la práctica interesa trabajar en un determinado *sistema de coordenadas* definido por una determinada *base*, por ejemplo por resultar los cálculos más sencillos de efectuar.

En esta sección vamos a ver cómo se sistematizan los cambios de base en espacios $\mathbb{K}^n$. Veamos primeramente un ejemplo sencillo sobre vectores expresados en diferentes bases.

Ejercicio 5.6. Sea un subespacio vectorial $W \subset \mathbb{R}^3$. Dado el vector $\vec{x} = [-1, 1, 1] \in W$, determinar los vectores de coordenadas del vector $\vec{x}$, en las siguientes dos bases de W :

$$\mathcal{B}_W = \{\vec{b}_1, \vec{b}_2\} = \{[1, 1, 0], [-2, 0, 1]\} \quad ; \quad \mathcal{C}_W = \{\vec{c}_1, \vec{c}_2\} = \{[1, 1, 0], [1, -1, -1]\}$$

Hay que resolver dos SEL, uno con cada base.

- a) Con respecto a la primera base $\mathcal{B}_W = \{\vec{b}_1, \vec{b}_2\} = \{[1, 1, 0], [-2, 0, 1]\}$, tenemos que resolver el sistema:

$$\alpha_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Cuya solución es $\alpha_1 = 1, \alpha_2 = 1$, por tanto el vector de coordenadas es: $[\vec{x}]_{\mathcal{B}} = [1, 1]$

- b) Con respecto a la segunda base $\mathcal{C}_W = \{\vec{c}_1, \vec{c}_2\} = \{[1, 1, 0], [1, -1, -1]\}$, tenemos que resolver el sistema:

$$\alpha_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Cuya solución es $\alpha_1 = 0, \alpha_2 = -1$, por tanto el vector de coordenadas es: $[\vec{x}]_{\mathcal{C}} = [0, -1]$.

De forma general, para cambiar un vector $\vec{v}$ perteneciente a un subespacio vectorial $W \subseteq \mathbb{K}^n$ expresado en una determinada base $\mathcal{B}_W$ de W , a través de su vector de coordenadas $[\vec{v}]_{\mathcal{B}}$, y expresarlo en otra base $\mathcal{C}_W$ de W de forma $[\vec{v}]_{\mathcal{C}}$ hay que seguir el siguiente proceso:

1. Conocido el vector de coordenadas $[\vec{v}]_{\mathcal{B}}$ y la base $\mathcal{B}_W = \{\vec{b}_1, \dots, \vec{b}_m\}$ con $m = \dim(W)$, obtener el vector $\vec{v}$ en la base canónica de forma¹⁰:

$$\vec{v} = P_{\mathcal{B}}[\vec{v}]_{\mathcal{B}}$$

2. Conocido $\vec{v}$ y otra base $\mathcal{C}_W = \{\vec{c}_1, \dots, \vec{c}_m\}$, determinamos el nuevo vector de coordenadas $[\vec{v}]_{\mathcal{C}}$ resolviendo el siguiente SEL:

$$P_{\mathcal{C}}[\vec{v}]_{\mathcal{C}} = \vec{v}$$

Ejercicio 5.7. Dado un vector $\vec{v}$ perteneciente a un subespacio vectorial $W \subset \mathbb{C}^4$ definido a través del vector de coordenadas $[\vec{v}]_{\mathcal{B}} = [1, i]$ correspondiente a una base $\mathcal{B}_W = \{[1, 0, i, 0], [-1, -1, 0, -i]\}$ de W , encontrar el vector de coordenadas de $\vec{v}$ en la base $\mathcal{C}_W = \{[1, 1, 0, i], [0, 1, -i, i]\}$ de W .

Siguiendo el proceso descrito anteriormente, calculamos primero el vector $\vec{v}$ en la base canónica de $\mathbb{C}^4$.

$$\vec{v} = P_{\mathcal{B}}[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \\ i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 1-i \\ -i \\ i \\ 1 \end{bmatrix}$$

¹⁰Ver la Definición 5.2, página 147.

Conocido $\vec{v} = [1 - i, -i, i, 1]$ calculamos el vector de coordenadas pedido $[\vec{v}]_{\mathcal{C}} = [x_1, x_2]$ resolviendo el SEL:

$$P_{\mathcal{C}}[\vec{v}]_{\mathcal{C}} = \vec{v} \quad ; \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & -i \\ i & i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 - i \\ -i \\ i \\ 1 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 1 & 0 & 1-i \\ 1 & 1 & -i \\ 0 & -i & i \\ i & i & 1 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 1-i \\ 0 & 1 & -1 \\ 0 & -i & i \\ 0 & i & -i \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 1-i \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right] \rightarrow \begin{cases} x_1 = 1 - i \\ x_2 = -1 \end{cases}$$

Por lo que finalmente el vector pedido es $[\vec{v}]_{\mathcal{C}} = [1 - i, -1]$, pudiéndose verificar que $\vec{v} = P_{\mathcal{C}}[\vec{v}]_{\mathcal{C}}$.

5.2.1. Matriz de Cambio de Base

En determinados casos es necesario realizar muchos cambios de variable entre dos bases distintas de un mismo subespacio vectorial. Cuando las dimensiones del espacio vectorial en el que se trabaja es grande¹¹ los múltiples y grandes SEL que hay resolver resultan bastante costos computacionalmente. En estos casos, resulta más apropiado realizar los múltiples cambios de variables utilizando un método alternativo que no necesite la resolución de ningún SEL. Para ello se utiliza la *matriz de cambio de base* que se define a continuación.

Definición 5.3. Matriz de Cambio de Base

Sean $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_k\}$ y $\mathcal{C} = \{\vec{c}_1, \dots, \vec{c}_k\}$ dos bases de un mismo subespacio vectorial $W \subseteq \mathbb{K}^n$ de dimensión $k \leq n$. Se denomina **matriz de cambio de base** de $\mathcal{B}$ a $\mathcal{C}$, y se denota por $P_{\mathcal{C} \leftarrow \mathcal{B}}$, a la matriz de $k \times k$ cuyas columnas son los vectores de coordenadas de $\mathcal{B}$ con respecto a $\mathcal{C}$:

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = [\vec{b}_1]_{\mathcal{C}}, \dots, [\vec{b}_k]_{\mathcal{C}}$$

Teorema 5.3. Cambio de Base

Sean $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_k\}$ y $\mathcal{C} = \{\vec{c}_1, \dots, \vec{c}_k\}$ dos bases de un espacio vectorial $W \subseteq \mathbb{K}^n$ de dimensión $k \leq n$, y $P_{\mathcal{C} \leftarrow \mathcal{B}}$ la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$. Entonces:

$$[\vec{x}]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}}[\vec{x}]_{\mathcal{B}} \quad ; \quad \forall \vec{x} \in W$$

Demostración 5.3.

Sea $\vec{x} \in W$ y $[\vec{x}]_{\mathcal{B}} = [\alpha_1, \dots, \alpha_k]^T$, entonces $\vec{x} = \alpha_1 \vec{b}_1 + \dots + \alpha_k \vec{b}_k$, y por tanto:

$$[\vec{x}]_{\mathcal{C}} = [\alpha_1 \vec{b}_1 + \dots + \alpha_k \vec{b}_k]_{\mathcal{C}} = \alpha_1 [\vec{b}_1]_{\mathcal{C}} + \dots + \alpha_k [\vec{b}_k]_{\mathcal{C}} = [\vec{b}_1]_{\mathcal{C}}, \dots, [\vec{b}_k]_{\mathcal{C}} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix} = P_{\mathcal{C} \leftarrow \mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

Toda *matriz de cambio de base* $P_{\mathcal{C} \leftarrow \mathcal{B}}$ en un subespacio vectorial de dimensión k , es una matriz cuadrada e invertible de dimensión $k \times k$.

¹¹En problemas prácticos de ciencia e ingeniería es fácil trabajar con sistemas definidos por cientos, miles e incluso millones de variables que dan lugar a espacios vectoriales de esas dimensiones.

Ejercicio 5.8. Dadas las dos siguientes bases de un subespacio vectorial $W \subset \mathbb{R}^3$, utilizadas en el anterior Ejemplo 5.6:

$$\mathcal{B}_W = \{\vec{b}_1, \vec{b}_2\} = \{[1, 1, 0], [-2, 0, 1]\} \quad ; \quad \mathcal{C}_W = \{\vec{b}_1, \vec{b}_2\} = \{[1, 1, 0], [1, -1, -1]\}$$

Se pide:

- Encontrar la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$, $P_{\mathcal{C} \leftarrow \mathcal{B}}$.
- Utilizar $P_{\mathcal{C} \leftarrow \mathcal{B}}$ para obtener el vector $[\vec{x}]_{\mathcal{C}}$, siendo $\vec{x} = [-1, 1, 1]$ y sabiendo que $[\vec{x}]_{\mathcal{B}} = [1, 1]$.
- Comprobar el resultado con el obtenido en el Ejemplo 5.6.

- Queremos encontrar la matriz:

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = \left[[\vec{b}_1]_{\mathcal{C}}, \dots, [\vec{b}_n]_{\mathcal{C}} \right]$$

Para ello calculamos por separado cada uno de sus dos vectores columna:

$$[\vec{b}_1]_{\mathcal{C}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} : P_{\mathcal{C}}[\vec{b}_1]_{\mathcal{C}} = \vec{b}_1 \implies \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \xrightarrow{\text{Sol.}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$[\vec{b}_2]_{\mathcal{C}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} : P_{\mathcal{C}}[\vec{b}_2]_{\mathcal{C}} = \vec{b}_2 \implies \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{\text{Sol.}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}$$

$$\text{b) } [\vec{x}]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}}[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

- El vector obtenido coincide con el calculado en el Ejemplo 5.6.

La ventaja que presenta la *matriz de cambio de base* con respecto al método tradicional en el que hay que resolver un SEL, es que una vez obtenida dicha matriz, los vectores de coordenadas se obtienen mediante una simple multiplicación matriz-vector, sin necesidad de resolver ningún sistema de ecuaciones.

Teorema 5.4.

Sean $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_m\}$ y $\mathcal{C} = \{\vec{c}_1, \dots, \vec{c}_m\}$ bases de un subespacio vectorial $W \subseteq \mathbb{K}^n$ con $\dim(W) = m \leq n$, y sea $P_{\mathcal{C} \leftarrow \mathcal{B}}$ la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$. Entonces:

- $P_{\mathcal{C} \leftarrow \mathcal{B}}$ es la única matriz M tal que: $M[\vec{x}]_{\mathcal{B}} = [\vec{x}]_{\mathcal{C}} \quad ; \quad \forall \vec{x} \in W$.
- $P_{\mathcal{C} \leftarrow \mathcal{B}}$ es invertible.
- $(P_{\mathcal{C} \leftarrow \mathcal{B}})^{-1} = P_{\mathcal{B} \leftarrow \mathcal{C}}$.

Demostración 5.4.

1. Sea M una matriz con la propiedad $M[\vec{x}]_{\mathcal{B}} = [\vec{x}]_{\mathcal{C}}$ para todo $\vec{x} \in W$. Tomando $\vec{x} = \vec{b}_i$, (i -ésima componente de la base $\mathcal{B}$), se tiene que $[\vec{x}]_{\mathcal{B}} = [\vec{b}_i]_{\mathcal{B}} = \vec{e}_i$ y por tanto:

$$M[\vec{x}]_{\mathcal{B}} = M[\vec{b}_i]_{\mathcal{B}} = M\vec{e}_i = [\vec{b}_i]_{\mathcal{C}}$$

que corresponde con la i -ésima columna de la matriz $P_{\mathcal{C} \leftarrow \mathcal{B}}$. Esto obliga a que $M = P_{\mathcal{C} \leftarrow \mathcal{B}}$.

2. La matriz $P_{\mathcal{C} \leftarrow \mathcal{B}} = \left[[\vec{b}_1]_{\mathcal{C}}, \dots, [\vec{b}_m]_{\mathcal{C}} \right]$ es cuadrada, ya que tiene m columnas (dimensión de W) y cada una de ellas formada por m valores (coeficientes de cada uno de los m vectores de la base $\mathcal{C}$). Por otra parte, como $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_m\}$ es un conjunto linealmente independiente, los vectores columna de la matriz $P_{\mathcal{C} \leftarrow \mathcal{B}}$, es decir $\{[\vec{b}_1]_{\mathcal{C}}, \dots, [\vec{b}_m]_{\mathcal{C}}\}$ también lo son, ya que la única posibilidad de que se cumpla:

$$\alpha_1[\vec{b}_1]_{\mathcal{C}} + \dots + \alpha_m[\vec{b}_m]_{\mathcal{C}} = [\alpha_1\vec{b}_1 + \dots + \alpha_m\vec{b}_m]_{\mathcal{C}} = \vec{0} \quad ; \quad \alpha_i \in \mathbb{K}$$

es que $\alpha_1 = \dots = \alpha_m = 0$, que es la definición de independencia lineal de vectores, (ver la [Sección 2.4.2, página 73](#)).

La matriz $P_{\mathcal{C} \leftarrow \mathcal{B}}$ invertible al ser cuadrada y con vectores columna linealmente independientes.

3. Como $P_{\mathcal{C} \leftarrow \mathcal{B}}$ es invertible y:

$$P_{\mathcal{C} \leftarrow \mathcal{B}}[\vec{x}]_{\mathcal{B}} = [\vec{x}]_{\mathcal{C}} \quad \implies \quad [\vec{x}]_{\mathcal{B}} = (P_{\mathcal{C} \leftarrow \mathcal{B}})^{-1}[\vec{x}]_{\mathcal{C}}$$

Debido a la unicidad de la matriz de cambio de base, tenemos que $P_{\mathcal{B} \leftarrow \mathcal{C}} = (P_{\mathcal{C} \leftarrow \mathcal{B}})^{-1}$.

Ejercicio 5.9. Sabiendo que la matriz de cambio de base $P_{\mathcal{C} \leftarrow \mathcal{B}}$ correspondiente al [Ejemplo 5.8](#) es:

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}$$

- Encontrar la matriz de cambio de base $P_{\mathcal{B} \leftarrow \mathcal{C}}$.
- Utilizar $P_{\mathcal{B} \leftarrow \mathcal{C}}$ para obtener el vector $[\vec{x}]_{\mathcal{B}}$, sabiendo que $[\vec{x}]_{\mathcal{C}} = [0, -1]$.
- Comprobar el resultado con el obtenido en el [Ejemplo 5.8](#).

- a) Tenemos que calcular la matriz inversa de la dada, ya que $P_{\mathcal{B} \leftarrow \mathcal{C}} = (P_{\mathcal{C} \leftarrow \mathcal{B}})^{-1}$.

$$\left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{array} \right]$$

En este caso la matriz inversa coincide con la original, (este tipo de matrices se denominan **involutivas**), por tanto:

$$P_{\mathcal{B} \leftarrow \mathcal{C}} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}$$

$$b) \quad [\vec{x}]_{\mathcal{B}} = P_{\mathcal{B} \leftarrow \mathcal{C}}[\vec{x}]_{\mathcal{C}} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- c) El vector obtenido coincide con el calculado en el [Ejemplo 5.8](#).

Teorema 5.5.

Sean $\mathcal{B} = [\vec{b}_1, \dots, \vec{b}_n]$ y $\mathcal{C} = [\vec{c}_1, \dots, \vec{c}_n]$ bases del espacio $\mathbb{K}^n$, entonces:

$$P_{\mathcal{B} \leftarrow \mathcal{C}} = (P_{\mathcal{B}})^{-1} P_{\mathcal{C}} \quad ; \quad P_{\mathcal{C} \leftarrow \mathcal{B}} = (P_{\mathcal{C}})^{-1} P_{\mathcal{B}}$$

Notar que esta fórmula solo es válida para bases de espacios $\mathbb{K}^n$, donde las matrices asociadas a las bases son invertibles (cuadradas y con vectores linealmente independientes), y no para bases de subespacios $W \subset \mathbb{K}^n$, donde las matrices $P_{\mathcal{B}} \in \mathbb{K}^{n \times m}$ al no ser cuadradas, no son invertibles.

Demostración 5.5.

$$\text{Sea } \vec{x} \in \mathbb{K}^n : \quad \left\{ \begin{array}{l} \vec{x} = P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} \\ \vec{x} = P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}} \end{array} \right\} \implies P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} = P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}} \implies \left\{ \begin{array}{l} [\vec{x}]_{\mathcal{B}} = (P_{\mathcal{B}})^{-1} P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}} = P_{\mathcal{B} \leftarrow \mathcal{C}}[\vec{x}]_{\mathcal{C}} \\ [\vec{x}]_{\mathcal{C}} = (P_{\mathcal{C}})^{-1} P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} = P_{\mathcal{C} \leftarrow \mathcal{B}}[\vec{x}]_{\mathcal{B}} \end{array} \right.$$

Ejercicio 5.10. La matriz de cambio de una base $\mathcal{B}$ a una base $\mathcal{C}$ de un subespacio W es:

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Sabiendo que expresados en la base canónica, los vectores de la base $\mathcal{B}$ son:

$$\mathcal{B} = \{\vec{b}_1 = [1, 0, 1], \vec{b}_2 = [1, 1, 0], \vec{b}_3 = [0, 0, 1]\}$$

encontrar los vectores de la base $\mathcal{C} = \{\vec{c}_1, \vec{c}_2, \vec{c}_3\}$.

La matriz de cambio de base cumple para cualquier vector $\vec{x}$ del subespacio:

$$[\vec{x}]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

donde $[\vec{x}]_{\mathcal{C}}$ y $[\vec{x}]_{\mathcal{B}}$ son los vectores de coordenadas del vector $\vec{x}$ en las bases $\mathcal{C}$ y $\mathcal{B}$ respectivamente.

La matriz de cambio de base se puede obtener como:

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} [\vec{b}_1]_{\mathcal{C}} & [\vec{b}_2]_{\mathcal{C}} & [\vec{b}_3]_{\mathcal{C}} \end{bmatrix}$$

donde $\vec{b}_i$ son los vectores de la base $\mathcal{B}$. Por tanto, los elementos de la columna i de la matriz $P_{\mathcal{C} \leftarrow \mathcal{B}}$ son los coeficientes de la combinación lineal de los vectores $\vec{c}_i$ de la base $\mathcal{C}$ para obtener $\vec{b}_i$.

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \implies \left\{ \begin{array}{l} \vec{c}_1 + 2\vec{c}_2 + \vec{c}_3 = \vec{b}_1 \\ \vec{c}_1 + \vec{c}_2 - \vec{c}_3 = \vec{b}_2 \\ \vec{c}_1 + \vec{c}_2 + \vec{c}_3 = \vec{b}_3 \end{array} \right.$$

De forma matricial lo podemos expresar como:

$$\begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \vec{c}_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix}$$

donde $\begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \vec{c}_3 \end{bmatrix}$ y $\begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix}$ son matrices con los correspondientes vectores como columnas.

Despejando la matriz con los vectores $\vec{c}_i$, se tiene:

$$\begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \vec{c}_3 \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \vec{b}_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}^{-1} \quad (5.1)$$

Por tanto basta con calcular la inversa de la matriz $P_{\mathcal{C} \leftarrow \mathcal{B}}$ y multiplicarla por una matriz con los vectores columna $\vec{b}_i$ y se obtendrá una matriz con los vectores columna $\vec{c}_i$.

Calculamos primero la matriz inversa:

$$\begin{aligned} & \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow[\frac{F_3-F_1}{F_2-2F_1}]{\frac{F_2-2F_1}{F_3-F_1}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & -2 & 0 & -1 & 0 & 1 \end{array} \right] \\ & \xrightarrow[\frac{F_3-2F_2}{F_1+F_2}]{\frac{F_1+F_2}{F_3-2F_2}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & 2 & 3 & -2 & 1 \end{array} \right] \xrightarrow[\frac{1}{2}F_3]{-\frac{F_2}{2}F_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 3/2 & -1 & 1/2 \end{array} \right] \\ & \xrightarrow{F_2-F_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1/2 & 0 & -1/2 \\ 0 & 0 & 1 & 3/2 & -1 & 1/2 \end{array} \right] \end{aligned}$$

Y ahora calculamos los vectores $\vec{c}_i$ con la fórmula (5.1):

$$[\vec{c}_1 \quad \vec{c}_2 \quad \vec{c}_3] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \\ 1/2 & 0 & -1/2 \\ 3/2 & -1 & 1/2 \end{bmatrix} = \begin{bmatrix} -1/2 & 1 & -1/2 \\ 1/2 & 0 & -1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

Y por tanto la solución es:

$$\vec{c}_1 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1/2 \end{bmatrix} ; \quad \vec{c}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} ; \quad \vec{c}_3 = \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$$

Ejercicio 5.11. Comprobar el resultado obtenido en el Ejemplo 5.10 para el vector $\vec{b}_1$.

El vector de coordenadas de $\vec{b}_1$ en la base $\mathcal{B}$ es $[\vec{b}_1]_{\mathcal{B}} = [1, 0, 0]$, ya que $\vec{b}_1$ es el primer vector de dicha base. Entonces el vector de coordenadas de $\vec{b}_1$ en la base $\mathcal{C}$ es:

$$[\vec{b}_1]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}} [\vec{b}_1]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Por tanto, $\vec{b}_1$ se puede obtener por combinación lineal de los vectores de la base $\mathcal{C}$ como:

$$\vec{b}_1 = P_{\mathcal{C}} [\vec{b}_1]_{\mathcal{C}} = \vec{c}_1 + 2\vec{c}_2 + \vec{c}_3 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1/2 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Esto comprueba que los vectores $\vec{c}_i$ están bien calculados y de igual forma se podría comprobar para los vectores $\vec{b}_2$ y $\vec{b}_3$.

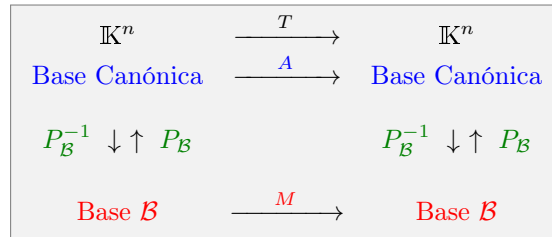
5.2.2. Cambio de Base en una TL

Sea una transformación lineal¹² $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ definida mediante una matriz $A \in \mathbb{K}^{n \times n}$ de la siguiente forma:

$$\vec{y} = T(\vec{x}) = A\vec{x}$$

Tanto los vectores del dominio $\vec{x}$ como los de la imagen $\vec{y}$ están en la *base canónica*¹³ de $\mathbb{K}^n$.

Supongamos que necesitamos trabajar en una determinada base $\mathcal{B}$ de $\mathbb{K}^n$ con matriz asociada $P_{\mathcal{B}}$ y queremos calcular cual es la forma que tendría la transformación lineal en esta base. De forma esquemática el proceso se puede expresar de la siguiente forma:



Sean $[\vec{x}]_{\mathcal{B}}$ y $[\vec{y}]_{\mathcal{B}}$ a los vectores del dominio y de la imagen en la nueva base $\mathcal{B}$, la nueva transformación lineal T' sería:

$$[\vec{y}]_{\mathcal{B}} = T'([\vec{x}]_{\mathcal{B}}) = M[\vec{x}]_{\mathcal{B}}$$

donde M es la matriz asociada a la misma transformación lineal T pero expresada en la base $\mathcal{B}$ y la cual queremos conocer. Con todos estos datos se tiene¹⁴:

$$\begin{cases} \vec{x} = P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} \\ \vec{y} = P_{\mathcal{B}}[\vec{y}]_{\mathcal{B}} \end{cases}$$

Por tanto se tiene que cumplir:

$$\vec{y} = A\vec{x} \quad \implies \quad P_{\mathcal{B}}[\vec{y}]_{\mathcal{B}} = AP_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} \quad \implies \quad [\vec{y}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

de forma que la matriz M buscada que define la transformación lineal T en la base $\mathcal{B}$ es:

$$M = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}}$$

- Las dos matrices A y M anteriores definen una misma TL pero en bases distintas.
- Como se verá en el [Tema 8, página 217](#), las matrices A y M se denominan *matrices semejantes*¹⁵ y tienen muchas propiedades interesantes¹⁶.

¹²En concreto es un *endomorfismo*, ver la [Sección 4.6.2, página 142](#).

¹³Ver la [Sección 2.5.3, página 83](#).

¹⁴Ver la [Definición 5.2, página 147](#).

¹⁵Ver la [Definición 8.6, página 230](#).

¹⁶Ver la [Sección 8.4.2, página 231](#).

Ejercicio 5.12. Dada la siguiente transformación lineal $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ y la siguiente base $\mathcal{B}$ de $\mathbb{R}^3$, se pide encontrar la matriz M asociada a T en la base $\mathcal{B}$.

$$T \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} 2x - y \\ 3y \\ x + 2z \end{bmatrix} \quad ; \quad \mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

La matriz $A \in \mathbb{R}^{3 \times 3}$ asociada al *isomorfismo*¹⁷ $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ en la *base canónica*¹⁸ de $\mathbb{R}^3$ es:

$$\vec{y} = T(\vec{x}) = A\vec{x} \quad ; \quad A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

Queremos encontrar la expresión de esta TL en la nueva base $\mathcal{B}$ con matriz asociada:

$$P_{\mathcal{B}} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3 & 1 \\ 1 & 2 & 0 \end{bmatrix}$$

de la forma:

$$[\vec{y}]_{\mathcal{B}} = T'([\vec{x}]_{\mathcal{B}}) = M[\vec{x}]_{\mathcal{B}}$$

donde M es la matriz asociada a la nueva transformación lineal y la cual queremos conocer. Entonces como hemos visto:

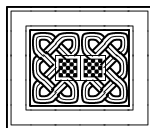
$$\vec{y} = A\vec{x} \quad \implies \quad P_{\mathcal{B}}[\vec{y}]_{\mathcal{B}} = AP_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} \quad \implies \quad [\vec{y}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} \quad \implies \quad \boxed{M = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}}}$$

Entonces calculamos primero $P_{\mathcal{B}}^{-1}$ mediante el método de Gauss-Jordan¹⁹:

$$\left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array} \right] \rightarrow \dots \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2/5 & 0 & 1/5 \\ 0 & 1 & 0 & -1/5 & 0 & 2/5 \\ 0 & 0 & 1 & 3/5 & 1 & E - 6/5 \end{array} \right] \implies P_{\mathcal{B}}^{-1} = \begin{bmatrix} 2/5 & 0 & 1/5 \\ -1/5 & 0 & 2/5 \\ 3/5 & 1 & -6/5 \end{bmatrix}$$

y por tanto:

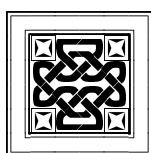
$$M = \begin{bmatrix} 2/5 & 0 & 1/5 \\ -1/5 & 0 & 2/5 \\ 3/5 & 1 & -6/5 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3 & 1 \\ 1 & 2 & 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 12 & -7 & -2 \\ 4 & 11 & 1 \\ -12 & 12 & 12 \end{bmatrix}$$



¹⁷Ver la Sección 4.6.2, página 142.

¹⁸Ver la Sección 2.5.3, página 83.

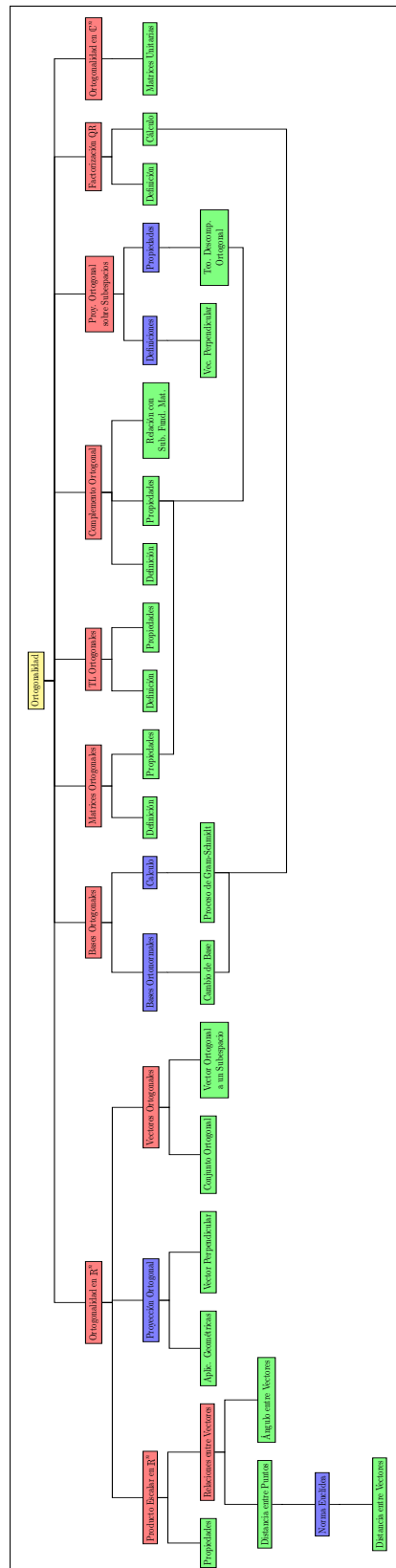
¹⁹Ver la Sección 3.7.1, página 99.



Ortogonalidad

Contenido

6.1	Introducción	163
6.2	Producto Escalar en $\mathbb{R}^n$	163
6.2.1	Distancia y Norma en $\mathbb{R}^n$	164
6.2.2	Relaciones entre Vectores	166
6.2.3	Vectores Ortogonales	167
6.3	Ortogonalidad en $\mathbb{R}^n$	168
6.3.1	Conjunto Ortogonal	168
6.3.2	Bases Ortogonales	169
6.3.3	Bases Ortonormales	170
6.3.4	Matrices Ortogonales	171
6.3.5	TL Ortogonales	173
6.4	Complemento Ortogonal	173
6.4.1	Vector Ortogonal a un Subespacio	173
6.4.2	Complemento Ortogonal	174
6.4.3	Características del Complemento Ortogonal	174
6.4.4	Cálculo del Complemento Ortogonal	174
6.4.5	Relaciones de Ortogonalidad entre los Subespacios de una Matriz Real	175
6.4.6	Dimensión del Complemento Ortogonal	177
6.5	Proyección Ortogonal	178
6.5.1	Interpretación Geométrica en $\mathbb{R}^2$ y $\mathbb{R}^3$	178
6.5.2	Vector Perpendicular	180
6.5.3	Proyección Ortogonal sobre un Subespacio	181
6.6	Teorema de la Descomposición Ortogonal	183
6.7	Proceso de Gram-Schmidt	184
6.8	Factorización QR	185
6.9	Ortogonalidad en Espacios $\mathbb{C}^n$	187
6.9.1	Producto Interno	187
6.9.2	Norma en $\mathbb{C}^n$	188
6.9.3	Vectores Ortogonales en $\mathbb{C}^n$	188
6.9.4	Complemento Ortogonal en $\mathbb{C}^n$	189
6.9.5	Proyección Ortogonal en $\mathbb{C}^n$	192
6.9.6	Proceso de Gram-Schmidt en $\mathbb{C}^n$	193
6.9.7	Matrices Unitarias	194



6.1. Introducción

El concepto de *ortogonalidad* es muy importante dentro de las matemáticas. Su significado es más amplio que el puramente geométrico y por ejemplo se habla de *polinomios ortogonales*, *campos ortogonales*, etc.

En esta sección nos vamos a centrar inicialmente en la ortogonalidad en espacios $\mathbb{R}^n$, por tener una interpretación geométrica en baja dimensión, y posteriormente en la [Sección 6.9, página 187](#) se generalizará a espacios $\mathbb{C}^n$.

6.2. Producto Escalar en $\mathbb{R}^n$

Definición 6.1. Producto Escalar en $\mathbb{R}^n$

Dados dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$, se define el **producto escalar** de $\vec{u}$ por $\vec{v}$ y lo representamos por $\vec{u} \cdot \vec{v}$ como el *escalar* resultante de la siguiente operación:

$$\vec{u} \cdot \vec{v} = [u_1, u_2, \dots, u_n] \cdot [v_1, v_2, \dots, v_n] = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{k=1}^n u_k v_k \quad ; \quad \vec{u} \cdot \vec{v} \in \mathbb{R}$$

Ejercicio 6.1. Calcular el producto escalar de los vectores de $\mathbb{R}^3$:

$$\vec{u} = [1, 2, -3] \quad ; \quad \vec{v} = [-3, 5, 2]$$

$$\vec{u} \cdot \vec{v} = [1, 2, -3] \cdot [-3, 5, 2] = 1(-3) + 2(5) + (-3)2 = -3 + 10 - 6 = \boxed{1}$$

Nota 6.1.

- No confundir el *producto escalar* de dos vectores, con el *producto por un escalar*¹, $c\vec{v}$.
- Los dos vectores tienen que tener igual dimensión (número de componentes).
- El resultado siempre es un escalar, **no** un vector.
- El producto escalar es un caso especial del producto más general denominado *producto interno*², definido para espacios vectoriales generales³.

El *producto escalar* puede definirse a través del **producto matricial**⁴ de una *matriz fila* por una *matriz columna*, de la siguiente forma:

- Si $\vec{u}$ y $\vec{v}$ son dos **vectores/matrices fila** de la misma dimensión, es decir, $\vec{u}, \vec{v} \in \mathbb{R}^{n \times 1}$ entonces:

$$\vec{u} \cdot \vec{v} = \vec{u} \vec{v}^T = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{k=1}^n u_k v_k$$

- Si u y v son dos **vectores/matrices columna** de la misma dimensión, es decir, $\vec{u}, \vec{v} \in \mathbb{R}^{1 \times n}$ entonces:

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{k=1}^n u_k v_k$$

¹Ver la [Sección 2.2.2, página 68](#).

²Ver la [Definición 6.18, página 187](#).

³Ver la [Definición 2.12, página 71](#).

⁴Ver la [Sección 3.3, página 93](#).

Propiedades del Producto Escalar

Dados los vectores $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ y $c \in \mathbb{R}$, se verifica:

- Conmutativa: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
- Distributiva: $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.
- Asociativa: $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$
- Si $\vec{u} \in \mathbb{R}^n$ entonces: $\vec{u} \cdot \vec{u} \geq 0$. Si $\vec{u} \cdot \vec{u} = 0 \iff \vec{u} = \vec{0}$.

6.2.1. Distancia y Norma en $\mathbb{R}^n$

La **distancia**⁵ entre dos puntos A y B de $\mathbb{R}^n$ viene definida por el número real no negativo:

$$\left. \begin{matrix} A(a_1, \dots, a_n) \\ B(b_1, \dots, b_n) \end{matrix} \right\} d(A, B) = d(B, A) = \sqrt{(a_1 - b_1)^2 + \dots + (a_n - b_n)^2} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

A esta *distancia* se le denomina **distancia euclídea**.

Propiedades de la Distancia

Dados dos puntos $A, B \in \mathbb{R}^n$, se verifica:

- $d(A, B) \in \mathbb{R}$. $d(A, B) \geq 0$
- $d(A, B) = 0 \iff A = B \implies d(X, X) = 0 \quad \forall X \in \mathbb{R}^n$
- $d(A, B) = d(B, A)$

Distancia de un Punto al Origen

Dado $A(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ con $x_i \in \mathbb{R}$, se tiene:

$$d(A, \vec{0}) = \sqrt{(x_1 - 0)^2 + \dots + (x_n - 0)^2} = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$$

Nota 6.2. Caso especial:

- Si $\mathbb{R}^n = \mathbb{R} \implies d(A, 0) = \sqrt{x^2} = |x|$ (valor absoluto de x).

Norma de un Vector en $\mathbb{R}^n$

Definición 6.2. Norma de un Vector en $\mathbb{R}^n$

Se define la **norma** de un vector $\vec{v} = [v_1, \dots, v_n] \in \mathbb{R}^n$ como la **distancia** del origen al extremo del vector y se identifica por $\|\vec{v}\|$.

$$\|\vec{v}\| = \sqrt{v_1^2 + \dots + v_n^2} = \sqrt{\sum_{i=1}^n v_i^2} = \sqrt{\sum_{i=1}^n v_i v_i} = \sqrt{\vec{v} \cdot \vec{v}}$$

Esta norma se la denomina **norma euclídea** o **norma 2**.

⁵Ver el Apéndice A.7, página 269.

Nota 6.3. Además de la *norma 2* o *euclídea* se definen otras *normas* muy utilizadas, como son la *norma 1*, la *norma n* y la *norma infinito*, las cuales tienen diferentes aplicaciones.

Propiedades de la Norma

Dado un vector $\vec{v} \in \mathbb{R}^n$ y un escalar $\alpha \in \mathbb{R}$, se verifica:

- $\|\vec{v}\| > 0$, $\forall \vec{v} \neq \vec{0}$; $\|\vec{0}\| = 0$

- $\|\alpha \vec{v}\| = |\alpha| \|\vec{v}\|$

Dem: $\|\alpha \vec{v}\| = \sqrt{(\alpha \vec{v}) \cdot (\alpha \vec{v})} = \sqrt{\alpha^2 (\vec{v} \cdot \vec{v})} = \sqrt{\alpha^2} \sqrt{\vec{v} \cdot \vec{v}} = |\alpha| \|\vec{v}\|$

- *Desigualdad de Cauchy-Schwarz* : $|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \|\vec{v}\|$

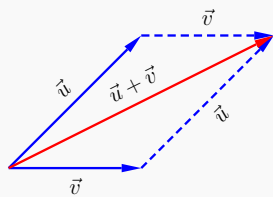
Dem: Ver **Ejemplo 6.10**, página 179.

- *Desigualdad Triangular*: $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Dem:

$$\begin{aligned} \|\vec{u} + \vec{v}\|^2 &= (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) = \vec{u} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{v} = \vec{u} \cdot \vec{u} + 2(\vec{u} \cdot \vec{v}) + \vec{v} \cdot \vec{v} = \\ &= \|\vec{u}\|^2 + 2(\vec{u} \cdot \vec{v}) + \|\vec{v}\|^2 \leq \|\vec{u}\|^2 + 2|\vec{u} \cdot \vec{v}| + \|\vec{v}\|^2 \stackrel{c-s}{\leq} \|\vec{u}\|^2 + 2\|\vec{u}\| \|\vec{v}\| + \|\vec{v}\|^2 = (\|\vec{u}\| + \|\vec{v}\|)^2 \\ &\implies \|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\| \end{aligned}$$

Nota 6.4.



Descripción gráfica de la desigualdad triangular en $\mathbb{R}^2$.

Vectores Unitarios y Normalización

Definición 6.3. Vector Unitario

Se dice que un vector $\vec{v} \in \mathbb{R}^n$ es **unitario** si su norma es la unidad.

$$\vec{v} \text{ vector unitario} \iff \|\vec{v}\| = 1$$

Cualquier vector $\vec{v} \in \mathbb{R}^n$ no nulo se puede transformar en unitario dividiendo por la norma del vector, de la siguiente forma:

$$\vec{v} = [x_1, x_2, \dots, x_n] \implies \vec{u} = \frac{1}{\|\vec{v}\|} \vec{v} = \frac{1}{\|\vec{v}\|} [x_1, x_2, \dots, x_n] = \left[\frac{x_1}{\|\vec{v}\|}, \dots, \frac{x_n}{\|\vec{v}\|} \right]$$

Comprobación:

$$\|\vec{u}\| = \sqrt{\frac{|x_1|^2}{\|\vec{v}\|^2} + \dots + \frac{|x_n|^2}{\|\vec{v}\|^2}} = \sqrt{\frac{|x_1|^2 + \dots + |x_n|^2}{\|\vec{v}\|^2}} = \sqrt{\frac{\|\vec{v}\|^2}{\|\vec{v}\|^2}} = \sqrt{1} = 1$$

- Se denomina **normalizar** un vector al proceso de obtener un vector unitario $\vec{u}$ a partir de un vector dado $\vec{v}$.
- El nuevo vector unitario $\vec{u}$ tiene la misma dirección y sentido que el vector original $\vec{v}$ pero con norma 1.

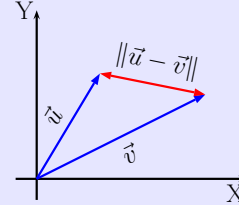
6.2.2. Relaciones entre Vectores

Distancia entre dos Vectores

Definición 6.4. Distancia entre dos Vectores

Dados dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$ se define la distancia entre ellos como:

$$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \|\vec{v} - \vec{u}\| = d(\vec{v}, \vec{u})$$



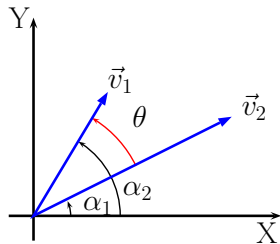
Nota 6.5. La definición de distancia entre vectores coincide con la definición de distancia entre puntos de $\mathbb{R}^n$, como se muestra a continuación:

$$A, B \in \mathbb{R}^n \implies d(A, B) = d(B, A) = \sqrt{(a_1 - b_1)^2 + \cdots + (a_n - b_n)^2} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

$$\vec{v}, \vec{u} \in \mathbb{R}^n \implies d(\vec{v}, \vec{u}) = d(\vec{u}, \vec{v}) = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2} = \sqrt{\sum_{i=1}^n (u_i - v_i)^2}$$

Ángulo entre dos Vectores en $\mathbb{R}^2$

Vamos a determinar el ángulo θ entre dos vectores $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^2$, siendo α_1 y α_2 sus respectivos ángulos.



$$\begin{cases} \vec{v}_1 = [x_1, y_1] \\ \vec{v}_2 = [x_2, y_2] \end{cases} \quad ; \quad \theta = |\alpha_1 - \alpha_2|$$

$$\sin \alpha_1 = \frac{y_1}{\|\vec{v}_1\|} \quad ; \quad \sin \alpha_2 = \frac{y_2}{\|\vec{v}_2\|}$$

$$\cos \alpha_1 = \frac{x_1}{\|\vec{v}_1\|} \quad ; \quad \cos \alpha_2 = \frac{x_2}{\|\vec{v}_2\|}$$

Teniendo en cuenta que: $\cos(a - b) = \cos a \cos b + \sin a \sin b$, se tiene:

$$\cos \theta = \cos(\alpha_1 - \alpha_2) = \cos \alpha_1 \cos \alpha_2 + \sin \alpha_1 \sin \alpha_2 = \frac{x_1}{\|\vec{v}_1\|} \frac{x_2}{\|\vec{v}_2\|} + \frac{y_1}{\|\vec{v}_1\|} \frac{y_2}{\|\vec{v}_2\|}$$

Despejando:

$$\|\vec{v}_1\| \|\vec{v}_2\| \cos \theta = x_1 x_2 + y_1 y_2 = \vec{v}_1 \cdot \vec{v}_2 \implies \cos \theta = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|}$$

Nota 6.6. El ángulo θ siempre se mide en *radianes*. La relación entre *grados sexagesimales* y *radianes* viene dada mediante la fórmula: $\text{radianes} = (\text{grados} \times 2\pi)/360$

Ángulo entre dos Vectores en $\mathbb{R}^n$

Generalizando la definición anterior a dos vectores no nulos $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$:

$$\cos \theta = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|}$$

Nota 6.7. Desde un punto de vista geométrico, solo tiene sentido el ángulo entre vectores de $\mathbb{R}^2$ y $\mathbb{R}^3$, aunque desde un punto de vista matemático, dicha expresión en $\mathbb{R}^n$ tiene múltiples aplicaciones.

Definición Alternativa del Producto Escalar en $\mathbb{R}^n$ **Definición 6.5. Producto Escalar en $\mathbb{R}^n$**

Dados dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$, se define el **producto escalar** de $\vec{u}$ por $\vec{v}$ como el número (escalar) resultado de la siguiente operación:

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos \theta$$

donde θ es el ángulo entre $\vec{u}$ y $\vec{v}$.

Recordamos que la primera definición dada⁶ del *producto escalar* en $\mathbb{R}^n$ fue:

$$\vec{u} \cdot \vec{v} = [u_1, u_2, \dots, u_n] \cdot [v_1, v_2, \dots, v_n] = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{k=1}^n u_k v_k$$

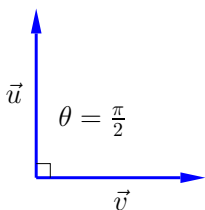
6.2.3. Vectores Ortogonales**Definición 6.6. Vectores Ortogonales**

Se dice que dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$ son **ortogonales** si su producto escalar es cero.

$$\vec{u} \cdot \vec{v} = 0 \iff \vec{u}, \vec{v} \text{ ortogonales}$$

Nota 6.8. El vector nulo $\vec{0} = [0, \dots, 0] \in \mathbb{R}^n$, es *ortogonal* a cualquier vector ya que: $\vec{v} \cdot \vec{0} = 0$; $\forall \vec{v} \in \mathbb{R}^n$

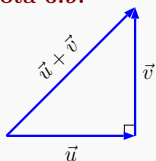
- En $\mathbb{R}^n$ dos vectores no nulos $\vec{u}$ y $\vec{v}$ son *ortogonales* si el ángulo entre ellos es $\pm \frac{\pi}{2} = \pm 90^\circ$.
- En $\mathbb{R}^2$ y $\mathbb{R}^3$, desde un punto de vista geométrico, los vectores *ortogonales* se dice que son *perpendiculares*.

Ejemplo 6.1. Ejemplo de vectores ortogonales en $\mathbb{R}^2$ y $\mathbb{R}^3$.

$$\cos \theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|} = \frac{0}{\|\vec{u}\| \|\vec{v}\|} = 0 \implies \theta = \arccos 0 = \pm \frac{\pi}{2}$$

$$\vec{u} \cdot \vec{v} = 0 \iff \vec{u} \perp \vec{v} \quad \text{En } \mathbb{R}^2 \text{ y } \mathbb{R}^3$$

Como se indicó en la introducción al tema, el concepto de *ortogonalidad* es muy importante dentro de las matemáticas. Su significado es más amplio que el puramente geométrico y por ejemplo se habla de *polinomios ortogonales*, *soluciones ortogonales*, *campos ortogonales*, etc.

Nota 6.9.

Si $\vec{u} \perp \vec{v} \implies$ Por el teorema de Pitágoras:

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$$

⁶Ver la **Definición 6.1**, página 163.

6.3. Ortogonalidad en $\mathbb{R}^n$

6.3.1. Conjunto Ortogonal

Definición 6.7. Conjunto Ortogonal

Se dice que un conjunto de vectores $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ de $\mathbb{R}^n$ es **ortogonal** si todos los pares de vectores *distintos* del conjunto son *ortogonales*⁷:

$$\vec{v}_i \cdot \vec{v}_j = 0 \quad \text{si } i \neq j \quad \forall i, j = 1, 2, \dots, k$$

Nota 6.10. Desde el punto de vista geométrico, los vectores de un *conjunto ortogonal* son mutuamente *perpendiculares*, es decir el ángulo entre ellos es $\pi/2$ rad = 90° .

Ejemplo 6.2. Base canónica.

La *base canónica*⁸ $\{\vec{e}_1, \dots, \vec{e}_n\}$ de $\mathbb{R}^n$ es un *conjunto ortogonal*.

Ejercicio 6.2. Demostrar que el conjunto $V = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \in \mathbb{R}^3$ es ortogonal, donde $\vec{v}_1 = [2, 1, -1]$, $\vec{v}_2 = [0, 1, 1]$ y $\vec{v}_3 = [1, -1, 1]$.

Tenemos que demostrar que todo producto escalar entre vectores distintos del conjunto es siempre **cero**.

$$\begin{cases} \vec{v}_1 \cdot \vec{v}_2 = 2(0) + 1(1) - 1(1) = 0 \\ \vec{v}_1 \cdot \vec{v}_3 = 2(1) - 1(1) - 1(1) = 0 \\ \vec{v}_2 \cdot \vec{v}_3 = 0(1) - 1(1) + 1(1) = 0 \end{cases}$$

Teorema 6.1.

Todo *conjunto ortogonal* de vectores no nulos $\{\vec{v}_1, \dots, \vec{v}_k\}$ de $\mathbb{R}^n$, es un conjunto *linealmente independiente*⁹.

Demostración 6.1.

Siguiendo la definición de conjunto *linealmente independiente* de vectores, vamos a determinar qué escalares $c_1, \dots, c_k \in \mathbb{R}$ verifican $c_1\vec{v}_1 + \dots + c_k\vec{v}_k = \vec{0}$.

Para cualquier vector $\vec{v}_i$ de un conjunto ortogonal se tendría:

$$(c_1\vec{v}_1 + \dots + c_k\vec{v}_k) \cdot \vec{v}_i = 0 \quad ; \quad c_1(\vec{v}_1 \cdot \vec{v}_i) + \dots + c_i(\vec{v}_i \cdot \vec{v}_i) + \dots + c_k(\vec{v}_k \cdot \vec{v}_i) = 0$$

Todos los productos $(\vec{v}_j \cdot \vec{v}_i)$ son **cero** salvo para $j = i$, por tanto

$$c_i(\vec{v}_i \cdot \vec{v}_i) = 0 \quad \implies \quad c_i = 0$$

ya que $\vec{v}_i \neq \vec{0}$ y por tanto $(\vec{v}_i \cdot \vec{v}_i) \neq 0$. Como esta expresión es válida para cualquier $i = 1, \dots, k$, implica que un *conjunto ortogonal* es un *conjunto linealmente independiente*.

⁷Ver la definición de *vectores ortogonales* en la **Sección 6.9.5, página 192**.

⁸Ver la **Sección 2.5.1, página 80**.

⁹Ver la **Sección 2.4.2, página 73**.

6.3.2. Bases Ortogonales

Definición 6.8. Base Ortogonal

Una **base ortogonal** de un subespacio¹⁰ W de $\mathbb{R}^n$ es una **base**¹¹ de W constituida por un *conjunto ortogonal*.

Ejemplo 6.3. Base ortogonal

En el **Ejemplo 6.2**, los tres vectores son *ortogonales* y por tanto son *linealmente independientes*. Como además su número (3) coincide con la *dimensión*¹² del espacio al que pertenecen $\mathbb{R}^3$, forman una *base ortogonal* de $\mathbb{R}^3$.

Ejercicio 6.3. Encontrar una base cualquiera y una base ortogonal del subespacio W de $\mathbb{R}^3$ dado por:

$$W = \{ [x, y, z] : x - y + 2z = 0 \}$$

Una base será un conjunto linealmente independiente de n vectores que cumplan esa ecuación, donde n es la dimensión de W . Despejando $x = y - 2z$ y tomando los parámetros $y = t$ y $z = s$ se tiene que el subespacio W formado por todos los vectores incluidos en el plano es:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t - 2s \\ t \\ s \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = t\vec{v}_1 + s\vec{v}_2$$

Como hay dos parámetros libre, se tiene¹³ que $\dim(W) = 2$. Como además $\vec{v}_1$ y $\vec{v}_2$ son vectores que pertenecen a W y son *linealmente independientes* (no *paralelos*), ambos forman una base del subespacio W :

$$\mathcal{B}_W = \{\vec{v}_1, \vec{v}_2\}$$

Esta base no es ortogonal ya que $\vec{v}_1 \cdot \vec{v}_2 = -2 \neq 0$. Una base ortogonal se formará con dos vectores que pertenezcan al plano y sean *ortogonales*. Sean esos vectores $\vec{u}_1 = [a, b, c]$ y $\vec{u}_2 = [d, e, f]$, entonces deben de cumplir:

$$\left\{ \begin{array}{lcl} a - b + 2c & = & 0 \\ d - e + 2f & = & 0 \\ ad + be + cf & = & 0 \end{array} \right\} \begin{array}{l} \text{Cumplen la ec. del plano.} \\ \text{Ortogonalidad} \end{array}$$

Este es un sistema no lineal de 3 ecuaciones y 6 incógnitas, por lo puede tener infinitas soluciones. Vamos a encontrar una solución simple, fijando algunos valores y determinando el resto:

$$\begin{array}{lll} a = 1, b = 1 & \xrightarrow{\text{Ec. 1}} & c = 0 \\ d = 1, e = -1 & \xrightarrow{\text{Ec. 3}} & f = -1 \end{array}$$

Por tanto, una posible *base ortogonal* de W será el conjunto:

$$\mathcal{B}_W = \{\vec{u}_1, \vec{u}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \right\}$$

¹⁰Ver la Definición 2.13, página 71.

¹¹Ver la Definición 2.25, página 80.

¹²Ver la Definición 2.26, página 80 y la Sección 2.5.2, página 81.

Teorema 6.2.

Sea $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ una *base ortogonal* de un subespacio $W \subseteq \mathbb{R}^n$ de dimensión $k \leq n$ y sea $\vec{x} \in W$. Entonces los escalares únicos¹⁴ $c_1, \dots, c_k$ tales que,

$$\vec{x} = c_1 \vec{v}_1 + \dots + c_k \vec{v}_k = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_m \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

están dados por:

$$c_i = \frac{\vec{v}_i \cdot \vec{x}}{\vec{v}_i \cdot \vec{v}_i} \quad ; \quad \forall i = 1, \dots, k$$
Demostración 6.2.

Como $\mathcal{B}$ es una base de W , existen los escalares únicos $c_1, \dots, c_k$ tales que $\vec{x} = c_1 \vec{v}_1 + \dots + c_k \vec{v}_k$, por tanto:

$$\begin{aligned} \vec{v}_i \cdot \vec{x} &= \vec{v}_i \cdot (c_1 \vec{v}_1 + \dots + c_k \vec{v}_k) = c_1 (\vec{v}_i \cdot \vec{v}_1) + \dots + c_i (\vec{v}_i \cdot \vec{v}_i) + \dots + c_k (\vec{v}_i \cdot \vec{v}_k) = \\ &= 0 + \dots + c_i (\vec{v}_i \cdot \vec{v}_i) + 0 = c_i (\vec{v}_i \cdot \vec{v}_i) \quad \implies \quad c_i = \frac{\vec{v}_i \cdot \vec{x}}{\vec{v}_i \cdot \vec{v}_i} \end{aligned}$$

Ejercicio 6.4. En la segunda base (ortogonal) del **Ejemplo 6.3**, $\mathcal{B} = \{\vec{u}_1, \vec{u}_2\} = \{[1, 1, 0], [1, -1, -1]\}$, determinar el vector de coordenadas del vector $\vec{x} = [-1, 1, 1] \in W$.

$$c_1 = \frac{\vec{u}_1 \cdot \vec{x}}{\vec{u}_1 \cdot \vec{u}_1} = \frac{0}{2} = 0 \quad ; \quad c_2 = \frac{\vec{u}_2 \cdot \vec{x}}{\vec{u}_2 \cdot \vec{u}_2} = \frac{-3}{3} = -1$$

Por tanto el vector de coordenadas es $[\vec{x}]_{\mathcal{B}} = [0, -1]$, igual que el obtenido en el **Ejemplo 5.6**, pero sin necesidad de resolver ningún SEL.

6.3.3. Bases Ortonormales**Definición 6.9. Conjunto Ortonormal y Base Ortonormal**

- Se dice que un **conjunto** de vectores $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ de $\mathbb{R}^n$ es **ortonormal** si es un *conjunto ortogonal*¹⁵ de vectores *unitarios*¹⁶.

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

- Una **base ortonormal** es aquella que está formada por un *conjunto ortonormal* de vectores.

Dada una *base ortogonal*¹⁷ $\{\vec{v}_1, \dots, \vec{v}_k\}$ se puede obtener fácilmente una *base ortonormal* $\{\vec{u}_1, \dots, \vec{u}_k\}$, simplemente *normalizando*¹⁸ cada uno de los vectores $\vec{v}_i$ de la base.

$$\vec{u}_i = \frac{1}{\|\vec{v}_i\|} \vec{v}_i = \frac{1}{\sqrt{\vec{v}_i \cdot \vec{v}_i}} \vec{v}_i$$

¹³Aplicando la *ecuación de equilibrio*, (ver la **Definición 1.15**, página 37.) a este caso obtenemos también que $n = 3 - 1 = 2$.

¹⁴Estos coeficientes corresponden con las coordenadas del vector $\vec{x}$ en la base $\mathcal{B}$. Ver **Sección 5.1.1**, página 146.

¹⁵Ver la **Definición 6.7**, página 168.

¹⁶Ver la **Definición 6.3**, página 165.

¹⁷Ver la **Definición 6.8**, página 169.

¹⁸Ver la **Sección 6.2.1**, página 165.

Teorema 6.3. Base Ortonormal

Sea $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_k\}$ una *base ortonormal* de un subespacio $W \subseteq \mathbb{R}^n$ de dimensión k y sea $\vec{x} \in W$ entonces:

$$\vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \dots + (\vec{u}_m \cdot \vec{x})\vec{u}_k$$

Demostración 6.3.

Trivial según el **Teorema 6.2** y teniendo en cuenta que cualquier vector ortonormal $\vec{u}$ cumple que $\vec{u} \cdot \vec{u} = 1$, ya que su norma es la unidad, $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = 1$.

Las *bases ortogonales* y *ortonormales* simplifican de forma considerable los cálculos de *cambios de base*¹⁹.

6.3.4. Matrices Ortogonales**Definición 6.10. Matriz Ortogonal**

Se dice que una matriz cuadrada $Q \in \mathbb{R}^{n \times n}$ es **ortogonal** si sus vectores columna forman un *conjunto ortonormal de vectores*²⁰.

Nota 6.11.

- No se utiliza el término *matriz ortonormal*, aunque quizás sería más *intuitivo*.
- Es habitual utilizar la notación Q para referirse a matrices ortogonales.

Teorema 6.4. Matriz Ortogonal y Transpuesta

Toda matriz **ortogonal** $Q_{n \times n}$ verifica:

$$Q^T Q = I_n$$

Demostración 6.4.

Si $\vec{u}, \vec{v} \in \mathbb{R}^n$ son vectores columna entonces²¹:

$$\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$$

Por tanto, el producto matricial $Q^T Q$ corresponde con todos los posibles productos escalares entre las columnas de Q . El resultado será 1 si coinciden columnas iguales dando lugar a los elementos de la diagonal, ó 0 si las columnas son distintas dando lugar al resto de los elementos.

$$Q^T Q = \begin{bmatrix} \boxed{1} \\ \boxed{2} \\ \vdots \\ \boxed{n} \end{bmatrix} \begin{bmatrix} \boxed{1} & \boxed{2} & \cdots & \boxed{n} \end{bmatrix} = I_n$$

¹⁹Ver la **Sección 5.2**, página 152.

²⁰Ver la **Definición 6.9**, página 170.

Teorema 6.5. Matriz Ortogonal e Inversa

Una matriz Q es **ortogonal** si y solo si:

$$Q^{-1} = Q^T \iff Q \text{ es una Matriz Ortogonal}$$

Demostración 6.5.

Según el **Teorema 6.4, página 171**, una *matriz ortogonal* Q verifica $Q^T Q = I_n$. Como los vectores columna de Q son *linealmente independientes*²² existe Q^{-1} , y por tanto multiplicando ambas partes de la igualdad anterior por la *inversa*, $Q^{-1} = Q^T$, se tiene:

$$Q^T Q = I_n \quad ; \quad Q^T \underbrace{Q Q^{-1}}_{I_n} = \underbrace{I_n Q^{-1}}_{Q^{-1}} \quad ; \quad \boxed{Q^T = Q^{-1}}$$

Teorema 6.6.

Los *vectores fila* de una **matriz ortogonal** Q forman un *conjunto ortonormal* de vectores.

Demostración 6.6.

$$Q^T = Q^{-1} \quad ; \quad \boxed{Q Q^T = I}$$

Propiedades de las Matrices Ortogonales

Sea $Q \in \mathbb{R}^{n \times n}$ una *matriz ortogonal* y $\vec{x}, \vec{y} \in \mathbb{R}^n$ dos vectores columna. Entonces:

- $|Q| = +1$ ó $|Q| = -1$.
- $Q\vec{x} \cdot Q\vec{y} = \vec{x} \cdot \vec{y}$
- $\|Q\vec{x}\| = \|\vec{x}\|$
- La *inversa* $Q^{-1} = Q^T$ también es una *matriz ortogonal*.
- Si A y B son *matrices ortogonales* de igual tamaño, entonces AB también es una *matriz ortogonal*.

Demostración

- $Q Q^T = I$; $|Q Q^T| = |I| = 1$; $|Q| |Q^T| = |Q|^2 = 1 \implies |Q| = \pm 1$
- Si Q es *ortogonal*, $Q^T Q = I$, y por tanto se tiene:

$$Q\vec{x} \cdot Q\vec{y} = (Q\vec{x})^T Q\vec{y} = \vec{x}^T Q^T Q\vec{y} = \vec{x}^T I\vec{y} = \vec{x}^T \vec{y} = \vec{x} \cdot \vec{y}$$
- $\|Q\vec{x}\| = \sqrt{Q\vec{x} \cdot Q\vec{x}} = \sqrt{(Q\vec{x})^T Q\vec{x}} = \sqrt{\vec{x}^T Q^T Q\vec{x}} = \sqrt{\vec{x}^T \vec{x}} = \sqrt{\vec{x} \cdot \vec{x}} = \|\vec{x}\|$
- Trivial.
- Si A y B son *matrices ortogonales* entonces:

$$(AB)^{-1} = B^{-1} A^{-1} = B^T A^T = (AB)^T \implies AB \text{ es una matriz ortogonal}$$

²¹Ver la **Sección 3.3.4, página 95**.

²²Al ser *ortogonales*, ver el **Teorema 6.1, página 168**.

6.3.5. TL Ortogonales

Definición 6.11. TL Ortogonal

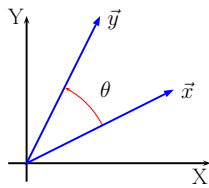
Una TL, $T: \mathbb{R}^n \mapsto \mathbb{R}^n$, $\vec{y} = T(\vec{x}) = A\vec{x}$, se dice que es **ortogonal** si la matriz $A \in \mathbb{R}^{n \times n}$ que la define es *ortogonal*.

Las TL ortogonales preservan la longitud de los vectores transformados, ya que²³:

$$\|\vec{y}\| = \|T(\vec{x})\| = \|Q\vec{x}\| = \|\vec{x}\|$$

Ejemplo 6.4. TL ortogonal

Una *rotación* de vectores en $\mathbb{R}^2$ con respecto al origen, corresponde con la TL²⁴:



$$\vec{y} = T(\vec{x}) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \vec{x}$$

Esta TL es *ortogonal* ya que la matriz que define la transformación es *ortogonal* (al ser los vectores columna y fila ortogonales). Claramente, la *rotación* preserva la longitud.

6.4. Complemento Ortogonal

6.4.1. Vector Ortogonal a un Subespacio

Definición 6.12. Vector Ortogonal a un Subespacio

Sea W un subespacio de $\mathbb{R}^n$. Se dice que un vector $\vec{v} \in \mathbb{R}^n$ es **ortogonal al subespacio W** si $\vec{v}$ es ortogonal a todos los vectores de W .

Teorema 6.7. Condición de Ortogonalidad a un Subespacio

Un vector $\vec{v} \in \mathbb{R}^n$ es **ortogonal a un subespacio $W \subset \mathbb{R}^n$** si y solo si es ortogonal a todos los vectores que generan²⁵ el subespacio.

Demostración 6.7.

Sea $\vec{v} \in \mathbb{R}^n$ y un subespacio $W \subset \mathbb{R}^n$ de dimensión $\dim(W) = k$ con una base cualquiera $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_k\}$. Todo vector $\vec{w} \in W$ se puede expresar entonces como:

$$\vec{w} = c_1\vec{u}_1 + \dots + c_k\vec{u}_k$$

Cualquier vector $\vec{v} \in \mathbb{R}^n$ será ortogonal a cualquier $\vec{w} \in W$ si:

$$\vec{v} \cdot \vec{w} = \vec{v} \cdot (c_1\vec{u}_1 + \dots + c_k\vec{u}_k) = c_1(\vec{v} \cdot \vec{u}_1) + \dots + c_k(\vec{v} \cdot \vec{u}_k)$$

De forma que: $\vec{v} \cdot \vec{u}_i = 0$; $1 \leq i \leq k$ $\iff$ $\vec{v} \cdot \vec{w} = 0$

²³Ver la Sección 6.3.4, página 172.

²⁴Ver la Sección B.3.2, página 288.

²⁵Es decir, a cualquier *conjunto generador* y más concretamente, a cualquiera de sus *bases*.

Ejercicio 6.5. Determinar si el vector $\vec{v} = [-1, -1, -1] \in \mathbb{R}^3$ es ortogonal al subespacio $W \subset \mathbb{R}^3$ definido por:

$$W : x + y + z = 0$$

Vamos a encontrar primero una base de W . Expresando W en forma vectorial:

$$\begin{cases} x = t \\ y = s \\ z = -t - s \end{cases} \implies \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \implies \mathcal{B}_W = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$$

Claramente podemos observar que el vector $\vec{v}$ es ortogonal a los dos vectores de la base y por tanto es ortogonal a W . Por otra parte, desde un punto de vista geométrico, es fácil comprobar que $\vec{v}$ es paralelo al vector normal al plano W , $\vec{n} = [1, 1, 1]$ y por tanto es perpendicular (ortogonal) a todo vector del plano.

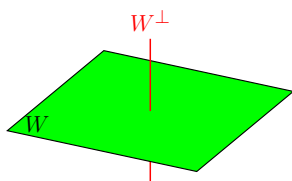
6.4.2. Complemento Ortogonal

Definición 6.13. Complemento Ortogonal

Sea W un subespacio de $\mathbb{R}^n$. El conjunto de todos los vectores $\vec{v} \in \mathbb{R}^n$ ortogonales²⁶ a W se denomina **complemento ortogonal de W** , y se denota por $W^\perp$.

$$W^\perp = \{\vec{v} \in \mathbb{R}^n : \vec{v} \cdot \vec{w} = 0 \quad ; \quad \forall \vec{w} \in W\}$$

Ejemplo 6.5. Complemento ortogonal



Sea W un plano en $\mathbb{R}^3$ que pasa por el origen²⁷. Una recta perpendicular al plano (es decir, paralela al vector normal²⁸ al plano), que pase por el origen será el **complemento ortogonal** de W , es decir $W^\perp$, donde todo vector de la recta $\vec{v} \in W^\perp$ es perpendicular a todo vector del plano $\vec{w} \in W$ y viceversa.

6.4.3. Características del Complemento Ortogonal

Sea W un subespacio de $\mathbb{R}^n$. Entonces:

- $W^\perp$ es un subespacio de $\mathbb{R}^n$
- $(W^\perp)^\perp = W$
- $W \cap W^\perp = \{\vec{0}\}$
- Si²⁹ $W = \text{gen}(\vec{w}_1, \dots, \vec{w}_k)$, entonces $\vec{v} \in W^\perp$ sii³⁰ $\vec{v} \cdot \vec{w}_i = 0, \quad \forall i = 1, \dots, k$

6.4.4. Cálculo del Complemento Ortogonal

Para determinar $W^\perp$ basta con encontrar todos los vectores que son ortogonales a todos los vectores de una base cualquiera de W .

$$\mathcal{B}_W = \{\vec{w}_1, \dots, \vec{w}_k\} \implies \vec{v} \in W^\perp \text{ sii } \vec{v} \cdot \vec{w}_i = 0, \quad \forall i = 1, \dots, k$$

Nota 6.12. En la Sección 6.4.5, página 175 se verá una forma alternativa de encontrar el complemento ortogonal de un subespacio vectorial, basado en las relaciones de ortogonalidad de los subespacios fundamentales de una matriz.

²⁶Ver la Sección 6.4.1, página 173.

²⁷El plano W es un subespacio vectorial de $\mathbb{R}^3$, (ver la Definición 2.13, página 71).

²⁸Ver el Tema 0, página 1.

²⁹Ver la Sección 2.5, página 79.

³⁰El acrónimo sii significa si y solo si.

Ejercicio 6.6. Dado el siguiente subespacio vectorial $W \subset \mathbb{R}^4$, encontrar una base del complemento ortogonal, $W^\perp$.

$$W : \begin{cases} x_1 + x_2 = 0 \\ x_2 + x_4 = 0 \end{cases}$$

Claramente $\dim(W) = 4 - 2 = 2$ y por tanto cualquiera de sus bases tienen dos vectores. Vamos a encontrar una de ellas resolviendo el sistema:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \rightarrow \begin{cases} x_1 = -x_2 = s \\ x_2 = -x_4 = -s \\ x_3 = t \\ x_4 = s \end{cases} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Por tanto una base del subespacio vectorial W es $\mathcal{B}_W = \{\vec{w}_1, \vec{w}_2\} = \{[0, 0, 1, 0], [1, -1, 0, 1]\}$. Los vectores $\vec{v} = [x_1, x_2, x_3, x_4] \in W^\perp$, deben ser ortogonales a los vectores de la base:

$$\begin{cases} \vec{w}_1 \cdot \vec{v} = 0 \\ \vec{w}_2 \cdot \vec{v} = 0 \end{cases} \Rightarrow \begin{cases} x_3 = 0 \\ x_1 - x_2 + x_4 = 0 \end{cases}$$

Resolviendo este nuevo SEL obtenemos el subespacio $W^\perp$:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix} \rightarrow \begin{cases} x_1 = x_2 - x_4 = t - s \\ x_2 = t \\ x_3 = 0 \\ x_4 = s \end{cases} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Una base del complemento ortogonal de W será por tanto:

$$\mathcal{B}_{W^\perp} = \{[1, 1, 0, 0], [-1, 0, 0, 1]\}$$

6.4.5. Relaciones de Ortogonalidad entre los Subespacios de una Matriz Real

Una matriz $A \in \mathbb{R}^{m \times n}$ tiene asociados cuatro *subespacios fundamentales*³¹:

- $\text{fil}(A), \ker(A) \subseteq \mathbb{R}^n$
- $\text{col}(A), \ker(A^T) \subseteq \mathbb{R}^m$

Teorema 6.8. Ortogonalidad en los Subespacios de una Matriz Real

Sea $A \in \mathbb{R}^{m \times n}$. Entonces³²:

$$[\text{fil}(A)]^\perp = \ker(A) \quad ; \quad [\text{col}(A)]^\perp = [\text{im}(A)]^\perp = \ker(A^T)$$

Demostración 6.8.

- Sea $\vec{x} \in \mathbb{R}^n$, entonces $\vec{x} \in (\text{fil}(A))^\perp$ sii es ortogonal a cada fila de A , y esto es lo mismo que $A\vec{x} = \vec{0}$ y por tanto $\vec{x} \in \ker(A)$.
- Igual que antes, sustituyendo A por A^T y teniendo en cuenta que $\text{fil}(A^T) = \text{col}(A)$.

³¹Ver las definiciones de $\text{fil}(A)$, $\text{col}(A)$ y $\ker(A)$ en la [Sección 3.10, página 111](#).

³²Debido a las propiedades del complemento ortogonal, también se pueden expresar como:

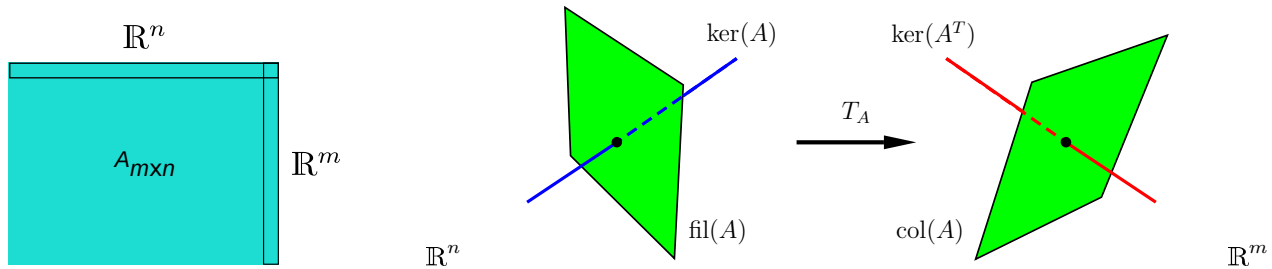
$$[\ker(A)]^\perp = \text{fil}(A) \quad ; \quad [\ker(A^T)]^\perp = \text{im}(A) = \text{col}(A)$$

En el [Teorema 6.11, página 189](#) se verá una generalización de este [Teorema 6.8](#) para matrices complejas.

Toda matriz $A \in \mathbb{R}^{m \times n}$ define una TL: $T : X \mapsto Y$

- El dominio de T es $X = \mathbb{R}^n$.
- La imagen de T es $Y = \text{im}(T) = \text{im}(A) = \text{col}(A)$, donde $\text{im}(T) \subseteq \mathbb{R}^m$.
- T envía $\ker(A) \subseteq X$ a $\vec{0} \in Y$.
- El complemento ortogonal de la imagen de T es $\ker(A^T) \subseteq \mathbb{R}^m$.

Esquemáticamente podemos relacionar estos conceptos de la siguiente forma:



Cuando se define un subespacio vectorial $W \subseteq \mathbb{R}^m$ a través de un SEL homogéneo de la forma:

$$W : \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \text{ en } \mathbb{R}^n ; \quad \boxed{A\vec{x} = \vec{0}}$$

El conjunto de vectores $\vec{x} = [x_1, \dots, x_n]$ que forman W coinciden con el núcleo de la matriz A asociada al SEL. Por el **Teorema 6.8, página 175** se tiene:

$$W = \ker(A) \implies W^\perp = \text{fil}(A)$$

Por tanto para encontrar una base de $W^\perp$ basta con determinar³³ una base del *espacio fila* de A .

Ejercicio 6.7. Resolver el anterior **Ejemplo 6.6** teniendo en cuenta el **Teorema 6.8**.

La definición del subespacio W la podemos definir matricialmente como:

$$W : \begin{cases} x_1 + x_2 = 0 \\ x_2 + x_4 = 0 \end{cases} \implies \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Por tanto, el subespacio W corresponde con el núcleo de la matriz:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Como hemos visto en el **Teorema 6.8**: $[\ker(A)]^\perp = \text{fil}(A)$, y como las dos filas de la matriz A son linealmente independientes se tiene:

$$\mathcal{B}_{\text{fil}(A)} = \{[1, 1, 0, 0], [0, 1, 0, 1]\} \implies \boxed{\mathcal{B}_{W^\perp} = \{[1, 1, 0, 0], [0, 1, 0, 1]\}}$$

Notar que esta solución no es igual que la obtenida en el **Ejemplo 6.6** aunque ambas son perfectamente válidas y corresponden a bases distintas del complemento ortogonal $W^\perp$.

³³Ver la **Sección 3.10.6, página 117**.

Esta segunda forma de calcular el complemento ortogonal de un subespacio vectorial, utilizando el **Teorema 6.8**, es mucho más rápida y práctica que la primera, ya que no necesita encontrar previamente una base de W .

Ejercicio 6.8. Sea W el subespacio de $\mathbb{R}^5$ generado por los vectores:

$$\vec{w}_1 = \begin{bmatrix} 1 \\ -3 \\ 5 \\ 0 \\ 5 \end{bmatrix} \quad ; \quad \vec{w}_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \\ -2 \\ 3 \end{bmatrix} \quad ; \quad \vec{w}_3 = \begin{bmatrix} 0 \\ -1 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Calcular una base de $W^\perp$.

Resolución de forma matricial. El espacio generado por los vectores $\vec{w}_1$, $\vec{w}_2$ y $\vec{w}_3$ es el mismo que $\text{col}(A)$ de:

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -3 & 1 & -1 \\ 5 & 2 & 4 \\ 0 & -2 & -1 \\ 5 & 3 & 5 \end{bmatrix}$$

Del **Teorema 6.8**: $W^\perp = (\text{col}(A))^\perp = \ker(A^T)$

$$[A^T | \vec{0}] = \left[\begin{array}{ccccc|c} 1 & -3 & 5 & 0 & 5 & 0 \\ -1 & 1 & 2 & -2 & 3 & 0 \\ 0 & -1 & 4 & -1 & 5 & 0 \end{array} \right] \longrightarrow \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 3 & 4 & 0 \\ 0 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \end{array} \right]$$

Por tanto, $\vec{y} \in W^\perp$ será $y_5 = t$, $y_4 = s$, $y_3 = -2t$, $y_2 = -3t - s$, $y_1 = -4t - 3s$. Finalmente:

$$W^\perp = \left\{ t \begin{bmatrix} -4 \\ -3 \\ -2 \\ 0 \\ 1 \end{bmatrix} + s \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} : t, s \in \mathbb{R} \right\} = \text{gen} \left(\begin{bmatrix} -4 \\ -3 \\ -2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right)$$

6.4.6. Dimensión del Complemento Ortogonal

- Sea un subespacio $W \subseteq \mathbb{R}^n$, entonces³⁴:

$$\dim(W) + \dim(W^\perp) = n$$

- En toda matriz $A \in \mathbb{R}^{m \times n}$, los subespacios vectoriales $\text{fil}(A)$ y $\text{col}(A)$ cumplen la relación anterior:

$$\begin{cases} \text{fil}(A) : & \text{rang}(A) + \text{nul}(A) = n \\ \text{col}(A) : & \text{rang}(A) + \text{nul}(A^T) = m \end{cases}$$

Nota 6.13. $\text{rang}(A) = \dim[\text{fil}(A)] = \dim[\text{col}(A)] = \dim[\text{im}(A)]$.

³⁴Notar que esto no quiere decir que se *rellene* todo el espacio $\mathbb{R}^n$ con W y $W^\perp$, sino que la suma de sus dimensiones tiene que ser igual a n .

Ejemplo 6.6. Dimensión del complemento ortogonal.

- En el **Ejemplo 6.6**, con respecto al subespacio $W \subset \mathbb{R}^4$, se tiene:

$$\dim(W) = \dim(W^\perp) = 2$$

cumpléndose la ecuación anterior, ya que:

$$\dim(W) + \dim(W^\perp) = 2 + 2 = 4 = n$$

- En el **Ejemplo 6.8**, con respecto al subespacio $W \subset \mathbb{R}^5$, se tiene:

$$\dim(W) = 3 \quad ; \quad \dim(W^\perp) = 2$$

cumpléndose igualmente la ecuación anterior:

$$\dim(W) + \dim(W^\perp) = 3 + 2 = 5 = n$$

6.5. Proyección Ortogonal

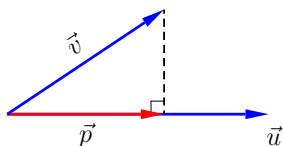
Definición 6.14. Proyección Ortogonal

Dados dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$ se denomina **proyección ortogonal**³⁵ de $\vec{v}$ sobre $\vec{u}$ al vector de $\mathbb{R}^n$ resultado de la siguiente operación:

$$\text{proy}_{\vec{u}}(\vec{v}) = \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \vec{u} \quad \left\{ \begin{array}{l} \vec{u} \cdot \vec{v} \in \mathbb{R} \\ \vec{u} \cdot \vec{u} \in \mathbb{R} \\ \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \in \mathbb{R} \\ \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \vec{u} \in \mathbb{R}^n \quad (\text{prod. por un escalar}) \end{array} \right.$$

6.5.1. Interpretación Geométrica en $\mathbb{R}^2$ y $\mathbb{R}^3$

La *proyección ortogonal* de un vector $\vec{v}$ sobre otro vector $\vec{u}$, $\text{proy}_{\vec{u}}(\vec{v})$, coincide geoméricamente con el vector *sombra* de $\vec{v}$ sobre $\vec{u}$.



Denominando $\vec{p}$ al vector *sombra* de $\vec{v}$ sobre $\vec{u}$, se cumple:

$$\vec{p} = \|\vec{p}\| \frac{1}{\|\vec{u}\|} \vec{u}$$

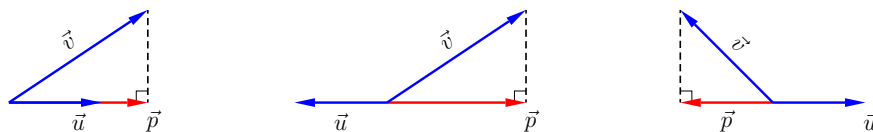
Por trigonometría: $\|\vec{p}\| = \|\vec{v}\| \cos \theta = \|\vec{v}\| \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|} = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|}$

Finalmente:

$$\vec{p} = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|} \frac{\vec{u}}{\|\vec{u}\|} = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|^2} \vec{u} = \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \vec{u}$$

³⁵La palabra *proyección* proviene del latín *proiectio*, de *proficere*; de *pro*=delante y *facere*=hacer. Desde de un punto de vista geométrico, la proyección es la representación gráfica de un objeto sobre una superficie plana, obtenida al unir las intersecciones sobre dicho plano de las líneas proyectantes de todos los puntos del objeto desde un determinado vértice. La proyección ortogonal es un tipo de proyección perpendicular.

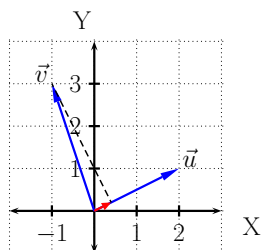
El vector proyección ortogonal, $\vec{p} = \text{proy}_{\vec{u}}(\vec{v})$, siempre tiene la *dirección* del vector $\vec{u}$ y su *sentido* viene dado por el vector $\vec{v}$.



Nota 6.14. En gráficos 3D las proyecciones ortogonales son muy útiles para determinar sombras, reflexiones, etc.

Ejercicio 6.9. En $\mathbb{R}^2$ encontrar la proyección ortogonal de $\vec{v} = [-1, 3]$ sobre $\vec{u} = [2, 1]$.

$$\text{proy}_{\vec{u}}(\vec{v}) = \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \vec{u} = \begin{cases} \vec{u} \cdot \vec{v} = [2, 1] \cdot [-1, 3] = 2(-1) + 1(3) = -2 + 3 = 1 \\ \vec{u} \cdot \vec{u} = [2, 1] \cdot [2, 1] = 2(2) + 1(1) = 4 + 1 = 5 \end{cases}$$



$$\text{proy}_{\vec{u}}(\vec{v}) = \frac{1}{5}[2, 1] = \left[\frac{2}{5}, \frac{1}{5} \right]$$

Ejercicio 6.10. De forma geométrica y por el teorema de Pitágoras es evidente que $\|\text{proy}_{\vec{u}}(\vec{v})\| \leq \|\vec{v}\|$. Demostrar que esta expresión es equivalente a la *desigualdad de Cauchy-Schwarz*³⁶: $|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \|\vec{v}\|$.

$$\|\text{proy}_{\vec{u}}(\vec{v})\| = \left\| \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \vec{u} \right\| = \left| \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right| \|\vec{u}\| = \left| \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|^2} \right| \|\vec{u}\| = \frac{|\vec{u} \cdot \vec{v}|}{\|\vec{u}\|} \|\vec{u}\| = \frac{|\vec{u} \cdot \vec{v}|}{\|\vec{u}\|} \implies \frac{|\vec{u} \cdot \vec{v}|}{\|\vec{u}\|} \leq \|\vec{v}\|$$

Ejercicio 6.11. En $\mathbb{R}^3$, encontrar un vector $\vec{v}$ cuya proyección ortogonal sobre el vector $\vec{u} = [1, -1, 1]$ es el vector $\vec{p} = [2, -2, 2]$.

Teniendo en cuenta la definición de proyección ortogonal se tiene:

$$\vec{p} = \text{proy}_{\vec{u}}(\vec{v}) = \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \vec{u} \implies (\vec{u} \cdot \vec{u})\vec{p} = (\vec{u} \cdot \vec{v})\vec{u}$$

Teniendo en cuenta que:

$$\vec{u} \cdot \vec{u} = [1, -1, 1] \cdot [1, -1, 1] = 3 \quad ; \quad (\vec{u} \cdot \vec{u})\vec{p} = 3[2, -2, 2] = [6, -6, 6]$$

y suponiendo que las coordenadas del vector $\vec{v}$ son $[x, y, z]$, (valores que queremos conocer):

$$\vec{u} \cdot \vec{v} = [1, -1, 1] \cdot [x, y, z] = x - y + z \quad ; \quad (\vec{u} \cdot \vec{v})\vec{u} = (x - y + z)[1, -1, 1] = [x - y + z, -x + y - z, x - y + z]$$

Por tanto:

$$(\vec{u} \cdot \vec{u})\vec{p} = (\vec{u} \cdot \vec{v})\vec{u} \quad ; \quad [6, -6, 6] = [x - y + z, -x + y - z, x - y + z]$$

Donde tenemos que las coordenadas de $\vec{v}$ deben cumplir la ecuación³⁷:

$$x - y + z = 6$$

Por tanto existen infinitos vectores $\vec{v}$ posibles, que están contenidos en el plano anterior. Una posible solución sería por ejemplo $\vec{v} = [1, 0, 5]$.

³⁶Ver la Sección 6.9.2, página 188.

Ejercicio 6.12. Encontrar una base del conjunto de vectores cuya proyección ortogonal sobre el vector $\vec{u} = [1, -1, 0] \in \mathbb{R}^3$, es ortogonal al vector $\vec{w} = [0, 1, 1]$.

Sea la solución buscada los vectores $\vec{v} = [x, y, z] \in \mathbb{R}^3$, entonces la proyección ortogonal de $\vec{v}$ sobre $\vec{u} = [1, -1, 0]$ será el vector:

$$\text{proy}_{\vec{u}}(\vec{v}) = \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \vec{u} = \frac{[1, -1, 0] \cdot [x, y, z]}{[1, -1, 0] \cdot [1, -1, 0]} [1, -1, 0] = \frac{x-y}{2} [1, -1, 0]$$

Para que estos vectores sean ortogonales al vector $\vec{w} = [0, 1, 1]$ se tiene que cumplir:

$$\frac{x-y}{2} [1, -1, 0] \cdot [0, 1, 1] = \frac{y-x}{2} = 0 \implies \boxed{x=y}$$

Por tanto el conjunto $S \subset \mathbb{R}^3$ de vectores $\vec{v}$ que verifican esa condición son:

$$S = \{[x, x, z] \in \mathbb{R}^3 : x, z \in \mathbb{R}\}$$

Como la ecuación que define este conjunto de vectores $x - y = 0$ es un plano de $\mathbb{R}^3$, una base de este conjunto serán dos vectores cualesquiera de ese plano que no sean paralelos. Por ejemplo:

$$\mathcal{B}_S = \{[1, 1, 0], [0, 0, 1]\}$$

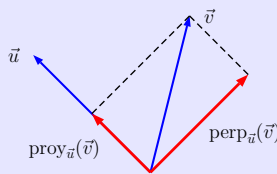
Nota 6.15. El Ejemplo 6.12 también se puede resolver mediante una interpretación geométrica sencilla: como la proyección de la solución $\vec{v}$ sobre $\vec{u}$ son vectores con dirección $\vec{u}$, todos los vectores ortogonales a este vector formarán un plano con vector normal $\vec{u}$, y por tanto de ecuación general $x - y = 0$. Una base de este subespacio (plano) serán dos vectores contenidos en el plano linealmente independientes, por ejemplo $[1, 1, 0]$ y $[0, 0, 2]$.

Nota 6.16. En el Apéndice B.2, página 284 se pueden ver algunas aplicaciones geométricas de la *proyección ortogonal*.

6.5.2. Vector Perpendicular

Definición 6.15. Vector Perpendicular

Dados dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$, se define el **vector perpendicular** a $\vec{u}$ según el vector $\vec{v}$ como:



$$\text{perp}_{\vec{u}}(\vec{v}) = \vec{v} - \text{proy}_{\vec{u}}(\vec{v}) = \vec{v} - \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \right) \vec{u}$$

Este vector es *ortogonal* a $\vec{u}$ con un tamaño *proporcional* a $\vec{v}$.

Todo vector $\vec{v} \in \mathbb{R}^n$ se puede definir en función de cualquier otro vector $\vec{u} \in \mathbb{R}^n$, no paralelo a $\vec{v}$, de la siguiente forma:

$$\vec{v} = \text{proy}_{\vec{u}}(\vec{v}) + \text{perp}_{\vec{u}}(\vec{v})$$

³⁷Obtenemos un sistema de tres ecuaciones linealmente dependientes y solo una de ellas da información útil.

Ejercicio 6.13. En $\mathbb{R}^2$ encontrar el vector perpendicular a $\vec{u} = [2, 1]$ según el vector $\vec{v} = [-1, 3]$ y representar gráficamente el resultado.

El vector pedido será:

$$\text{perp}_{\vec{u}}(\vec{v}) = \vec{v} - \text{proy}_{\vec{u}}(\vec{v})$$

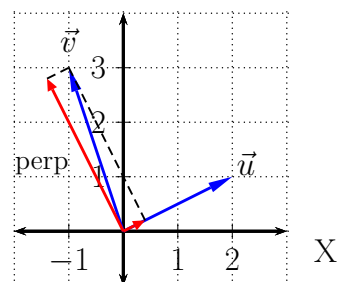
El vector perpendicular ya se calculó en el [Ejemplo 6.9](#), siendo:

$$\text{proy}_{\vec{u}}(\vec{v}) = \left[\frac{2}{5}, \frac{1}{5} \right]$$

Por tanto:

$$\text{perp}_{\vec{u}}(\vec{v}) = \vec{v} - \text{proy}_{\vec{u}}(\vec{v}) = [-1, 3] - [2/5, 1/5] = \boxed{[-7/5, 14/5]}$$

En la figura de la derecha se muestran una representación gráfica del problema y los resultados.



6.5.3. Proyección Ortogonal sobre un Subespacio

Definición 6.16. Proyección Ortogonal sobre un Subespacio

Sea un subespacio $W \subseteq \mathbb{R}^n$ y $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_k\}$ una *base ortogonal*³⁸ de W .

Entonces, para cualquier vector $\vec{v} \in \mathbb{R}^n$ se define la **proyección ortogonal**³⁹ de $\vec{v}$ sobre W como:

$$\text{proy}_W(\vec{v}) = \text{proy}_{\vec{u}_1}(\vec{v}) + \dots + \text{proy}_{\vec{u}_k}(\vec{v}) = \left(\frac{\vec{u}_1 \cdot \vec{v}}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 + \dots + \left(\frac{\vec{u}_k \cdot \vec{v}}{\vec{u}_k \cdot \vec{u}_k} \right) \vec{u}_k$$

Nota 6.17. Si la base $\mathcal{B}$ es ortonormal $\implies \vec{u}_i \cdot \vec{u}_i = 1$.

$$\text{proy}_W(\vec{v}) = \text{proy}_{\vec{u}_1}(\vec{v}) + \dots + \text{proy}_{\vec{u}_k}(\vec{v}) = \left(\frac{\vec{u}_1 \cdot \vec{v}}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 + \dots + \left(\frac{\vec{u}_k \cdot \vec{v}}{\vec{u}_k \cdot \vec{u}_k} \right) \vec{u}_k = (\vec{u}_1 \cdot \vec{v}) \vec{u}_1 + \dots + (\vec{u}_k \cdot \vec{v}) \vec{u}_k$$

Definición 6.17. Vector Perpendicular a un Subespacio

Sea un subespacio $W \subseteq \mathbb{R}^n$. Para cualquier vector no nulo $\vec{v} \in \mathbb{R}^n$ se define el **vector perpendicular**⁴⁰ de $\vec{v}$ sobre W al vector *proyección ortogonal* de $\vec{v}$ sobre el *complemento ortogonal* de W :

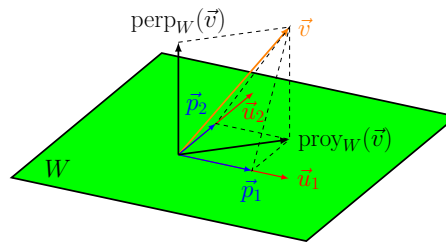
$$\text{perp}_W(\vec{v}) = \text{proy}_{W^\perp}(\vec{v}) = \vec{v} - \text{proy}_W(\vec{v})$$

³⁸La base debe ser obligatoriamente ortogonal para que la fórmula sea válida.

³⁹Ver la [Sección 6.5, página 178](#).

⁴⁰Ver la [Sección 6.5.2, página 180](#).

Interpretación Geométrica



En la figura superior se muestra la *proyección ortogonal* de un vector $\vec{v} \in \mathbb{R}^3$ (en naranja) sobre un subespacio $W \subset \mathbb{R}^3$ de dimensión 2 (plano verde), con *base ortogonal* $\mathcal{B}_{oW} = \{\vec{u}_1, \vec{u}_2\}$ (vectores rojos).

Los vectores (azules) $\vec{p}_1 = \text{proy}_{\vec{u}_1}(\vec{v})$ y $\vec{p}_2 = \text{proy}_{\vec{u}_2}(\vec{v})$ son las *proyecciones ortogonales* de $\vec{v}$ sobre cada uno de los vectores de la *base ortogonal* de W . Gráficamente podemos comprobar como la suma de los dos vectores azules es igual al vector negro $\text{proy}_W(\vec{v})$ correspondiente a la proyección ortogonal de $\vec{v}$ sobre W : $\text{proy}_W(\vec{v}) = \vec{p}_1 + \vec{p}_2$. Ortonogonal al plano (en el $W^\perp$) está el vector $\text{perp}_W(\vec{v})$, cumpliéndose: $\text{proy}_W(\vec{v}) + \text{perp}_W(\vec{v}) = \vec{v}$.

Ejercicio 6.14. En $\mathbb{R}^3$ sea el vector $\vec{v} = [3, -1, 2]$ y el plano W definido por:

$$W : x - y + 2z = 0$$

Encontrar la proyección ortogonal de $\vec{v}$ sobre W y la componente de $\vec{v}$ ortogonal a W .

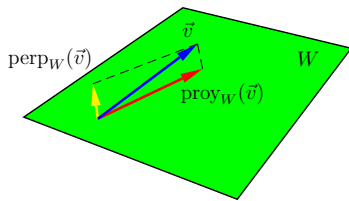
Podemos resolverlo de dos formas:

- Forma 1: Utilizando razonamientos geométricos.
- Forma 2: Utilizando la fórmula de proyección ortogonal sobre subespacios.

Veamos a continuación dichas formas de resolución.

Forma 1

El vector $\text{perp}_W(\vec{v})$ lo podemos calcular como la proyección ortogonal del vector $\vec{v}$ sobre un vector normal al plano $\vec{n}$, $\text{perp}_W(\vec{v}) = \text{proy}_{\vec{n}}(\vec{v})$. Basándonos en la ecuación del plano, $\vec{n} = [1, -1, 2]$.



$$\text{proy}_W(\vec{v}) = \vec{v} + [-\text{perp}_W(\vec{v})] = \vec{v} - \text{perp}_W(\vec{v})$$

$$\text{perp}_W(\vec{v}) = \text{proy}_{\vec{n}}(\vec{v}) = \frac{\vec{n} \cdot \vec{v}}{\vec{n} \cdot \vec{n}} \vec{n} = \frac{[1, -1, 2] \cdot [3, -1, 2]}{[1, -1, 2] \cdot [1, -1, 2]} [1, -1, 2] = \frac{4}{3} [1, -1, 2]$$

$$\text{proy}_W(\vec{v}) = \vec{v} - \text{perp}_W(\vec{v}) = [3, -1, 2] - \frac{4}{3} [1, -1, 2] = \frac{1}{3} [5, 1, -2]$$

Forma 2

En el **Ejemplo 6.3** se encontró una base ortogonal a este plano, $B = \{\vec{u}_1, \vec{u}_2\}$ con $\vec{u}_1 = [1, 1, 0]$ y $\vec{u}_2 = [1, -1, -1]$.

$$\text{proy}_W(\vec{v}) = \vec{p}_1 + \vec{p}_2 = \text{proy}_{\vec{u}_1}(\vec{v}) + \text{proy}_{\vec{u}_2}(\vec{v}) = \left(\frac{\vec{u}_1 \cdot \vec{v}}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 + \left(\frac{\vec{u}_2 \cdot \vec{v}}{\vec{u}_2 \cdot \vec{u}_2} \right) \vec{u}_2 \quad ; \quad \text{perp}_W(\vec{v}) = \vec{v} - \text{proy}_W(\vec{v})$$

Por tanto:

$$\text{proy}_W(\vec{v}) = \frac{2}{2} [1, 1, 0] + \frac{2}{3} [1, -1, -1] = [1, 1, 0] + \frac{2}{3} [2, -2, -2] = \frac{1}{3} [5, 1, -2]$$

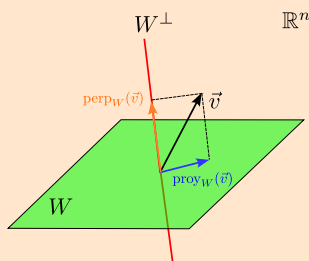
$$\text{perp}_W(\vec{v}) = [3, -1, 2] - \frac{1}{3} [5, 1, -2] = \frac{4}{3} [1, -1, 2]$$

Notar que el vector $\text{perp}_W(\vec{v})$ es paralelo al vector normal al plano $\vec{n} = [1, -1, 2]$.

6.6. Teorema de la Descomposición Ortogonal

Teorema 6.9. Teorema de la Descomposición Ortogonal

Sea $W \subseteq \mathbb{R}^n$ y sea un vector $\vec{v}$ cualquiera de $\mathbb{R}^n$. Entonces existen vectores únicos $\vec{w} \in W$ y $\vec{w}^\perp \in W^\perp$ tales que:



$$\vec{v} = \vec{w} + \vec{w}^\perp = \text{proy}_W(\vec{v}) + \text{perp}_W(\vec{v})$$

Demostración 6.9.

$$\vec{w} + \vec{w}^\perp = \text{proy}_W(\vec{v}) + \text{perp}_W(\vec{v}) = \text{proy}_W(\vec{v}) + (\vec{v} - \text{proy}_W(\vec{v})) = \vec{v}$$

Ejercicio 6.15. En $\mathbb{R}^4$, encontrar la descomposición ortogonal del vector $\vec{v} = [1, 1, 1, 0]$ en el subespacio vectorial:

$$W : \begin{cases} x_1 + x_2 = 0 \\ x_1 + x_4 = 0 \end{cases}$$

Tenemos que encontrar la proyección ortogonal y el vector perpendicular de $\vec{v}$ sobre W . Para ello, previamente tenemos que encontrar una base ortogonal de W :

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \rightarrow \begin{cases} x_1 = -x_2 = -t_2 \\ x_2 = x_4 = t_2 \\ x_3 = t_1 \\ x_4 = t_2 \end{cases} \rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t_1 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Claramente la base $\mathcal{B}_W = \{\vec{u}_1, \vec{u}_2\} = \{[0, 0, 1, 0], [-1, 1, 0, 1]\}$ ya es ortogonal. Calculamos ahora la proyección ortogonal de $\vec{v}$ sobre W .

$$\begin{aligned} \text{proy}_W(\vec{v}) &= \text{proy}_{\vec{u}_1}(\vec{v}) + \text{proy}_{\vec{u}_2}(\vec{v}) = \\ &= \frac{[0, 0, 1, 0] \cdot [1, 1, 1, 0]}{[0, 0, 1, 0] \cdot [0, 0, 1, 0]} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \frac{[-1, 1, 0, 1] \cdot [1, 1, 1, 0]}{[-1, 1, 0, 1] \cdot [-1, 1, 0, 1]} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{1} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \Rightarrow \vec{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

Ahora hay que calcular el vector perpendicular de $\vec{v}$ con respecto a W :

$$\text{perp}_W(\vec{v}) = \vec{v} - \text{proy}_W(\vec{v}) = [1, 1, 1, 0] - [0, 0, 1, 0] = [1, 1, 0, 0] \Rightarrow \vec{w}^\perp = [1, 1, 0, 0]$$

Por tanto, el vector $\vec{v}$ se puede descomponer en la suma del vector $\vec{w}$ que pertenece al subespacio vectorial W y del vector $\vec{w}^\perp$ que pertenece al complemento ortogonal de W .

$$\vec{v} = \vec{w} + \vec{w}^\perp \quad ; \quad \vec{w} \in W, \quad \vec{w}^\perp \in W^\perp$$

6.7. Proceso de Gram-Schmidt

El **proceso de Gram-Schmidt** es un método sencillo y sistemático para construir una *base ortogonal* (u *ortonormal* si se *normaliza*) a partir de cualquier base de un subespacio de $\mathbb{R}^n$.

Sea $\mathcal{B} = \{\vec{x}_1, \dots, \vec{x}_m\}$ una base de un $W \subseteq \mathbb{R}^n$ y definimos:

$$\begin{aligned}
 \vec{u}_1 &= \vec{x}_1 & W_1 &= \text{gen}(\vec{x}_1) = \text{gen}(\vec{u}_1) \\
 \vec{u}_2 &= \text{perp}_{W_1}(\vec{x}_2) = \vec{x}_2 - \text{proy}_{W_1}(\vec{x}_2) = \vec{x}_2 - \left(\frac{\vec{u}_1 \cdot \vec{x}_2}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 & W_2 &= \text{gen}(\vec{x}_1, \vec{x}_2) = \\
 & & &= \text{gen}(\vec{u}_1, \vec{u}_2) \\
 \vec{u}_3 &= \text{perp}_{W_2}(\vec{x}_3) = \vec{x}_3 - \text{proy}_{W_2}(\vec{x}_3) = \vec{x}_3 - \left(\frac{\vec{u}_1 \cdot \vec{x}_3}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 - \left(\frac{\vec{u}_2 \cdot \vec{x}_3}{\vec{u}_2 \cdot \vec{u}_2} \right) \vec{u}_2 & W_3 &= \text{gen}(\vec{x}_1, \vec{x}_2, \vec{x}_3) = \\
 & & &= \text{gen}(\vec{u}_1, \vec{u}_2, \vec{u}_3) \\
 & \vdots & & \vdots \\
 \vec{u}_m &= \text{perp}_{W_{m-1}}(\vec{x}_m) = \vec{x}_m - \text{proy}_{W_{m-1}}(\vec{x}_m) = \\
 & & &= \vec{x}_m - \left(\frac{\vec{u}_1 \cdot \vec{x}_m}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 - \dots - \left(\frac{\vec{u}_{m-1} \cdot \vec{x}_m}{\vec{u}_{m-1} \cdot \vec{u}_{m-1}} \right) \vec{u}_{m-1} & W &= W_m = \text{gen}(\vec{x}_1, \dots, \vec{x}_m) \\
 & & & & &= \text{gen}(\vec{u}_1, \dots, \vec{u}_m)
 \end{aligned}$$

Entonces para cada $i = 1, \dots, m$ el conjunto de vectores $\{\vec{u}_1, \dots, \vec{u}_i\}$ forma una *base ortogonal* del espacio W_i . En particular el conjunto $\{\vec{u}_1, \dots, \vec{u}_m\}$ es una *base ortogonal* de $W_m = W$.

Ejercicio 6.16. Encontrar una base ortogonal de $\mathbb{R}^3$ que contenga el vector $\vec{x}_1 = [1, 2, 3]$.

Para poder aplicar el proceso de Gram-Schmidt necesitamos partir de una base cualquiera, por tanto elegimos dos vectores $\vec{x}_2$ y $\vec{x}_3$ que no sean paralelos entre ellos ni entre $\vec{x}_1$. Por ejemplo $\vec{x}_2 = [0, 1, 0]$ y $\vec{x}_3 = [0, 0, 1]$.

Ahora aplicamos Gram-Schmidt a esta base:

$$\begin{aligned}
 \vec{u}_1 &= \vec{x}_1 & &= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\
 \vec{u}_2 &= \vec{x}_2 - \left(\frac{\vec{u}_1 \cdot \vec{x}_2}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1/7 \\ 5/7 \\ -3/7 \end{bmatrix} \implies \begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix} \\
 \vec{u}_3 &= \vec{x}_3 - \left(\frac{\vec{u}_1 \cdot \vec{x}_3}{\vec{u}_1 \cdot \vec{u}_1} \right) \vec{u}_1 - \left(\frac{\vec{u}_2 \cdot \vec{x}_3}{\vec{u}_2 \cdot \vec{u}_2} \right) \vec{u}_2 = \\
 &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{3}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{-3}{35} \begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix} = \begin{bmatrix} -3/10 \\ 0 \\ 1/10 \end{bmatrix} \implies \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}
 \end{aligned}$$

Por tanto, $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una posible base ortogonal para $\mathbb{R}^3$ que contiene al vector $[1, 2, 3]$. Normalizando estos vectores se podría obtener una base *ortonormal* de $\mathbb{R}^3$.

6.8. Factorización QR

Si $A \in \mathbb{R}^{m \times n}$ es una matriz con columnas linealmente independientes⁴¹ y se le aplica el proceso de Gram-Schmidt a los *vectores columnas*, se obtiene una importante *descomposición matricial*⁴² de A conocida con el nombre de **factorización QR**.

Teorema 6.10. Factorización QR

Toda matriz $A \in \mathbb{R}^{m \times n}$ con columnas linealmente independientes⁴³ puede ser factorizada como $A = QR$, donde $Q_{m \times n}$ es una matriz con columnas ortonormales⁴⁴ y $R_{n \times n}$ es una matriz triangular superior invertible⁴⁵, de la forma:

$$A_{m \times n} = Q_{m \times n} R_{n \times n}$$

Demostración 6.10.

Sean $\vec{c}_1, \dots, \vec{c}_n$ los vectores columna de la matriz $A_{m \times n}$ y sean $\vec{q}_1, \dots, \vec{q}_n$ los vectores ortonormales obtenidos al aplicar el proceso Gram-Schmidt a los vectores columna⁴⁶ de A y posterior normalización. Entonces, existen escalares α_{ij} y r_{ij} tales que:

$$\left\{ \begin{array}{l} \vec{q}_1 = \alpha_{11}\vec{c}_1 \\ \vec{q}_2 = \alpha_{12}\vec{c}_1 + \alpha_{22}\vec{c}_2 \\ \vec{q}_3 = \alpha_{13}\vec{c}_1 + \alpha_{23}\vec{c}_2 + \alpha_{33}\vec{c}_3 \\ \vdots \\ \vec{q}_n = \alpha_{1n}\vec{c}_1 + \alpha_{2n}\vec{c}_2 + \dots + \alpha_{nn}\vec{c}_n \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \vec{c}_1 = r_{11}\vec{q}_1 \\ \vec{c}_2 = r_{12}\vec{q}_1 + r_{22}\vec{q}_2 \\ \vec{c}_3 = r_{13}\vec{q}_1 + r_{23}\vec{q}_2 + r_{33}\vec{q}_3 \\ \vdots \\ \vec{c}_n = r_{1n}\vec{q}_1 + r_{2n}\vec{q}_2 + \dots + r_{nn}\vec{q}_n \end{array} \right.$$

Lo cual se puede expresar matricialmente como:

$$A = [\vec{c}_1 \quad \vec{c}_2 \quad \dots \quad \vec{c}_n] = [\vec{q}_1 \quad \vec{q}_2 \quad \dots \quad \vec{q}_n] \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix} = QR$$

La matriz R se puede obtener fácilmente una vez determinada Q , ya que al tener columnas ortogonales⁴⁷, $Q^T Q = I_n$ y por tanto:

$$A = QR \quad ; \quad Q^T A = Q^T QR \quad ; \quad R = Q^T A$$

La **Demostración 6.10** presenta de forma detallada el procedimiento de cálculo de la de la factorización QR de una matriz.

Nota 6.18.

- La factorización QR se puede extender a matrices arbitrarias (no necesariamente con columnas linealmente independientes), mediante una forma ligeramente modificada, utilizando **matrices de Householder**.
- Cuando el proceso de Gram-Schmidt se implementa en un ordenador, los errores de redondeo propios del cálculo, generan una pérdida de ortogonalidad de los vectores, que puede llegar a ser una importante causa de errores. En la práctica se suele utilizar una versión *modificada* de la factorización QR mostrada aquí, para minimizar estos errores de redondeo.

⁴¹Para lo cual es necesario que $n \leq m$.

⁴²Una *descomposición* de una matriz es una *factorización* de la misma en un producto de otras matrices, de la forma siguiente: $A = M_1 M_2 \dots M_k$.

⁴³Si A es *cuadrada* esto implica que es *invertible*.

⁴⁴La matriz Q será una *matriz ortogonal* si A , y por tanto Q , es cuadrada.

⁴⁵Lo que implica que sus *elementos diagonales* son no nulos.

⁴⁸Ver el **Tema 8, página 217**.

⁴⁹Ver la **Sección 7.4, página 204**.

La factorización QR de una matriz tiene multitud de aplicaciones en:

- Resolución de SEL de forma numérica eficiente.
- Aproximación numérica de *autovalores*⁴⁸.
- Aproximación del problema de *mínimos cuadrados*⁴⁹.

Ejercicio 6.17. Encontrar, en caso de ser posible, la factorización QR de la siguiente matriz:

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$$

Claramente los dos vectores columna de la matriz son linealmente independientes y por tanto sí es posible realizar la descomposición QR de la matriz. Vamos a seguir los pasos dados en la **Demostración 6.10**.

1. Ortogonalizar los vectores columna de la matriz mediante el proceso de Gram-Schmidt.

$$\begin{aligned} \vec{u}_1 &= \vec{v}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \\ \vec{u}_2 &= \vec{v}_2 - \frac{\vec{u}_1 \cdot \vec{v}_2}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \begin{bmatrix} -1 & 4 & 2 & 1 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 4 & 2 & 1 \end{bmatrix}}{\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} -\frac{5}{2} & \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \end{aligned}$$

2. Normalizando los vectores ortogonales obtenidos anteriormente construimos la matriz ortogonal Q :

$$Q = [\vec{u}_1 \quad \vec{u}_2] = \begin{bmatrix} 1/2 & -5/\sqrt{52} \\ 1/2 & 5/\sqrt{52} \\ 1/2 & 1/\sqrt{52} \\ 1/2 & -1/\sqrt{52} \end{bmatrix}$$

3. La matriz R se obtiene aplicando la fórmula $R = Q^T A$.

$$R = Q^T A = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -5/\sqrt{52} & 5/\sqrt{52} & 1/\sqrt{52} & -1/\sqrt{52} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & \sqrt{13} \end{bmatrix}$$

La matriz R obtenida es triangular superior y fácilmente se puede comprobar como $A = QR$.

Ejercicio 6.18. Encontrar, en caso de ser posible, la factorización QR de la matriz:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Como esta matriz tiene dos filas, como mucho existirán dos vectores de dos componentes linealmente independientes, y por tanto los tres vectores columna de A son *linealmente dependientes*.

En conclusión, esta matriz no tiene factorización QR .

⁴⁸Siempre en el mismo orden en que aparecen, es decir $\vec{q}_1 = \vec{c}_1$, $\vec{q}_2 = \text{perp}_{\vec{q}_1}(\vec{c}_2)$, etc. Ya que de otra forma, la matriz R no será necesariamente triangular superior.

⁴⁹No necesariamente Q es invertible. Solo si A es cuadrada, entonces Q es ortogonal y tiene sentido hablar de la inversa de Q , siendo además $Q^{-1} = Q^T$.

6.9. Ortogonalidad en Espacios $\mathbb{C}^n$

En esta sección vamos a extender los conceptos sobre *ortogonalidad* que se han visto en este tema en espacios $\mathbb{R}^n$ a espacios $\mathbb{C}^n$.

6.9.1. Producto Interno

El producto interno⁵⁰ es una generalización del concepto de *producto escalar* visto en la [Definición 6.1, página 163](#) para espacios vectoriales $\mathbb{R}^n$. Para ello empezaremos generalizando el concepto de producto escalar de vectores.

Definición 6.18. Producto Interno en $\mathbb{C}^n$

Dados dos vectores *columna* $\vec{u}, \vec{v} \in \mathbb{C}^n$, se define el **producto interno** de $\vec{u}$ por $\vec{v}$ y lo representamos por $\langle \vec{u}, \vec{v} \rangle$, al *escalar* resultante de la siguiente operación:

$$\langle \vec{u}, \vec{v} \rangle = \vec{u}^H \vec{v} = [\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \bar{u}_1 v_1 + \bar{u}_2 v_2 + \dots + \bar{u}_n v_n = \sum_{k=1}^n \bar{u}_k v_k \quad ; \quad \vec{u} \cdot \vec{v} \in \mathbb{C}$$

donde $\bar{u}_i$ corresponde con el complejo conjugado de la componente i -ésima del vector $\vec{u}$.

A través de la definición del *producto escalar* de dos vectores⁵¹, el **producto interno** de dos vectores $\mathbf{u}, \mathbf{v} \in \mathbb{C}$ se puede definir como el **producto escalar** del **conjugado del primer vector** por el segundo:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \bar{\mathbf{u}} \cdot \mathbf{v}$$

Nota 6.19. Si los vectores son reales, el *producto interno* coincide con el *producto escalar*: $\vec{u}, \vec{v} \in \mathbb{R}^n \implies \langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v}$

Ejercicio 6.19. Calcular el producto interno de los vectores de $\mathbb{C}^2$:

$$\vec{u} = [1 + i, 2 - i] \quad ; \quad \vec{v} = [2 - 3i, -1 - i]$$

$$\begin{aligned} \langle \vec{u}, \vec{v} \rangle &= \langle [1 + i, 2 - i], [2 - 3i, -1 - i] \rangle = \overline{[1 + i, 2 - i]} \cdot [2 - 3i, -1 - i] = \\ &= \overline{(1 + i)}(2 - 3i) + \overline{(2 - i)}(-1 - i) = (1 - i)(2 - 3i) + (2 + i)(-1 - i) = (-1 - 5i) + (-1 - 3i) = \boxed{-2 - 8i} \end{aligned}$$

Propiedades del Producto Interno

Dados los vectores $\vec{x}, \vec{y}, \vec{z} \in \mathbb{C}^n$ y los escalares $a, b \in \mathbb{C}$, se verifica:

- **Hermiticidad:** $\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}$
- **Linealidad**
 - Por la izquierda: $\langle a\vec{x} + b\vec{y}, \vec{z} \rangle = \bar{a} \langle \vec{x}, \vec{z} \rangle + \bar{b} \langle \vec{y}, \vec{z} \rangle$
 - Por la derecha: $\langle \vec{x}, a\vec{y} + b\vec{z} \rangle = a \langle \vec{x}, \vec{y} \rangle + b \langle \vec{x}, \vec{z} \rangle$
- **Definida positiva:** $\langle \vec{x}, \vec{x} \rangle \in \mathbb{R} \quad , \quad \langle \vec{x}, \vec{x} \rangle \geq 0$. Si $\langle \vec{x}, \vec{x} \rangle = 0 \implies \vec{x} = \vec{0}$.

⁵⁰También denominado *producto escalar* o *producto punto*.

⁵¹Ver la [Definición 6.1, página 163](#).

6.9.2. Norma en $\mathbb{C}^n$

Definición 6.19. Norma de un Vector en $\mathbb{C}^n$

Se define la **norma**⁵² de un vector $\mathbf{v} = [v_1, \dots, v_n] \in \mathbb{C}^n$ con $v_i = a_i + i b_i$, y se denota por $\|\mathbf{v}\|$, al número real no negativo dado por:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\bar{\mathbf{v}} \cdot \mathbf{v}} = \sqrt{\bar{v}_1 v_1 + \dots + \bar{v}_n v_n} = \sqrt{|v_1|^2 + \dots + |v_n|^2} = \sqrt{(a_1^2 + b_1^2) + \dots + (a_n^2 + b_n^2)}$$

donde $|v_i|$ representa el *módulo*⁵³ del número complejo v_i (componente i -ésima del vector $\vec{v}$).

Notar la siguiente equivalencia para todo vector complejo $\vec{v} = [v_1, \dots, v_n] = [a_1 + i b_1, \dots, a_n + i b_n]$.

$$\|\vec{v}\|^2 = \langle \vec{v}, \vec{v} \rangle = \sum_{i=1}^n (a_i^2 + b_i^2)$$

Propiedades de la Norma

Dado un vector $\vec{v} \in \mathbb{C}^n$ y un escalar $\alpha \in \mathbb{C}$, se verifica:

- $\|\vec{v}\| > 0$, $\forall \vec{v} \neq \vec{0}$; $\|\vec{0}\| = 0$
- $\|\alpha \vec{v}\| = |\alpha| \|\vec{v}\|$
- *Desigualdad de Cauchy-Schwarz* : $|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$
- *Desigualdad Triangular*: $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Definición 6.20. Vector Unitario

Se dice que un vector $\vec{v} \in \mathbb{C}^n$ es **unitario** si su norma es la unidad.

$$\vec{v} \text{ vector unitario} \iff \|\vec{v}\| = 1$$

Definición 6.21. Distancia entre Vectores en $\mathbb{C}^n$

Se define la distancia entre dos vectores $\vec{u}, \vec{v} \in \mathbb{C}^n$ al número real positivo dado por:

$$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \|\vec{v} - \vec{u}\|$$

6.9.3. Vectores Ortogonales en $\mathbb{C}^n$

Definición 6.22. Vectores Ortogonales en $\mathbb{C}^n$

Se dice que dos vectores $\vec{u}, \vec{v} \in \mathbb{C}^n$ son **ortogonales** si su *producto interno* es cero.

$$\langle \vec{u}, \vec{v} \rangle = 0 \iff \vec{u}, \vec{v} \text{ ortogonales}$$

⁵²Esta es una de las infinitas normas que se pueden definir a partir del producto interno, denominada *norma 2* o *euclídea*. En algunos textos también se le conoce con el nombre de *módulo* o *longitud del vector*.

⁵³Ver la [Definición C.13](#), página 303.

Los conceptos de ortogonalidad vistos en espacios $\mathbb{R}^n$ en la [Sección 6.3](#) y [Sección 6.4](#) se extienden fácilmente a los espacios $\mathbb{C}^n$ a través de la definición de ortogonalidad en espacios complejos, [Definición 6.22](#).

- Se dice que un **conjunto de vectores** $\{\vec{v}_1, \dots, \vec{v}_n\}$ con $\vec{v}_i \in \mathbb{C}^n$ es **ortogonal** si los vectores son *ortogonales* a pares, es decir, $\langle \vec{v}_i, \vec{v}_j \rangle = 0 \quad \forall i \neq j$.
- Se dice que un vector $\vec{v} \in \mathbb{C}^n$ es **ortogonal a un subespacio vectorial** $W \subseteq \mathbb{C}^n$ si $\vec{v}$ es ortogonal a todos los vectores de W , es decir, $\langle \vec{v}, \vec{w} \rangle = 0 \quad \forall \vec{w} \in W$.
- Para que un vector $\vec{v} \in \mathbb{C}^n$ sea **ortogonal a un subespacio** $W \subseteq \mathbb{C}^n$ basta con que sea ortogonal a todos los vectores de una base de W (extensión del [Teorema 6.7](#), [página 173](#)).

6.9.4. Complemento Ortogonal en $\mathbb{C}^n$

Definición 6.23. Complemento Ortogonal en $\mathbb{C}^n$

Sea W un subespacio de $\mathbb{C}^n$. El subespacio vectorial formado por todos los vectores $\vec{v} \in \mathbb{C}^n$ ortogonales a W se denomina **complemento ortogonal de W** , y se denota por $W^\perp$.

$$W^\perp = \{\vec{v} \in \mathbb{C}^n : \langle \vec{v}, \vec{w} \rangle = 0 \quad ; \quad \forall \vec{w} \in W\}$$

Cumpléndose que:

$$\dim(W) + \dim(W^\perp) = n$$

Todos los vectores de la una base $\mathcal{B}_W$ de un subespacio vectorial $W \subseteq \mathbb{C}^n$ son **ortogonales** a todos los vectores de una base $\mathcal{B}_{W^\perp}$ del complemento ortogonal de W , $W^\perp$.

Subespacios Fundamentales de una Matriz Compleja

Teorema 6.11. Ortogonalidad en los Subespacios de una Matriz Compleja

Una matriz $A \in \mathbb{C}^{m \times n}$ tiene asociados cuatro *subespacios fundamentales*⁵⁴:

- $\text{fil}(A), \overline{\ker(A)} \subseteq \mathbb{C}^n$
- $\text{col}(A), \ker(A^H) \subseteq \mathbb{C}^m$

Cumpléndose⁵⁵:

$$[\text{fil}(A)]^\perp = \overline{\ker(A)} \quad ; \quad [\text{col}(A)]^\perp = \ker(A^H)$$

donde $\overline{\ker(A)}$ es el subespacio vectorial formado por los vectores conjugados de $\ker(A)$.

$$\forall \mathbf{v} \in \ker(A) \quad \implies \quad \bar{\mathbf{v}} \in \overline{\ker(A)}$$

Nota 6.20. Este teorema es una generalización del [Teorema 6.8](#), [página 175](#).

⁵⁴Ver las definiciones de $\text{fil}(A)$, $\text{col}(A)$ y $\ker(A)$ en la [Sección 3.10](#), [página 111](#).

⁵⁵Debido a las propiedades del complemento ortogonal, también se pueden expresar como:

$$[\ker(A)]^\perp = \overline{\text{fil}(A)} \quad ; \quad [\ker(A^H)]^\perp = \text{im}(A) = \text{col}(A)$$

Ejercicio 6.20. Dado la siguiente matriz compleja, comprobar la relación de ortogonalidad de sus subespacios vectoriales fundamentales $\text{fil}(A)$, $\text{col}(A)$, $\ker(A)$, $\ker(A^H)$.

$$A = \begin{bmatrix} 1 & 0 & i & 1 \\ 0 & -i & 1 & 0 \\ 1 & -i & 1+i & 1 \end{bmatrix} \in \mathbb{C}^{3 \times 4}$$

Siguiendo el procedimiento descrito en la [Sección 3.10.6, página 117](#), encontramos bases de los cuatro subespacios vectoriales:

$$\begin{bmatrix} 1 & 0 & i & 1 \\ 0 & -i & 1 & 0 \\ 1 & -i & 1+i & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & i & 1 \\ 0 & -i & 1 & 0 \\ 0 & -i & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & i & 1 \\ 0 & -i & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{cases} x_1 = -ix_3 - x_4 = t - s \\ x_2 = t \\ x_3 = ix_2 = it \\ x_4 = s \end{cases} \quad \forall t, s \in \mathbb{C}$$

$$\mathcal{B}_{\text{fil}(A)} = \{[1, 0, i, 1], [0, -i, 1, 0]\} \quad ; \quad \mathcal{B}_{\text{col}(A)} = \{[1, 0, 1], [0, 1, 1]\} \quad ; \quad \mathcal{B}_{\ker(A)} = \{[1, 1, i, 0], [-1, 0, 0, 1]\}$$

Comprobamos la relación de ortogonalidad entre los vectores de $\mathcal{B}_{\text{fil}(A)}$ y $\overline{\mathcal{B}_{\ker(A)}}$, teniendo en cuenta que $\overline{\mathcal{B}_{\ker(A)}} = \{[1, 1, -i, 0], [-1, 0, 0, 1]\}$:

$$\begin{aligned} \langle [1, 0, i, 1], [1, 1, -i, 0] \rangle &= \overline{[1, 0, i, 1]} \cdot [1, 1, -i, 0] = [1, 0, -i, 1] \cdot [1, 1, -i, 0] = 1 - 1 = 0 \\ \langle [1, 0, i, 1], [-1, 0, 0, 1] \rangle &= \overline{[1, 0, i, 1]} \cdot [-1, 0, 0, 1] = [1, 0, -i, 1] \cdot [-1, 0, 0, 1] = -1 + 1 = 0 \\ \langle [0, -i, 1, 0], [1, 1, -i, 0] \rangle &= \overline{[0, -i, 1, 0]} \cdot [1, 1, -i, 0] = [0, i, 1, 0] \cdot [1, 1, -i, 0] = i - i = 0 \\ \langle [0, -i, 1, 0], [-1, 0, 0, 1] \rangle &= \overline{[0, -i, 1, 0]} \cdot [-1, 0, 0, 1] = [0, i, 1, 0] \cdot [-1, 0, 0, 1] = 0 \end{aligned}$$

Notar que $\dim[\text{fil}(A)] + \dim[\overline{\ker(A)}] = 2 + 2 = 4$, cumpliéndose la relación entre las dimensiones de los espacios vectoriales ortogonales, ya que $\text{fil}(A), \ker(A) \subset \mathbb{C}^4$.

Calculamos ahora una base de $\ker(A^H)$:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & i & i \\ -i & 1 & 1-i \\ 1 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & i & i \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{cases} x_1 = -x_3 = -t \\ x_2 = -x_3 = -t \\ x_3 = t \end{cases} \quad \forall t \in \mathbb{C}$$

$$\mathcal{B}_{\ker(A^H)} = \{[1, 1, -1]\}$$

Comprobamos la relación de ortogonalidad entre los vectores de $\mathcal{B}_{\text{col}(A)}$ y $\mathcal{B}_{\ker(A^H)}$:

$$\begin{aligned} \langle [1, 0, 1], [1, 1, -1] \rangle &= \overline{[1, 0, 1]} \cdot [1, 1, -1] = [1, 0, 1] \cdot [1, 1, -1] = 1 - 1 = 0 \\ \langle [0, 1, 1], [1, 1, -1] \rangle &= \overline{[0, 1, 1]} \cdot [1, 1, -1] = [0, 1, 1] \cdot [1, 1, -1] = 1 - 1 = 0 \end{aligned}$$

Notar que $\dim[\text{col}(A)] + \dim[\ker(A^H)] = 2 + 1 = 3$, cumpliéndose la relación entre las dimensiones de los espacios vectoriales ortogonales, ya que $\text{col}(A), \ker(A^H) \subset \mathbb{C}^3$.

Cálculo del Complemento Ortogonal de un Subespacio Vectorial Complejo

Dado un subespacio vectorial complejo $W \subseteq \mathbb{C}^n$ la determinación de su complemento ortogonal $W^\perp$ puede hacerse por dos procedimientos generales:

1. Encontrando primero una base cualquiera de W , $\mathcal{B}_W$ y posteriormente determinando los vectores que son ortogonales a todos los vectores de $\mathcal{B}_W$.

$$\mathcal{B}_W = \{\vec{u}_1, \dots, \vec{u}_k\} \implies \vec{v} \in W^\perp \text{ sii } \langle \vec{v}, \vec{u}_i \rangle = 0, \quad \forall i = 1, \dots, k$$

Esto dará lugar a un SEL homogéneo de $\mathbb{C}^n$ que define $W^\perp$, formado por tantas ecuaciones como dimensión k tenga W (número de vectores de cualquier base de W). Resolviendo dicho SEL se puede encontrar una base de $W^\perp$.

2. Si el subespacio vectorial $W \subseteq \mathbb{C}^m$ viene definido a través de un SEL homogéneo de la forma:

$$W : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0 \end{cases} ; \quad A\vec{x} = \vec{0}$$

El conjunto de vectores $\vec{x} = [x_1, \dots, x_n]$ que forman W coinciden con el núcleo de la matriz A asociada al SEL. Por el [Teorema 6.11, página 189](#) se tiene:

$$W = \ker(A) \implies W^\perp = \overline{\text{fil}(A)}$$

Por tanto, en este caso, para encontrar una base de $W^\perp$ basta con determinar (ver la [Sección 3.10.6, página 117](#)) una base del espacio fila de A y **conjugarla**.

Ejercicio 6.21. Dado el siguiente subespacio vectorial $W \subset \mathbb{C}^4$, encontrar una base de su complemento ortogonal, $W^\perp$.

$$W : \begin{cases} x_1 + ix_2 = 0 \\ x_2 - ix_3 + x_4 = 0 \end{cases}$$

Vamos a hacerlo por los dos procedimientos anteriores.

1. Encontramos primero una base de W , para ello resolvemos el sistema que define W :

$$\begin{bmatrix} 1 & i & 0 & 0 \\ 0 & 1 & -i & 1 \end{bmatrix} \rightarrow \begin{cases} x_1 = -ix_2 = -it \\ x_2 = t \\ x_3 = s \\ x_4 = -x_2 + ix_3 = -t + is \end{cases} \implies \mathcal{B}_W = \left\{ \begin{bmatrix} -i \\ 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ i \end{bmatrix} \right\}$$

Encontramos ahora $W^\perp$ determinando todos los vectores que son ortogonales (producto interno nulo) a los vectores de $\mathcal{B}_W$.

$$\begin{cases} \overline{[-i, 1, 0, -1]} \cdot [x_1, x_2, x_3, x_4] = ix_1 + x_2 - x_4 = 0 \\ \overline{[0, 0, 1, i]} \cdot [x_1, x_2, x_3, x_4] = x_3 - ix_4 = 0 \end{cases} \rightarrow \begin{cases} x_1 = t \\ x_2 = -ix_1 + x_4 = -it + s \\ x_3 = ix_4 = is \\ x_4 = s \end{cases}$$

Por tanto:

$$\mathcal{B}_{W^\perp} = \{[1, -i, 0, 0], [0, 1, i, 1]\}$$

Se puede comprobar como cada vector de la base $\mathcal{B}_W$ es ortogonal (producto intrerno nulo) a cada uno de los vectores de la base $\mathcal{B}_{W^\perp}$.

2. El SEL homogéneo que define W se puede poner de forma matricial:

$$\begin{bmatrix} 1 & i & 0 & 0 \\ 0 & 1 & -i & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad ; \quad A\vec{x} = \vec{0} \quad ; \quad A = \begin{bmatrix} 1 & i & 0 & 0 \\ 0 & 1 & -i & 1 \end{bmatrix}$$

Como $W = \ker(A)$ entonces $W^\perp = \overline{\text{fil}(A)}$. Como la matriz A ya está escalonada, vemos que los dos únicos vectores fila de A son linealmente independientes por lo que representan una base de $\text{fil}(A)$, por tanto calculando sus conjugados se tiene:

$$B_{W^\perp} = \{[1, -i, 0, 0], [0, 1, i, 1]\}$$

Notar que este segundo procedimiento es más sencillo que el primero, pudiéndose calcular una base de $W^\perp$ sin necesidad de calcular⁵⁶ W . Notar también que en este segundo caso se ha obtenido exactamente la misma base de $W^\perp$ del procedimiento 1, aunque de forma general no va ser éste el caso ya que existen infinitas bases de un subespacio vectorial.

6.9.5. Proyección Ortogonal en $\mathbb{C}^n$

Definición 6.24. Proyección Ortogonal en $\mathbb{C}^n$

Dados dos vectores $\vec{u}, \vec{v} \in \mathbb{C}^n$ se denomina **proyección ortogonal** de $\vec{v}$ sobre $\vec{u}$ al vector de $\mathbb{C}^n$ resultado de la siguiente operación:

$$\text{proy}_{\vec{u}}(\vec{v}) = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u}$$

Definición 6.25. Vector Perpendicular en $\mathbb{C}^n$

Dados $\vec{u}, \vec{v} \in \mathbb{C}^n$, se define el **vector perpendicular** a $\vec{u}$ según el vector $\vec{v}$ como:

$$\text{perp}_{\vec{u}}(\vec{v}) = \vec{v} - \text{proy}_{\vec{u}}(\vec{v}) = \vec{v} - \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u}$$

Verificándose:

$$\vec{v} = \text{proy}_{\vec{u}}(\vec{v}) + \text{perp}_{\vec{u}}(\vec{v})$$

Nota 6.21. Estas definiciones generalizan las dadas por la Definición 6.14, página 178 y la Definición A.10.1, página 274 para vectores reales, a través del *producto interno* de vectores (generalización del producto escalar).

⁵⁶Sin embargo, en muchos casos prácticos será necesario calcular una base del subespacio vectorial original para otros requerimientos.

6.9.6. Proceso de Gram-Schmidt en $\mathbb{C}^n$

El **proceso de Gram-Schmidt** es un método sencillo y sistemático para construir una *base ortogonal/ortonormal* a partir de cualquier base de un subespacio de $\mathbb{C}^n$.

Sea $\mathcal{B} = \{\vec{x}_1, \dots, \vec{x}_m\}$ una base de un $W \subseteq \mathbb{C}^n$ y definimos:

$$\begin{aligned}
 \vec{u}_1 &= \vec{x}_1 & W_1 &= \text{gen}(\vec{x}_1) = \text{gen}(\vec{u}_1) \\
 \vec{u}_2 &= \text{perp}_{W_1}(\vec{x}_2) = \vec{x}_2 - \text{proy}_{W_1}(\vec{x}_2) = \vec{x}_2 - \frac{\langle \vec{u}_1, \vec{x}_2 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \vec{u}_1 & W_2 &= \text{gen}(\vec{x}_1, \vec{x}_2) = \text{gen}(\vec{u}_1, \vec{u}_2) \\
 \vec{u}_3 &= \text{perp}_{W_2}(\vec{x}_3) = \vec{x}_3 - \text{proy}_{W_2}(\vec{x}_3) = \vec{x}_3 - \frac{\langle \vec{u}_1, \vec{x}_3 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \vec{u}_1 - \frac{\langle \vec{u}_2, \vec{x}_3 \rangle}{\langle \vec{u}_2, \vec{u}_2 \rangle} \vec{u}_2 & W_3 &= \text{gen}(\vec{x}_1, \vec{x}_2, \vec{x}_3) = \\
 & & &= \text{gen}(\vec{u}_1, \vec{u}_2, \vec{u}_3) \\
 & \vdots & & \vdots \\
 \vec{u}_m &= \text{perp}_{W_{m-1}}(\vec{x}_m) = \vec{x}_m - \text{proy}_{W_{m-1}}(\vec{x}_m) = \\
 &= \vec{x}_m - \frac{\langle \vec{u}_1, \vec{x}_m \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \vec{u}_1 - \dots - \frac{\langle \vec{u}_{m-1}, \vec{x}_m \rangle}{\langle \vec{u}_{m-1}, \vec{u}_{m-1} \rangle} \vec{u}_{m-1} & W &= W_m = \text{gen}(\vec{x}_1, \dots, \vec{x}_m) \\
 & & &= \text{gen}(\vec{u}_1, \dots, \vec{u}_m)
 \end{aligned}$$

Entonces para cada $i = 1, \dots, m$ el conjunto de vectores $\{\vec{u}_1, \dots, \vec{u}_i\}$ forma una *base ortogonal* del espacio W_i . En particular el conjunto $\{\vec{u}_1, \dots, \vec{u}_m\}$ es una *base ortogonal* de $W_m = W$.

Ejercicio 6.22. Dada la base del subespacio vectorial $W \subset \mathbb{C}^4$, $\mathcal{B}_W = \{[-i, 1, 0, -1] \ [0, 0, 1, i]\}$ y de su complemento ortogonal $W^\perp$, $\mathcal{B}_{W^\perp} = \{[1, -i, 0, 0] \ [0, 1, i, 1]\}$, obtenidas en el [Ejercicio 6.21, página 191](#), encontrar una base ortogonal de W , $\mathcal{B}_{oW}$, y otra base ortogonal del complemento ortogonal $W^\perp$, $\mathcal{B}_{oW^\perp}$.

Comprobamos previamente que ninguna de las dos bases son ortogonales (usando el producto interno, no el producto escalar) y aplicamos el proceso de G-S a ambos casos.

1. Aplicamos el proceso de G-S a los vectores de $\mathcal{B}_W$, eligiendo el vector más sencillo como $\vec{v}_1$:

$$\vec{u}_1 = \vec{v}_1 = [0, 0, 1, i]$$

$$\begin{aligned}
 \vec{u}_2 &= \vec{v}_2 - \text{proy}_{\vec{u}_1}(\vec{v}_2) = [-i, 1, 0, -1] - \frac{\langle [0, 0, 1, i], [-i, 1, 0, -1] \rangle}{\langle [0, 0, 1, i], [0, 0, 1, i] \rangle} [0, 0, 1, i] = [-i, 1, 0, -1] - \frac{i}{2} [0, 0, 1, i] = \\
 &= [-i, 1, -\frac{i}{2}, -\frac{1}{2}] \implies [2i, -2, i, 1]
 \end{aligned}$$

Por tanto, por ejemplo:

$$\mathcal{B}_{oW} = \{[2i, -2, i, 1], [0, 0, 1, i]\}$$

Podemos comprobar como ambos vectores son ahora ortogonales:

$$\langle [2i, -2, i, 1], [0, 0, 1, i] \rangle = [-2i, -2, -i, 1] \cdot [0, 0, 1, i] = -i + i = 0$$

2. Repetimos el proceso con la otra base

$$\vec{u}_1 = \vec{v}_1 = [1, -i, 0, 0]$$

$$\begin{aligned}
 \vec{u}_2 &= \vec{v}_2 - \text{proy}_{\vec{u}_1}(\vec{v}_2) = [0, 1, i, 1] - \frac{\langle [1, -i, 0, 0], [0, 1, i, 1] \rangle}{\langle [1, -i, 0, 0], [1, -i, 0, 0] \rangle} [1, -i, 0, 0] = [0, 1, i, 1] - \frac{i}{2} [1, -i, 0, 0] = \\
 &= [-\frac{i}{2}, \frac{1}{2}, i, 1] \implies [-i, 1, 2i, 2]
 \end{aligned}$$

Por tanto, por ejemplo:

$$\mathcal{B}_{oW^\perp} = \{[1, -i, 0, 0], [-i, 1, 2i, 2]\}$$

Podemos comprobar como también ambos vectores son ahora ortogonales:

$$\langle [1, -i, 0, 0], [-i, 1, 2i, 2] \rangle = [1, i, 0, 0] \cdot [-i, 1, 2i, 2] = -i + i = 0$$

De igual forma se puede comprobar como se conservan las relaciones de ortogonalidad entre los vectores de $\mathcal{B}_{oW}$ y $\mathcal{B}_{oW^\perp}$.

$$\begin{aligned} \langle [2i, -2, i, 1], [1, -i, 0, 0] \rangle &= -2i + 2i = 0 \\ \langle [2i, -2, i, 1], [-i, 1, 2i, 2] \rangle &= -2 - 2 + 2 + 2 = 0 \\ \langle [0, 0, 1, i], [1, -i, 0, 0] \rangle &= 0 \\ \langle [0, 0, 1, i], [-i, 1, 2i, 2] \rangle &= 2i - 2i = 0 \end{aligned}$$

6.9.7. Matrices Unitarias

Definición 6.26. Matriz Unitaria en $\mathbb{C}^n$

Se dice que una matriz cuadrada $U \in \mathbb{C}^{n \times n}$ es **unitaria** si verifica:

$$U^{-1} = U^H \quad \implies \quad U^H U = U U^H = I_n$$

- Es común usar la letra U para designar a las *matrices unitarias*, de igual forma que se suele usar Q para designar a las *matrices ortogonales*⁵⁷.
- Las *matrices unitarias* son una generalización en el campo complejo de la *matrices ortogonales*. Recordar que $Q^{-1} = Q^T$.
- En la **Sección 8.8, página 243** se verá la importancia de las *matrices unitarias* a través de la *diagonalización unitaria*⁵⁸, (generalización compleja de la *diagonalización ortogonal*⁵⁹).

Ejercicio 6.23. Comprobar que la siguiente matriz es unitaria.

$$A = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ -1-i & 1 \end{bmatrix}$$

Calculamos $A^H A$ y $A A^H$ y comprobamos si son iguales:

$$\left. \begin{aligned} A^H A &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & -1+i \\ 1+i & 1 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ -1-i & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \\ A A^H &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ -1-i & 1 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & -1+i \\ 1+i & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \end{aligned} \right\} \implies A^H A = A A^H$$

⁵⁷Ver la **Definición 6.10, página 171**.

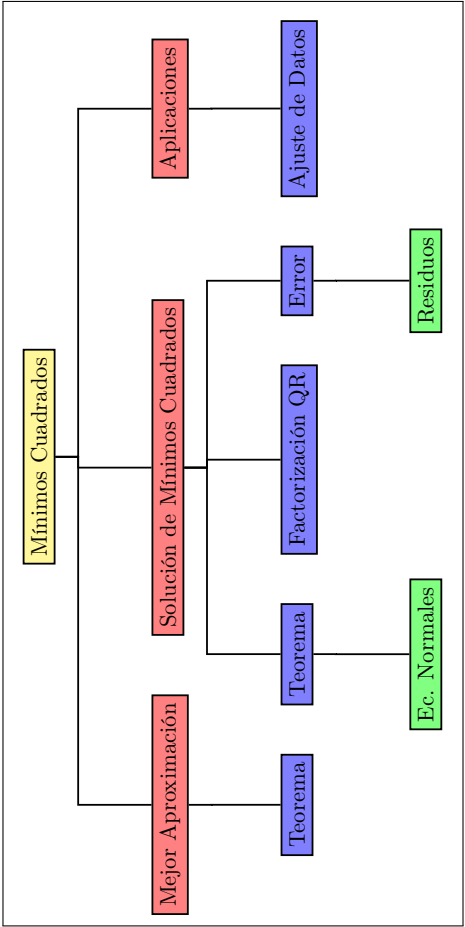
⁵⁸Ver la **Definición 8.12, página 243**.

⁵⁹Ver la **Definición 8.8, página 238**.

Mínimos Cuadrados

Contenido

7.1	Introducción	197
7.2	Mejor Aproximación	197
7.3	Aproximación mediante Mínimos Cuadrados	198
7.3.1	SEL Incompatibles	198
7.3.2	Solución de Mínimos Cuadrados	199
7.3.3	Ecuaciones Normales	200
7.4	Mínimos Cuadrados mediante la Factorización QR	204
7.5	Error en la Aproximación por Mínimos Cuadrados: Residuos	206
7.6	Resumen	206
7.7	Ajuste de Datos mediante Mínimos Cuadrados	207
7.7.1	Ejercicios Resueltos sobre Mínimos Cuadrados	211
7.8	Resumen de Fórmulas de Mínimos Cuadrados	216



7.1. Introducción

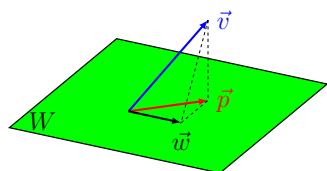
En este tema se presenta una aplicación muy importante de las ideas vistas sobre el *complemento ortogonal*¹ y la *proyección ortogonal*² sobre subespacios vectoriales vistos en el tema anterior.

7.2. Mejor Aproximación

Definición 7.1. Mejor Aproximación

Si W es un subespacio³ de $\mathbb{K}^n$, $W \subseteq \mathbb{K}^n$, y si $\vec{v} \in \mathbb{K}^n$, entonces la **mejor aproximación de $\vec{v}$ en W** es el vector $\vec{p} \in W$ tal que:

$$\|\vec{v} - \vec{p}\| < \|\vec{v} - \vec{w}\| \quad ; \quad \forall \vec{w} \in W \text{ con } \vec{w} \neq \vec{p}$$

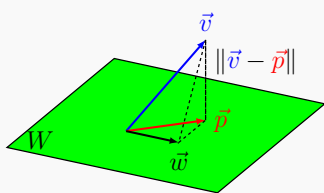


Interpretación gráfica en $\mathbb{R}^3$ de la *mejor aproximación* de un vector $\vec{v}$ a un plano W que pasa por el origen. El resultado $\vec{p}$ coincide con la idea intuitiva del vector del plano más cercano al vector $\vec{v}$, correspondiendo la *distancia más corta* como la *línea recta perpendicular*.

Teorema 7.1. Teorema de la Mejor Aproximación

Si W es un subespacio de dimensión finita de un espacio $\mathbb{K}^n$ con producto interno⁴ y si $\vec{v} \in \mathbb{K}^n$, entonces $\text{proy}_W(\vec{v})$ es la *mejor aproximación* de $\vec{v}$ en W .

Demostración 7.1.



Vamos a ver una intuitiva demostración geométrica en $\mathbb{R}^3$. Debido a la *ortogonalidad* de la mejor aproximación, podemos identificar un *triángulo rectángulo*, donde $\|\vec{v} - \vec{w}\| \neq 0$ es la *hipotenusa* y $\|\vec{v} - \vec{p}\|$ es uno de los *catetos*. Por el teorema de Pitágoras:

$$\Rightarrow \|\vec{v} - \vec{p}\| < \|\vec{v} - \vec{w}\| \quad ; \quad \forall \vec{w} \in W \text{ con } \vec{w} \neq \vec{p}$$

Ejercicio 7.1. Sea W el plano en $\mathbb{R}^3$ con ecuación $x - y + 2z = 0$, y sea $\vec{v} = [3, -1, 2]$. Encontrar la *mejor aproximación* de $\vec{v}$ en W y la *distancia euclídea* desde $\vec{v}$ hasta W .

La *mejor aproximación* de $\vec{v}$ en W será el vector $\text{proy}_W(\vec{v})$ calculado en el [Ejemplo 6.14](#).

$$\text{proy}_W(\vec{v}) = \begin{bmatrix} 5/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

¹Ver la [Sección 6.4](#), página 173.

²Ver la [Sección 6.5.3](#), página 181.

³Debe ser un *subespacio normado*, es decir, en el cual se ha definido una *norma* determinada, como por ejemplo, la norma euclídea.

⁴Generalización de producto escalar.

La *distancia* de $\vec{v}$ a W es la *distancia* de $\vec{v}$ al punto de W más cercano a $\vec{v}$ y ésta es precisamente $\|\text{perp}_W(\vec{v})\| = \|\vec{v} - \text{proy}_W(\vec{v})\|$. El vector $\text{perp}_W(\vec{v})$ también se calculó en [Ejemplo 6.14](#) siendo:

$$\text{perp}_W(\vec{v}) = \begin{bmatrix} 4/3 \\ -4/3 \\ 8/3 \end{bmatrix}$$

por tanto:

$$d = \|\text{perp}_W(\vec{v})\| = \sqrt{\frac{16}{9} + \frac{16}{9} + \frac{64}{9}} = \sqrt{\frac{96}{9}} = \frac{4\sqrt{2}}{\sqrt{3}}$$

7.3. Aproximación mediante Mínimos Cuadrados

Hasta ahora los sistemas de ecuaciones lineales (SEL) que hemos considerado útiles son los compatibles, es decir, aquellos que tienen solución. Si el SEL era incompatible poco se podía hacer ya que no tiene solución. Sin embargo en la práctica, muchos SEL incompatibles representan casos que merecen la pena considerar y aproximar, como veremos posteriormente.

Vamos a analizar primero los SEL incompatibles.

7.3.1. SEL Incompatibles

Un sistema de ecuaciones lineales de la forma⁵:

$$A\vec{x} = \vec{b}$$

es *incompatible*, es decir sin solución, según el teorema de Rouché-Fröbenius⁶ si $\text{rang}(A) < \text{rang}(A|B)$.

Vamos a ver una forma alternativa de entender las condiciones bajo las cuales un SEL tiene solución. Como se ha estudiado⁷, el SEL se puede expresar de la siguiente forma:

$$A\vec{x} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \vec{b}$$

Esto implica que para que exista una solución $\vec{x}$ que cumpla el sistema, debe de existir al menos una combinación lineal de los vectores columna de la matriz A que produzca el vector de términos independientes $\vec{b}$. Dicho de otra forma:

La condición necesaria y suficiente para que un SEL sea compatible es que el vector de términos independientes pertenezca al espacio columna $\text{col}(A)$ de la matriz de coeficientes.

Vamos a ver un caso particular en el podemos hacer una interpretación gráfica de esta condición. Sea un SEL de 3 ecuaciones y 3 incógnitas de la siguiente forma:

$$A\vec{x} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{212} \\ a_{32} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = x_1 \vec{c}_1 + x_2 \vec{c}_2 + x_3 \vec{c}_3 = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \vec{b}$$

Sea W el subespacio vectorial generado por la combinación lineal de los vectores columna $\vec{c}_i$ de la matriz de coeficientes A , es decir $W = \text{col}(A)$.

- Si los tres vectores columna de A son linealmente independientes entonces $W = \mathbb{R}^3$ y cualquier vector $\vec{b} \in \text{col}(A)$, entonces existirá una única combinación lineal de los vectores columna $\vec{c}_i$ de A que genere $\vec{b}$ y por tanto el sistema será compatible y determinado (SCD).

⁵Ver la [Sección 3.4](#), página 96.

⁶Ver la [Sección 1.5.6](#), página 51.

⁷Ver por ejemplo la [Sección 4.2.1](#), página 126 y el [Ejemplo 3.13](#), página 112.

- Si los tres vectores columna de A son linealmente dependientes, por ejemplo si sólo dos de ellos son linealmente independientes⁸ entonces $\dim(W) = 2$ (un plano de $\mathbb{R}^3$) y podría ocurrir dos cosas en función del vector $\vec{b}$:
 - Si $\vec{b} \notin W$ no habrá forma de expresar $\vec{b}$ como combinación de los vectores columna de A y por tanto el SEL será incompatible (SI).
 - Si $\vec{b} \in W$ entonces existirán infinitas formas de combinar linealmente los tres vectores columna $\vec{c}_i$ de A contenidos en el plano W para obtener $\vec{b}$ que también pertenece al mismo plano, y por tanto el sistema será compatible indeterminado (SCI).

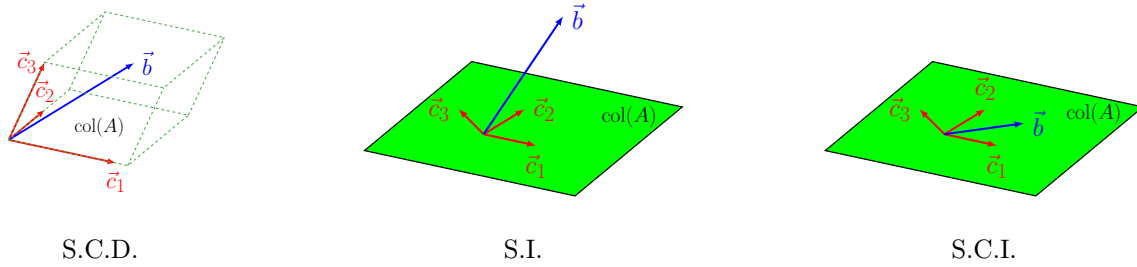


Figura 7.1: Interpretación geométrica de las condiciones de compatibilidad e incompatibilidad de los SEL en un caso sencillo de $\mathbb{R}^3$ (SEL 3×3). A la izquierda se representa un caso compatible determinado donde los 3 vectores columna $\vec{c}_i$ de A son linealmente independientes. En el centro se representa un caso incompatible donde los 3 vectores columna $\vec{c}_i$ de A son linealmente dependientes y el vector de términos independientes $\vec{b}$ no pertenece al $\text{col}(A)$. A la derecha se representa un caso compatible indeterminado donde los 3 vectores columna $\vec{c}_i$ de A son linealmente dependientes y el vector de términos independientes $\vec{b}$ sí pertenece al $\text{col}(A)$.

7.3.2. Solución de Mínimos Cuadrados

Claramente los sistemas incompatibles no tienen solución. Sin embargo en muchos casos estamos interesados en encontrar soluciones aproximadas, que si bien no cumplen el sistema, sí es una buena aproximación al mismo. Esto se suele denotar de la siguiente forma:

$$A\vec{s} \approx \vec{b}$$

en la cual la solución aproximada $\vec{s}$ no cumple el SEL pero sí es una *buena* aproximación. Para encontrarla podemos sustituir el SEL incompatible original por otro *parecido* pero que sea compatible.

Esto se puede hacer de la siguiente forma: como hemos visto en la [Sección 7.3.1](#) la compatibilidad de un SEL depende de la pertenencia o no del vector de términos independientes al espacio columna de la matriz de coeficientes. En los casos incompatibles podemos sustituir el vector $\vec{b}$ por la *mejor aproximación*⁹ del mismo al subespacio $\text{col}(A)$. Como hemos visto en el [Teorema 7.1, página 197](#), la mejor aproximación de un vector a un subespacio viene dado por su proyección ortogonal sobre dicho subespacio.

De esta forma se puede sustituir un sistema incompatible por otro *similar* pero compatible, cuya solución sea una aproximación a la *solución* del sistema original¹⁰.

$$\begin{array}{ccccc}
 \boxed{A\vec{x} = \vec{b}} & \implies & \boxed{A\vec{s} = \text{proy}_{\text{col}(A)} \vec{b}} & \implies & \boxed{A\vec{s} \approx \vec{b}} \\
 \text{Incompatible} & & \text{Compatible} & & \text{Sol. Aproximada}
 \end{array}$$

Esta nueva solución aproximada $\vec{s}$ se denomina *solución de mínimos cuadrados* de un SEL y es la que tiene un menor¹¹ *error*, definiendo el error de la siguiente forma

$$\boxed{A\vec{x} = \vec{b}} \quad \left\{ \begin{array}{ll} \bullet \text{ Compatible: } A\vec{x} - \vec{b} = \vec{0} & \implies \text{error} = \|A\vec{x} - \vec{b}\| = 0 \\ \bullet \text{ Incompatible: } A\vec{x} - \vec{b} \neq \vec{0} & \implies \text{error} = \|A\vec{x} - \vec{b}\| > 0 \end{array} \right.$$

⁸Si solo un vector columna fuera linealmente independiente ocurriría lo mismo pero en este caso $\dim(W) = 1$ (recta en $\mathbb{R}^3$).

⁹Ver la [Sección 7.2, página 197](#).

¹⁰Que realmente no existe al ser un sistema incompatible.

¹¹Consecuencia de las propiedades de la mejor aproximación de un vector en un subespacio.

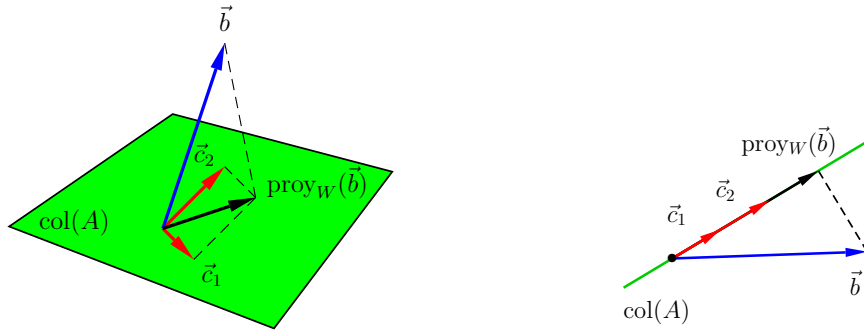


Figura 7.2: Interpretación geométrica de la sustitución del vector de términos independientes $\vec{b}$ por la mejor aproximación del mismo en el espacio columna de la matriz de coeficientes, $\text{proy}_{\text{col}(A)}(\vec{b})$, en un caso con vectores columna $\vec{c}_i$ de A linealmente independientes (figura izquierda) y linealmente dependientes (figura derecha).

Sin embargo, la utilización de la $\text{proy}_{\text{col}(A)}(\vec{b})$ en la determinación de la solución de mínimos cuadrados no es la forma de cálculo más apropiada y en la siguiente sección se verá una mejor alternativa.

Definición 7.2. Solución de Mínimos Cuadrados

Si $A \in \mathbb{K}^{m \times n}$ y $\vec{b} \in \mathbb{K}^m$, una *solución de mínimos cuadrados* del SEL $A\vec{x} = \vec{b}$ es un vector $\vec{s} \in \mathbb{K}^n$ tal que:

$$A\vec{s} \approx \vec{b} : \quad \|\vec{b} - A\vec{s}\| \leq \|\vec{b} - A\vec{x}\| \quad ; \quad \forall \vec{x} \in \mathbb{K}^n$$

Características de la *solución de mínimos cuadrados* de un SEL:

- La solución de mínimos cuadrados de un SEL siempre existe.
- La solución de mínimos cuadrados no siempre es única.
- La solución de mínimos cuadrados de un SEL compatible es igual a su solución.

7.3.3. Ecuaciones Normales

Vamos a ver ahora un teorema que establece un método sencillo de cálculo de la *solución de mínimos cuadrados* de un SEL. Este método, conocido con el nombre de **ecuaciones normales**, es muy conveniente al evitar la necesidad de calcular la proyección ortogonal de un vector sobre un subespacio vectorial¹².

Teorema 7.2. Teorema de los Mínimos Cuadrados

Sea $A \in \mathbb{R}^{m \times n}$ y $\vec{b} \in \mathbb{R}^m$. Entonces¹³:

- El SEL $A\vec{x} = \vec{b}$, siempre tiene al menos una *solución de mínimos cuadrados* $\vec{s} \in \mathbb{R}^n$ que viene dada por las denominadas **ecuaciones normales**:

$$A^T A \vec{s} = A^T \vec{b}$$

- Si la matriz A tiene *vectores columnas linealmente independientes* entonces la matriz $A^T A$ es invertible y la *solución de mínimos cuadrados* es única, dada por:

$$\vec{s} = (A^T A)^{-1} A^T \vec{b}$$

¹²Ver la Sección 6.5.3, página 181.

¹³En el caso general de matrices y vectores en espacios $\mathbb{C}^n$ hay que sustituir la matriz transpuesta A^T por la matriz transpuesta conjugada A^H .

Demostración 7.2.

La solución de mínimos cuadrados de un SEL $A\vec{x} = \vec{b}$ viene dada por la solución del siguiente SEL siempre compatible:

$$A\vec{s} = \text{proy}_{\text{col}(A)}(\vec{b})$$

En lugar de calcular primero el vector proyección y después resolver el SEL anterior, vamos a utilizar una forma más eficiente de resolverlo.

Ya que el siguiente vector:

$$\vec{b} - A\vec{s} = \vec{b} - \text{proy}_{\text{col}(A)}(\vec{b}) = \text{perp}_{\text{col}(A)}(\vec{b})$$

es *ortogonal* a $\text{col}(A)$, entonces para cualquier vector columna $\vec{c}_i$ de A , se tiene:

$$\vec{c}_i \cdot (\vec{b} - A\vec{s}) = \vec{c}_i^T (\vec{b} - A\vec{s}) = 0$$

Para todos los vectores columna de A , esta expresión se puede expresar de forma matricial como:

$$A^T(\vec{b} - A\vec{s}) = \vec{0} \quad ; \quad A^T\vec{b} - A^T A\vec{s} = \vec{0} \quad \implies \quad \boxed{A^T A\vec{s} = A^T \vec{b}}$$

Este nuevo SEL siempre es compatible (determinado o indeterminado) y se conoce con el nombre de **ecuaciones normales**.

Además si su solución es única (sistema compatible determinado), el *rango* de la matriz $A^T A$ debe ser n (número de incógnitas del SEL) y entonces según el **Teorema 3.7, página 116**, las columnas de A son linealmente independientes, siendo la solución:

$$\boxed{\vec{s} = (A^T A)^{-1} A^T \vec{b}}$$

Ejercicio 7.2. Encontrar la solución de mínimos cuadrados para el sistema *incompatible* $A\vec{x} = \vec{b}$, donde:

$$A = \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \quad ; \quad \vec{b} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

Tenemos que resolver el sistema (ecuaciones normales) $A^T A\vec{s} = A^T \vec{b}$, de forma que:

$$A^T A = \begin{bmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 30 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 16 \end{bmatrix}$$

Por tanto el nuevo SEL a resolver es:

$$\begin{bmatrix} 6 & 0 \\ 0 & 30 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 16 \end{bmatrix} \quad \implies \quad \boxed{\vec{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 8/15 \end{bmatrix}}$$

En este caso la solución de mínimos cuadrados es única. Esto es debido a que las columnas de A son linealmente independientes y por el **Teorema 7.2** la solución de mínimos cuadrados es única.

Ejercicio 7.3. Encontrar la solución de mínimos cuadrados para el sistema *incompatible*:

$$A\vec{x} = \vec{b} \quad ; \quad A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix} \quad ; \quad \vec{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Las ecuaciones normales $A^T A \vec{s} = A^T \vec{b}$ son en este caso:

$$A^T A = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -3 & 1 \\ -3 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \implies \begin{bmatrix} 4 & -3 & 1 \\ -3 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Resolviendo el SEL por Gauss:

$$\left[\begin{array}{ccc|c} 4 & -3 & 1 & 1 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ -3 & 3 & 0 & 0 \\ 4 & -3 & 1 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ -3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right] \rightarrow \begin{cases} s_1 = 1 - t \\ s_2 = 1 - t \\ s_3 = t \end{cases}$$

En este caso existen infinitas soluciones de mínimos cuadrados dadas por:

$$\vec{s} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ -1 \\ t \end{bmatrix} \quad ; \quad \forall t \in \mathbb{R}$$

consecuencia de que la matriz A tiene vectores columna linealmente dependientes.

Ejercicio 7.4. Dar la solución de mínimos cuadrados del siguiente SEL *compatible determinado*:

$$A\vec{x} = \vec{b} \quad ; \quad A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \quad ; \quad \vec{b} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

Las ecuaciones normales $A^T A \vec{s} = A^T \vec{b}$, son de la forma:

$$A^T A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix} \quad ; \quad A^T \vec{b} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 7 \end{bmatrix} = \begin{bmatrix} 8 \\ 13 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 13 \end{bmatrix} \implies \boxed{\vec{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}}$$

Por otra parte, si resolvemos el sistema compatible y determinado original por Gauss:

$$\left[\begin{array}{cc|c} 1 & -1 & 1 \\ 1 & 2 & 7 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 3 & 6 \end{array} \right] \rightarrow \begin{cases} x_1 = 1 + x_2 = 3 \\ x_2 = 6/3 = 2 \end{cases}$$

comprobándose que su solución coincide con la de mínimos cuadrados al ser un SEL compatible.

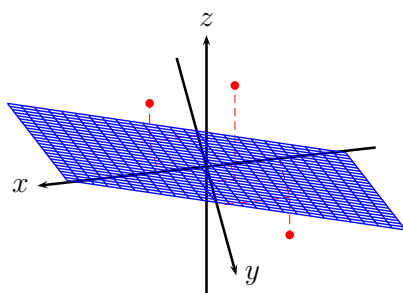
Ejercicio 7.5. Crear dos enunciados de problemas que den dos posibles interpretaciones geométricas distintas del **Ejemplo 7.2**.

1. Determinar el plano que pasa por el origen: $z = ax + by$ que mejor ajuste al conjunto de puntos de $\mathbb{R}^3$: $\{(1, 5, 3), (2, -2, 2), (-1, 1, 5)\}$.

Queremos encontrar valores a y b tales que verifiquen:

$$\begin{cases} a + 5b = 3 \\ 2a - 2b = 2 \\ -a + b = 5 \end{cases} \implies \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

Ya que conocemos la solución de mínimos cuadrados del problema, establecemos que el plano buscado es:



$$z = \frac{1}{3}x + \frac{8}{15}y$$

2. Determinar el vector de $\mathbb{R}^3$ perteneciente al plano definido por los vectores $\vec{v}_1 = [1, 2, -1]$ y $\vec{v}_2 = [5, -2, 1]$ que es más cercano al vector $\vec{u} = [3, 2, 5]$.

Como $\vec{v}_1$ y $\vec{v}_2$ son linealmente independientes definen un plano. Si el vector $\vec{u}$ pertenece a dicho plano, el vector buscado es el mismo $\vec{u}$.

Por otra parte si $\vec{u}$ no pertenece al plano, existirá un determinado vector del mismo, que puede obtenerse por combinación lineal de $\vec{v}_1$ y $\vec{v}_2$, que sea el más cercano a $\vec{u}$.

Vamos a crear una ecuación vectorial suponiendo que $\vec{u}$ pertenece al plano.

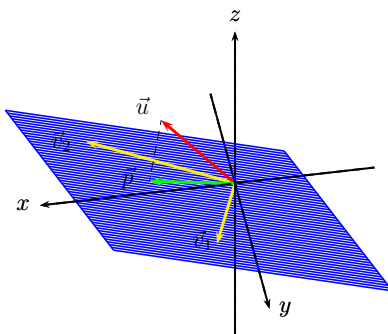
$$a\vec{v}_1 + b\vec{v}_2 = \vec{u} \quad ; \quad a \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + b \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} \implies \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

En caso de que dicho SEL fuera compatible, efectivamente $\vec{u}$ pertenecería al plano y $\vec{u}$ sería la solución. Sin embargo si es incompatible, como ocurre aquí, $\vec{u}$ no pertenece al plano y habrá un vector del mismo que sea lo más parecido a posible a $\vec{u}$ que será el más cercano.

Sabiendo la solución de mínimos cuadrados del problema es $a = 1/3$ y $b = 8/15$, el vector buscado es:

$$\vec{p} = \frac{1}{3}\vec{v}_1 + \frac{8}{15}\vec{v}_2 = \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \frac{8}{15} \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -2/5 \\ 1/5 \end{bmatrix}$$

En la siguiente figura se muestra una representación gráfica de la solución.



7.4. Mínimos Cuadrados mediante la Factorización QR

Definición 7.3. Condicionamiento

Se dice que un SEL (o la matriz que lo define) está **mal condicionado**, si *pequeñas* variaciones en los valores que lo definen¹⁴ producen *grandes* variaciones en la solución del mismo.

En la práctica, las *ecuaciones normales* que aparecen en muchos problemas de *mínimos cuadrados* están *mal condicionadas*, lo cual significa que los errores numéricos que se producen en las operaciones matemáticas del método (Gauss, Gauss-Jordan, etc.) como consecuencia de la utilización de un ordenador¹⁵ en su resolución, se van acumulando y pueden producir grandes errores en la solución final del problema.

La *factorización*¹⁶ QR produce un método de resolución que mejora sensiblemente la *precisión* de los resultados¹⁷, aunque tiene un coste computacional superior.

No obstante, la obtención de la factorización QR a través del método de ortogonalización de Gram-Schmidt¹⁸ aplicado a las columnas de la matriz A , es un método inestable numéricamente.

Existen métodos alternativos¹⁹ para obtener la descomposición QR que evitan este problema, pero incrementan considerablemente el coste computacional.

Por otra parte, la utilización del método QR en la resolución del método de los mínimos cuadrados, tiene el inconveniente que solo es aplicable cuando la matriz de coeficientes A tiene vectores columnas linealmente independientes²⁰.

Vamos a ver a continuación en que consiste este método.

Teorema 7.3.

Sea $A \in \mathbb{R}^{m \times n}$ con *columnas linealmente independientes*²¹ y sea $\vec{b} \in \mathbb{R}^m$. Si $A = QR$ es una factorización QR de A , entonces la única solución del problema de *mínimos cuadrados* $A\vec{x} = \vec{b}$ viene dada por:

$$A\vec{x} = \vec{b} \quad ; \quad QR\vec{s} = \vec{b} \quad ; \quad R\vec{s} = Q^T\vec{b} \quad ; \quad \vec{s} = R^{-1}Q^T\vec{b}$$

¹⁴Debidas a errores experimentales, de redondeo en el cálculo, etc.

¹⁵El cual tiene una precisión finita en los cálculos.

¹⁶Ver la [Sección 6.8, página 185](#).

¹⁷Ya que el procedimiento llevado a cabo involucra el producto de matrices ortogonales, y este proceso conserva la norma matricial y el número de condición.

¹⁸Ver la [Sección 6.7, página 184](#).

¹⁹Por ejemplo el método de Householder.

²⁰Ver la [Sección 6.8, página 185](#).

²¹Lo que produce una solución *única* en el problema de mínimos cuadrados.

Demostración 7.3.

La solución de mínimos cuadrados es:

$$A^T A \vec{s} = A^T \vec{b}$$

Si $A = QR$ entonces:

$$(QR)^T (QR) \vec{s} = (QR)^T \vec{b} \quad ; \quad R^T Q^T Q R \vec{s} = R^T Q^T \vec{b} \quad ; \quad R^T R \vec{s} = R^T Q^T \vec{b}$$

Como R es invertible, multiplicando por la izquierda ambas partes de la igualdad por $(R^T)^{-1}$ se obtiene:

$$R \vec{s} = Q^T \vec{b}$$

En la práctica y teniendo en cuenta que la matriz R tiene forma triangular, resulta mucho más sencillo resolver el SEL:

$$R \vec{s} = Q^T \vec{b}$$

por el método de sustitución hacia atrás, que calcular R^{-1} y después $\vec{s} = R^{-1} Q^T \vec{b}$.

Ejercicio 7.6. Encontrar la solución de mínimos cuadrados del **Ejemplo 7.2** utilizando la factorización QR de la matriz A .

$$A = \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \quad ; \quad \vec{b} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

Como se vio en **Sección 6.8**, para calcular la factorización QR de una matriz se necesita ortonormalizar los vectores columna de A mediante el método de Gram-Schmidt. Llamando $\vec{v}_i$ a los vectores columna de A y $\vec{u}_i$ a los nuevos vectores ortogonales, se tiene:

$$\left. \begin{aligned} \vec{u}_1 &= \vec{v}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \\ \vec{u}_2 &= \vec{v}_2 - \frac{\vec{u}_1 \cdot \vec{v}_2}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \begin{bmatrix} 5 & -2 & 1 \end{bmatrix} \end{aligned} \right\} \quad Q = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & 5/\sqrt{30} \\ 2/\sqrt{6} & -2/\sqrt{30} \\ -1/\sqrt{6} & 1/\sqrt{30} \end{bmatrix}$$

La matriz R se obtiene de la siguiente forma²²:

$$R = Q^T A = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 5/\sqrt{30} & -2/\sqrt{30} & 1/\sqrt{30} \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{30} \end{bmatrix}$$

Finalmente, la solución de mínimos cuadrados vendrá dada mediante la solución del sistema:

$$R \vec{s} = Q^T \vec{b} \quad ; \quad \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{30} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 5/\sqrt{30} & -2/\sqrt{30} & 1/\sqrt{30} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 16/\sqrt{30} \end{bmatrix}$$

Cuya solución se puede calcular fácilmente:

$$\vec{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 8/15 \end{bmatrix}$$

Esta solución coincide con la obtenida mediante las ecuaciones normales en el **Ejemplo 7.2**.

²²Ver la **Sección 6.8**, página 185.

7.5. Error en la Aproximación por Mínimos Cuadrados: Residuos

Sea un sistema *incompatible* $A\vec{x} = \vec{b}$, con $A \in \mathbb{R}^{m \times n}$. La *solución de mínimos cuadrados* del sistema, $\vec{s}$, cumple²³:

$$A\vec{s} \approx \vec{b}$$

siendo la *mejor aproximación* al problema²⁴. En función de cada caso particular, este valor puede resultar una buena o mala aproximación del problema. Para determinar la *calidad* de esta aproximación particular, estamos interesados en calcular los *errores* cometidos en cada una de las variables cuando se utiliza la aproximación $\vec{s}$ en lugar de la solución exacta²⁵ $\vec{x}$. Estos errores se pueden expresar en forma vectorial mediante un *vector de errores*, $\vec{e}$, denominado también *vector de residuos*, de la siguiente forma²⁶:

$$\vec{e} = \vec{b} - A\vec{s}$$

Cada una de las coordenadas de $\vec{e} = [e_1, \dots, e_m]$ corresponde con los errores o *residuos* cometidos en cada una de las incógnitas $\vec{x} = [x_1, \dots, x_m]$. Desde un punto de vista práctico interesa determinar un único *error global* que cuantifique la *calidad* de la aproximación, y éste se puede obtener calculando la *norma*²⁷ del *vector de residuos*:

$$e = \|\vec{e}\| = \|\vec{b} - A\vec{s}\| = \|\vec{b} - \text{proy}_{\text{col}(A)}(\vec{b})\| = \|\text{perp}_{\text{col}(A)}(\vec{b})\|$$

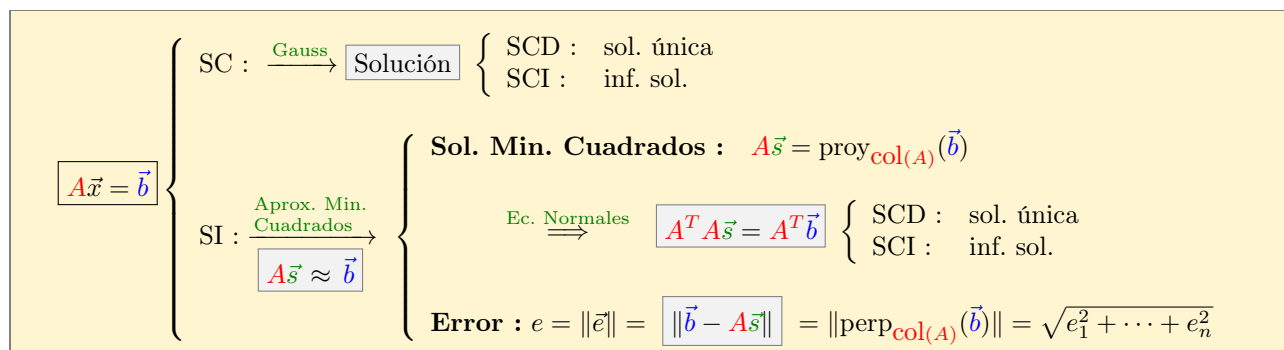
El *método de mínimos cuadrados* busca una solución $\vec{s}$ que minimiza la *norma euclídea* del *vector de residuos* (errores) definido por:

$$\|\vec{e}\| = \sqrt{e_1^2 + \dots + e_n^2} \quad \Rightarrow \quad \|\vec{e}\|^2 = e_1^2 + \dots + e_n^2$$

y por tanto, minimiza la suma de los cuadrados de los errores, de ahí el nombre de **mínimos cuadrados**.

7.6. Resumen

En el siguiente cuadro se muestra a modo de resumen las posibilidades que hay a la hora de resolver un SEL, en sus dos modalidades: sistema compatible (SC) y sistema incompatible (SI).



²³Ver la Sección 7.3, página 198.

²⁴Ver la Sección 7.2, página 197.

²⁵Que no existe al ser un sistema *incompatible*.

²⁶Esta definición corresponde con el *error absoluto* de la aproximación y también se puede definir a través de su valor absoluto, $\vec{e} = |\vec{b} - A\vec{s}| = |A\vec{s} - \vec{b}|$.

²⁷Ver la Sección 6.2.1, página 164.

Ejercicio 7.7. Encontrar los errores cometidos (residuos) en la aproximación de mínimos cuadrados del [Ejemplo 7.2](#) y calcular el valor del error global. En base a ellos, determinar si es una *buena* o *mala* aproximación.

$$A = \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} ; \quad \vec{b} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} ; \quad \vec{s} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 8/15 \end{bmatrix}$$

$$\vec{e} = \vec{b} - A\vec{s} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1/3 \\ 8/15 \end{bmatrix} = \begin{bmatrix} 0 \\ 14/5 \\ 24/5 \end{bmatrix}$$

La primera ecuación se cumple ($e_1 = 0$) y el máximo error se produce en la tercera ecuación ($e_3 > e_2$). Por tanto el error *global* es:

$$\|\vec{e}\| = \sqrt{0^2 + \left(\frac{14}{5}\right)^2 + \left(\frac{24}{5}\right)^2} \approx 5.6$$

Este valor indica que el error cometido en la aproximación es considerablemente alto. Podemos estimar el error *relativo* en cada ecuación dividiendo los términos b_i entre los correspondientes e_i y multiplicando por 100, obteniéndose 0% en la primera ecuación, 71% en la segunda y 100% en la tercera. Se puede observar que el error global es un valor promedio (en tanto por uno) de los errores relativos. Según la naturaleza del problema, estos altos errores podrían indicar una revisión del planteamiento del problema en busca de unos mejores resultados.

7.7. Ajuste de Datos mediante Mínimos Cuadrados

El método de los *mínimos cuadrados* es una técnica matemática ampliamente utilizada en el campo de la ingeniería. Una de las aplicaciones más extendidas es en el *ajuste de datos* a curvas, superficies, etc.

Definición 7.4. Ajuste de Datos

- El *ajuste de datos* consiste en encontrar una determinada función que aproxime a unos determinados datos.
- El error del ajuste determinará la precisión del ajuste.

Hay diferentes técnicas para realizar el ajuste de datos, siendo la más ampliamente utilizada el método de los mínimos cuadrados.

²⁸En este curso se estudia únicamente el método de los mínimos cuadrados *lineales*, ya que en el caso *no lineal* aparece un sistema de ecuaciones no lineales, tema del que no se trata aquí.

²⁹Para el caso de los mínimos cuadrados lineales, la función debe ser lineal con respecto a estos parámetros o en caso contrario se debe poder linealizar para transformarla en una función lineal en sus parámetros (ver un caso típico en el [Ejemplo 7.12, página 215](#)).

³⁰Ver la [Sección 7.5, página 206](#).

- El ajuste de datos por mínimos cuadrados encuentra una *función de ajuste* (curva, superficie, etc.) que minimiza la suma de los cuadrados de los errores (distancia de los puntos a la *función*).
- Para ello se plantea un SEL²⁸ (siempre compatible y normalmente indeterminado) y se encuentra su solución de mínimos cuadrados.
- La *función de ajuste* siempre viene dada a través de un número determinado de *parámetros*²⁹, definiendo una *familia de infinitas funciones*.
- El ajuste da como resultado unos valores concretos de los parámetros que determinan una única función de esta familia, la cual minimiza el error con respecto a la solución de los mínimos cuadrados, (ver la [Sección 7.3, página 198](#)).
- Para determinar la *calidad* del ajuste es necesario calcular los errores del problema de mínimos cuadrados. El *vector de errores*³⁰ $\vec{e} = [e_1, \dots, e_n]$ contiene la siguiente información:
 - Cada *error individual* e_i representa el error cometido en cada ecuación y geoméricamente se puede interpretar como la distancia del punto (x_i, y_i) a la función de ajuste.
 - El *error global* $e = \|\vec{e}\|$ da la norma del vector error y representa la acumulación de los errores en todos los puntos, siendo un indicador de la calidad global del ajuste.

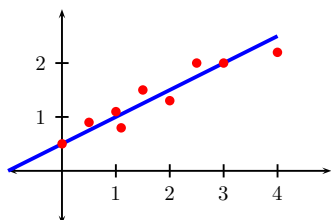
Vamos a estudiar los casos más representativos del ajuste de datos por mínimos cuadrados a funciones lineales en sus parámetros.

Ajuste de datos a una Recta en $\mathbb{R}^2$

Dado un conjunto de datos $(x, y) \in \mathbb{R}^2$, normalmente dados en forma de tabla, se pide encontrar la recta que mejor ajuste a estos datos. En este caso la familia de funciones será:

$$y = f(x) = a + bx$$

siendo a y b los *parámetros* a determinar.



x	y
x_1	y_1
$\vdots$	$\vdots$
x_n	y_n

$$y = a + bx$$

Sustituyendo cada uno de los pares (x_i, y_i) de la tabla en la familia de funciones se obtiene el SEL, siendo los parámetros a y b las *incógnitas* a determinar.

$$\begin{cases} a + bx_1 = y_1 \\ a + bx_2 = y_2 \\ \vdots \\ a + bx_n = y_n \end{cases} \quad \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \Rightarrow \quad A\vec{x} = \vec{b}$$

Claramente si la tabla tiene más de dos datos, el SEL será incompatible y no tendrá solución³¹. Tendremos que buscar entonces la mejor solución de mínimos cuadrados para encontrar la función, perteneciente a la familia de funciones, que mejor ajuste a los datos dados.

Como siempre existe al menos una solución de mínimos cuadrados³², siempre existirá al menos una función de ajuste. El problema es que esta función no siempre será una buena solución del problema.

³¹Lo cual indica que no existe ninguna función de la familia de funciones propuestas que pase o contenga a todos los datos de la tabla.

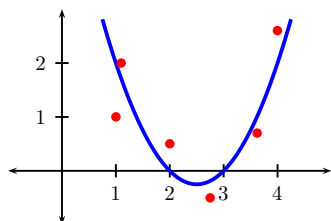
³²Ver el [Teorema 7.2, página 200](#).

Parábola en $\mathbb{R}^2$

Este caso es análogo al ajuste de rectas, cambiando únicamente la familia de funciones del ajuste, siendo en este caso:

$$y = f(x) = a + bx + cx^2$$

con a , b y c los parámetros del ajuste (incógnitas a determinar).



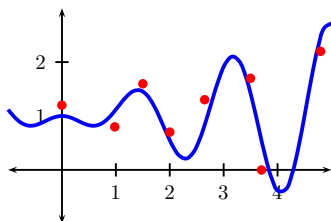
x	y
x_1	y_1
$\vdots$	$\vdots$
x_n	y_n

$$y = a + bx + cx^2$$

$$\begin{cases} a + bx_1 + cx_1^2 = y_1 \\ a + bx_2 + cx_2^2 = y_2 \\ \vdots \\ a + bx_n + cx_n^2 = y_n \end{cases} ; \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \Rightarrow A\vec{x} = \vec{b}$$

Polinomio General en $\mathbb{R}^2$

Se puede utilizar un polinomio de cualquier grado para el ajuste, siendo el procedimiento a seguir análogo a los anteriores.



x	y
x_1	y_1
$\vdots$	$\vdots$
x_n	y_n

$$y = c_0 + c_1x + \cdots + c_mx^m$$

$$\begin{cases} c_0 + c_1x_1 + \cdots + c_mx_1^m = y_1 \\ c_0 + c_1x_2 + \cdots + c_mx_2^m = y_2 \\ \vdots \\ c_0 + c_1x_n + \cdots + c_mx_n^m = y_n \end{cases} ; \begin{bmatrix} 1 & x_1 & \cdots & x_1^m \\ 1 & x_2 & \cdots & x_2^m \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \Rightarrow A\vec{x} = \vec{b}$$

Naturalmente algunos polinomios darán un mejor ajuste que otros, pudiéndose determinar su calidad mediante el cálculo de los errores de mínimos cuadrados. De forma general, al aumentar el grado del polinomio, (mayor número de parámetros de ajuste), disminuirá el error cometido en el ajuste, pero no necesariamente aumentará la calidad del ajuste. Esto se debe a que al aumentar el grado de un polinomio aumentan las oscilaciones en su gráfica, obteniéndose resultados no deseados.

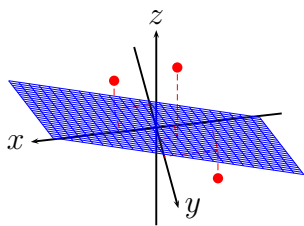
En un ajuste polinómico se buscará siempre el polinomio de menor grado que mejor ajuste a los datos.

Plano en $\mathbb{R}^3$

Este es el caso más sencillo de ajuste de datos en 3D. Dado un conjunto de puntos (x_i, y_i, z_i) queremos encontrar el plano que mejor los ajusta:

$$z = f(x, y) = a + bx + cy$$

Siguiendo el planteamiento de los casos anteriores se obtendría el correspondiente SEL del ajuste.



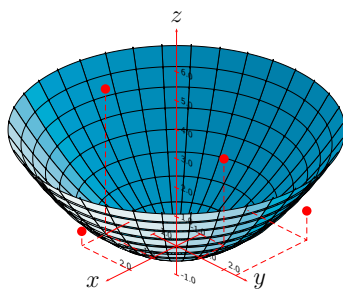
x	y	z
x_1	y_1	z_1
$\vdots$	$\vdots$	$\vdots$
x_n	y_n	z_n

$$z = a + bx + cy$$

$$\begin{cases} a + bx_1 + cy_1 = z_1 \\ a + bx_2 + cy_2 = z_2 \\ \vdots \\ a + bx_n + cy_n = z_n \end{cases} ; \quad \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \Rightarrow \boxed{A\vec{x} = \vec{b}}$$

Superficie Cuadrática en $\mathbb{R}^3$

El ajuste se puede realizar a una superficie más general como por ejemplo una *superficie cuadrática* o *cuádrica* (elipsoide, paraboloides e hiperboloides).



x	y	z
x_1	y_1	z_1
$\vdots$	$\vdots$	$\vdots$
x_n	y_n	z_n

$$z = a + bx + cy + dx^2 + ey^2 + fxy$$

$$\begin{cases} a + bx_1 + cy_1 + dx_1^2 + ey_1^2 + fx_1y_1 = z_1 \\ a + bx_2 + cy_2 + dx_2^2 + ey_2^2 + fx_2y_2 = z_2 \\ \vdots \\ a + bx_n + cy_n + dx_n^2 + ey_n^2 + fx_ny_n = z_n \end{cases}$$

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & y_1^2 & x_1y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n^2 & y_n^2 & x_ny_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \Rightarrow \boxed{A\vec{x} = \vec{b}}$$

Hiperplano en $\mathbb{R}^m$

Se puede realizar un ajuste de datos en cualquier dimensión espacial, aunque para dimensiones altas no existirá una interpretación geométrica del problema.

El caso más sencillo corresponde al ajuste de n datos del tipo $(x_1, \dots, x_m) \in \mathbb{R}^m$ a un hiperplano:

$$y = f(x_1, \dots, x_m) = c_0 + c_1x_1 + \dots + c_mx_m$$

siendo $\{c_0, c_1, \dots, c_m\}$ el conjunto de parámetros a ajustar. De forma análoga a los casos anteriores, el planteamiento del problema es el siguiente:

x_1	x_2	$\dots$	x_m	y
x_{11}	x_{12}	$\dots$	x_{1m}	y_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{n1}	x_{n2}	$\dots$	x_{nm}	y_n

$$y = c_0 + c_1x_1 + \dots + c_mx_m$$

$$\begin{cases} c_0 + c_1 x_{11} + \cdots + c_m x_{1m} = y_1 \\ c_0 + c_1 x_{21} + \cdots + c_m x_{2m} = y_2 \\ \vdots \\ c_0 + c_1 x_{n1} + \cdots + c_m x_{nm} = y_n \end{cases} ; \begin{bmatrix} 1 & x_{11} & \cdots & x_{1m} \\ 1 & x_{21} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nm} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \Rightarrow A\vec{x} = \vec{b}$$

En todos los casos la solución de *mínimos cuadrados* del problema puede obtenerse mediante la resolución del sistema de las *ecuaciones normales*³³:

$$A^T A \vec{s} = A^T \vec{b} \quad ; \quad A \vec{s} \approx \vec{b}$$

7.7.1. Ejercicios Resueltos sobre Mínimos Cuadrados

Ejercicio 7.8. Encontrar la ecuación de la recta en $\mathbb{R}^2$ que mejor aproxime por mínimos cuadrados a los datos siguientes:

x	-2	-1	1	2
y	1	0	1	-1

La ecuación del ajuste es:

$$y = a + bx$$

siendo a y b los parámetros a determinar.

El sistema de ecuaciones lineales correspondiente es en este caso:

$$\begin{cases} a - 2b = 1 \\ a - b = 0 \\ a + b = 1 \\ a + 2b = -1 \end{cases} ; \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} ; A\vec{c} = \vec{b}$$

La solución de mínimos cuadrados se puede encontrar resolviendo las ecuaciones normales:

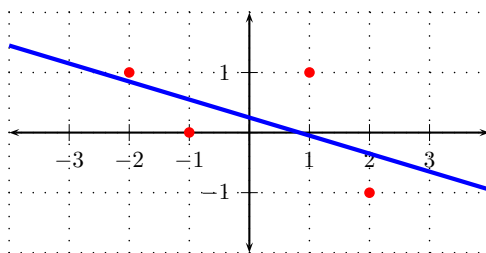
$$A^T A = \begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix} ; A^T \vec{b} = \begin{bmatrix} 1 \\ -3 \end{bmatrix} ; A^T A \vec{s} = A^T \vec{b}$$

La solución final es:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1/4 \\ -3/10 \end{bmatrix} ; y = \frac{1}{4} - \frac{3}{10}x$$

Ejercicio 7.9. Realizar una interpretación gráfica del **Ejemplo 7.8** anterior y calcular el error cometido en el ajuste.

Vamos a dibujar como puntos rojos los cuatro datos (x, y) dados en la tabla del enunciado, y de color azul dibujamos la recta solución dada por el método de mínimos cuadrados.



Podemos apreciar que el ajuste no es bueno, ya que hay una separación significativa entre algunos puntos y la recta que los aproxima.

³³Ver el **Teorema 7.2**, página 200.

Para determinar de forma más precisa el error del ajuste, calculamos el vector de residuos $\vec{e}$ y el error global $e = \|\vec{e}\|$.

$$\vec{e} = \vec{b} - A\vec{s} = \begin{bmatrix} 3/20 \\ -11/20 \\ 21/20 \\ -13/20 \end{bmatrix} ; \quad \text{error} = \|\vec{e}\| = \sqrt{\frac{37}{20}} \approx 1.36$$

Las componentes del vector de residuos indican los errores individuales de cada uno de los puntos. Geométricamente, esto se puede interpretar como la distancia del punto a la recta, indicando su signo si está por encima o por debajo de la recta. A través de estos datos comprobamos lo que muestra la gráfica, es decir que el punto más cercano a la recta es el primero y el más alejado es el tercero.

El error global es una forma de cuantificar de forma unificada estos errores individuales, correspondiendo a un buen ajuste los valores pequeños del mismo. Este valor hay que analizarlo en función del orden de magnitud de los datos analizados. En nuestro caso, los valores (x, y) son de orden unidad y por tanto un buen ajuste correspondería a valores del error global mucho más pequeños de la unidad.

Como podemos ver, en este ejemplo el error global es de orden unidad, igual que los datos de entrada, lo que indica claramente que no es un buen ajuste.

Ejercicio 7.10. Dada la siguiente tabla de valores:

x	-1	-1	1	1
y	-1	-1	-1	1

Determinar cuál de las dos curvas siguientes es más apropiada para representarlos, teniendo en cuenta el posible error de cada uno de los ajustes.

$$a) \quad y = a + \frac{b}{x}$$

$$b) \quad y = a + bx^2$$

Vamos a calcular cada uno de los ajustes y sus correspondientes errores.

a) Planteamos el sistema de ecuaciones de los datos a la curva $y = a + b/x$.

$$\begin{cases} a - b = -1 \\ a - b = -1 \\ a + b = -1 \\ a + b = 1 \end{cases} ; \quad \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix} ; \quad A\vec{x} = \vec{b}$$

Notar que hay dos datos que se repiten. Esto puede ser habitual en muchos casos prácticos y es importante recordar que siempre hay que considerar todos los datos, se repitan o no, ya que desde un punto de vista práctico, el hecho de que se repitan datos puede corresponder, por ejemplo, a distintos experimentos en los que aparecen múltiples veces condiciones y resultados. Cuantas más veces se repitan, más importancia tendrán esos datos ya que se tiene una mayor seguridad sobre la veracidad de los mismos, y por tanto, en el sistema de ecuaciones tienen que aparecer tantas veces como corresponda.

La solución de mínimos cuadrados se puede encontrar resolviendo las ecuaciones normales:

$$A^T A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} ; \quad A^T \vec{b} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} ; \quad A^T A \vec{s} = A^T \vec{b}$$

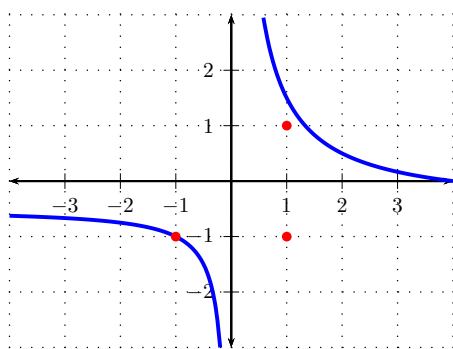
cuya solución es:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix} ; \quad y_1 = -\frac{1}{2} + \frac{1}{2x}$$

El error en este caso es:

$$\vec{e}_1 = \vec{b} - A\vec{s} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} ; \quad \text{error}_1 = \|\vec{e}_1\| = \sqrt{2} \approx 1.41$$

A continuación se muestra una interpretación gráfica de este caso:



Observar como la curva (en azul) pasa justamente por el punto $(-1, -1)$, ya que tiene más importancia que el resto, al aparecer repetido dos veces en los datos.

b) Planteamos ahora el sistema de ecuaciones de los datos a la segunda curva $y = a + bx^2$.

$$\begin{cases} a + b = -1 \\ a + b = -1 \\ a + b = -1 \\ a + b = 1 \end{cases} ; \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix} ; \quad A\vec{x} = \vec{b}$$

La solución de mínimos cuadrados se puede encontrar resolviendo las correspondientes ecuaciones normales:

$$A^T A = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} ; \quad A^T \vec{b} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} ; \quad A^T A \vec{s} = A^T \vec{b}$$

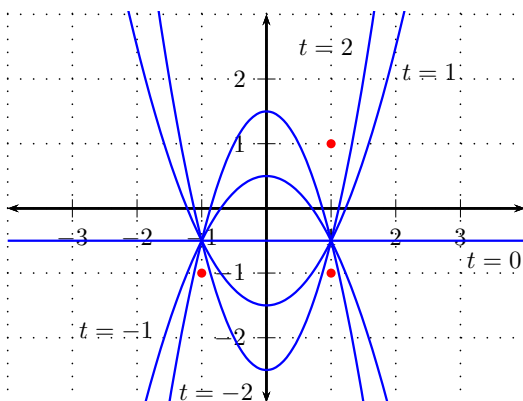
que en este caso tienen infinitas soluciones:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -1/2 - t \\ t \end{bmatrix} ; \quad y_2 = -\left(\frac{1}{2} + t\right) + tx^2 ; \quad \forall t \in \mathbb{R}$$

Esto corresponde a una *familia de parábolas*, (para el caso particular $t = 0$ es una recta horizontal), que ajustan a estos datos con error mínimo, siendo en todos ellos el mismo error, con valor:

$$\vec{e}_1 = \vec{b} - A\vec{s} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} ; \quad \text{error}_1 = \|\vec{e}_1\| = \sqrt{2} \approx 1.41$$

A continuación se muestra una interpretación gráfica de la solución del problema:



En la gráfica se han representado solo 5 de los infinitos posibles casos que hay, para valores del parámetro $t = \{-2, -1, 0, 1, 2\}$. Todas, a pesar de ser curvas distintas, tienen el mismo error por ser soluciones del problema de mínimos cuadrados. Notar que para el caso $t = 0$ la parábola es en realidad una recta horizontal, por tener el coeficiente cuadrático nulo.

Análisis de la elección de la mejor curva

Para determinar cual de las curvas es la más apropiada hay que tener en cuenta principalmente factores físicos³⁴, error del ajuste y simplicidad de las curvas.

En nuestro caso no tenemos factores físicos y lo importante será el menor error de ajuste y la mayor simplicidad de la curva. Los dos tipos de curvas, el primero de tipo hipérbole y el segundo de tipo parabólico, tienen el mismo error, $\sqrt{2}$, y por tanto podría resultar más apropiada la curva más simple, que corresponde al segundo caso con valor del parámetro $t = 0$, la cual como hemos visto es una recta horizontal.

Ejercicio 7.11. Dada la siguiente tabla de valores:

x	0	1	1	0
y	1	0	1	1
z	1	-1	1	1

a) Encontrar el ajuste de mínimos cuadrados de dichos datos a la función:

$$z = a + bx + cxy^2 \quad ; \quad a, b, c \in \mathbb{R}$$

b) Encontrar el error del ajuste.

Notar que la tabla contiene 4 colecciones de datos (x, y, z) dispuestos en columnas y que se repiten el primero y cuarto con valores $(0, 1, 1)$. Bajo ningún concepto se puede eliminar alguno de ellos pensando que sobra por estar repetido, ya que el número de veces que aparece un valor es importante en el resultado de mínimos cuadrados.

a) Planteamos el sistema de ecuaciones de los datos a la curva $z = a + bx + cxy^2$.

$$\begin{cases} a &= 1 \\ a + b &= -1 \\ a + b + c &= 1 \\ a &= 1 \end{cases} \quad ; \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \quad ; \quad A\vec{x} = \vec{b}$$

Como vemos se repite la ecuación primera y cuarta, siendo un hecho importante en el método y no se debe eliminar ninguna. Por otra parte, no confundir la incógnita b a determinar con el vector de términos independientes $\vec{b}$.

Resolvemos ahora las ecuaciones normales: $A^T A \vec{s} = A^T \vec{b}$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

Resolvemos el SEL:

$$\left[\begin{array}{ccc|c} 4 & 2 & 1 & 2 \\ 2 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 0 & -2 & -3 & -2 \\ 0 & 0 & -1 & -2 \\ 1 & 1 & 1 & 1 \end{array} \right] \rightarrow \begin{cases} a = 1 - b - c = 1 \\ b = \frac{-2 + 3c}{-2} = -2 \\ c = 2 \end{cases}$$

³⁴Determinados problemas reales pueden imponer una determinada función al ajuste debido a las leyes físicas que rigen el problema.

La función del ajuste de mínimos cuadrados es entonces:

$$z = 1 - 2x + 2xy^2$$

b) Encontramos el error del ajuste:

$$e = \|\vec{e}\| = \|A\vec{s} - \vec{b}\| = \left\| \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\| = 0$$

En este caso el error es cero, lo que indica que la superficie $z = f(x, y)$ encontrada pasa exactamente por todos los puntos (x, y, z) dados en la tabla de valores.

Ejercicio 7.12. Encontrar el ajuste de mínimos cuadrados de la siguiente tabla de datos a la función $y(x)$ dada, en la cual a , b y c son constantes a determinar.

x	0	1	2	3	4
y	0.2	0.25	0.3	0.5	1

$$y(x) = \frac{a}{b + cx} \quad ; \quad a, b, c \in \mathbb{R}$$

La función $y(x)$ no es lineal en los parámetros b y c y es necesario *linealizar*³⁵ dicha función antes de realizar el ajuste por mínimos cuadrados.

Primeramente vamos a simplificar la función:

$$y = \frac{a}{b + cx} = \frac{1}{\frac{b}{a} + \frac{c}{a}x} = \frac{1}{A + Bx}$$

Claramente podemos reducir los tres parámetros a , b y c dados inicialmente a solo dos: $A = b/a$ y $B = c/a$. Ahora linealizamos la función resultante, con respecto a los parámetros A y B :

$$y = \frac{1}{A + Bx} \quad ; \quad \frac{1}{y} = A + Bx$$

De esta forma la nueva función es lineal en los parámetros A y B a determinar. Recalculamos los valores de la tabla para $1/y$:

x	0	1	2	3	4
y	0.2	0.25	0.3	0.5	1
$1/y$	5	4	3	2	1

Formamos el SEL y las ecuaciones normales correspondientes:

$$\left. \begin{array}{l} A = 5 \\ A + x = 4 \\ A + 2x = 3 \\ A + 3x = 2 \\ A + 4x = 1 \end{array} \right\} \Rightarrow \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 1 & 1 & 4 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \\ 1 & 4 & 1 \end{array} \right] \Rightarrow \left\{ \begin{array}{l} A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 30 \end{bmatrix} \\ A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 15 \\ 20 \end{bmatrix} \end{array} \right.$$

Resolvemos el sistema $A^T A \vec{p} = A^T \vec{b}$,

$$\left[\begin{array}{cc|c} 5 & 10 & 15 \\ 10 & 30 & 20 \end{array} \right] \xrightarrow{\frac{1}{5}F_1} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 1 & 3 & 2 \end{array} \right] \xrightarrow{F_2 - F_1} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 1 & -1 \end{array} \right] \rightarrow \left\{ \begin{array}{l} A = 3 - 2B = 5 \\ B = -1 \end{array} \right.$$

Por tanto la función buscada es:

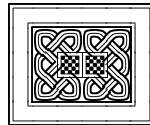
$$y(x) = \frac{1}{5-x}$$

En este caso comprobamos como el ajuste produce una solución exacta, siendo por tanto el error de mínimos cuadrados cero.

7.8. Resumen de Fórmulas de Mínimos Cuadrados

$$A\vec{x} = \vec{b} \xrightarrow[A\vec{p} \approx \vec{b}]{\text{Aprox. Min. Cuadrados}} \begin{cases} \text{Solución} & : A\vec{s} = \text{proy}_{\text{col}(A)}(\vec{b}) \xrightarrow{\text{Ec. Normales}} A^T A\vec{s} = A^T \vec{b} \\ \text{Sol. } QR & : A(\text{col. lin. indep}) = QR \implies R\vec{s} = \vec{b}Q^T \\ \text{Error} & : \text{error} = \|\vec{e}\| = \|\vec{b} - A\vec{s}^*\| = \|\text{perp}_{\text{col}(A)}(\vec{b})\| = \sqrt{e_1^2 + \dots + e_n^2} \end{cases}$$

Nota 7.1. Existen otros procedimientos de cálculo en la determinación del problema de mínimos cuadrados.

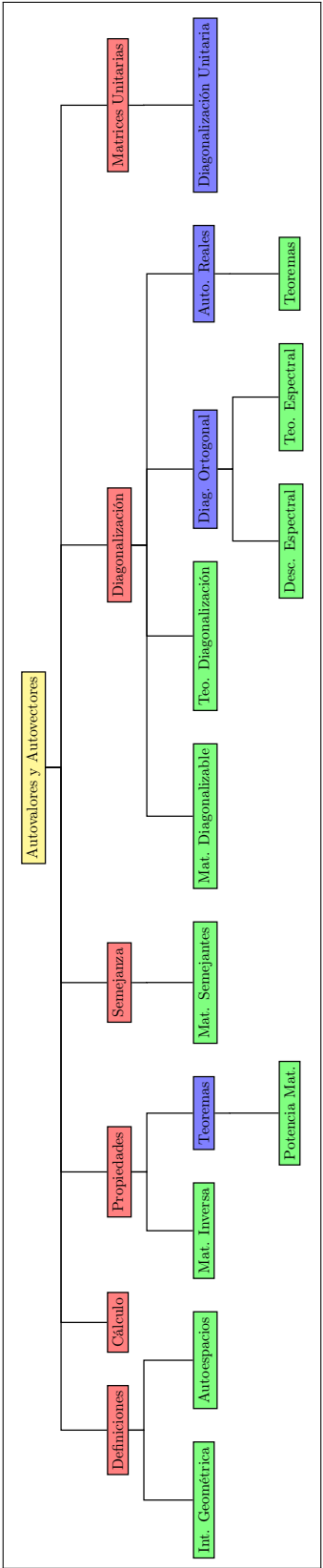


³⁵Es posible resolver este problema sin linealizarlo, mediante un método de *mínimos cuadrados no lineal*, (que no se verá en este curso), aunque con una mayor complejidad, un coste computacional mucho mayor y seguramente una menor precisión en los resultados.

Autovalores y Autovectores

Contenido

8.1	Introducción: Sistemas Dinámicos Discretos	219
8.2	Autovalores y Autovectores	220
8.2.1	Interpretación Geométrica de los Autovalores y Autovectores	220
8.2.2	Determinación de Autovalores y Autovectores	221
8.2.3	Autoespacios	223
8.2.4	Multiplicidades Algebraica y Geométrica	225
8.2.5	Autovalores y Matrices Inversas	226
8.2.6	Teoremas sobre Autovalores y Autovectores	227
8.3	Propiedades de los Autovalores de una Matriz	229
8.4	Matrices Semejantes	230
8.4.1	Matrices Triangulares	230
8.4.2	Matrices Semejantes	230
8.4.3	Interpretación de la Semejanza de Matrices	232
8.5	Diagonalización	234
8.5.1	Autovalores Reales y Complejos	236
8.6	Diagonalización Ortogonal	237
8.6.1	Interpretación Geométrica de la Diagonalización Ortogonal	238
8.6.2	Teorema Espectral	241
8.6.3	Descomposición Espectral	241
8.7	Teoremas Avanzados Adicionales	242
8.8	Diagonalización Unitaria	243
8.9	Resumen Final del Tema	245



8.1. Introducción: Sistemas Dinámicos Discretos

Definición 8.1. Sistema Dinámico Discreto

Dada una TL de la forma,

$$T(\vec{x}) = A\vec{x}$$

se denomina *Sistema Dinámico Discreto* a un proceso del estilo:

$$\vec{x}(t+1) = A\vec{x}(t)$$

donde $\vec{x}(t)$ es un **vector de estado** a tiempo t (valor entero) y $\vec{x}(t+1)$ es el *vector de estado* al tiempo siguiente ($t+1$). La matriz A se denomina **matriz de transición** del sistema¹.

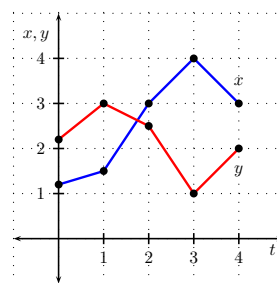
Vamos a analizar como ejemplo el caso bidimensional. Supongamos que $\vec{x} \in \mathbb{R}^2$, $A \in \mathbb{R}^{2 \times 2}$ y partimos de un *vector inicial* a tiempo $t=0$:

$$\vec{x}_0 = \begin{bmatrix} x_0 & y_0 \end{bmatrix}$$

Entonces:

$$\begin{aligned} \vec{x}_1 &= A\vec{x}_0 \\ \vec{x}_2 &= A\vec{x}_1 \\ \vec{x}_3 &= A\vec{x}_2 \\ &\vdots \\ \vec{x}_n &= A\vec{x}_{n-1} \end{aligned}$$

$\vec{x}_i$		
t	x_i	y_i
0	x_0	y_0
1	x_1	y_1
$\vdots$	$\vdots$	$\vdots$
n	x_n	y_n



Por otra parte, también se pueden obtener los valores temporales $x(t)$ de la siguiente forma:

$$\begin{aligned} \vec{x}_0 &= \vec{x}_0 \\ \vec{x}_1 &= A\vec{x}_0 \\ \vec{x}_2 &= A\vec{x}_1 = AA\vec{x}_0 = A^2\vec{x}_0 \\ \vec{x}_3 &= A\vec{x}_2 = AA^2\vec{x}_0 = A^3\vec{x}_0 \\ &\vdots \\ \vec{x}_n &= A\vec{x}_{n-1} = AA^{n-1}\vec{x}_0 = A^n\vec{x}_0 \end{aligned}$$

Lo cual demuestra que en todo *sistema dinámico discreto* con una matriz A fija, la *evolución* del sistema depende únicamente del *estado inicial* $\vec{x}_0$.

Es interesante conocer la *evolución* a tiempos grandes y para ello será necesario calcular A^n , cuestión complicada si A es grande². Sin embargo, si existiera un determinado vector inicial $\vec{v}$ tal que cumpliera:

$$A\vec{v} = \lambda\vec{v}$$

siendo λ un escalar cualquiera, se tendría:

$$\begin{aligned} \vec{x}_0 &= \vec{v} \\ \vec{x}_1 &= A\vec{x}_0 = A\vec{v} = \lambda\vec{v} \\ \vec{x}_2 &= A\vec{x}_1 = A\lambda\vec{v} = \lambda A\vec{v} = \lambda^2\vec{v} \\ &\vdots \\ \vec{x}_n &= A\vec{x}_{n-1} = A\lambda^{n-1}\vec{v} = \lambda^{n-1}A\vec{v} = \lambda^n\vec{v} \end{aligned}$$

De esta forma se consigue cambiar el cálculo de A^n mediante multiplicaciones matriciales (lentas), por el mucho más sencillo cálculo de λ^n (potencia de un escalar).

¹Si la matriz A representa *probabilidades*, los procesos se denominan **cadena de Markov**.

²El número de operaciones (multiplicaciones y sumas) necesarias para obtener el producto matricial AA con $A \in \mathbb{R}^{m \times m}$ es de orden $2m^3$ y por tanto A^n es de orden $2nm^3$, (ver [Sección 3.3, página 93](#)). Por ejemplo, A^{10} siendo $A \in \mathbb{R}^{50 \times 50}$ necesita unos dos millones y medio de operaciones.

Este proceso solo es válido para el caso especial en que $\vec{x}_0 = \vec{v}$. Sin embargo, si existiera otro vector $\vec{v}_2$ con propiedades análogas al anterior $\vec{v}$ que llamaremos a partir de ahora $\vec{v}_1$, tendríamos:

$$A\vec{v}_2 = \mu\vec{v}_2$$

Si además $\vec{v}_2$ fuera *linealmente independiente*³ de $\vec{v}_1$, entonces cualquier otro vector $\vec{x}_0$ se podría poner como *combinación lineal*⁴ de ambos vectores:

$$\vec{x}_0 = \alpha\vec{v}_1 + \beta\vec{v}_2$$

y en este caso se tendría:

$$\begin{aligned}\vec{x}_0 &= \alpha\vec{v}_1 + \beta\vec{v}_2 \\ \vec{x}_1 &= A\vec{x}_0 = A(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha\lambda\vec{v}_1 + \beta\mu\vec{v}_2 \\ \vec{x}_2 &= A\vec{x}_1 = A(\alpha\lambda\vec{v}_1 + \beta\mu\vec{v}_2) = \alpha\lambda^2\vec{v}_1 + \beta\mu^2\vec{v}_2 \\ &\vdots \\ \vec{x}_n &= A\vec{x}_{n-1} = A(\alpha\lambda^{n-1}\vec{v}_1 + \beta\mu^{n-1}\vec{v}_2) = \alpha\lambda^n\vec{v}_1 + \beta\mu^n\vec{v}_2\end{aligned}$$

Se reemplaza una potencia matricial por dos potencias escalares.

Nota 8.1. Si la matriz A fuera de orden m , se necesitarían m vectores linealmente independientes con estas propiedades especiales y la potencia A^n se transformaría en la suma de m potencias escalares.

Este tipo de situaciones se producen con bastante asiduidad en la práctica y se utilizan técnicas de este estilo para resolver problemas de forma analítica y numérica⁵.

8.2. Autovalores y Autovectores

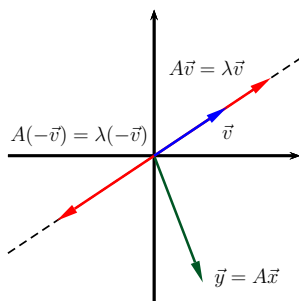
Definición 8.2. Autovalores y Autovectores

Sea una matriz *cuadrada* $A \in \mathbb{K}^{n \times n}$. Un *escalar* (real o complejo) $\lambda \in \mathbb{K}$ se llama **autovalor** o **valor propio**⁶ de A si existe un vector (real o complejo) no nulo $\vec{v} \in \mathbb{K}^n$ tal que:

$$A\vec{v} = \lambda\vec{v}$$

Al vector $\vec{v}$ se le denomina **autovector** o **vector propio**⁷ de A correspondiente o asociado al *autovalor* λ , y viceversa.

8.2.1. Interpretación Geométrica de los Autovalores y Autovectores



Sea una TL, $\vec{y} = T(\vec{x}) = A\vec{x}$, con $A \in \mathbb{R}^{n \times n}$. Un autovector de la matriz A será aquel vector de entrada $\vec{v} \in \mathbb{R}^n$ cuyo vector de salida produzca un múltiplo, (vector paralelo), del propio vector $\vec{v}$:

$$T(\vec{v}) = A\vec{v} = \lambda\vec{v}$$

El correspondiente *autovalor* λ asociado a $\vec{v}$ será dicho múltiplo. Realmente el *autovector* indica una determinada *dirección*, y cualquier vector de esa recta, (subespacio vectorial), será un autovector de A , por ejemplo $(-\vec{v})$.

En el gráfico se puede ver una interpretación en el espacio $\mathbb{R}^2$.

³Ver la Sección 2.4.2, página 73.

⁴Ver la Sección 2.4.1, página 73.

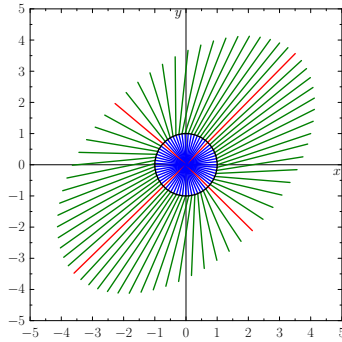
⁵En caso de que el tiempo no sea *discreto* sino *continuo* se denominan **sistemas dinámicos continuos**, apareciendo *sistemas de ecuaciones diferenciales* en su formulación.

⁶En algunos libros lo denominan **eigenvalor**, término derivado de las terminologías inglesa y alemana. La palabra *eigen* significa *propio* o *característico de*, indicando que contiene información importante acerca de la naturaleza de la matriz A .

⁷También denominado **eigenvector** en las terminologías inglesa y alemana.

Vamos a determinar gráficamente⁸ los *autovalores* y *autovectores* de dos casos correspondientes a diferentes TL $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$.

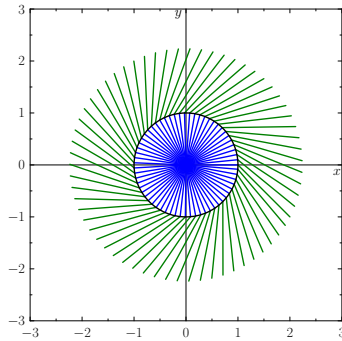
Caso 1



$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

En la figura se muestran los resultados obtenidos al aplicar como entradas vectores unitarios $\vec{x}$, (círculo pequeño en azul), y las correspondientes salidas $\vec{y}$, (elipse grande verde). En el gráfico se ha desplazado la salida $\vec{y}$ y se ha colocado a continuación de la entrada $\vec{x}$. Como se puede apreciar, existen cuatro *autovectores*, (mostrados en rojo), con longitudes iguales dos a dos. Por tanto, existen dos direcciones privilegiadas, (*autovectores*), *autovalores* y en cada una de ellas hay una determinada ampliación (*autovalores*).

Caso 2



$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

En este caso y siguiendo el mismo procedimiento anterior, se puede observar como no existe ninguna dirección privilegiada (*autovector*) y por tanto ningún *autovalor* asociado.

Nota 8.2. La representación gráfica de los *autovalores* de una matriz real 3×3 estaría formada por figuras 3D de esferas y elipsoides. Para dimensiones mayores de 3 no se podría hacer representaciones gráficas.

Ejercicio 8.1. Mostrar que el vector $\vec{x} = [1, 1]$ es un *autovector* de la matriz A y encontrar el *autovalor* asociado al mismo.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$$A\vec{x} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4\vec{x}$$

Por tanto, $\vec{x}$ es un *autovector* de A con *autovalor* asociado $\lambda = 4$.

8.2.2. Determinación de Autovalores y Autovectores

Los *autovalores* λ y *autovectores* $\vec{v}$ de una matriz cuadrada A , son aquellos que verifican $A\vec{v} = \lambda\vec{v}$. Operando matricialmente se tiene:

$$A\vec{v} = \lambda \vec{v} \implies \vec{v}A\vec{v} - \lambda\vec{v} = \vec{0} \implies (A - \lambda I_n)\vec{v} = \vec{0} \quad (8.1)$$

Los *autovectores* $\vec{v}$ de la matriz A son el *espacio nulo* o *núcleo*⁹ de la matriz $(A - \lambda I)$, que tiene la forma:

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix} = \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

⁸En la Sección 8.6.1, página 238 se ampliará esta interpretación gráfica.

⁹Ver la Definición 3.25, página 115.

Como la ecuación (8.1) es un *sistema homogéneo*, (ver la [Sección 1.8, página 59](#)), la única forma de que tenga soluciones $\vec{v}$ distintas de la nula, es que sea un sistema *compatible indeterminado* (infinitas soluciones) y por tanto con una *matriz escalonada reducida equivalente*¹⁰ con al menos una fila de ceros. Matricialmente esto es equivalente a que la matriz $(A - \lambda I)$ sea *singular*¹¹ y por tanto, su *determinante* igual a cero, $\det(A - \lambda I) = 0$, siendo ésta la condición necesaria y suficiente para la existencia de autovalores en una matriz.

Los *autovalores* de una matriz *cuadrada* A son las soluciones λ de la ecuación:

$$|A - \lambda I| = 0$$

Definición 8.3. Ecuación y Polinomio Característicos

- La ecuación $\det(A - \lambda I) = 0$ se conoce con el nombre de **ecuación característica**.
- Cuando se desarrolla el $\det(A - \lambda I)$ se obtiene un *polinomio* en λ que se denomina **polinomio característico**. El *polinomio característico* de una matriz $A \in \mathbb{R}^{n \times n}$ será siempre de orden n y por el *Teorema Fundamental del Álgebra*¹², tendrá n soluciones reales o complejas.

Ejercicio 8.2. Calcular el *polinomio característico* de una matriz general 2×2 :

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$$

Ejercicio 8.3. Calcular la *ecuación y polinomio característico* de la matriz:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

El *polinomio característico* es:

$$\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 0 & 1 \\ 1 & 1 - \lambda & 0 \\ -1 & 0 & -\lambda \end{vmatrix} = -\lambda(1 - \lambda)^2 + (1 - \lambda) = (1 - \lambda)(\lambda^2 - \lambda + 1)$$

La *ecuación característica* por tanto es:

$$(1 - \lambda)(\lambda^2 - \lambda + 1) = 0$$

Procedimiento de Cálculo de Autovalores y Autovectores

Sea $A \in \mathbb{K}^{n \times n}$.

1. Calcular el *polinomio característico* $\det(A - \lambda I)$.
2. Calcular los *autovalores* de A resolviendo la *ecuación característica* $\det(A - \lambda I) = 0$ para λ .
3. Para cada *autovalor* λ , encontrar el *núcleo* de la correspondiente matriz $A - \lambda I$. Las soluciones no nulas son los *autovectores* $\vec{v}$ asociados al *autovalor* λ de la matriz A .

¹⁰Ver la [Definición 1.36, página 57](#).

¹¹Ver la [Definición 3.16, página 99](#).

¹²Un polinomio $P(x)$ tiene tantas *raíces*, (soluciones de la ecuación $P(x) = 0$), como indica su grado, teniendo en cuenta sus *multiplicidades*, (número de veces que se repite una raíz).

8.2.3. Autoespacios

Definición 8.4. Autoespacios

Se denomina **autoespacio** o **subespacio propio** de una matriz $A \in \mathbb{K}^{n \times n}$ asociado a un *autovalor* λ de A , al *núcleo* de la matriz $(A - \lambda I)$, formado por los *autovectores* de A asociados al *autovalor* λ .

$$\ker(A - \lambda I) = \{\vec{v} \in \mathbb{K}^n : A\vec{v} = \lambda\vec{v}\}$$

El *autoespacio* de λ se suele dar a través de una de sus *bases* y se representa por E_λ .

Ejercicio 8.4. Encontrar los *autovalores* y *autovectores* de la matriz:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 2 & -5 & 4 - \lambda \end{vmatrix} = \lambda^2(4 - \lambda) + 0 + 2 - 0 - 5\lambda - 0 = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = 0$$

Fácilmente se ve que $\lambda = 2$ es solución, por tanto factorizando:

$$(\lambda - 2)(-\lambda^2 + 2\lambda - 1) = (\lambda - 1)^2(\lambda - 2) = 0 \implies \begin{cases} \lambda_1 = \lambda_2 = 1 & (\text{raíz doble}) \\ \lambda_3 = 2 \end{cases}$$

Vamos a encontrar los *autovectores* asociados a los *autovalores*.

$$\underline{\lambda_1 = \lambda_2 = 1}$$

$$[A - 1I|0] = \left[\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 2 & -5 & 3 & 0 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \implies \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \implies E_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$$\underline{\lambda_3 = 2}$$

$$[A - 2I|0] = \left[\begin{array}{ccc|c} -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 2 & -5 & 2 & 0 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1/4 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \implies \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1/4 \\ 1/2 \\ 1 \end{bmatrix} \implies E_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \right\}$$

Notar que los autovectores solo indican una *dirección* determinada y por tanto se puede tomar cualquier múltiplo del mismo¹³.

En resumen la matriz A tiene 3 *autovalores* (dos de ellos iguales) y asociados a cada *valor* hay un *autovector*.

Los SEL que resultan en el cálculo de los *autovectores* son siempre *homogéneos*¹⁴ y por tanto podemos obviar el vector de términos independientes en su resolución.

Ejercicio 8.5.

Encontrar los autovalores y autoespacios asociados de la siguiente matriz:

$$B = \begin{bmatrix} 4 & 4 & -4 \\ 4 & 4 & -4 \\ -4 & -4 & 4 \end{bmatrix}$$

¹³Es decir, cualquier vector *paralelo*.

¹⁴Ver la [Sección 1.8, página 59](#).

Claramente esta matriz se puede expresar como:

$$B = 4 \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} = 4A$$

Resulta más sencillo trabajar con A , siempre que se conozca la relación entre los autovalores y autovectores de A y B . Teniendo en cuenta que:

$$A\vec{v} = \lambda\vec{v} \implies \alpha A\vec{v} = \alpha\lambda\vec{v} \implies B\vec{v} = (\alpha\lambda)\vec{v}$$

llegamos a la conclusión de que la matriz B tiene los mismos autovectores que A y los autovalores están multiplicados por α .

Vamos entonces a calcular los autovalores y autovectores de A .

$$\begin{vmatrix} 1-\lambda & 1 & -1 \\ 1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{vmatrix} = (1-\lambda)^3 + 2 - 3(1-\lambda) = 0$$

Esta ecuación se puede resolver desarrollando las potencias y obteniendo el polinomio de tercer grado en λ . Sin embargo resulta más sencillo realizar un cambio de variable, resolver una ecuación más sencilla y deshacer el cambio. En este caso, haciendo el cambio de variable $t = 1 - \lambda$ la ecuación queda:

$$(1-\lambda)^3 - 3(1-\lambda) + 2 = 0 \xrightarrow{t=1-\lambda} t^3 - 3t + 2 = 0$$

Una solución es claramente $t = 1$ y factorizando por Ruffini, (ver el [Apéndice A.6.1, página 266](#)) se obtiene:

$$t^3 - 3t + 2 = (t-1)(t^2 + t - 2) = (t-1)(t-1)(t+2)$$

Por tanto las raíces son $t_1 = 1$, $t_2 = 1$ y $t_3 = -2$. Deshaciendo el cambio de variable se obtienen los autovalores de A : $\lambda_1 = 0$, $\lambda_2 = 0$ y $\lambda_3 = 3$. Finalmente sabiendo que los autovalores de B son iguales que los de A multiplicados por cuatro se tiene:

$$\text{Autovalores matriz } B: \lambda_{1,2} = 0 \quad ; \quad \lambda_3 = 12$$

Vamos a calcular ahora los autovectores de la matriz A , que coinciden con los de la matriz B .

$$\underline{\lambda_{1,2} = 0}$$

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \xrightarrow[\substack{F_2-F_1 \\ F_3+F_1}]{F_2-F_1} \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \implies \begin{cases} v_1 = -v_2 + v_3 = -t + s \\ v_2 = t \\ v_3 = s \end{cases}$$

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \implies E_0 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

En este caso el autovalor 0 de la matriz A corresponde con el también autovalor 0 en la matriz B , por tanto el autoespacio de la matriz B asociado a este autovalor es:

$$E_0 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$\underline{\lambda_3 = 3}$ (Para simplificar los cálculos utilizamos el autovalor de A , $\lambda_3 = 3$, y no de B , $\lambda_3 = 12$)

$$\begin{bmatrix} -2 & 1 & -1 \\ 1 & -2 & -1 \\ -1 & -1 & -2 \end{bmatrix} \xrightarrow[\substack{F_1+2F_2 \\ F_3+F_2}]{F_1+2F_2} \begin{bmatrix} 0 & -3 & -3 \\ 1 & -2 & -1 \\ 0 & -3 & -3 \end{bmatrix} \xrightarrow{F_3+F_1} \begin{bmatrix} 0 & -3 & -3 \\ 1 & -2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \implies \begin{cases} v_2 = -v_3 = -t \\ v_1 = 2v_2 + v_3 = -t \\ v_3 = t \end{cases}$$

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = t \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \implies E_3 = \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

Con respecto a la matriz B , que tiene los mismos autovectores asociados a los autovalores correspondientes, se tiene:

$$E_{12} = \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

8.2.4. Multiplicidades Algebraica y Geométrica

En una matriz puede haber varios *autovalores* iguales, y asociados a ellos pueden existir uno o más *autovectores*. La cantidad de autovalores y autovectores asociados a un determinado valor se denominan respectivamente *multiplicidad algebraica* y *multiplicidad geométrica*.

Definición 8.5. Multiplicidad Algebraica y Geométrica

- **Multiplicidad algebraica** de un *autovalor* es la *multiplicidad*¹⁵ de su raíz en el *polinomio característico*.
- **Multiplicidad geométrica** de un *autovalor* es la *dimensión*¹⁶ del *autoespacio* asociado al *autovalor*.

Todo autovalor, siempre tiene al menos un autovector asociado (multiplicidad geométrica 1).

En el caso de que un autovalor aparezca repetido $k > 1$ veces, (multiplicidad algebraica k), el número de autovectores asociados a él (multiplicidad geométrica del autovector) siempre será mayor o igual que 1 y menor o igual que la multiplicidad algebraica.

$$1 \leq \text{multiplicidad geométrica} \leq \text{multiplicidad algebraica}$$

Ejemplo 8.1. Multiplicidad algebraica y geométrica.

Las siguientes matrices tienen un único autovalor con multiplicidad algebraica 3, pero multiplicidades geométricas 1, 2 y 3 respectivamente.

$$\begin{aligned} \bullet \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \quad \begin{cases} \lambda_{1,2,3} = 1 & \rightarrow \text{mul. alg.} = 3 \\ E_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} & \rightarrow \text{mul. geo.} = 1 \end{cases} \\ \\ \bullet \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \quad \begin{cases} \lambda_{1,2,3} = 1 & \rightarrow \text{mul. alg.} = 3 \\ E_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} & \rightarrow \text{mul. geo.} = 2 \end{cases} \\ \\ \bullet \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \quad \begin{cases} \lambda_{1,2,3} = 1 & \rightarrow \text{mul. alg.} = 3 \\ E_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} & \rightarrow \text{mul. geo.} = 3 \end{cases} \end{aligned}$$

¹⁵La *multiplicidad* de una solución en una ecuación es el número de veces que aparece ese valor como solución. Geométricamente tiene que ver con la *concavidad* de la función que define la ecuación cerca de esa raíz.

¹⁶Número de vectores linealmente independientes.

Ejemplo 8.2. Multiplicidades algebraicas y geométricas de los autovalores del **Ejemplo 8.4** y **Ejemplo 8.5**.

- **Ejemplo 8.4:** el autovalor $\lambda = 1$ tiene *multiplicidad algebraica 2* y *multiplicidad geométrica 1*, mientras que el autovalor $\lambda = 2$ tiene *multiplicidad algebraica 1* y *multiplicidad geométrica 1*.
- **Ejemplo 8.5:** el autovalor $\lambda = 0$ tiene *multiplicidad algebraica 2* y *multiplicidad geométrica 2*, mientras que el autovalor $\lambda = 12$ tiene *multiplicidad algebraica 1* y *multiplicidad geométrica 1*.

Ejercicio 8.6. Encontrar los autovalores y sus multiplicidades de la siguiente matriz:

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$$

Empezamos calculando los autovalores de la matriz.

$$\det(A - \lambda I) = \begin{vmatrix} -1 - \lambda & 0 & 1 \\ 3 & -\lambda & -3 \\ 1 & 0 & -1 - \lambda \end{vmatrix} = -\lambda(1 + \lambda)^2 + 0 + 0 + \lambda - 0 - 0 = -\lambda^2(\lambda + 2) = 0$$

$$-\lambda^2(\lambda + 2) = 0 \implies \begin{cases} \lambda_1 = \lambda_2 = 0 & (\text{multiplicidad algebraica 2}) \\ \lambda_3 = -2 & (\text{multiplicidad algebraica 1}) \end{cases}$$

Vamos a calcular los autoespacios asociados a estos autovalores.

$$\underline{\lambda_1 = \lambda_2 = 0}$$

$$\begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix} \xrightarrow[F_3 + F_1]{F_2 + 3F_1} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \implies \begin{cases} v_1 = v_3 = t \\ v_2 = s \\ v_3 = t \end{cases} \implies \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$E_0 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \implies \text{La multiplicidad algebraica del autovalor } \lambda = 0 \text{ es } 2$$

$$\underline{\lambda_3 = -2}$$

$$\begin{bmatrix} 1 & 0 & 1 \\ 3 & 2 & -3 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow[F_3 + F_1]{F_2 + 3F_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{cases} v_1 = -v_3 = -t \\ v_2 = 3v_3 = 3t \\ v_3 = t \end{cases} \implies \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = t \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$$

$$E_{-2} = \left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right\} \implies \text{La multiplicidad algebraica del autovalor } \lambda = -2 \text{ es } 1$$

8.2.5. Autovalores y Matrices Inversas

Teorema 8.1.

Una matriz *cuadrada* A es *invertible* si y solo si 0 **no** es un *autovalor* de A .

Demostración 8.1.

Una matriz A es invertible si y solo si $\det(A) \neq 0$. Si 0 fuera un *autovalor* de esta matriz se verificaría:

$$\det(A - 0I) = \det(A) = 0 \quad \xrightarrow{\text{¡Contradicción!}} \quad \lambda = 0 \text{ no puede ser autovalor de una matriz invertible.}$$

Teorema Fundamental de las Matrices Invertibles

Vamos a ampliar el *Teorema Fundamental de las Matrices Invertibles* que vimos en el [Teorema 3.8, página 117](#).

Teorema 8.2. Teorema Fundamental de las Matrices Invertibles

Sea $A \in \mathbb{K}^{n \times n}$ y sea $T : V \longrightarrow W$ una transformación lineal definida por A . Entonces las siguientes afirmaciones son equivalentes:

- A es invertible.
- $A\vec{x} = \vec{b}$ tiene *solución única* $\forall \vec{b} \in \mathbb{K}^n$.
- $A\vec{x} = \vec{0}$ tiene solo la *solución trivial* $\vec{x} = \vec{0} \in \mathbb{K}^n$.
- La forma *escalonada reducida equivalente* de A es I_n .
- $\text{rang}(A) = n$.
- $\ker(T) = \ker(A) = \{\vec{0}\}$.
- T es biyectiva.
- Los vectores *columna* de A son *linealmente independientes* y forman una *base* del espacio $\mathbb{K}^n$.
- Los vectores *fila* de A son *linealmente independientes* y forman una *base* del espacio $\mathbb{K}^n$.
- $\det(A) \neq 0$.
- 0 no es un *autovalor* de A .

8.2.6. Teoremas sobre Autovalores y Autovectores

Teorema 8.3.

Sea $A \in \mathbb{K}^{n \times n}$ con *autovalor* λ y $\vec{v}$ un *autovector* asociado.

- Si A es invertible, entonces $1/\lambda$ es un *autovalor* de A^{-1} con *autovector* asociado $\vec{v}$.
- Para todo entero k , λ^k es un *autovalor* de la matriz A^k con *autovector* asociado $\vec{v}$.

Demostración 8.3.

- Sea $A\vec{v} = \lambda\vec{v}$, con A invertible.

$$A^{-1}A\vec{v} = A^{-1}\lambda\vec{v} \quad ; \quad \vec{v} = A^{-1}\lambda\vec{v} = A^{-1}\vec{v}\lambda \quad ; \quad \lambda^{-1}\vec{v} = A^{-1}\vec{v}$$

Por tanto, λ^{-1} es un *autovalor* de A^{-1} con *autovector* asociado $\vec{v}$.

- Sea $A\vec{v} = \lambda\vec{v}$. Entonces multiplicando reiteradamente ambas partes de la igualdad por A se tiene:

$$A^2\vec{v} = \lambda A\vec{v} = \lambda^2\vec{v} \quad ; \quad \dots \quad ; \quad A^k\vec{v} = \lambda^k\vec{v}$$

Por tanto, λ^k es un *autovalor* de A^k con $\vec{v}$ *autovector* asociado. Como también se cumple $A^{-1}\vec{v} = \lambda^{-1}\vec{v}$ se puede realizar la misma demostración para $A^{-k}\vec{v} = \lambda^{-k}\vec{v}$, siendo $A^{-k} = (A^{-1})^k$.

Teorema 8.4.

Sea $A \in \mathbb{K}^{n \times n}$ y sean $\lambda_1, \dots, \lambda_m$ *autovalores distintos* de A con *autovectores* asociados $\vec{v}_1, \dots, \vec{v}_m$. Entonces $\vec{v}_1, \dots, \vec{v}_m$ son *linealmente independientes*.

Teorema 8.5.

Sea $A \in \mathbb{K}^{n \times n}$ y sean $\lambda_1, \dots, \lambda_m$ *autovalores* de A con *autovectores* asociados $\vec{v}_1, \dots, \vec{v}_m$. Sea un vector $\vec{x}$ combinación lineal de los *autovectores*, tal que:

$$\vec{x} = c_1 \vec{v}_1 + \dots + c_m \vec{v}_m$$

Entonces para cualquier entero k se tiene:

$$A^k \vec{x} = c_1 \lambda_1^k \vec{v}_1 + \dots + c_m \lambda_m^k \vec{v}_m$$

Nota 8.3.

- La demostración del **Teorema 8.4** no verá aquí debido a su longitud. Puede consultarse en la bibliografía.
- La demostración del **Teorema 8.5** fue vista en la introducción del presente tema.

Ejercicio 8.7. Dada la matriz:

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} \in \mathbb{R}^{3 \times 3}.$$

Calcular $A^{23} \vec{x}$ donde $\vec{x} = [0, -1, 2]^T$.

Calculamos los autovalores y autoespacios asociados de A :

$$\begin{vmatrix} 2-\lambda & 1 & -1 \\ 1 & 2-\lambda & 1 \\ -1 & 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^3 - 2 - 3(2-\lambda) = 0$$

$$t^3 - 3t - 2 = 0 \xrightarrow{\text{Ruffini}} t = \{-1, -1, 2\} \implies \lambda = 2 - t = \{3, 3, 0\}$$

$$\underline{\lambda=3}: \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix} \rightarrow \begin{cases} v_1 = t \\ v_2 = v_1 + v_3 = t + s \\ v_3 = s \end{cases} ; E_3 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$$\underline{\lambda=0}: \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 3 & 3 \\ 1 & 2 & 1 \\ 0 & 3 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{cases} v_1 = -2v_2 - v_3 = t \\ v_2 = -v_3 = -t \\ v_3 = t \end{cases} ; E_0 = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

Encontramos ahora la combinación lineal de los autovectores para dar el vector $\vec{x}$.

$$\alpha \vec{v}_1 + \beta \vec{v}_2 + \gamma \vec{v}_3 = \vec{x}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & -1 \\ 0 & 1 & 1 & 2 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & -1 \\ 0 & 1 & 1 & 2 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 3 & 3 \end{array} \right] \rightarrow \begin{cases} \alpha = -\gamma = -1 \\ \beta = -1 + 2\gamma = 1 \\ \gamma = 1 \end{cases}$$

Por tanto:

$$A^{23} \vec{x} = A^{23}(\alpha \vec{v}_1 + \beta \vec{v}_2 + \gamma \vec{v}_3) = \alpha A^{23} \vec{v}_1 + \beta A^{23} \vec{v}_2 + \gamma A^{23} \vec{v}_3 = \alpha \lambda_1^{23} \vec{v}_1 + \beta \lambda_2^{23} \vec{v}_2 + \gamma \lambda_3^{23} \vec{v}_3 =$$

$$= -3^{23} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 3^{23} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 3^{23} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

8.3. Propiedades de los Autovalores de una Matriz

Vamos a resumir las propiedades¹⁷ más interesantes de los autovalores de una matriz $A \in \mathbb{R}^{n \times n}$.

- Las matrices A y A^T tienen los mismos autovalores¹⁸
- El número de autovalores nulos de A es mayor o igual a $n - r$, siendo $r = \text{rang}(A)$.
- La matriz A es invertible si y solo si ninguno de sus autovalores es nulo.
- Si λ es autovalor de A entonces $\lambda - k$ es autovalor¹⁹ de $A - kI$, siendo $k \in \mathbb{R}$.
- El *producto* de todos los autovalores de A coincide con el *determinante* de A .
- La *suma* de todos los autovalores de A coincide con la *traza*²⁰ de A .
- Los autovalores de A^{-1} son los inversos de los autovalores²¹ de A .
- Los autovalores de cA para todo $c \in \mathbb{R}$ son los autovalores²² de A multiplicados por c .
- Los autovalores de A^k con $k \in \mathbb{R}$ son las potencias k -ésimas de los autovalores de A .
- Si A es triangular sus autovalores coinciden con los elementos de su diagonal principal.
- Si A es simétrica sus autovalores son siempre números reales.
- Si A es ortogonal sus autovalores son 1 ó -1 .

Ejercicio 8.8. Encontrar los autovalores de una matriz A de dimensión 3, tal que $\text{rang}(A) = 3$, $\text{tr}(A) = 3$ y $|A| = -4$.

- Como es una matriz cuadrada de 3×3 , A posee tres autovalores (iguales o distintos): λ_1 , λ_2 y λ_3 .
- Como el rango de A coincide con su dimensión, dicha matriz es invertible y por tanto, no posee autovalores nulos.
- Como la traza de una matriz coincide con la suma de todos sus autovalores:

$$\lambda_1 + \lambda_2 + \lambda_3 = 3$$

- Como el determinante de una matriz coincide con el producto de sus autovalores se tiene:

$$\lambda_1 \lambda_2 \lambda_3 = -4$$

Por tanto, una posibilidad podría ser:

$$\lambda_1 = \lambda_2 = 2, \quad \lambda_3 = -1$$

¹⁷No se verán las correspondientes demostraciones debido a la extensión y profundidad de alguna de ellas.

¹⁸Aunque generalmente sus autovectores son distintos.

¹⁹Si $\vec{v}$ es el autovector asociado al autovalor λ de A , entonces también $\vec{v}$ es el autovector asociado al autovalor $\lambda - k$ de $A - kI$.

²⁰Ver la **Definición 3.3**, página 90.

²¹Los autovectores de A y A^{-1} coinciden.

²²Sin embargo los autovectores de A y cA coinciden para $c \neq 0$.

8.4. Matrices Semejantes

8.4.1. Matrices Triangulares

Teorema 8.6.

Los *autovalores* de una *matriz triangular*²³ (superior, inferior o diagonal) son los elementos de su *diagonal principal*.

Demostración 8.6.

Si A es triangular, su determinante es el producto de sus elementos diagonales²⁴. Como la matriz $(A - \lambda I)$ sigue siendo triangular, su determinante es:

$$\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = \prod_{i=1}^n (a_{ii} - \lambda) = 0 \quad \begin{cases} \lambda_1 = a_{11} \\ \vdots \\ \lambda_n = a_{nn} \end{cases}$$

y por tanto, las soluciones de la *ecuación característica* serán los n elementos diagonales de la matriz.

Ya que las *matrices triangulares* muestran los *autovalores* de una matriz en su diagonal principal, es muy interesante poder relacionar una matriz *cuadrada* cualquiera con una matriz *triangular equivalente* que tenga los mismos *autovalores*.

Mediante el método de eliminación gaussiana se puede obtener una *matriz triangular equivalente* de una matriz dada²⁵. Sin embargo, esta matriz *equivalente por filas* **no** conserva el valor del *determinante* y por tanto tampoco conserva los *autovalores*.

En esta sección vamos a estudiar una nueva *transformación matricial* que sí conserva los *autovalores*.

8.4.2. Matrices Semejantes

Definición 8.6. Matrices Semejantes

Sean $A, B \in \mathbb{K}^{n \times n}$. Se dice que A es **semejante a** B si existe una matriz (real o compleja) $P \in \mathbb{K}^{n \times n}$ invertible tal que,

$$AP = PB \quad \implies \quad \begin{cases} A = PBP^{-1} \\ B = P^{-1}AP \end{cases}$$

La semejanza de A a B se representa²⁶ por $A \sim B$.

Nota 8.4. La matriz P que aparece en la fórmula de semejanza de dos matrices **no** es única.

Teorema 8.7. Características de la Semejanza

Sean $A, B, C \in \mathbb{K}^{n \times n}$, entonces:

- $A \sim A$.
- Si $A \sim B$ entonces $B \sim A$.
- Si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

²³Ver la Sección 3.1.1, página 89.

²⁵Ver la Sección 1.5.4, página 50.

²⁶El símbolo $\sim$ se suele utilizar también para representar matrices *equivalentes por filas*.

²⁶Ver la Sección 3.9.7, página 107.

Demostración 8.7.

- Esta propiedad es consecuencia de: $I^{-1}AI = A$.
- Si $A \sim B$ entonces:

$$AP = PB \implies A = PBP^{-1} \implies P^{-1}A = BP^{-1} \implies B \sim A$$

- $\begin{cases} A \sim B \implies AP = PB & ; & A = PBP^{-1} \\ B \sim C \implies BM = MC & ; & B = MCM^{-1} \end{cases}$

$$A = PBP^{-1} = P(MCM^{-1})P^{-1} = (PM)C(PM)^{-1} = RCR^{-1} \implies AR = RC \implies A \sim C$$

Propiedades de las Matrices Semejantes**Teorema 8.8. Propiedades de las Matrices Semejantes**

Sean $A, B \in \mathbb{K}^{n \times n}$ con $A \sim B$, entonces:

1. $\det(A) = \det(B)$.
2. A y B tienen el mismo *polinomio característico* y por tanto los mismos *autovalores*.
3. A es *invertible* si y solo si B también lo es.
4. $\text{nul}(A) = \text{nul}(B) \implies \text{rang}(A) = \text{rang}(B)$.

Nota 8.5. Los autovectores $\vec{v}$ de A y B no tienen porqué ser iguales.

Demostración 8.8.

1. Como $A = PBP^{-1}$,

$$\det(A) = \det(P) \det(B) \det(P^{-1}) = \det(B) \det(P) \det(P^{-1}) = \det(B)$$

2. Sea $p_A(\lambda)$ el polinomio característico de A y como $A = PBP^{-1}$ se tiene,

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det(PBP^{-1} - \lambda I) = \det(PBP^{-1} - \lambda PIP^{-1}) = \det(P(B - \lambda I)P^{-1}) = \\ &= \det(P) \det(B - \lambda I) \det(P^{-1}) = \det(P) \det(P^{-1}) \det(B - \lambda I) = \det(B - \lambda I) = p_B(\lambda) \end{aligned}$$

3. Como $A = PBP^{-1}$, si existe A^{-1} debe ser: $A^{-1} = (PBP^{-1})^{-1} = PB^{-1}P^{-1}$

4. Sea un vector $\vec{x} \in \ker(B) \implies B\vec{x} = \vec{0}$, entonces como $AP = PB$:

$$AP\vec{x} = PB\vec{x} \ ; \ A(P\vec{x}) = P(B\vec{x}) = \vec{0} \implies P\vec{x} \in \ker(A)$$

Esto indica que por cada vector $\vec{x}$ que anula $B\vec{x}$ existe al menos otro vector $P\vec{x}$ que anula $A(P\vec{x})$. Por tanto, $\dim(\ker(B)) = p \leq \dim(\ker(A))$. Por otra parte, sea un vector $\vec{y} \in \ker(A) \implies A\vec{y} = \vec{0}$, entonces como $BM = MA$, siendo $M = P^{-1}$ se tiene:

$$BM\vec{y} = MA\vec{y} \ ; \ B(M\vec{y}) = M(A\vec{y}) = \vec{0} \implies M\vec{y} \in \ker(B)$$

Por tanto, $\dim(\ker(A)) = r \leq \dim(\ker(B))$. En definitiva, $\text{nul}(A) = \dim(\ker(A)) = \dim(\ker(B)) = \text{nul}(B)$. Como $\text{rang}(X) + \text{nul}(X) = n$, finalmente:

$$\text{rang}(A) = \text{rang}(B)$$

Dos matrices *semejantes* son siempre matrices *equivalentes por filas*, aunque no necesariamente al contrario.

8.4.3. Interpretación de la Semejanza de Matrices

Como hemos visto, dos matrices $A, B \in \mathbb{K}^{n \times n}$ son *semejantes* si existe una matriz *invertible* $P \in \mathbb{K}^{n \times n}$ tal que cumple $AP = PB$. También hemos comprobado que ambas matrices A y B tienen muchas propiedades en común. Pero uno podría preguntarse qué representa la matriz P y porqué en la ecuación multiplica a una matriz por la izquierda y a otra por la derecha.

Vamos a dar un sentido a esta fórmula y vamos a explicar qué pueden representar cada una de las matrices A , B y P , utilizando para ello las *transformaciones lineales*²⁷ y los *cambios de base*²⁸.

Como se vio en la [Sección 5.2.2, página 158](#), si tenemos una transformación lineal²⁹ $T : \mathbb{K}^n \rightarrow \mathbb{K}^n$ definida mediante una matriz $A \in \mathbb{K}^{n \times n}$ de la siguiente forma:

$$\vec{y} = T(\vec{x}) = A\vec{x}$$

la cual está definida en la base canónica. Queremos encontrar la forma que tiene esta TL en una nueva base $\mathcal{B}$ de $\mathbb{K}^n$ con matriz asociada³⁰ P de forma:

$$\vec{y}' = T'(\vec{x}') = B\vec{x}'$$

donde:

- T' es la nueva transformación lineal equivalente a T en la nueva base.
- $\vec{x}'$ y $\vec{y}'$ son respectivamente los vectores del dominio y de la imagen en la nueva base $\mathcal{B}$
- B es la matriz asociada a la nueva transformación lineal T' .

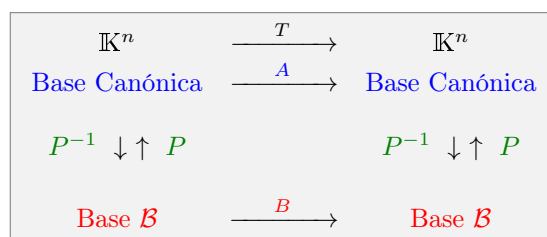
Entonces se tiene que cumplir:

$$\vec{y} = A\vec{x} \quad \implies \quad P\vec{y}' = AP\vec{x}' \quad \implies \quad \vec{y}' = P^{-1}AP\vec{x}' \quad \implies \quad \boxed{B = P^{-1}AP}$$

Las matrices A y B que definen una transformación lineal (endomorfismo) en bases distintas, A en la base canónica y B en una base con matriz asociada P , son *matrices semejantes*, siendo la relación entre ellas:

$$AP = PB$$

De forma esquemática el proceso se puede expresar de la siguiente forma:



²⁷Ver el [Tema 4, página 119](#).

²⁸Ver el [Tema 5, página 143](#).

²⁹En concreto es un *endomorfismo*, ver la [Sección 4.6.2, página 142](#).

³⁰Ver la [Sección 5.2.1, página 153](#).

Ejercicio 8.9. Sea el endomorfismo³¹ de $\mathbb{R}^3$:

$$F(x, y, z) = [x + y - z, -z, y + z]$$

- a) Encontrar la matriz A asociada a la transformación lineal F .
 b) Encontrar la matriz B asociada a F en la base $\mathcal{B} = \{[1, 0, 1], [1, 1, 0], [0, 1, 1]\}$.

$$\vec{x} \xrightarrow{A} \vec{y}$$

Base Canónica

$$[\vec{x}]_{\mathcal{B}} \xrightarrow{B} [\vec{y}]_{\mathcal{B}}$$

Base $\mathcal{B}$

- c) ¿Qué relación hay entre las matrices A y B ?

- a) Vamos a expresar la función en forma de sistema de ecuaciones:

$$\vec{y} = F(\vec{x}) = A\vec{x} \quad ; \quad \begin{cases} y_1 = x + y - z \\ y_2 = -z \\ y_3 = y + z \end{cases} \implies \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

- b) Para cualquier vector $\vec{v} \in \mathbb{R}^3$ se cumple $\vec{v} = P_{\mathcal{B}}[\vec{v}]_{\mathcal{B}}$. Sean $\vec{x}, \vec{y} \in \mathbb{R}^3$ dos vectores pertenecientes al dominio e imagen de F , entonces:

$$\left. \begin{array}{l} \vec{x} = P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} \\ \vec{y} = P_{\mathcal{B}}[\vec{y}]_{\mathcal{B}} \end{array} \right\} \implies \vec{y} = A\vec{x} \implies P_{\mathcal{B}}[\vec{y}]_{\mathcal{B}} = A(P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}) \implies [\vec{y}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

Por tanto la matriz B asociada a la función F en la base $\mathcal{B}$ es:

$$[\vec{y}]_{\mathcal{B}} = B[\vec{x}]_{\mathcal{B}} \implies B = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}}$$

donde $P_{\mathcal{B}}$ es la matriz asociada a la base $\mathcal{B}$:

$$P_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Tenemos que calcular la inversa de esta matriz:

$$\begin{aligned} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 & 1 & 1 \end{array} \right] \rightarrow \\ &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1/2 & 1/2 & 1/2 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1/2 & 1/2 & -1/2 \\ 0 & 0 & 1 & -1/2 & 1/2 & 1/2 \end{array} \right] \rightarrow \\ &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 1 & 0 & 1/2 & 1/2 & -1/2 \\ 0 & 0 & 1 & -1/2 & 1/2 & 1/2 \end{array} \right] \implies P_{\mathcal{B}}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \end{aligned}$$

Por tanto:

$$B = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}} = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & 3 & 3 \\ -2 & 1 & -3 \\ 0 & -1 & 1 \end{bmatrix}$$

- c) Son matrices semejantes

8.5. Diagonalización

Las matrices *diagonales* presentan grandes ventajas desde el punto de vista matricial:

- Son sencillas y necesitan poca memoria para su almacenamiento en un ordenador³².
- Si existe, su *inversa* es muy fácil de calcular (invertir los elementos diagonales).
- Su *determinante* es inmediato: $|D| = \prod d_{ii}$.
- Sus *autovalores* coinciden con los elementos *diagonales*.

Por tanto, es interesante encontrar una matriz *diagonal* que sea *semejante* a una matriz dada: $A \sim D$.

Definición 8.7. Matriz Diagonalizable

Una matriz $A \in \mathbb{K}^{n \times n}$ se dice que es **diagonalizable** si existe una matriz *diagonal* $D \in \mathbb{K}^{n \times n}$ tal que A sea *semejante* a D , es decir, que exista una matriz $P \in \mathbb{K}^{n \times n}$ *invertible* tal que:

$$AP = PD \quad \implies \quad A = PDP^{-1}$$

Teorema 8.9. Diagonalización de una Matriz

Una matriz $A \in \mathbb{K}^{n \times n}$ es *diagonalizable* si y solo si tiene n *autovectores linealmente independientes*, verificándose:

$$P^{-1}AP = D$$

donde los *elementos diagonales* de D son los *autovalores* de A y las *columnas* de P son los n *autovectores* asociados *linealmente independientes*, en el mismo orden que los *autovalores* correspondientes.

Demostración 8.9.

Sea A semejante a una matriz diagonal D mediante $AP = PD$. Sean $\vec{v}_1, \dots, \vec{v}_n$ las columnas de P y $\lambda_1, \dots, \lambda_n$ los elementos diagonales de A , entonces:

$$AP = PD \implies A \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$\begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \end{bmatrix} \quad \left\{ \begin{array}{l} A\vec{v}_1 = \lambda_1\vec{v}_1 \\ \vdots \\ A\vec{v}_n = \lambda_n\vec{v}_n \end{array} \right.$$

Igualando las columnas: $A\vec{v}_i = \lambda_i\vec{v}_i, \forall i = 1, \dots, n$. Lo que demuestra que los vectores columna de P son los *autovectores* de A siendo los *autovalores* asociados, los elementos diagonales de D en el mismo orden.

Como P es invertible, sus columnas son linealmente independientes. De igual forma se demuestra $\Leftarrow$, comenzando por los *autovalores* y *autovectores* de A .

³¹Ver la Sección 4.6.2, página 142.

³²Basta con almacenar únicamente los elementos de su diagonal.

Ejercicio 8.10. En caso de ser posible, encontrar una matriz P que *diagonalice* la siguiente matriz A :

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$$

Del ejemplo [Ejemplo 8.4](#) sabemos que sus *autovalores* son $\lambda_1 = \lambda_2 = 1$ y $\lambda_3 = 2$, siendo sus *autoespacios*:

$$E_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad ; \quad E_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \right\}$$

Como A no tiene 3 autovectores *linealmente independientes*, por el [Teorema 8.9](#), A **no es diagonalizable** y por tanto **no existe** la matriz P pedida.

Ejercicio 8.11. Si es posible, encontrar una matriz P que *diagonalice* la siguiente matriz A :

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$$

Del [Ejemplo 8.6](#) sabemos que sus *autovalores* son $\lambda_1 = \lambda_2 = 0$ y $\lambda_3 = -2$, siendo sus *autovectores*:

$$E_0 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \quad ; \quad E_{-2} = \left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right\}$$

Es fácil comprobar que los tres *autovectores* son linealmente independientes y se pueden formar con ellos una matriz P **invertible** de la forma:

$$P = [\vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3] = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix}$$

Esta matriz sirve para diagonalizar la matriz A :

$$AP = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & -6 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} = PD$$

Nota 8.6.

- El orden no importa, pero las matrices P y D serán distintas en función del orden elegido. Por ejemplo, en el caso anterior se podría haber elegido:

$$P = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad ; \quad D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$P = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad ; \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Verificar $AP = PD$ a mano es mucho más sencillo que verificar $A = PDP^{-1}$ ya que no es necesario calcular P^{-1} .

³³El conjunto de todos los *autoespacios* de A contiene n *autovectores*.

³⁴Ver la [Definición 8.5](#), página 225.

Teorema 8.10. Teorema de la Diagonalización

Sea $A \in \mathbb{K}^{n \times n}$ con *autovalores* distintos $\lambda_1, \dots, \lambda_k$ con $k \leq n$. Entonces, los siguientes enunciados son equivalentes:

- A es *diagonalizable*.
- A tiene en total n *autovectores linealmente independientes*³³.
- La *multiplicidad algebraica* de cada *autovalor* es igual a su *multiplicidad geométrica*³⁴.

Demostración 8.10.

Trivial ya que con el conjunto de autovectores formamos una matriz P **invertible**.

8.5.1. Autovalores Reales y Complejos

Los *autovalores* de una matriz *real* pueden ser reales o complejos, al igual que sus *autovectores*.

Ejercicio 8.12. Calcular los *autovalores* y *autovectores* de la matriz *real*:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Calculamos primero los autovalores de la matriz.

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0 \implies \lambda = \pm\sqrt{-1} = \pm i \quad \left\{ \begin{array}{l} \lambda_1 = i \\ \lambda_2 = -i \end{array} \right.$$

Con ellos calculamos los autovectores asociados.

$$\underline{\lambda_1 = i}$$

$$\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \xrightarrow{F_1 + iF_2} \begin{bmatrix} 0 & 0 \\ 1 & -i \end{bmatrix} \implies \left\{ \begin{array}{l} v_2 = t \\ v_1 = iv_2 = it \end{array} \right\} \implies \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = t \begin{bmatrix} i \\ 1 \end{bmatrix} \implies E_i = \left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}$$

$$\underline{\lambda_1 = -i}$$

$$\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \xrightarrow{F_1 - iF_2} \begin{bmatrix} 0 & 0 \\ 1 & i \end{bmatrix} \implies \left\{ \begin{array}{l} v_2 = t \\ v_1 = -iv_2 = -it \end{array} \right\} \implies \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = t \begin{bmatrix} -i \\ 1 \end{bmatrix} \implies E_{-i} = \left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\}$$

Ejercicio 8.13. Encontrar, en caso de que exista, una base de $\mathbb{R}^3$ donde la matriz A asociada a la transformación lineal F del **Ejemplo 8.9** sea diagonal.

$$F(x, y, z) = [x + y - z, -z, y + z]$$

La matriz asociada a esta transformación lineal era:

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

Como se vio en la solución del problema, la matriz B asociada a la función F en una base $\mathcal{B}$ cumplía:

$$B = P_{\mathcal{B}}^{-1} A P_{\mathcal{B}}$$

Por tanto, en caso de que exista dicha matriz (base), corresponderá con la diagonalización de la matriz A . Vamos a ver si es posible, encontrando los autovalores y autovectores de A .

$$\begin{vmatrix} 1-\lambda & 1 & -1 \\ 0 & -\lambda & -1 \\ 0 & 1 & 1-\lambda \end{vmatrix} = -\lambda(1-\lambda)^2 + (1-\lambda) = 0 \quad ; \quad \lambda(1-\lambda)^2 = (1-\lambda) \quad \begin{cases} \lambda = 1 \\ \lambda(1-\lambda) = 1 \end{cases}$$

$$\lambda^2 - \lambda + 1 = 0 \quad ; \quad \lambda = \frac{1 \pm \sqrt{1-4}}{2} \implies \text{No todas las soluciones son reales.}$$

Como algunos autovalores son complejos, sus autovectores asociados también serán complejos y por tanto, no existe una matriz **real** y diagonal semejante a A .

Por tanto, no existe una base de $\mathbb{R}^3$ donde la matriz asociada a F sea diagonal.

Hay un tipo de matrices reales que siempre tienen *autovalores* y *autovectores* reales: las matrices **reales simétricas**³⁵.

Teorema 8.11.

En toda matriz cuadrada **real y simétrica** sus *autovalores* y *autovectores* asociados son **reales**.

Demostración 8.11.

Un número complejo³⁶ $z = a + bi$ es real si $z = \bar{z} \Leftrightarrow$ parte imaginaria nula, $b = 0$.

Sea una matriz A real y simétrica, entonces $\bar{A} = A$ y sea λ un *autovalor* con *autovector* asociado $\mathbf{v}$ de A . Entonces $A\mathbf{v} = \lambda\mathbf{v}$.

Tomando complejos conjugados en ambas partes se tiene: $\bar{A}\bar{\mathbf{v}} = \bar{A}\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$.

Posteriormente, tomando transpuestas y teniendo en cuenta el hecho de que A es simétrica:

$$(A\bar{\mathbf{v}})^T = (\bar{\lambda}\bar{\mathbf{v}})^T \quad ; \quad \bar{\mathbf{v}}^T A = \bar{\lambda}\bar{\mathbf{v}}^T$$

Calculemos entonces cuanto vale $\lambda(\bar{\mathbf{v}}^T \mathbf{v})$,

$$\lambda(\bar{\mathbf{v}}^T \mathbf{v}) = \bar{\mathbf{v}}^T (\lambda\mathbf{v}) = \bar{\mathbf{v}}^T (A\mathbf{v}) = (\bar{\mathbf{v}}^T A)\mathbf{v} = (\bar{\lambda}\bar{\mathbf{v}}^T)\mathbf{v} = \bar{\lambda}(\bar{\mathbf{v}}^T \mathbf{v}) \implies (\lambda - \bar{\lambda})(\bar{\mathbf{v}}^T \mathbf{v}) = 0$$

Como $\mathbf{v}$ es un *autovector* y por tanto no es nulo, $\bar{\mathbf{v}}^T \mathbf{v} \neq 0$ y por tanto,

$$\lambda - \bar{\lambda} = 0 \quad ; \quad \lambda = \bar{\lambda} \implies \lambda \in \mathbb{R}$$

Si los *autovalores* son reales y la matriz A también lo es, las soluciones del sistema homogéneo (*autovectores* asociados) también serán reales.

8.6. Diagonalización Ortogonal

Como hemos visto, la diagonalización de una matriz permite encontrar matrices diagonales semejantes³⁷:

$$A = PDP^{-1} \quad ; \quad A \sim D$$

Frente a las claras ventajas que presenta sustituir una matriz diagonal D con respecto a una matriz cuadrada A , aparece además como contrapartida una matriz P y su inversa P^{-1} .

Resultaría muy interesante poder encontrar matrices que fueran diagonalizables a través de matrices P que fueran *ortogonales*, con todas las ventajas que ello conlleva³⁸.

³⁵Ver la Definición 3.13, página 97.

³⁷Siempre que ésta sea diagonalizable

³⁸Ver la Sección 6.3.4, página 171.

Definición 8.8. Diagonalización Ortogonal

Una matriz $A \in \mathbb{R}^{n \times n}$ se dice que es *diagonalizable ortogonalmente* si existe una matriz *ortogonal* Q y una matriz *diagonal* D tales que:

$$Q^T A Q = D$$

$$A = Q D Q^T \iff D = Q^T A Q$$

Teorema 8.12.

Si A es *simétrica* todos los *autovectores* asociados a *autovalores* distintos son *ortogonales* entre sí.

Demostración 8.12.

Sean $\vec{v}_1$ y $\vec{v}_2$ dos *autovectores* asociados a dos *autovalores* distintos $\lambda_1 \neq \lambda_2$, $A\vec{v}_1 = \lambda_1\vec{v}_1$, $A\vec{v}_2 = \lambda_2\vec{v}_2$. Teniendo en cuenta que A es simétrica, $A^T = A$ y que $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$, se tiene:

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1\vec{v}_1) \cdot \vec{v}_2 = (A\vec{v}_1) \cdot \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T (A\vec{v}_2) = \vec{v}_1^T (\lambda_2\vec{v}_2) = \lambda_2(\vec{v}_1 \cdot \vec{v}_2)$$

Por tanto, $(\lambda_1 - \lambda_2)(\vec{v}_1 \cdot \vec{v}_2) = 0$ y como $\lambda_1 \neq \lambda_2$ por tanto $\vec{v}_1 \cdot \vec{v}_2 = 0$.

Ejercicio 8.14. Verificar el resultado del **Teorema 8.12** con la matriz,

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

El *polinomio característico* de A es:

$$p_A(\lambda) = \det(A - \lambda I) = -\lambda^3 + 6\lambda^2 - 9\lambda + 4 = -(\lambda - 4)(\lambda - 1)^2 = 0$$

Los *autovalores* son $\lambda_1 = 4$, $\lambda_{2,3} = 1$ y los *autoespacios* correspondientes:

$$E_4 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad ; \quad E_1 = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Donde fácilmente se comprueba que todo *autovector* de E_4 , (en este caso solo uno), es ortogonal a todo *autovector* de E_1 .

Nota 8.7. En el **Ejemplo 8.6** se obtiene un conjunto de *autovectores* que son *linealmente independientes*, pero **no** son *ortogonales*. Por tanto, es *diagonalizable* pero **no** *ortogonalmente*.

8.6.1. Interpretación Geométrica de la Diagonalización Ortogonal

Vamos a ver desde un punto de vista geométrico en $\mathbb{R}^3$ la diferencia entre matrices no diagonalizables, diagonalizables no ortogonalmente y diagonalizables ortogonalmente.

Para ello, vamos a representar gráficamente el conjunto de autovectores y los correspondientes autoespacios asociados a los autovalores.

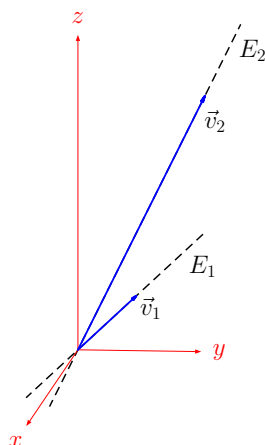
³⁸El *conjugado* de un número complejo $z = a + bi$ es el complejo $\bar{z} = a - bi$ (parte imaginaria cambiada de signo).

- Matriz no diagonalizable.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$$

Esta matriz se estudió en el [Ejemplo 8.10, página 235](#) y vimos que tenía tres autovalores $\lambda_1 = \lambda_2 = 1$ y $\lambda_3 = 2$ con autoespacios asociados:

$$E_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad ; \quad E_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \right\}$$



Cada uno de los autoespacios corresponde a una recta (dimensión 1), siendo los respectivos autovectores una base de estos subespacios vectoriales de $\mathbb{R}^3$.

Claramente los autovectores son linealmente independientes por estar asociados a autovalores distintos³⁹, pero como solamente hay dos, no tenemos un conjunto completo de vectores para formar una base de todo el espacio $\mathbb{R}^3$ y por tanto la matriz no es diagonalizable.

Por otra parte, como la matriz no es simétrica estos autovectores no son ortogonales⁴⁰.

En la gráfica se muestran los autoespacios (rectas) E_1 y E_2 y los autovectores (bases) $\vec{v}_1$ y $\vec{v}_2$.

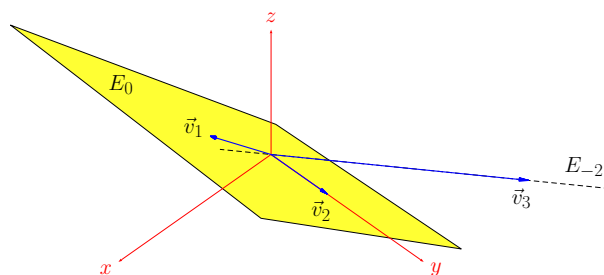
- Matriz diagonalizable no ortogonalmente.

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$$

Esta matriz se estudió en el [Ejemplo 8.6, página 226](#) y vimos que tenía tres autovalores $\lambda_1 = \lambda_2 = 0$ y $\lambda_3 = -2$ con autoespacios asociados:

$$E_0 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \quad ; \quad E_{-2} = \left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right\}$$

En este caso el autoespacio asociado al autovalor 0 es un *plano* y los dos autovectores correspondientes a este autovalor, $(\vec{v}_1$ y $\vec{v}_2)$, son una base de este subespacio vectorial de dimensión 2 en $\mathbb{R}^3$. Estos dos vectores pueden ser cualesquiera del plano, siempre que sean linealmente independientes, de forma que aunque no sean inicialmente ortogonales⁴¹ siempre se pueden ortogonalizar⁴² para formar un conjunto (base) ortogonal.



³⁹Ver el [Teorema 8.4, página 227](#).

⁴⁰Ver el [Teorema 8.12, página 238](#).

⁴¹En este caso sí lo son.

⁴²Mediante el proceso de Gram-Schmidt.

Por otra parte, el autoespacio asociado al autovalor -2 es una recta, siendo el autovector asociado a este autovalor, $(\vec{v}_3)$, una base de este subespacio vectorial de dimensión 1 en $\mathbb{R}^3$. Se puede tomar como autovector asociado a este autovalor, cualquier vector no nulo perteneciente a la recta.

Como en total se tiene un conjunto de tres autovectores linealmente independientes (no coplanarios) en $\mathbb{R}^3$, la matriz sí es diagonalizable.

Sin embargo, como la recta es fija y no ortogonal al plano⁴³, no forman un conjunto ortogonal de vectores y por tanto la matriz no es diagonalizable ortogonalmente.

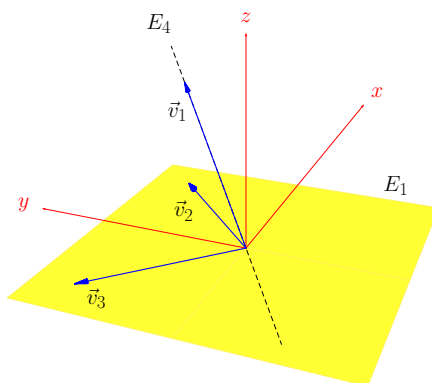
• **Matriz diagonalizable ortogonalmente.**

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Esta matriz simétrica se estudió en el [Ejemplo 8.14, página 238](#) y vimos que tenía tres autovalores $\lambda_1 = 4$, $\lambda_{2,3} = 1$ con autoespacios asociados:

$$E_4 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad ; \quad E_1 = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

En este caso el autoespacio asociado al autovalor 4 es una recta, siendo el autovector correspondiente, $(\vec{v}_1)$, una base de este subespacio vectorial. Este vector puede ser cualquiera de la recta.



Por otra parte, el autoespacio asociado al autovalor 1 es un plano y sus dos autovectores $\vec{v}_2$ y $\vec{v}_3$ son una base del mismo. Estos dos autovectores pueden ser cualesquiera del plano, siempre que sean linealmente independientes, de forma que aunque inicialmente no sean ortogonales⁴⁴ siempre se pueden ortogonalizar mediante el proceso de Gram-Schmidt, para formar un conjunto (base) ortogonal. Por ejemplo, $\vec{v}_3$ podría ser:

$$\vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_2 \cdot \vec{v}_3}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ -1/2 \end{bmatrix}$$

que pertenece al plano E_1 y es ortogonal a $\vec{v}_1$ y $\vec{v}_2$.

Como en total se tiene un conjunto de tres autovectores linealmente independientes, (no coplanarios), en $\mathbb{R}^3$, la matriz sí es diagonalizable.

Además como la matriz es simétrica, la recta E_4 siempre es ortogonal al plano⁴⁵ E_1 , formando un conjunto ortogonal de vectores. Por tanto toda matriz simétrica como esta es diagonalizable ortogonalmente⁴⁶.

⁴³O lo que es lo mismo, el autovector asociado al autovalor -2, $(\vec{v}_3)$, no es ortogonal a los autovectores asociados al autovalor 0, $(\vec{v}_1$ y $\vec{v}_2)$.

⁴⁴Como ocurre en este caso.

⁴⁵O lo que es lo mismo, el autovector asociado al autovalor 4, $(\vec{v}_1)$, es ortogonal a los autovectores asociados al autovalor 1, $(\vec{v}_2$ y $\vec{v}_3)$.

⁴⁶Ver el [Teorema 8.12, página 238](#).

8.6.2. Teorema Espectral

Definición 8.9. Espectro de una Matriz

Al conjunto de los *autovalores* de una matriz A se le conoce con el nombre de **espectro**⁴⁷ de la matriz.

Teorema 8.13. Teorema Espectral de las Matrices Ortogonales

Una matriz $A \in \mathbb{R}^{n \times n}$ es **ortogonalmente diagonalizable** si y solo si es **simétrica**.

Demostración 8.13.

Si A es ortogonalmente diagonalizable, existe una matriz ortogonal Q tal que:

$$AQ = QD \quad ; \quad A = QDQ^T \quad \implies \quad A^T = (QDQ^T)^T = QD^TQ^T = QDQ^T = A$$

De forma análoga, se puede demostrar que toda matriz simétrica es ortogonalmente diagonalizable.

8.6.3. Descomposición Espectral

Definición 8.10. Descomposición Espectral

Cualquier matriz *simétrica* $A \in \mathbb{R}^{n \times n}$ se puede expresar a través de una matriz *ortogonal* Q y una matriz *diagonal* D como:

$$\begin{aligned} A = QDQ^T &= [\vec{q}_1 \quad \cdots \quad \vec{q}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{q}_1^T \\ \vdots \\ \vec{q}_n^T \end{bmatrix} = \\ &= [\lambda_1 \vec{q}_1 \quad \cdots \quad \lambda_n \vec{q}_n] \begin{bmatrix} \vec{q}_1^T \\ \vdots \\ \vec{q}_n^T \end{bmatrix} = \boxed{\lambda_1 \vec{q}_1 \vec{q}_1^T + \lambda_2 \vec{q}_2 \vec{q}_2^T + \cdots + \lambda_n \vec{q}_n \vec{q}_n^T} \end{aligned}$$

A la obtención de este resultado se le denomina **descomposición espectral**⁴⁸ de A .

Ejercicio 8.15. Diagonalizar ortogonalmente la matriz siguiente y obtener su *descomposición espectral*.

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

En el ejemplo **Ejemplo 8.14** determinamos los *autovalores* y *autovectores* de A , siendo $\lambda_1 = 4$ y $\lambda_2 = \lambda_3 = 1$ (doble) y los *autoespacios* correspondientes:

$$E_4 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad ; \quad E_1 = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

⁴⁷La palabra **espectro** proviene del latín *spectrum* que significa *imagen*. Cuando los átomos son excitados, estos vibran y emiten luz, que es mezcla de distintas *frecuencias* y que corresponden a diferentes líneas brillantes en su *espectro*. Estas *frecuencias* se pueden obtener de forma matemática calculando los *autovalores* de un determinado *operador matricial*. De aquí deriva el nombre de *espectro* al conjunto de *autovalores* de una determinada matriz.

⁴⁸Cada uno de los términos $\vec{q}_i \vec{q}_i^T$ corresponde con la matriz asociada a la TL definida por la *proyección ortogonal* sobre el subespacio generado por $\vec{q}_i$, por esta razón a la *descomposición espectral* se la denomina a veces **forma de proyección del teorema espectral**.

Como necesitamos tres *autovectores* ortogonales, aplicando el proceso de Gram-Schmidt⁴⁹ a E_1 se obtiene:

$$E_1 = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ -1/2 \end{bmatrix} \right\}$$

Normalizando estos vectores podemos obtener una matriz Q *ortogonal* de la forma:

$$Q = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix} ; \quad \vec{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \vec{q}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \vec{q}_3 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$$

Haciendo las operaciones oportunas se puede comprobar que:

$$Q^T A Q = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Finalmente, su *descomposición espectral* será:

$$A = \lambda_1 \vec{q}_1 \vec{q}_1^T + \lambda_2 \vec{q}_2 \vec{q}_2^T + \lambda_3 \vec{q}_3 \vec{q}_3^T$$

8.7. Teoremas Avanzados Adicionales

Los dos teoremas que se presentan a continuación se han hecho famosos por formar parte del algoritmo de Google para la determinación rápida del *ranking* de páginas web.

Teorema 8.14. Teorema de Perron

Sea $A \in \mathbb{R}^{n \times n}$ una matriz *positiva* ($a_{ij} > 0$). Entonces A tiene un *autovalor* λ_1 real con las siguientes propiedades:

- $\lambda_1 > 0$.
- λ_1 tiene un *autovector* asociado positivo ($v_i > 0$).
- Para cualquier otro *autovalor* de A se verifica $|\lambda| < \lambda_1$.

Teorema 8.15. Teorema de Perron-Frobenius

Sea $A \in \mathbb{R}^{n \times n}$ una matriz *irreducible no negativa*. Entonces A tiene un *autovalor* λ_1 real con las siguientes propiedades:

- $\lambda_1 > 0$.
- λ_1 tiene un *autovector* asociado positivo.
- Para cualquier otro *autovalor* de A se verifica $|\lambda| \leq \lambda_1$.
- Cualquier *autovalor* (real o complejo) λ de A que verifique $|\lambda| = \lambda_1$, entonces λ es una raíz (compleja) de la ecuación $\lambda^n - \lambda_1^n = 0$.
- λ_1 tiene *multiplicidad algebraica* 1.

⁴⁹Ver la Sección 6.7, página 184.

8.8. Diagonalización Unitaria

Vamos a generalizar varios de los conceptos importantes vistos en los temas anteriores a matrices cuadradas complejas $A \in \mathbb{C}^{n \times n}$.

Definición 8.11. Matriz Normal

Una matriz $A \in \mathbb{C}^{n \times n}$ se dice que es **normal** si verifica:

$$AA^H = A^H A$$

Dentro de las *matrices normales* se encuentran:

- **Matrices hermíticas**⁵⁰ : $A^H = A$
- **Matrices simétricas reales**⁵¹ : $A^H = A^T = A$
- **Matrices unitarias**⁵² : $A^H = A^{-1}$
- **Matrices ortogonales**⁵³ : $A^H = A^T = A^{-1}$

Ejercicio 8.16. Comprobar que la siguiente matriz es *normal* pero que no es *unitaria*.

$$A = \begin{bmatrix} -i & -i \\ -i & i \end{bmatrix}$$

Todas las matrices unitarias son normales pero no necesariamente se cumple el recíproco. Comprobamos en este caso, primero que A es normal para ello debe cumplirse que $AA^H = A^H A$.

$$AA^H = \begin{bmatrix} -i & -i \\ -i & i \end{bmatrix} \begin{bmatrix} i & i \\ i & -i \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} i & i \\ i & -i \end{bmatrix} \begin{bmatrix} -i & -i \\ -i & i \end{bmatrix} = A^H A$$

Por tanto A sí es normal. Para que sea unitaria se tiene que cumplir que $A^{-1} = A^H$. Calcemos entonces la inversa de A .

$$\left[\begin{array}{cc|cc} -i & -i & 1 & 0 \\ -i & i & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & i & 0 \\ 0 & 2i & -1 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & i & 0 \\ 0 & 1 & i/2 & -i/2 \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & i/2 & i/2 \\ 0 & 1 & i/2 & -i/2 \end{array} \right]$$

$$A^{-1} = \frac{i}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \neq \begin{bmatrix} i & i \\ i & -i \end{bmatrix} = A^H$$

Por tanto A no es unitaria.

Definición 8.12. Diagonalización Unitaria

Una matriz $A \in \mathbb{C}^{n \times n}$ se dice que es *diagonalizable unitariamente* si existe una matriz *unitaria* U y una matriz *diagonal* D tales que:

$$A = U D U^H \iff D = U^H A U$$

⁵⁰Ver la Definición 3.15, página 99.

⁵¹Ver la Definición 3.13, página 97.

⁵²Ver la Definición 6.26, página 194.

⁵³Ver la Definición 6.10, página 171.

La *diagonalización ortogonal*⁵⁴ es un caso particular de la *diagonalización unitaria* para matrices reales.

Teorema 8.16. Teorema Espectral de las Matrices Unitarias

Una matriz $A \in \mathbb{C}^{n \times n}$ es **unitariamente diagonalizable** si y solo si es **normal**.

- En las *matrices normales* las multiplicidades algebraica y geométrica de sus autovalores siempre coinciden y por tanto sus autovectores asociados forman un conjunto ortogonal.
- En las *matrices hermiticas* sus autovalores son siempre reales.
- En las *matrices unitarias* sus autovalores siempre tienen módulo 1.

Ejercicio 8.17. En caso de ser posible, diagonalizar unitariamente la siguiente matriz $A \in \mathbb{C}^{2 \times 2}$.

$$A = \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix}$$

Comprobar el resultado obtenido.

Para que una matriz A sea diagonalizable unitariamente^a es necesario que A sea *normal*, es decir que $AA^H = A^H A$. Lo comprobamos:

$$AA^H = \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} = \begin{bmatrix} 2 & -2i \\ 2i & 2 \end{bmatrix} = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} = A^H A$$

Como efectivamente la matriz A es normal puede ser diagonalizable unitariamente, por tanto existe una matriz unitaria U (que cumple $U^{-1} = U^H$) y una matriz diagonal D que verifican:

$$AU = UD \implies A = UDU^H$$

La matriz U está formada por los autovectores de A normalizados y dispuestos como columnas, y la matriz D tiene en su diagonal principal los correspondientes autovalores de A .

Determinamos entonces los autovalores y autoespacios normalizados de A :

$$|A - \lambda I| = \begin{vmatrix} i - \lambda & 1 \\ -1 & i - \lambda \end{vmatrix} = (i - \lambda)^2 + 1 = 0 \quad ; \quad i - \lambda = \pm i \quad ; \quad \lambda = \{0, 2i\}$$

$$\begin{aligned} \lambda_1 = 0 : \quad & \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 \\ 1 & -i \end{bmatrix} \implies \begin{cases} v_1 = iv_2 = it \\ v_2 = t \end{cases} \implies \vec{v}_1 = t \begin{bmatrix} i \\ 1 \end{bmatrix} \implies \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix} \\ \lambda_2 = 2i : \quad & \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 \\ 1 & i \end{bmatrix} \implies \begin{cases} v_1 = -iv_2 = it \\ v_2 = t \end{cases} \implies \vec{v}_1 = t \begin{bmatrix} -i \\ 1 \end{bmatrix} \implies \frac{1}{\sqrt{2}} \begin{bmatrix} -i \\ 1 \end{bmatrix} \end{aligned}$$

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \quad ; \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 2i \end{bmatrix} \quad ; \quad AU = UD$$

Para verificar el resultado, tiene que cumplirse que $A = UDU^H$. Comprobamos:

$$UDU^H = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 2i \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} -i & 1 \\ i & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 2 \\ 0 & 2i \end{bmatrix} \begin{bmatrix} -i & 1 \\ i & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2i & 2 \\ -2 & 2i \end{bmatrix} = \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} = A$$

⁵⁴Ver la Definición 8.8, página 238.

^aVer la Definición 8.12, página 243.

8.9. Resumen Final del Tema

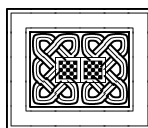
- Toda matriz *cuadrada* siempre tiene *autovalores* y *autovectores* asociados (reales o complejos).
- Si A es **real y simétrica** sus *autovalores* y *autovectores* siempre serán reales.
- En las matrices **triangulares** (superiores, inferiores o diagonales), los *autovalores* corresponden con los elementos de la *diagonal principal*: $\lambda_i = a_{ii}$.
- A es **invertible** $\iff$ **cero** no es un *autovalor* de A .
- Tres teoremas importantes:
 1. Los *autovectores* asociados a *autovalores* distintos son **linealmente independientes**.
 2. $A^k \vec{x} = \alpha_1 \lambda_1^k \vec{v}_1 + \dots + \alpha_n \lambda_n^k \vec{v}_n$ siendo $\vec{x} = \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n$.
 3. Si $A\vec{v} = \lambda\vec{v} \implies \begin{cases} A^{-1} & : \text{ autovectores } 1/\lambda, \text{ mismos autovectores de } A. \\ A^n & : \text{ autovalores } \lambda^n, \text{ mismos autovalores de } A. \end{cases}$
- **Matrices Semejantes**: $A \sim B \iff \exists P \text{ invertible} : AP = PB$.
Propiedades:
 - $p_A(\lambda) = p_B(\lambda) \implies |A| = |B|$
 - $\text{rang}(A) = \text{rang}(B) \quad ; \quad \text{nul}(A) = \text{nul}(B)$.
 - Si $\exists A^{-1} \implies B^{-1}$.
- **Diagonalización**. $\exists P : AP = PD$ si y solo si existen n *autovectores linealmente independientes* (multiplicidades $m_a = m_g$), donde $D = [\lambda_1 \dots \lambda_n]$ y $P = [\vec{v}_1 \dots \vec{v}_n]$. Entonces:

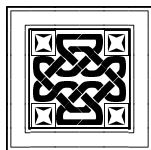
$$A = PDP^{-1} \quad ; \quad D = P^{-1}AP$$
- **Diagonalización Ortogonal (Teorema Espectral Ortogonal)**: $\exists Q$ *ortogonal* tal que: $AQ = QD$ si y solo si A es *simétrica*. Entonces:

$$A = QDQ^T \quad ; \quad D = Q^T A Q$$
- **Descomposición Espectral**. Si A es *real y simétrica*, existe Q *ortogonal* tal que:

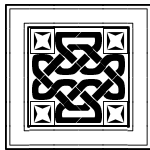
$$A = QDQ^T = [\vec{q}_1 \dots \vec{q}_n][\lambda_1 \dots \lambda_n][\vec{q}_1 \dots \vec{q}_n]^T = \lambda_1 \vec{q}_1 \vec{q}_1^T + \dots + \lambda_n \vec{q}_n \vec{q}_n^T$$
- **Diagonalización Unitaria (Teorema Espectral Unitario)**: $\exists U$ *unitaria* tal que: $AU = UD$ si y solo si A es *unitaria*. Entonces:

$$A = UDU^H \quad ; \quad D = U^H A U$$





APÉNDICES



Temas de Repaso

Contenido

A.1	Trigonometría	250
A.1.1	Medida de Ángulos: Radianes	250
A.1.2	Funciones Trigonométricas	251
A.2	Funciones Hiperbólicas	260
A.3	Productos Notables	262
A.4	Sumatorios	262
A.5	Binomio de Newton	263
A.6	Polinomios	264
A.6.1	Regla de Ruffini	266
A.7	Distancia Euclídea	269
A.7.1	Valor Absoluto	269
A.7.2	Distancia en $\mathbb{R}$	269
A.7.3	Distancia en $\mathbb{R}^2$	269
A.7.4	Distancia en $\mathbb{R}^3$	270
A.7.5	Distancia en $\mathbb{R}^n$	270
A.8	Vectores en el Plano y en el Espacio	270
A.8.1	Vectores de Posición	271
A.8.2	Vectores que no son de posición: Vectores Generales	271
A.8.3	Vectores Elementales	272
A.9	Vectores en Espacios $\mathbb{R}^n$, $\mathbb{C}^n$ y $\mathbb{K}^n$	272
A.9.1	Caracterización de Vectores	273
A.10	Producto Escalar y Vectorial en $\mathbb{R}^3$	274
A.10.1	Producto Escalar	274
A.10.2	Producto Vectorial	274
A.11	Fórmulas Geométricas sobre Distancias	275
A.11.1	Distancia de un Punto a un Recta en $\mathbb{R}^2$	275
A.11.2	Distancia entre dos Rectas paralelas en $\mathbb{R}^2$	275
A.11.3	Distancia de un Punto a un Plano en $\mathbb{R}^3$	275
A.11.4	Distancia entre dos Planos Paralelos en $\mathbb{R}^3$	275
A.11.5	Distancia de un Punto a un Recta en $\mathbb{R}^3$	276
A.11.6	Distancia entre dos Rectas paralelas en $\mathbb{R}^3$	276
A.12	Introducción a las Funciones	277
A.12.1	Tipos de Funciones	277
A.12.2	Representación Gráfica de Funciones Reales	278
A.12.3	Tipos de Funciones	279
A.12.4	Función Inversa (o Recíproca) $f^{-1}(x)$	280
A.12.5	Composición de Funciones	281

A.1. Trigonometría

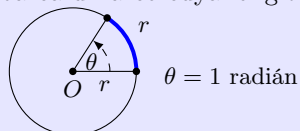
A.1.1. Medida de Ángulos: Radianes

En diversos campos de las matemáticas, los ángulos conviene medirlos en *radianes*.

Definición A.1. Radián

Un *radián* es el ángulo central de una circunferencia que abarca un arco cuya longitud es el radio.

$$1 \text{ radián} = \frac{360}{2\pi} \text{ grados} \approx 57.3^\circ$$



Relación Grados-Radianes

$$2\pi \text{ radianes} = 360^\circ \quad \Rightarrow \quad \left\{ \begin{array}{l} 360^\circ \rightarrow 2\pi \\ \text{grad} \rightarrow \text{rad} \end{array} \right\} \quad \Rightarrow \quad \boxed{\text{rad} = \frac{(\text{grad}) \times \pi}{180}}$$

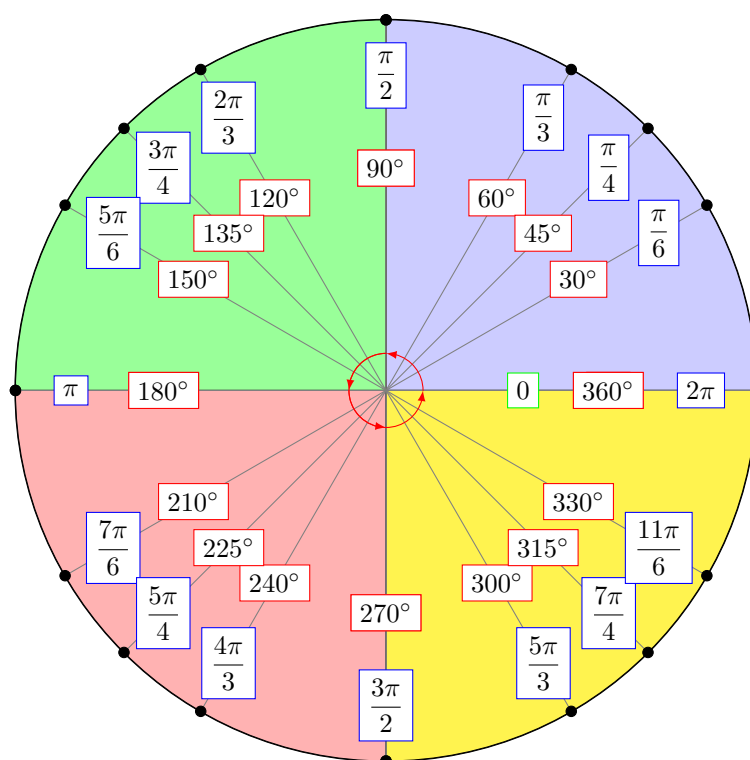


Figura A.1: Relación entre grados sexagesimales (marco rojo) y radianes (marco azul) de los ángulos más usuales en los cuatro cuadrantes.

Ejercicio A.1. Determinar cuantos radianes son 15° y cuantos grados son $\frac{5\pi}{12}$ radianes.

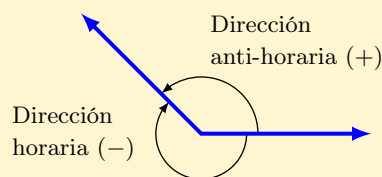
$$\bullet \quad 15^\circ = \frac{15 \times \pi}{180} = \frac{\pi}{12} \text{ rad}$$

$$\bullet \quad \frac{5\pi}{12} \text{ rad} = \frac{5 \times 180}{12} = 75^\circ$$

El signo del ángulo depende del sentido del giro:

- Signo positivo: sentido contrario a las agujas del reloj.
- Signo negativo: sentido igual a las agujas del reloj.

$$-\alpha = +(2\pi - \alpha) \quad ; \quad \alpha \in [0, \pi]$$



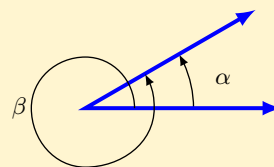
Ejercicio A.2. Determinar en sentido positivo el ángulo $\alpha = -\frac{\pi}{12}$ rad y en sentido negativo el ángulo $\beta = \frac{5\pi}{12}$ rad.

$$\bullet \alpha = 2\pi - \frac{\pi}{12} = \frac{23\pi}{12} \text{ rad}$$

$$\bullet \beta = -2\pi + \frac{5\pi}{12} = -\frac{19\pi}{12} \text{ rad}$$

Los ángulos cuyo valor absoluto es mayor que 2π radianes (360°) indican más de una vuelta en una determinada dirección.

$$\left. \begin{aligned} \beta &= 2k\pi + \alpha = \alpha \\ -\beta &= -2k\pi - \alpha = -\alpha \end{aligned} \right\} \quad \alpha \in [0, 2\pi) \quad ; \quad k \in \mathbb{Z}^+$$



Ejercicio A.3. Simplificar el ángulo $\alpha = 63\pi$ rad y el ángulo $\beta = -\frac{41\pi}{2}$ rad.

$$\bullet \alpha = 63\pi = 2\pi * 31 + \pi = \pi \text{ rad}$$

$$\bullet \beta = -\frac{41\pi}{2} = -20,5\pi = -2\pi * 10 - \frac{\pi}{2} = -\frac{\pi}{2} \text{ rad}$$

A.1.2. Funciones Trigonómicas

Definición de Seno, Coseno y Tangente

Dado un triángulo rectángulo $\triangle ABC$ con su ángulo recto en C y longitudes a , b y h , como indica la **Figura A.2**.

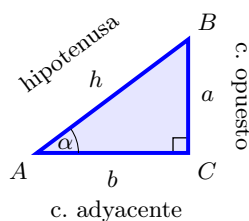


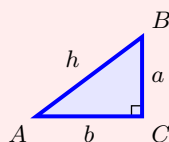
Figura A.2: Triángulo rectángulo de hipotenusa h y catetos a y b .

Se definen la siguientes funciones trigonométricas del ángulo agudo $\alpha = \widehat{BAC}$:

Nombre Función	Abreviatura	Definición
seno α	$\text{sen } \alpha$	$= \frac{\text{cateto opuesto}}{\text{hipotenusa}} = \frac{a}{h}$
coseno α	$\text{cos } \alpha$	$= \frac{\text{cateto adyacente}}{\text{hipotenusa}} = \frac{b}{h}$
tangente α	$\text{tan } \alpha$	$= \frac{\text{cateto opuesto}}{\text{cateto adyacente}} = \frac{\text{sen } \alpha}{\text{cos } \alpha} = \frac{a}{b}$

Tabla A.1: Funciones Trigonómicas Directas: Seno, Coseno y Tangente

Ejercicio A.4. Determinar los valores del seno, coseno y tangente correspondientes al ángulo $\widehat{ABC}$ del siguiente triángulo rectángulo cuyos valores son: $a = 3$ y $b = 4$.



Calculamos primero el valor de la hipotenusa aplicando el teorema de Pitágoras:

$$h = \sqrt{a^2 + b^2} = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

Por tanto, llamando $\beta = \widehat{ABC}$ y teniendo en cuenta las definiciones de la [Tabla A.1](#) se tiene:

$$\sin \beta = \frac{\text{c. op.}}{\text{hipot.}} = \frac{b}{h} = \frac{4}{5} \quad ; \quad \cos \beta = \frac{\text{c. ady.}}{\text{hipot.}} = \frac{a}{h} = \frac{3}{5} \quad ; \quad \tan \beta = \frac{\text{c. op.}}{\text{c. ady.}} = \frac{b}{a} = \frac{4}{3}$$

Valores Trigonómicos: Seno, Coseno y Tangente

En la [Figura A.3, página 253](#) se muestran en el exterior de la circunferencia unidad, los valores trigonométricos seno (en negro), coseno (en azul) y tangente (en rojo) de algunos de los ángulos más usuales.

Nota A.1. Tener en cuenta que los valores de la tangente de $\frac{\pi}{2} = 90^\circ$ y $\frac{3\pi}{2} = 270^\circ$ son respectivamente $+\infty$ y $-\infty$.

Ejercicio A.5. Determinar los siguientes valores trigonométricos, donde los ángulos están expresados en radianes.

a) $\text{sen } \frac{5\pi}{3}$

b) $\text{cos } \frac{3\pi}{4}$

c) $\text{tan } \frac{4\pi}{3}$

Utilizando los valores de la [Figura A.3](#), se tiene:

a) $\frac{5\pi}{3} \text{ rad} = 300^\circ = -\frac{\pi}{3} \text{ rad} = -60^\circ \quad ; \quad \text{sen } \frac{5\pi}{3} = -\frac{\sqrt{3}}{2}$

b) $\frac{3\pi}{4} \text{ rad} = 135^\circ = -\frac{5\pi}{4} \text{ rad} = -225^\circ \quad ; \quad \text{cos } \frac{3\pi}{4} = -\frac{\sqrt{2}}{2}$

c) $\frac{4\pi}{3} \text{ rad} = 240^\circ = -\frac{2\pi}{3} \text{ rad} = -120^\circ \quad ; \quad \text{tan } \frac{4\pi}{3} = \sqrt{3}$

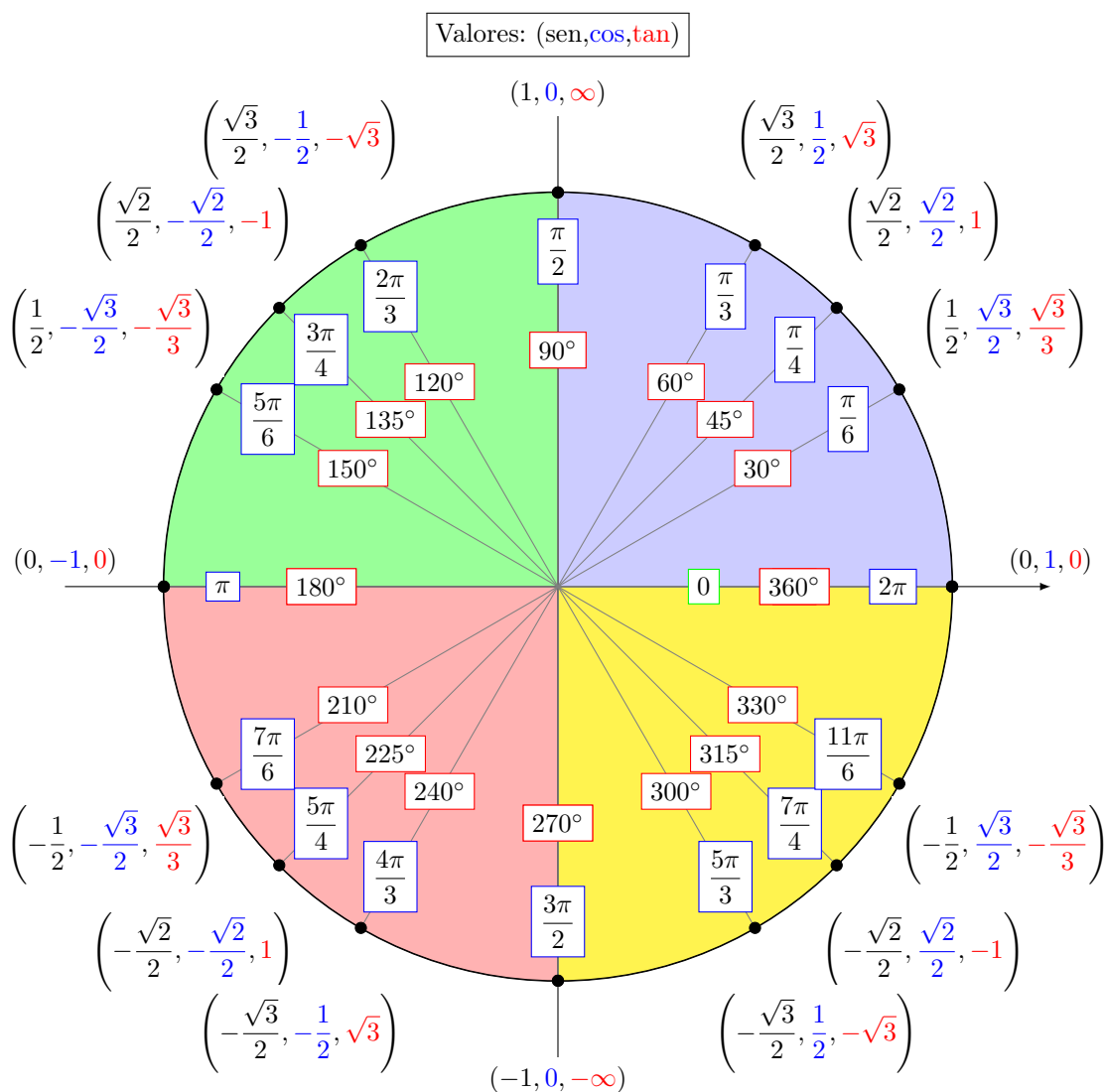


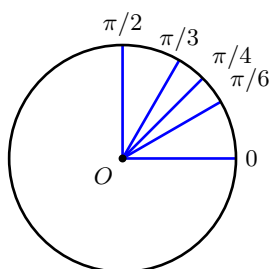
Figura A.3: Valores trigonométricos del seno, coseno y tangente.

Determinación Simplificada

Recordar todos los valores de la **Figura A.3** no es fácil. Resulta más práctico conocer una pequeña lista de ángulos simples del primer cuadrante y mediante ellos determinar otros valores del resto de cuadrantes.

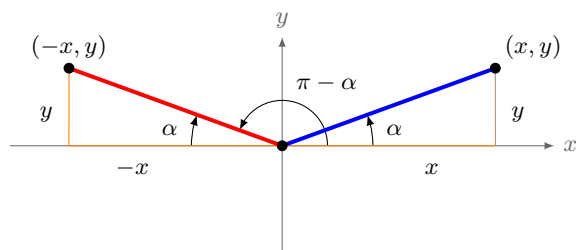
Primer Cuadrante ($[0^\circ, 90^\circ] = [0, \pi/2]$)

Vamos a ver una pequeña tabla de valores del seno, coseno y tangente de un conjunto de valores $\alpha \in [0, \pi/2]$. Con ellos se puede obtener fácilmente el valor de otros ángulos del resto de cuadrantes.



	Primer Cuadrante				
Grados	0°	30°	45°	60°	90°
Radianes	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
sen	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1
cos	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0
tan	0	$1/\sqrt{3} = \sqrt{3}/3$	1	$\sqrt{3}$	$+\infty$

Segundo Cuadrante ($[90^\circ, 180^\circ] = [\pi/2, \pi]$)



$$\boxed{\operatorname{sen}(\pi - \alpha) = \operatorname{sen}(\alpha)}$$

;

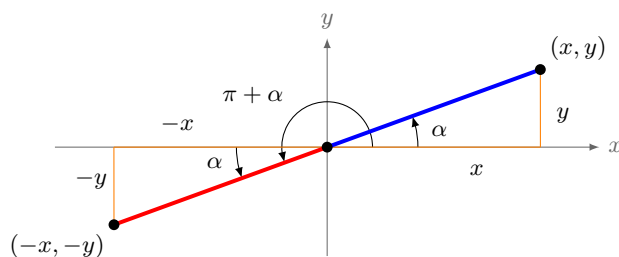
$$\boxed{\cos(\pi - \alpha) = -\cos(\alpha)}$$

;

$$\boxed{\operatorname{tg}(\pi - \alpha) = -\operatorname{tg}(\alpha)}$$

;

Tercer Cuadrante ($[180^\circ, 270^\circ] = [\pi, 3\pi/2]$)



$$\boxed{\operatorname{sen}(\pi + \alpha) = -\operatorname{sen}(\alpha)}$$

;

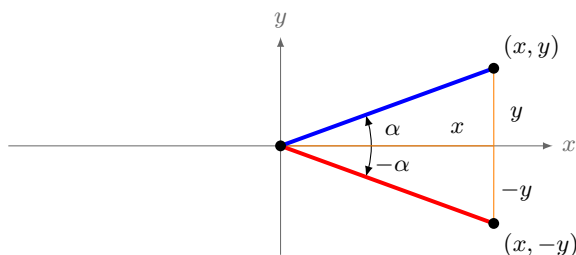
$$\boxed{\cos(\pi + \alpha) = -\cos(\alpha)}$$

;

$$\boxed{\operatorname{tg}(\pi + \alpha) = \operatorname{tg}(\alpha)}$$

;

Cuarto Cuadrante ($[270^\circ, 360^\circ] = [3\pi/2, 2\pi] = [0, -\pi/2]$)



$$\boxed{\operatorname{sen}(-\alpha) = -\operatorname{sen}(\alpha)}$$

;

$$\boxed{\cos(-\alpha) = \cos(\alpha)}$$

;

$$\boxed{\operatorname{tg}(\pi - \alpha) = -\operatorname{tg}(\alpha)}$$

;

Ejercicio A.6. Determinar los siguientes valores trigonométricos, relacionando el correspondiente ángulo con un valor en el primer cuadrante.

a) $\operatorname{sen} \frac{11\pi}{6}$

b) $\cos \frac{3\pi}{4}$

c) $\tan \frac{4\pi}{3}$

Utilizando las relaciones vistas en esta sección se tiene:

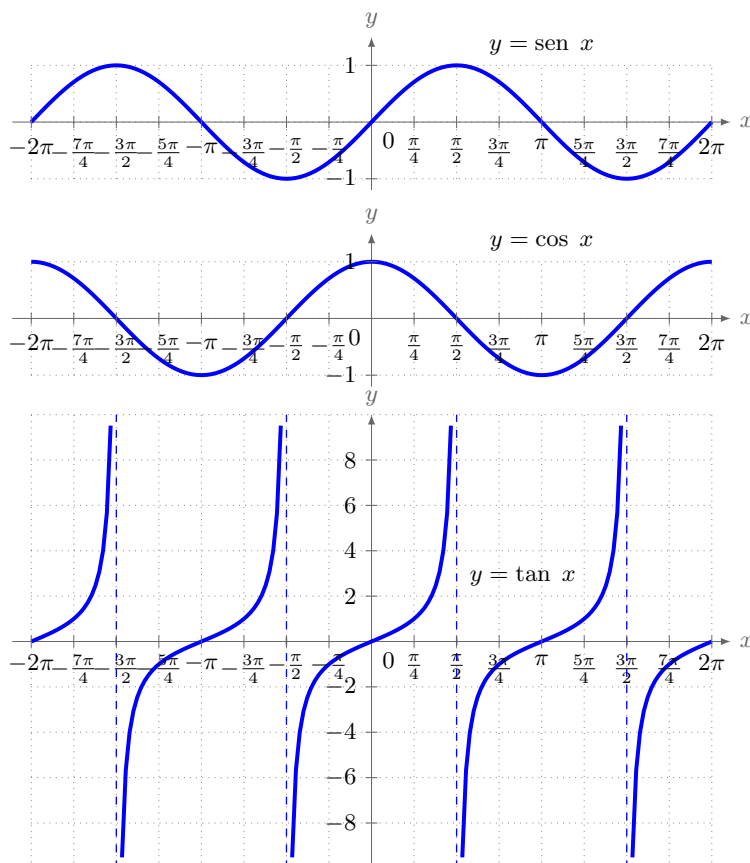
$$\text{a) } \operatorname{sen} \frac{11\pi}{6} = \operatorname{sen} \left(\frac{12\pi}{6} - \frac{\pi}{6} \right) = \operatorname{sen} \left(2\pi - \frac{\pi}{6} \right) = \operatorname{sen} \left(-\frac{\pi}{6} \right) = -\operatorname{sen} \left(\frac{\pi}{6} \right) = \boxed{-\frac{1}{2}}$$

$$\text{b) } \cos \frac{3\pi}{4} = \cos \left(\pi - \frac{\pi}{4} \right) = -\cos \left(\frac{\pi}{4} \right) = -\frac{\sqrt{2}}{2}$$

$$\text{c) } \tan \frac{4\pi}{3} = \tan \left(\pi + \frac{\pi}{3} \right) = \tan \left(\frac{\pi}{3} \right) = \sqrt{3}$$

Utilizando los valores de la [Figura A.3, página 253](#) se pueden verificar los resultados.

Representación Gráfica de las Funciones Trigonométricas Directas



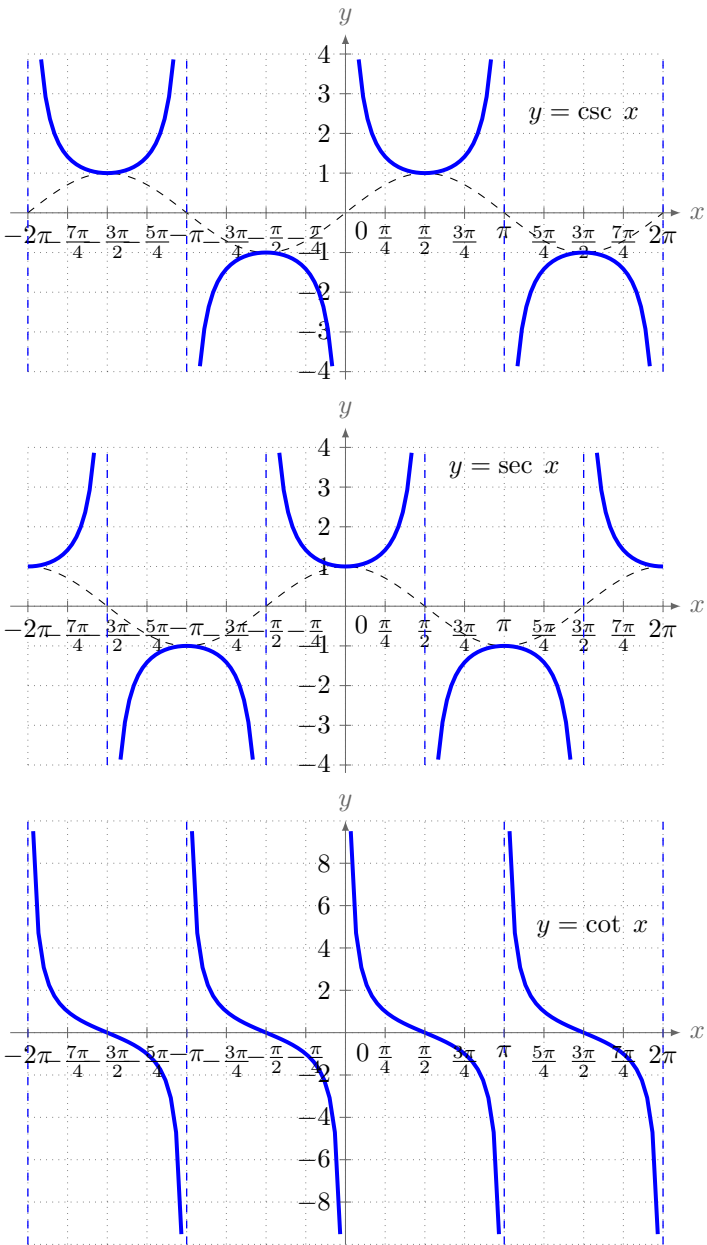
Inversas de las Funciones Trigonométricas

Las inversas de las funciones trigonométricas seno, coseno y tangente definen unas nuevas funciones denominadas cosecante, secante y cotangente respectivamente.

Basándonos en los datos de la [Figura A.2, página 251](#) estas funciones se definen de la siguiente forma:

Nombre Función	Abreviatura	Definición
cosecante α	$\csc \alpha$	$= \frac{\text{hipotenusa}}{\text{cateto opuesto}} = \frac{1}{\sin \alpha} = (\sin \alpha)^{-1} = \frac{h}{a}$
secante α	$\sec \alpha$	$= \frac{\text{hipotenusa}}{\text{cateto adyacente}} = \frac{1}{\cos \alpha} = (\cos \alpha)^{-1} = \frac{h}{b}$
cotangente α	$\cot \alpha$	$= \frac{\text{cateto adyacente}}{\text{cateto opuesto}} = \frac{1}{\tan \alpha} = (\tan \alpha)^{-1} = \frac{b}{a}$

Representación Gráfica de las Inversas de las Funciones Trigonómicas

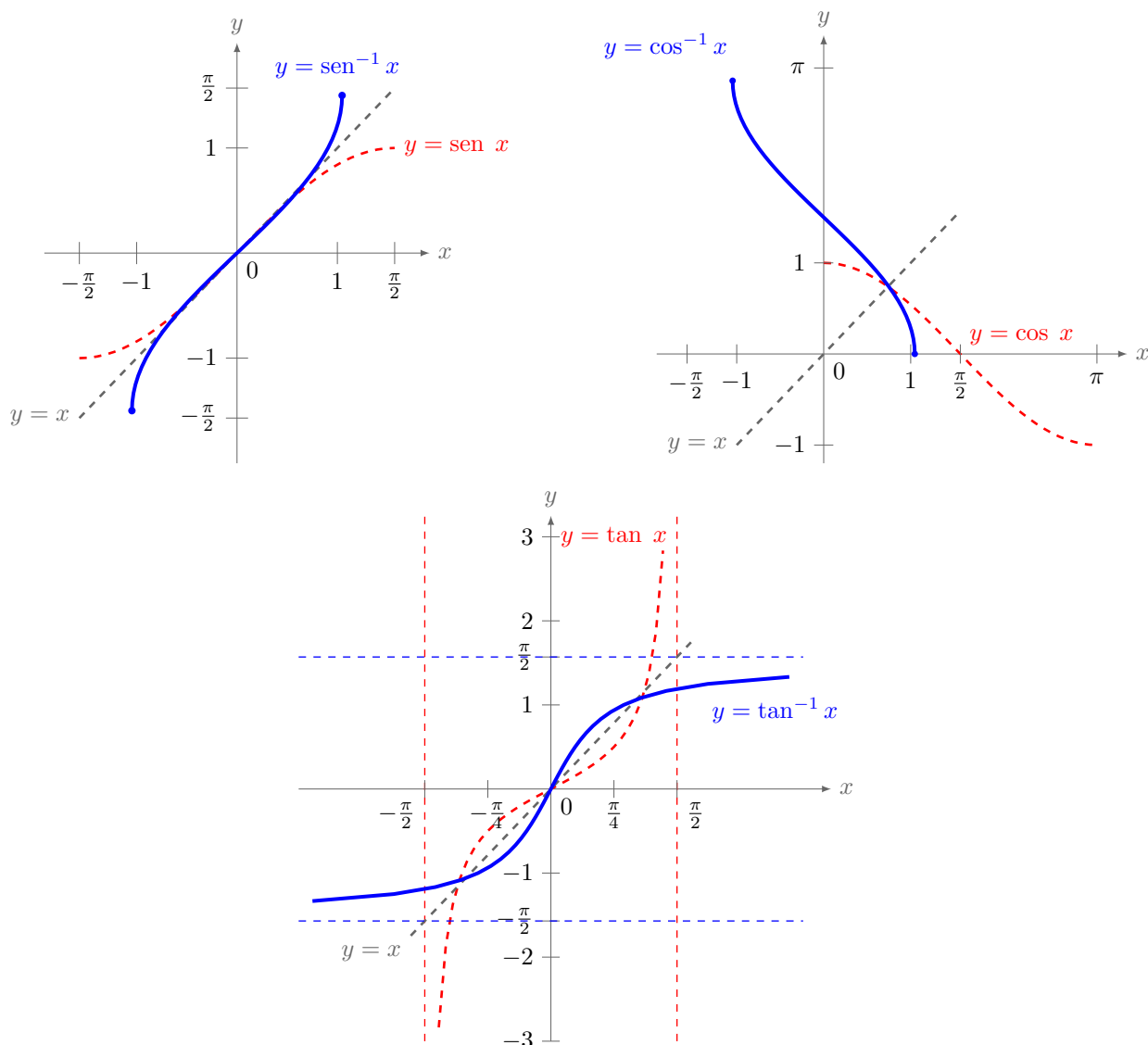


Funciones Inversas de las Funciones Trigonómicas

Las funciones inversas de las funciones trigonométricas de un determinado valor x , corresponden con los ángulos que tienen por seno, coseno o tangente dicho valor x . La siguiente tabla define las funciones inversas más usuales.

Nombre Función	Abreviatura	Consecuencia	Dominio
arcoseno	$\arcsen(x) = \sen^{-1}(x) = \alpha \implies \sen \alpha = x$	$\sen \alpha = x$	$x \in [-1, 1]$
arccoseno	$\arccos(x) = \cos^{-1}(x) = \alpha \implies \cos \alpha = x$	$\cos \alpha = x$	$x \in [-1, 1]$
arcotangente	$\arctan(x) = \tan^{-1}(x) = \alpha \implies \tan \alpha = x$	$\tan \alpha = x$	$x \in \mathbb{R}$

Representación Gráfica de las Funciones Trigonométricas Inversas



Ejercicio A.7. Determinar los siguientes valores de las funciones trigonométricas inversas:

a) $\arcsen \frac{1}{2}$

b) $\arccos -\frac{1}{2}$

c) $\arctan \sqrt{3}$

Utilizando los valores de la **Figura A.3, página 253**, comprobamos que hay dos posibles soluciones simples en cada caso. La solución apropiada dependerá de consideraciones geométricas adicionales.

a) $\sen \alpha = \frac{1}{2} \Rightarrow \alpha = \left\{ \frac{\pi}{6}, \frac{5\pi}{6} \right\} \quad ; \quad \text{b) } \cos \alpha = \frac{1}{2} \Rightarrow \alpha = \left\{ \frac{\pi}{3}, \frac{2\pi}{3} \right\}$

c) $\tan \alpha = \sqrt{3} \Rightarrow \alpha = \left\{ \frac{\pi}{3}, \frac{4\pi}{3} \right\}$

Relaciones Trigonométricas Básicas

Relación Fundamental y Derivadas de ella

- $a^2 + b^2 = h^2$ (Teorema Pitágoras)
- $\sin^2 x + \cos^2 x = 1$
- $1 + \tan^2 x = \frac{1}{\cos^2 x}$
- $1 + \cot^2 x = \frac{1}{\sin^2 x}$

Ángulo Suma, Resta y Doble

- $\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$
- $\sin(2x) = 2 \sin x \cos x$
- $\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$
- $\cos(2x) = \cos^2 x - \sin^2 x$
- $\tan(x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \tan y}$
- $\tan(2x) = \frac{2 \tan x}{1 - \tan^2 x}$

Ángulo Mitad

- $\sin \frac{1}{2}\theta = \pm \sqrt{\frac{1 - \cos \theta}{2}}$
- $\cos \frac{1}{2}\theta = \pm \sqrt{\frac{1 + \cos \theta}{2}}$
- $\tan \frac{1}{2}\theta = \pm \sqrt{\frac{1 - \cos \theta}{1 + \cos \theta}}$
- $\tan \frac{1}{2}\theta = \frac{1 - \cos \theta}{\sin \theta} = \frac{\sin \theta}{1 + \cos \theta}$

Transformaciones de Sumas en Productos

- $\sin \alpha + \sin \beta = 2 \sin \frac{\alpha + \beta}{2} \cdot \cos \frac{\alpha - \beta}{2}$
- $\cos \alpha + \cos \beta = 2 \cos \frac{\alpha + \beta}{2} \cdot \cos \frac{\alpha - \beta}{2}$
- $\sin \alpha - \sin \beta = 2 \cos \frac{\alpha + \beta}{2} \cdot \sin \frac{\alpha - \beta}{2}$
- $\cos \alpha - \cos \beta = -2 \sin \frac{\alpha + \beta}{2} \cdot \sin \frac{\alpha - \beta}{2}$

Transformaciones de Productos en Sumas

- $\sin \alpha \cdot \sin \beta = -\frac{1}{2}[\cos(\alpha + \beta) - \cos(\alpha - \beta)]$
- $\cos \alpha \cdot \cos \beta = \frac{1}{2}[\cos(\alpha + \beta) + \cos(\alpha - \beta)]$
- $\sin \alpha \cdot \cos \beta = \frac{1}{2}[\sin(\alpha + \beta) + \sin(\alpha - \beta)]$

Equivalencias Útiles en Integración

- $\cos 2\alpha = 1 - 2 \sin^2 \alpha$
- $\cos 2\alpha = 2 \cos^2 \alpha - 1$
- $\sin^2 \alpha = \frac{1 - \cos 2\alpha}{2}$
- $\cos^2 \alpha = \frac{1 + \cos 2\alpha}{2}$
- $\tan^2 \alpha = \frac{1 - \cos 2\alpha}{1 + \cos 2\alpha}$

Ejercicio A.8. Determinar los siguientes valores:

a) $\sin \frac{\pi}{12}$

b) $\cos -\frac{5\pi}{12}$

c) $\tan \frac{7\pi}{12}$

Utilizando las relaciones trigonométricas vistas anteriormente se tiene:

a) Como $\frac{\pi}{12} = \frac{1}{2} \frac{\pi}{6}$, utilizando la fórmula del seno del ángulo mitad¹ se tiene:

$$\sin \frac{\theta}{2} = \pm \sqrt{\frac{1 - \cos \theta}{2}} \implies \sin \frac{\pi}{12} = \pm \sqrt{\frac{1 - \cos(\pi/6)}{2}} = \pm \sqrt{\frac{1 - \sqrt{3}/2}{2}} = \pm \sqrt{\frac{2 - \sqrt{3}}{4}}$$

El signo depende del cuadrante donde se encuentre el ángulo θ . Como $\pi/12 = 15^\circ$ está en el primer cuadrante y ahí el seno siempre es positivo, se tiene:

$$\sin \frac{\pi}{12} = + \frac{\sqrt{2 - \sqrt{3}}}{2}$$

b) Como $\frac{5\pi}{12} = \frac{6\pi}{12} - \frac{\pi}{12} = \frac{\pi}{2} - \frac{\pi}{12}$, utilizando la fórmula del coseno de una diferencia se tiene:

$$\cos(x - y) = \cos x \cos y + \sin x \sin y \implies$$

$$\cos(\pi/2 - \pi/12) = \cos(\pi/2) \cos(\pi/12) + \sin(\pi/2) \sin(\pi/12) = 0 \cos(\pi/12) + 1 \sin(\pi/12) =$$

$$= + \sin(\pi/12) = + \frac{\sqrt{2 - \sqrt{3}}}{2}$$

donde el valor de $\sin(\pi/12)$ se ha obtenido del apartado a).

c) Como $\frac{7\pi}{12} = \pi - \frac{5\pi}{12}$, utilizando la fórmula de la tangente de una diferencia se tiene:

$$\tan(x - y) = \frac{\tan x - \tan y}{1 + \tan x \tan y} \implies \tan(\pi - 5\pi/12) = \frac{\tan(\pi) - \tan(5\pi/12)}{1 + \tan(\pi) \tan(5\pi/12)} =$$

$$= \frac{0 - \tan(5\pi/12)}{1 + 0 \tan(5\pi/12)} = -\tan(5\pi/12) = -\frac{\sin(5\pi/12)}{\cos(5\pi/12)} = -\frac{\sin(5\pi/12)}{\sqrt{1 - \sin^2(5\pi/12)}} =$$

$$= -\frac{\sqrt{2 - \sqrt{3}}}{\sqrt{4 - 2 + \sqrt{3}}} = -\sqrt{\frac{2 - \sqrt{3}}{2 + \sqrt{3}}}$$

donde el valor de $\sin(5\pi/12)$ se ha obtenido del apartado b).

¹Ver el Apéndice A.1.2, página 258.

A.2. Funciones Hiperbólicas

Definición A.2. Funciones Hiperbólicas

Las **funciones hiperbólicas** están basadas en la función exponencial $f(x) = e^x$ a través de operaciones racionales y presentan unas propiedades análogas a las *funciones trigonométricas*.

Seno hiperbólico

$$\sinh x = \frac{e^x - e^{-x}}{2}$$

Coseno hiperbólico

$$\cosh x = \frac{e^x + e^{-x}}{2}$$

Tangente hiperbólica

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Inversar de las Funciones Hiperbólicas

Secante hiperbólica

$$\operatorname{sech} x = \frac{1}{\cosh x}$$

Cosecante hiperbólica

$$\operatorname{cosech} x = \frac{1}{\sinh x}$$

Cotangente hiperbólica

$$\operatorname{cotanh} x = \frac{1}{\tanh x}$$

Funciones Hiperbólicas Inversas

Arcoseno hiperbólico

$$\operatorname{arsinh} x = \ln \left(x + \sqrt{x^2 + 1} \right)$$

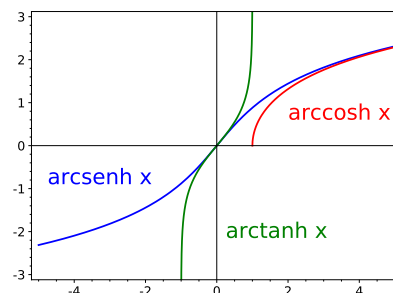
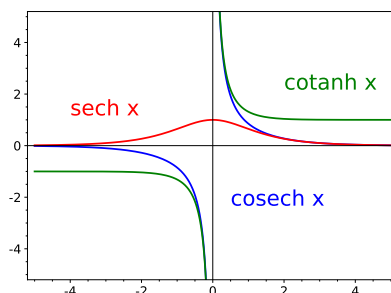
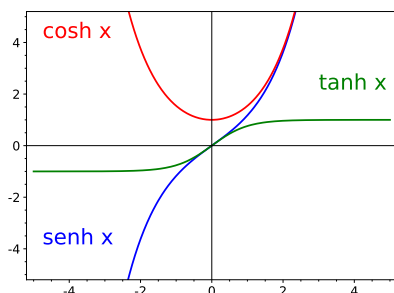
Arcocoseno hiperbólico ($x \geq 1$)

$$\operatorname{arcosh} x = \ln \left(x + \sqrt{x^2 - 1} \right)$$

Arcotangente hiperbólica ($|x| < 1$)

$$\operatorname{artanh} x = \frac{1}{2} \ln \left(\frac{1+x}{1-x} \right)$$

Representación Gráfica de las Funciones Hiperbólicas



Relaciones Hiperbólicas Básicas

- Ecuación Fundamental: $\cosh^2 z - \sinh^2 z = 1$
- $1 - \tanh^2 x = \operatorname{sech}^2 x$
- $\sinh(x \pm y) = \sinh x \cosh y \pm \cosh x \sinh y$
- $\cosh(x \pm y) = \cosh x \cosh y \pm \sinh x \sinh y$
- $\tanh(x \pm y) = \frac{\tanh x \pm \tanh y}{1 \mp \tanh x \tanh y}$
- $\operatorname{cotanh}^2 x - 1 = \operatorname{cosech}^2 x$
- $\sinh(2x) = 2 \sinh x \cosh x$
- $\cosh(2x) = \cosh^2 x + \sinh^2 x$
- $\tanh(2x) = \frac{2 \tanh x}{1 + \tanh^2 x}$

- $e^x = \cosh x + \sinh x$
- $e^{-x} = \cosh x - \sinh x$
- $\frac{d}{dx} \sinh x = \cosh x$
- $\frac{d}{dx} \cosh x = \sinh x$

Ejercicio A.9. Demostrar las fórmulas correspondientes a las funciones hiperbólicas inversas.

$$\text{a) } \operatorname{arcsenh} x = \ln(x + \sqrt{x^2 + 1}) \quad \text{b) } \operatorname{arccosh} x = \ln(x + \sqrt{x^2 - 1}) \quad \text{c) } \operatorname{arctanh} x = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right)$$

- a) Viendo la gráfica de la función $\sinh(x)$ vemos que es biyectiva y por tanto existe su función inversa en todo $\mathbb{R}$.

Buscamos un valor A tal que:

$$\operatorname{arcsenh} x = A \quad ; \quad x = \sinh A = \frac{e^A - e^{-A}}{2} \quad ; \quad 2x = e^A - e^{-A}$$

Llamando $B = e^A$ se tiene:

$$2x = B - \frac{1}{B} \quad ; \quad 2xB = B^2 - 1 \quad ; \quad B^2 - 2xB - 1 = 0 \quad ; \quad B = \frac{2x \pm \sqrt{4x^2 + 4}}{2} = x \pm \sqrt{x^2 + 1}$$

Como:

$$\ln B = A \quad \Rightarrow \quad A = \ln(x \pm \sqrt{x^2 + 1})$$

Notar que $\sqrt{x^2 + 1} > x$ y como no existe en $\mathbb{R}$ el logaritmo de números negativos se tiene:

$$\boxed{\operatorname{arcsenh} x = \ln(x + \sqrt{x^2 + 1})}$$

- b) Viendo la gráfica de función $\cosh(x)$ vemos que no es inyectiva. Para encontrar su función inversa tomamos solo valores $x \geq 0$, donde este tramo sí es una función inyectiva. Claramente para estos valores $x \geq 0$ el resultado de la función inversa de $\cosh(x)$ tiene que ser mayor o igual que 0.

Siguiendo el mismo proceso del apartado anterior se tiene:

$$\operatorname{arccosh} x = A \quad ; \quad x = \cosh A = \frac{e^A + e^{-A}}{2} \quad ; \quad 2x = e^A + e^{-A}$$

Llamando $B = e^A$ se tiene:

$$2x = B + \frac{1}{B} \quad ; \quad 2xB = B^2 + 1 \quad ; \quad B^2 - 2xB + 1 = 0 \quad ; \quad B = \frac{2x \pm \sqrt{4x^2 - 4}}{2} = x \pm \sqrt{x^2 - 1}$$

Como:

$$\ln B = A \quad \Rightarrow \quad A = \ln(x \pm \sqrt{x^2 - 1})$$

Notar que para que exista el valor de la raíz es necesario que $|x| \geq 1$ y en este caso $\sqrt{x^2 - 1} < x$, para que el argumento del logaritmo sea positivo es necesario que $x \geq 1$.

En principio los dos signos delante de la raíz serían válidos, sin embargo para que los valores obtenidos sean mayores o iguales que cero es necesario que el argumento del logaritmo sea mayor que 1 y por tanto solo se toma el signo positivo, de forma que:

$$\boxed{\operatorname{arccosh} x = \ln(x + \sqrt{x^2 - 1})} \quad ; \quad x \geq 1$$

- c) Siguiendo el mismo proceso que en los apartados anteriores se tiene:

$$\operatorname{arctanh} x = A \quad ; \quad x = \tanh A = \frac{e^A - e^{-A}}{e^A + e^{-A}} \quad ; \quad xe^A + xe^{-A} = e^A - e^{-A}$$

Llamando $B = e^A$ se tiene:

$$xB + \frac{x}{B} = B - \frac{1}{B} \quad ; \quad xB^2 + x = B^2 - 1 \quad ; \quad B^2(1 - x) = 1 + x \quad ; \quad B = \sqrt{\frac{1+x}{1-x}}$$

Como:

$$\ln B = A \quad \Rightarrow \quad \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) \quad \Rightarrow \quad \boxed{\operatorname{arctanh} x = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right)} \quad ; \quad |x| < 1$$

A.3. Productos Notables

Definición A.3. Producto Notable

Se llaman **productos notables** a ciertas expresiones muy comunes cuyo resultado es conocido.

Algunos de los productos notables más sencillos y útiles son:

- $(a + b)(a - b) = a^2 - b^2$
- $(a + b)^2 = a^2 + 2ab + b^2$; $(a - b)^2 = a^2 - 2ab + b^2$
- $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$; $(a - b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$

Nota A.2. Estas expresiones se pueden generalizar a través del **binomio de Newton** (ver [Apéndice A.5, página 263](#)).

A.4. Sumatorios

- Al símbolo $\sum$ se le denomina **sumatorio** y representa una suma de términos.
- La expresión $\sum_{j=1}^m$ se lee “sumatorio desde j igual a 1 hasta m ”.
- La variable j es una variable *variable muda* que puede designarse por cualquier otra letra o símbolo.

Ejemplo A.1. Sumatorios.

- Sumatorio simple:

$$\sum_{j=1}^n P_j = P_1 + P_2 + \cdots + P_n$$

- Independencia de la variable muda:

$$\sum_{j=1}^n P_j = \sum_{i=1}^n P_i = \sum_{\alpha=1}^n P_{\alpha}$$

- Variables independientes del sumatorio:

$$\sum_{j=1}^m a_k x_j = a_k x_1 + a_k x_2 + \cdots + a_k x_m = a_k (x_1 + x_2 + \cdots + x_m) = a_k \sum_{j=1}^m x_j$$

- Sumatorios combinados:

$$\begin{aligned} \sum_{a=5}^n \sum_{b=-2}^m s_a w_b &= \sum_{a=5}^n \left(s_a \sum_{b=-2}^m w_b \right) = \sum_{a=5}^n s_a (w_{-2} + w_{-1} + \cdots + w_m) = \\ &= a_5 (w_{-2} + w_{-1} + \cdots + w_m) + a_6 (w_{-2} + w_{-1} + \cdots + w_m) + \cdots + a_n (w_{-2} + w_{-1} + \cdots + w_m) = \\ &= (a_5 + a_6 + \cdots + a_n) (w_{-2} + w_{-1} + \cdots + w_m) = \left(\sum_{a=5}^n s_a \right) \left(\sum_{b=-2}^m w_b \right) \end{aligned}$$

A.5. Binomio de Newton

El **binomio de Newton**, también conocido como **teorema del binomio**, establece la forma de expandir en sumas cualquier potencia de un *binomio* $x + y$.

Recordamos previamente la definición de **factorial de un número natural** y el **cociente binomial** o **números combinatorios**.

Factorial : $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n$; **Números Combinatorios :** $\binom{n}{k} = \frac{n!}{(n-k)! k!}$

Nota A.3. Por definición, el factorial de cero es la unidad: $0! = 1$

Teorema A.1. Binomio de Newton

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k = \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \dots + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n =$$

$$= x^n + n x^{n-1} y + \frac{n(n-1)}{2!} x^{n-2} y^2 + \dots + n x y^{n-1} + y^n$$

Ejercicio A.10. Utilizando el binomio de Newton, desarrollar la siguiente potencia:

$$(1+x)^5$$

$$(1+x)^5 = \sum_{k=0}^5 \binom{5}{k} 1^{5-k} x^k = \sum_{k=0}^5 \binom{5}{k} x^k = \binom{5}{0} x^0 + \binom{5}{1} x^1 + \binom{5}{2} x^2 + \binom{5}{3} x^3 + \binom{5}{4} x^4 + \binom{5}{5} x^5 =$$

$$= \frac{5!}{5!0!} + \frac{5!}{4!1!} x + \frac{5!}{3!2!} x^2 + \frac{5!}{2!3!} x^3 + \frac{5!}{1!4!} x^4 + \frac{5!}{0!5!} x^5 = 1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5$$

Ejercicio A.11. Encontrar el valor del coeficiente del término xy^2z^7 de la siguiente potencia

$$(3x - y + 2z)^{10}$$

Como en la fórmula del binomio de Newton solo aparecen dos términos y aquí tenemos tres, vamos a agrupar los dos primeros y aplicamos dos veces dicha fórmula.

$$(3x - y + 2z)^{10} = [(3x - y) + 2z]^{10} = \sum_{k=0}^{10} \binom{10}{k} (3x - y)^k (2z)^{10-k} =$$

$$= \sum_{k=0}^{10} \binom{10}{k} \left[\sum_{i=0}^k \binom{k}{i} (3x)^i (-y)^{k-i} \right] (2z)^{10-k}$$

Como la potencia de z tiene que ser 7, por tanto $10 - k = 7$ y $k = 3$. Por otra parte como la potencia de x debe ser 1, $i = 1$ y de ahí la potencia de y será $k - i = 3 - 1 = 2$ como indica el enunciado. Por tanto:

$$\binom{10}{3} \binom{3}{1} (3x)(-y)^2 (2z)^7 = \frac{10!}{7! \cdot 3!} \frac{3!}{2! \cdot 1!} 3 \cdot 2^7 xy^2 z^7 = \frac{10 \cdot 9 \cdot 8}{2} 384 xy^2 z^7 = 138240 xy^2 z^7$$

A.6. Polinomios

Definición A.4. Polinomio

Un **polinomio** es una suma y/o resta de términos formados por productos de **coeficientes** (números reales o complejos) y potencias enteras de una o más variables **variables independientes**.

Ejemplo A.2. Polinomios de una y varias variables

- $P(x) = x^5 - 4x^3 + 12$
- $P(x, y, z) = -1 + x^2z + xy^2z^3$

Nos vamos a centrar en los polinomios reales de una sola *variable independiente* x , de *grado* n (potencia mayor):

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \quad ; \quad a_i \in \mathbb{R} \quad \left\{ \begin{array}{ll} n & : \text{Grado del polinomio } (a_n \neq 0) \\ a_i & : \text{Coeficientes} \\ a_n & : \text{Coeficiente principal o director} \\ a_0 & : \text{Término independiente} \\ \text{Monomio} & : \text{Cada uno de los términos } a_i x^i \end{array} \right.$$

Raíces o Ceros de un Polinomio

Definición A.5. Raíz o Cero de un Polinomio

Se dice que $r \in \mathbb{C}$ es una **raíz** o **cero** del polinomio $P(x)$ si: $P(r) = 0$.

Teorema A.2. Teorema Fundamental del Álgebra

- Todo polinomio de grado n tiene exactamente n raíces reales y/o complejas, iguales y/o distintas.
- En caso de que un número complejo $a + bi$ sea raíz de un polinomio, entonces siempre su complejo conjugado $a - bi$ también será raíz del polinomio, de forma que las raíces complejas siempre aparecen a pares $a \pm bi$.

Definición A.6. Multiplicidad de una Raíz

La **multiplicidad** de una raíz de un polinomio de grado n es el número $m \leq n$ de veces que se repite esa raíz en el conjunto de las n raíces del polinomio.

Definición A.7. Factorización de un Polinomio

Factorizar un polinomio consiste en descomponer el polinomio en un *producto* de **polinomios irreducibles**² del tipo:

$$P(x) = a_n (x - a)^n (x - b)^m \dots [(x - \alpha)^2 + \beta^2]^k \dots \quad ; \quad a, b, \alpha, \beta \in \mathbb{R} \quad ; \quad n, m, k \in \mathbb{N}$$

donde a_n es el **coeficiente principal** del polinomio, $a, b, \alpha \pm \beta i, \dots$ son las **raíces** reales y/o complejas y $n, m, k, \dots$ sus correspondientes **multiplicidades**.

²Ver la Definición A.8, página 268.

Ejercicio A.12. Encontrar el polinomio que tiene por raíces:

$$\begin{cases} -2 & \text{con multiplicidad } 3 \\ 3 & \text{con multiplicidad } 2 \\ 1 \pm 2i & \text{con multiplicidad } 1 \end{cases}$$

El polinomio buscado $P(x)$ tiene 5 raíces reales (una se repite tres veces y otra dos) y una raíz compleja (formada por dos raíces $1 + 2i$ y $1 - 2i$) y por tanto el grado de $P(x)$ debe ser $n = 3 + 2 + 2 = 7$. Lo calculamos a través de su factorización:

$$P(x) = (x + 2)^3(x - 3)^2[(x - 1)^2 + 2^2] = (x^3 + 6x^2 + 12x + 8) \cdot 3(x^2 - 6x + 9) \cdot (x^2 - 2x + 5) =$$

$$x^7 - 2x^6 - 10x^5 + 20x^4 + 5x^3 - 98x^3 + 156x + 360$$

División entre Polinomios

Todo polinomio $P(x)$ se puede dividir entre otro $Q(x)$ de grado menor o igual de la siguiente forma:

Dividendo	$P(x)$	$Q(x)$	Divisor
Resto	$R(x)$	$C(x)$	Cociente

donde $C(x)$ y $R(x)$ son polinomios apropiados, de forma que $P(x)$ se puede **descomponer** como:

$$P(x) = Q(x) \cdot C(x) + R(x) \quad \begin{cases} \text{grado } P(x) \geq \text{grado } Q(x) < \text{grado } R(x) \\ \text{grado } P(x) = \text{grado } Q(x) + \text{grado } C(x) \end{cases}$$

Ejercicio A.13. Dividir el polinomio $P(x) = 6x^4 + 4x^2 + x - 4$ entre el polinomio $Q(x) = 2x^2 - 1$ y descomponer $P(x)$ en base al resultado.

$$\begin{array}{r} \cancel{6x^4} + 4x^2 + x - 5 \\ -\cancel{6x^4} + 3x^2 \\ \hline 0 \quad 7x^2 + x - 5 \\ \quad \cancel{-7x^2} + 7/2 \\ \quad \hline 0 \quad x - 3/2 \end{array}$$

$$\begin{array}{r} 2x^2 - 1 \\ \hline 3x^2 + 7/2 \end{array}$$

Por tanto se tiene:

$$P(x) = Q(x)C(x) + R(x)$$

$$6x^4 + 4x^2 + x - 4 = (2x^2 - 1) \left(3x^2 + \frac{7}{2} \right) + \left(x - \frac{3}{2} \right)$$

Se dice que un polinomio $P(x)$ es **divisible** por otro polinomio $Q(x)$ si el resto de la división $P(x)/Q(x)$ es cero, y por tanto:

$$P(x) = Q(x) \cdot C(x) \quad \text{grado } P(x) \geq \text{grado } Q(x)$$

Ejemplo A.3.

El polinomio $P(x) = 2x^3 + 3x^2 + 3x + 1$ es *divisible* por el polinomio $Q(x) = 2x + 1$ ya que:

$$P(x) = (2x + 1)(x^2 + x + 1)$$

Vamos a repasar una forma sencilla de dividir un polinomio $P(x)$ a través de la **Regla de Ruffini** que nos permitirá encontrar sus **raíces** y a través de ellas su **factorización**.

A.6.1. Regla de Ruffini

La **regla de Ruffini**³ permite dividir fácilmente polinomios de orden n $P(x)$ entre polinomios de grado uno $Q(x)$ de la forma:

$$\begin{cases} P(x) = a_n x^n + \cdots + a_0 \\ Q(x) = x - r \end{cases} \implies \frac{P(x)}{Q(x)} \implies \boxed{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \mid x - r}$$

donde⁴ $r \in \mathbb{C}$. El procedimiento es el siguiente:

1. Se trazan dos líneas a modo de ejes y se ordenan todos los coeficientes de $P(x)$ de mayor a menor en la parte superior, poniendo ceros en caso de que algún monomio entre a_n y a_0 no aparezca. El valor de r se coloca abajo en la zona izquierda, siendo positivo si $Q(x) = x - r$ y negativo si $Q(x) = x + r$.

$$\begin{array}{r|rrrrrr} r & a_n & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ \hline \end{array}$$

2. El primer coeficiente a_n se baja directamente.

$$\begin{array}{r|rrrrrr} r & a_n & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ \hline & a_n & & & & & \end{array}$$

3. Se multiplica a_n por r y se coloca debajo de a_n . Posteriormente se suman ambos obteniéndose $b_{n-2} = a_{n-1} + a_n r$.

$$\begin{array}{r|rrrrrr} r & a_n & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ \hline & & +a_n r & & & & \\ & a_n & b_{n-2} & & & & \end{array}$$

4. Se repite el paso 3 obteniéndose: $b_{n-3} = a_{n-2} + b_{n-2}r$, $b_{n-4} = a_{n-3} + b_{n-2}r$, etc. hasta obtener $b = a_0 + b_0 r$.

$$\begin{array}{r|rrrrrr} r & a_n & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ \hline & & +a_n r & +b_{n-2} r & \cdots & +b_1 r & +b_0 r \\ & a_n & b_{n-2} & b_{n-3} & \cdots & b_0 & b \end{array}$$

5. Llamando $\boxed{b_{n-1} = a_n}$, al cociente de la división $\boxed{C(x)}$ y al resto $\boxed{R(x) = b}$, se tiene:

$$\begin{array}{r|rrrrrr} r & a_n & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ \hline & & +a_n r & +b_{n-2} r & \cdots & +b_1 r & +b_0 r \\ & b_{n-1} & b_{n-2} & b_{n-3} & \cdots & b_0 & b \end{array} \implies \begin{cases} C(x) = b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + \cdots + b_1 x + b_0 \\ R(x) = b \end{cases}$$

6. Con estos resultados el polinomio $P(x)$ se puede descomponer como:

$$\boxed{P(x) = C(x)(x - r) + R(x)}$$

El uso de la regla de Ruffini para la búsqueda de raíces de un polinomio interesa hacerlo cuando el grado del mismo es mayor o igual que tres. Para grado 1 es trivial encontrar la raíz y para grado dos es preferible usar la fórmula de la ecuación cuadrática:

$$ax^2 + bx + c = 0 \implies x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

³Paolo Ruffini (1765-1822) fue un matemático, profesor y médico italiano.

⁴El caso más sencillo es cuando r es un número real entero.

Ejercicio A.14. Calcular el siguiente cociente de polinomios:

$$3x^3 - 5x^2 + 2x - 7 \quad \underline{x + 2}$$

Aplicando la *regla de Ruffini*:

$$\begin{array}{r|rrrr} & 3 & -5 & 2 & -7 \\ -2 & & -6 & 22 & -48 \\ \hline & 3 & -11 & 24 & -55 \end{array} \quad \Rightarrow \quad \begin{cases} C(x) = 3x^2 - 11x + 24 \\ R(x) = -55 \end{cases}$$

Por tanto: $3x^3 - 5x^2 + 2x - 7 = (3x^2 - 11x + 24)(x + 2) - 55$

Teorema A.3. Teorema del Resto

El *resto* R de la división del polinomio $P(x)$ entre $x - \alpha$ coincide con el valor $P(\alpha)$, donde $\alpha \in \mathbb{C}$.

Nota A.4. Demostración: $P(x) = (x - \alpha)C(x) + R$; Si $x = \alpha \Rightarrow P(\alpha) = R$

Teorema A.4. Divisibilidad de un Polinomio

El polinomio $P(x)$ es **divisible** por $(x - \alpha)$ si y solo si $\alpha \in \mathbb{C}$ es una **raíz** de $P(x)$.

Cálculo de las Raíces de un Polinomio

Sea un polinomio $P(x)$ con todos sus coeficientes a_i enteros. Entonces:

1. Si tiene una raíz *entera* r , ésta debe ser un divisor del *término independiente* a_0 del polinomio.
2. Si tiene una raíz *racional* $r = p/q$ (irreducible), entonces el numerador p debe ser un divisor de a_0 y el denominador q debe ser un divisor del *coeficiente principal* a_n .

Descomposición de un Polinomio en Factores

Si un polinomio $P(x)$ de grado n tiene raíces reales $r_1, \dots, r_m$, entonces se puede **descomponer** de la forma:

$$P(x) = (x - r_1)(x - r_2) \cdots (x - r_m)C_k(x)$$

donde $C_k(x)$ es un polinomio de grado $k = n - m$.

En particular si $P(x)$ tiene tantas raíces reales como su grado entonces se puede hacer una **descomposición en factores simples** (factorización⁵) de la forma:

$$P(x) = a_n(x - r_1)(x - r_2) \cdots (x - r_m)$$

siendo a_n el **coeficiente principal** del polinomio.

⁵Ver la Definición A.7, página 264.

Definición A.8. Polinomios Irreducibles

Un polinomio se dice **irreducible** si no puede descomponerse en producto de dos o más polinomios de grado mayor o igual que uno.

- Todos los polinomios de grado cero (constantes) y de grado uno $ax + b$, son irreducibles.
- No existen polinomios irreducibles de grado mayor o igual que tres.
- Los polinomios de grado dos irreducibles son los que no tienen raíces reales.

Ejercicio A.15. En caso de ser posible, descomponer en factores simples el polinomio:

$$P(x) = -2x^3 + 4x^2 + 10x - 12$$

Al ser un polinomio de grado 3, tiene exactamente 3 raíces (reales y/o complejas, iguales o distintas). Buscando posibles raíces enteras entre los divisores de -12, $\{-12, -6, -3, -2, -1, 1, 2, 3, 6, 12\}$, encontramos que $P(r) = 0$ para $r = \{-2, 1, 3\}$, por tanto:

$$P(x) = -2(x+2)(x-1)(x-3)$$

Ejercicio A.16. Descomponer en factores simples el siguiente polinomio:

$$P(x) = x^5 - x^4 - 6x^3 - 50x^2 - 100x$$

Como el polinomio no tiene coeficiente independiente una raíz de $P(x)$ es $x = 0$. Sacando factor común:

$$P(x) = x(x^4 - x^3 - 6x^2 - 50x - 100)$$

Buscamos entonces las raíces del polinomio $Q(x) = x^4 - x^3 - 6x^2 - 50x - 100$ entre los divisores de 100.

$$Q(-2) = (-2)^4 - (-2)^3 - 6 \cdot (-2)^2 - 50 \cdot (-2) - 100 = 16 + 8 - 24 + 100 - 100 = 0$$

Por tanto $x = -2$ es otra raíz del polinomio $P(x)$. Usamos Ruffini para factorizar.

$$\begin{array}{r|rrrrr} & 1 & -1 & -6 & -50 & -100 \\ -2 & & -2 & 6 & 0 & 100 \\ \hline & 1 & -3 & 0 & -50 & 0 \end{array} \Rightarrow \begin{cases} C(x) = x^3 - 3x^2 - 50 \\ R(x) = 0 \end{cases} \Rightarrow P(x) = x(x+2)(x^3 - 3x^2 - 50)$$

Buscamos entonces las raíces del nuevo polinomio $Q_2(x) = x^3 - 3x^2 - 6x^2 - 50$ entre los divisores de -50.

$$Q_2(5) = (5)^3 - 3 \cdot 5^2 - 50 = 125 - 75 - 50 = 0$$

Por tanto $x = 5$ es otra raíz del polinomio $P(x)$. Usamos nuevamente Ruffini para factorizar.

$$\begin{array}{r|rrrr} & 1 & -3 & 0 & -50 \\ 5 & & 5 & 10 & 50 \\ \hline & 1 & 2 & 10 & 0 \end{array} \Rightarrow \begin{cases} C(x) = x^2 + 2x + 10 \\ R(x) = 0 \end{cases} \Rightarrow P(x) = x(x+2)(x-5)(x^2 + 2x + 10)$$

Aplicando la fórmula de la ecuación cuadrática en el polinomio $Q_3(x) = x^2 + 2x + 10$ se tiene:

$$x = \frac{-2 \pm \sqrt{4 - 40}}{2} = \frac{-2 \pm \sqrt{-36}}{2} = \frac{-2 \pm \sqrt{(-1)36}}{2} = \frac{-2 \pm \sqrt{-1}\sqrt{36}}{2} = \frac{-2 \pm i\sqrt{36}}{2} = \frac{-2 \pm 6i}{2} = -1 \pm 3i$$

Que indica que es irreducible, por tanto la factorización del polinomio pedido se puede expresar de estas dos formas:

$$P(x) = x(x+2)(x-5)(x^2 + 2x + 10) = x(x+2)(x-5)[(x+1)^2 + 3^2]$$

A.7. Distancia Euclídea

La **distancia euclídea** o **euclidiana** es la considerada como distancia ordinaria y habitualmente se le suele denominar simplemente como *distancia*.

A.7.1. Valor Absoluto

Definición A.9. Valor Absoluto

Dado un número real $x \in \mathbb{R}$ se define su **valor absoluto** como:

$$|x| = \sqrt{x^2} = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}$$

El concepto de valor absoluto se puede generalizar para número complejos $\mathbb{C}$ a través del **módulo de un número complejo**⁶:

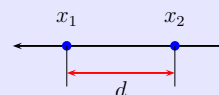
$$z = a + bi \in \mathbb{C} \quad \implies \quad |z| = \sqrt{a^2 + b^2}$$

A.7.2. Distancia en $\mathbb{R}$

Definición A.10. Distancia en $\mathbb{R}$

La **distancia entre dos números reales** x_1 y x_2 se representa mediante $d(x_1, x_2)$ y se define como:

$$d(x_1, x_2) = |x_1 - x_2| = \sqrt{(x_1 - x_2)^2} = \sqrt{(x_2 - x_1)^2} = |x_2 - x_1| = d(x_2, x_1)$$



Nota A.5.

$$\left. \begin{aligned} (a-b)^2 &= a^2 + b^2 - 2ab \\ (b-a)^2 &= b^2 + a^2 - 2ab \end{aligned} \right\} \implies (a-b)^2 = (b-a)^2$$

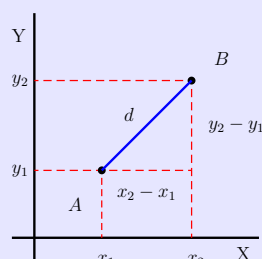
Ejemplo A.4. Distancia euclídea entre dos números.

$$d(2, -5) = |2 - (-5)| = |7| = 7 \quad ; \quad d(-5, 2) = |-5 - 2| = |-7| = 7$$

A.7.3. Distancia en $\mathbb{R}^2$

Definición A.11. Distancia en $\mathbb{R}^2$

En el plano $\mathbb{R}^2$ la **distancia entre dos puntos** $A = (x_1, y_1)$ y $B = (x_2, y_2)$, se define a través del Teorema de Pitágoras.



$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = d(B, A)$$

Teorema de Pitágoras : $d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$

⁶Ver el Apéndice C, página 295.

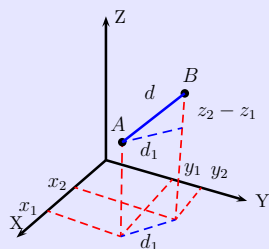
Ejemplo A.5. Distancia euclídea entre dos puntos del plano.

$$d[(1, 2), (-1, -1)] = \sqrt{[1 - (-1)]^2 + [2 - (-1)]^2} = \sqrt{4 + 9} = \sqrt{13}$$

A.7.4. Distancia en $\mathbb{R}^3$

Definición A.12. Distancia en $\mathbb{R}^3$

En el espacio $\mathbb{R}^3$ la **distancia entre dos puntos** $A = (x_1, y_1, z_1)$ y $B = (x_2, y_2, z_2)$ se define como:



$$d(A, B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} = d(B, A)$$

$$d_1^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$$

$$d(A, B) = \sqrt{d_1^2 + (z_2 - z_1)^2}$$

A.7.5. Distancia en $\mathbb{R}^n$

Definición A.13. Distancia en $\mathbb{R}^n$

De forma general y aunque no siempre se pueda representar gráficamente, en el espacio $\mathbb{R}^n$ la **distancia entre dos puntos** $A = (x_1, x_2, \dots, x_n)$ y $B = (y_1, y_2, \dots, y_n)$ se define como:

$$d(A, B) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} = d(B, A)$$

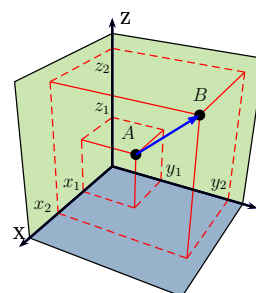
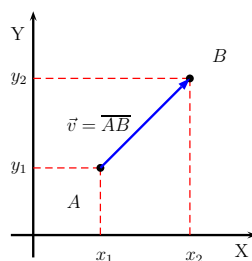
A.8. Vectores en el Plano y en el Espacio

Definición A.14. Vector Geométrico en $\mathbb{R}^n$

Un **vector⁷ geométrico** en $\mathbb{R}^n$ es un segmento de recta dirigido que corresponde a un desplazamiento desde un punto $A(x_1, x_2, \dots, x_n)$ a otro $B(y_1, y_2, \dots, y_n)$.

$$\vec{v} = \overrightarrow{AB} = [y_1 - x_1, y_2 - x_2, \dots, y_n - x_n] \begin{cases} A & : \text{ punto inicial} \\ B & : \text{ punto final} \end{cases}$$

Particularizando al plano euclídeo $\mathbb{R}^2$ y al espacio 3D $\mathbb{R}^3$, los vectores geométricos pueden representarse gráficamente de la siguiente forma.

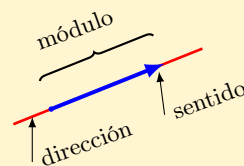


⁷La palabra *vector* proviene de la raíz latina *vector, vectoris* que significa “que conduce”.

Características de los Vectores geométricos

Los vectores geométricos se caracterizan por tres elementos:

- **Módulo**⁸: es la longitud del segmento.
- **Dirección**: es la orientación de la recta.
- **Sentido**: indica cual el punto inicial y el punto final.



A.8.1. Vectores de Posición

Definición A.15. Vector Posición

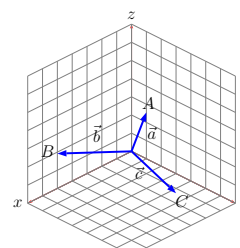
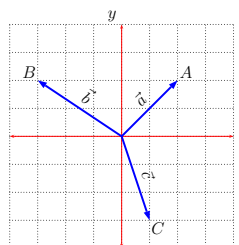
Se denomina **vector posición** al vector geométrico $\overrightarrow{AB}$ en el que el punto inicio es el origen $A = O = (0, 0, 0)$.

Es natural representar los vectores posición utilizando **coordenadas** encerradas entre *paréntesis cuadrados*, $[\]$. A las coordenadas se le denominan **componentes** del vector.

$$\text{Plano (2D)} \quad \{ A = (x, y) \} \implies \vec{a} = \overrightarrow{OA} = [x, y]$$

$$\text{Espacio (3D)} \quad \{ B = (x, y, z) \} \implies \vec{b} = \overrightarrow{OB} = [x, y, z]$$

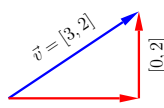
Vectores posición en el plano y en el espacio 3D. En ambos casos los vectores tienen su punto de inicio en el origen de coordenadas.



A.8.2. Vectores que no son de posición: Vectores Generales

Definición A.16. Vectores Generales

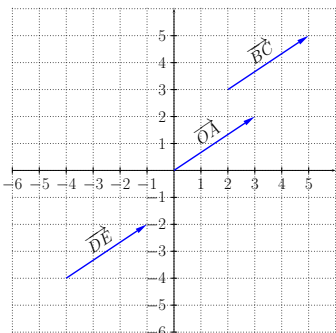
Un vector se puede considerar como un *ente* que dirige una cierta **cantidad** en una cierta **dirección** y **sentido**.



Por ejemplo, el vector $\vec{v} = [3, 2]$ se puede interpretar como: “Estando en un punto cualquiera, muévete 3 unidades hacia la derecha y después 2 unidades hacia arriba”.

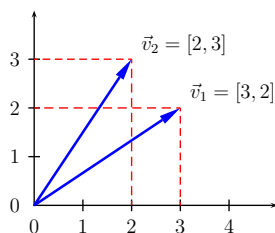
$$\vec{v} = [3, 2] = [3, 0] + [0, 2]$$

⁸La *norma*, (ver la Sección 6.2.1, página 164), generaliza el concepto de *módulo* de un vector en el espacio euclídeo.



En este sentido, los *vectores geométricos* $\overrightarrow{OA}$, $\overrightarrow{BC}$ y $\overrightarrow{DE}$ mostrados en esta figura, corresponden con el mismo *vector (general)* $[3, 2]$.

$$\begin{aligned}\overrightarrow{OA} &= [3 - 0, 2 - 0] &&= [3, 2] \\ \overrightarrow{BC} &= [5 - 2, 5 - 3] &&= [3, 2] \\ \overrightarrow{DE} &= [-1 - (-4), -2 - (-4)] &&= [3, 2]\end{aligned}$$



Las *componentes* o *coordenadas* de un vector son un conjunto de números *ordenados* (pares en el plano, triadas en el espacio, etc.)

$$\vec{v}_1 = [3, 2] \neq \vec{v}_2 = [2, 3]$$

A.8.3. Vectores Elementales

- Vector nulo o cero : $\vec{0} = [0, 0, 0]$
- Vector fila: $\vec{v} = [x, y, z]$
- Vector columna : $\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

Nota A.6.

- El vector nulo $\vec{0} = [0, 0, 0]$ no se puede dibujar, ya que corresponde con el origen.
- La distinción entre vector fila y vector columna quedará clara en la formación de matrices, (ver [Tema 3](#)).
- Los vectores tienen siempre más de una componente. Se denominan **escalares** a los números (una única componente) reales o complejos.

A.9. Vectores en Espacios $\mathbb{R}^n$, $\mathbb{C}^n$ y $\mathbb{K}^n$

Definición A.17. Vectores en $\mathbb{R}^n$

El conjunto de todos los vectores de n componentes reales se representa por $\mathbb{R}^n$.

$$\mathbb{R}^n = \{[x_1, x_2, \dots, x_n] : x_i \in \mathbb{R}, 1 \leq i \leq n\}$$

Ejemplo A.6. Vectores en espacios $\mathbb{R}^n$.

$$\begin{aligned}\mathbb{R}^2 &= \{[x, y] : x, y \in \mathbb{R}\} &&\longrightarrow [1, 0] \in \mathbb{R}^2 \\ \mathbb{R}^3 &= \{[x, y, z] : x, y, z \in \mathbb{R}\} &&\longrightarrow [1, -1, 1/3] \in \mathbb{R}^3\end{aligned}$$

Nota A.7. Los vectores $\mathbb{R}^n$ para $n > 3$ **no** se pueden representar gráficamente.

Definición A.18. Vectores en $\mathbb{C}^n$

El conjunto de todos los vectores de n componentes complejas⁹ se representa por $\mathbb{C}^n$.

$$\mathbb{C}^n = \{[x_1, x_2, \dots, x_n] : x_i \in \mathbb{C}, 1 \leq i \leq n\}$$

Ejemplo A.7. Vectores en espacios $\mathbb{C}^n$.

$$\begin{aligned} \mathbb{C}^2 &= \{[x, y] : x, y \in \mathbb{C}\} & \longrightarrow & [0, i] \in \mathbb{C}^2 \\ \mathbb{C}^3 &= \{[x, y, z] : x, y, z \in \mathbb{C}\} & \longrightarrow & [2 + i, i, 1] \in \mathbb{C}^3 \end{aligned}$$

Nota A.8. En general los vectores $\mathbb{C}^n$ **no** se pueden representar gráficamente para ningún valor de $n > 1$.

Definición A.19. Vectores en $\mathbb{K}^n$

De forma general, se puede hablar de vectores en espacios $\mathbb{K}^n$ donde $\mathbb{K}$ puede ser o bien $\mathbb{R}$ o bien $\mathbb{C}$.

Nomenclatura

Generalmente la notación utilizada para referirse a los vectores es la siguiente:

$$\begin{cases} 1. \vec{v} \in \mathbb{K} \\ 2. \mathbf{v} \in \mathbb{K} \end{cases}$$

La segunda forma, utilizando letras en *negrita*, se suele usar con vectores complejos para que se pueda distinguir claramente el conjugado a través de su *rayita*.

A.9.1. Caracterización de Vectores

Además de por sus coordenadas los vectores de $\mathbb{R}^n$ pueden caracterizarse por su longitud y ángulo.

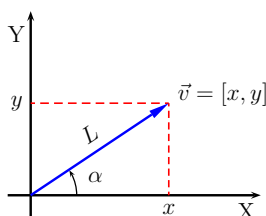
Norma de un Vector en $\mathbb{R}^n$ **Definición A.20. Norma de un Vector en $\mathbb{R}^n$**

Dado un vector $\vec{v} = [x_1, x_2, \dots, x_n]$ de un espacio cualquier $\mathbb{R}^n$, se define la **norma euclídea** y se representa por $\|\vec{v}\|$, a la distancia desde el origen al extremo del vector:

$$\|\vec{v}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

Longitud y Ángulo en $\mathbb{R}^2$

Dado un vector $\vec{v} = [x, y]$ en $\mathbb{R}^2$ se define su longitud L y su ángulo¹⁰ α como los valores:



$$L = \sqrt{x^2 + y^2} \quad (\text{Teorema de Pitágoras})$$

$$\begin{cases} \alpha = \arctan \frac{y}{x} & \text{si } x \neq 0 \\ \alpha = \pm \frac{\pi}{2} & \text{si } x = 0 \end{cases} \longrightarrow \begin{cases} y > 0 & \alpha = \pi/2 \\ y < 0 & \alpha = -\pi/2 \end{cases}$$

⁹Ver el **Apéndice C**, página 295.

¹⁰Para $\mathbb{R}^3$ existen 3 *vectores directores*, uno por cada dimensión.

A.10. Producto Escalar y Vectorial en $\mathbb{R}^3$

A.10.1. Producto Escalar

Definición A.21. Producto Escalar de Vectores en $\mathbb{R}^3$

Dados dos vectores $\vec{u} = [u_1, u_2, u_3]$ y $\vec{v} = [v_1, v_2, v_3]$ de $\mathbb{R}^3$, se define el producto escalar de ambos, y se representa por $\vec{u} \cdot \vec{v}$, al número real definido por:

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + u_3v_3$$

Definición A.22. Ángulo entre dos Vectores de $\mathbb{R}^n$

Dados dos vectores $\vec{u}$ y $\vec{v}$ de $\mathbb{R}^n$, se define el ángulo¹¹ α formado entre ellos mediante la siguiente expresión:

$$\cos \alpha = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|}$$

Definición A.23. Vectores Perpendiculares

Se dice que dos vectores de $\mathbb{R}^2$ ó $\mathbb{R}^3$ son **perpendiculares**, si el ángulo formado entre ellos es un ángulo recto¹². En estos casos, el producto escalar entre ellos es cero.

A.10.2. Producto Vectorial

Definición A.24. Producto Vectorial en $\mathbb{R}^3$

Dados dos vectores $\vec{u} = [u_1, u_2, u_3]$ y $\vec{v} = [v_1, v_2, v_3]$ de $\mathbb{R}^3$, se define el producto vectorial de ambos, y se representa por $\vec{u} \times \vec{v}$ o por $\vec{u} \wedge \vec{v}$, al vector de $\mathbb{R}^3$ definido por:

$$\vec{u} \times \vec{v} = \vec{u} \wedge \vec{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = (u_2v_3 - u_3v_2)\vec{i} + (u_3v_1 - u_1v_3)\vec{j} + (u_1v_2 - u_2v_1)\vec{k}$$

El vector $\vec{u} \times \vec{v}$ tiene la **dirección perpendicular** al plano formado por los vectores $\vec{u}$ y $\vec{v}$ y **sentido** siguiendo la *regla de la mano derecha*¹³.

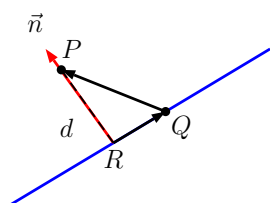
¹¹Medido en radianes.

¹²Es decir su valor es $\pm\pi/2$ rad = 90° .

¹³Con los dedos de la mano derecha hacemos un movimiento circular del primer vector $\vec{u}$ al segundo vector $\vec{v}$, entonces el sentido será el apuntado por el dedo pulgar.

A.11. Fórmulas Geométricas sobre Distancias

A.11.1. Distancia de un Punto a una Recta en $\mathbb{R}^2$



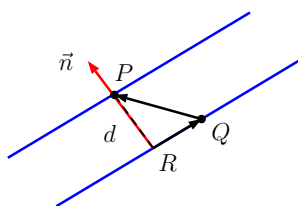
$$\begin{cases} \text{Punto} & P = (x_1, y_1) \\ \text{Recta} & Ax + By + C = 0 \end{cases}$$

$$d = \frac{|Ax_1 + By_1 + C|}{\sqrt{A^2 + B^2}}$$

Demostración

$$\begin{aligned} \vec{RP} &= \vec{RQ} + \vec{QP} ; \quad \vec{n} \cdot \vec{RP} = \vec{n} \cdot \vec{RQ} + \vec{n} \cdot \vec{QP} ; \quad \vec{n} \cdot \vec{RP} = 0 + \vec{n} \cdot \vec{QP} ; \\ |\vec{n} \cdot \vec{RP}| &= |\vec{n} \cdot \vec{QP}| ; \quad \|\vec{n}\| \|\vec{RP}\| = |\vec{n} \cdot \vec{QP}| ; \quad d = \|\vec{RP}\| = \frac{|\vec{n} \cdot \vec{QP}|}{\|\vec{n}\|} \end{aligned}$$

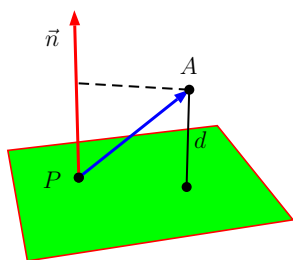
A.11.2. Distancia entre dos Rectas paralelas en $\mathbb{R}^2$



$$\begin{cases} \text{Punto Recta 1} & P = (x_1, y_1) \\ \text{Recta 2} & Ax + By + C = 0 \end{cases}$$

$$d = \frac{|Ax_1 + By_1 + C|}{\sqrt{A^2 + B^2}}$$

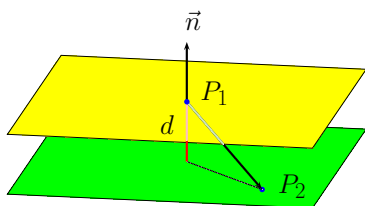
A.11.3. Distancia de un Punto a un Plano en $\mathbb{R}^3$



$$\begin{cases} \text{Punto} & P = (x_1, y_1, z_1) \\ \text{Plano} & Ax + By + Cz + D = 0 \end{cases}$$

$$d = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

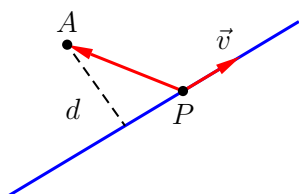
A.11.4. Distancia entre dos Planos Paralelos en $\mathbb{R}^3$



$$\begin{cases} \text{Punto del Plano 1} & P = (x_1, y_1, z_1) \\ \text{Plano 2} & Ax + By + Cz + D = 0 \end{cases}$$

$$d = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

A.11.5. Distancia de un Punto a una Recta en $\mathbb{R}^3$



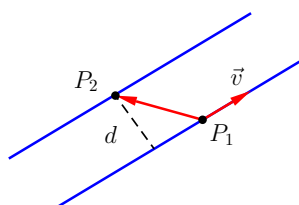
$$\begin{cases} \text{Punto} & A = (x_1, y_1, z_1) \\ \text{Recta} & \vec{x} = \vec{p} + t\vec{v} \end{cases}$$

$$d = \frac{\|\vec{v} \times \overrightarrow{PA}\|}{\|\vec{v}\|}$$

Demostración

$$\|\vec{v} \times \overrightarrow{PA}\| = \|\vec{v}\| \|\overrightarrow{PA}\| \sin(\theta) = \|\vec{v}\| d$$

A.11.6. Distancia entre dos Rectas paralelas en $\mathbb{R}^3$



$$\begin{cases} \text{Punto Recta 2} & P_2 = (x_2, y_2, z_2) \\ \text{Recta 1} & \vec{x} = \vec{p}_2 + t\vec{v} \end{cases}$$

$$d = \frac{\|\vec{v} \times \overrightarrow{P_1P_2}\|}{\|\vec{v}\|}$$

Ejercicio A.17. Dado el plano $\Pi : x - 2y - 2z = 3$, encontrar un plano paralelo a Π que diste 4 unidades de él.

Un *vector normal*^a al plano Π viene dado por los coeficientes de las variables x , y y z :

$$\vec{n} = [1, -2, -2]$$

Todo plano Π' paralelo a Π tiene un vector normal que es paralelo a $\vec{n}$, por tanto podemos representar su ecuación general como:

$$\Pi' : Ax + By + Cz + D = 0 \quad \implies \quad x - 2y - 2z + \alpha = 0$$

Siendo α el valor a determinar para que dicho plano Π' diste 4 unidades de Π .

Como la distancia entre dos planos paralelos se puede expresar mediante la ecuación^b:

$$d = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}} = \frac{|x_1 - y_1 - z_1 + \alpha|}{\sqrt{1^2 + (-2)^2 + (-2)^2}} = \frac{|x_1 - y_1 - z_1 + \alpha|}{\sqrt{9}} = 4$$

donde $P = (x_1, y_1, z_1)$ son las coordenadas de un punto cualquiera del plano Π , por ejemplo $P = (3, 0, 0)$:

$$4 = \frac{|3 - 0 - 0 + \alpha|}{3} \quad ; \quad 12 = |3 + \alpha| \quad ; \quad 3 + \alpha = \pm 12 \quad \implies \quad \alpha = \pm 12 - 3 = \{9, -15\}$$

Aparecen dos valores posibles de α correspondientes a los dos planos (por encima y por debajo de Π) que distan 4 unidades.

Por tanto las dos posibles soluciones del problema son los siguientes planos:

$$x - 2y - 2z = -9 \quad ; \quad x - 2y - 2z = 15$$

A.12. Introducción a las Funciones

Definición A.25. Función o Aplicación

Dados dos conjuntos X e Y , se denomina **función** o **aplicación** a toda regla que hace corresponder a cada elemento x del conjunto inicial X (**conjunto dominio**), un y solo un elemento y del conjunto final Y (**conjunto imagen**).

Se representa por:

$$f : X \mapsto Y \quad ; \quad x \xrightarrow{f} y$$

A.12.1. Tipos de Funciones

Funciones Reales y Complejas

- Si $X \subseteq \mathbb{R}$ se habla de **variable real**.
- Si $X \subseteq \mathbb{C}$ se habla de **variable compleja**.
- Si $Y \subseteq \mathbb{R}$ se habla de **funciones reales**.
- Si $Y \subseteq \mathbb{C}$ se habla de **funciones complejas**.

Ejemplo A.8. Tipos de funciones.

$f : \mathbb{R} \mapsto \mathbb{R}$	Función Real de Variable Real.	Ej: $f(x) = x^2$	; $x \in \mathbb{R}$
$f : \mathbb{C} \mapsto \mathbb{R}$	Función Real de Variable Compleja.	Ej: $f(x) = x $	; $x \in \mathbb{C}$
$f : \mathbb{R} \mapsto \mathbb{C}$	Función Compleja de Variable Real.	Ej: $f(x) = ix$	; $x \in \mathbb{R}$
$f : \mathbb{C} \mapsto \mathbb{C}$	Función Compleja de Variable Compleja.	Ej: $f(x) = x^2$	; $x \in \mathbb{C}$

Funciones Escalares y Vectoriales

Dada una función general:

$$f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$$

- Si $\begin{cases} n = 1 & \text{Función de una variable} \\ n > 1 & \text{Función de varias variables.} \end{cases}$
- Si $\begin{cases} m = 1 & \text{Función Escalar} \\ m > 1 & \text{Función Vectorial.} \end{cases}$

$$f(x_1, x_2, \dots, x_n) = [f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_m(x_1, x_2, \dots, x_n)]$$

Ejemplo A.9. Ejemplos de Funciones Reales.

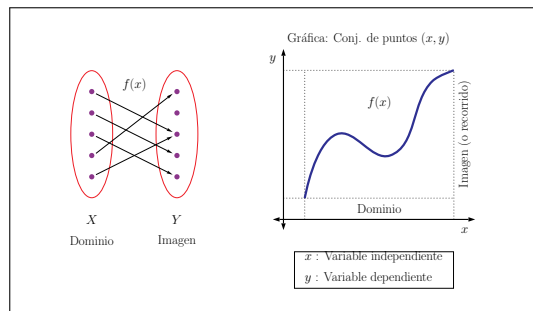
$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$	$F(x, y) = [x, \text{sen}(x, y)]$	Función Vectorial de varias variables.
$g : \mathbb{R}^3 \longrightarrow \mathbb{R}$	$g(x, y, z) = x + yz$	Función Escalar de varias variables.
$h : \mathbb{R} \longrightarrow \mathbb{R}^3$	$H(x) = [x, x^2, x^3]$	Función Vectorial de una variable.

^aVer la Sección 0.3.3, página 13.

^bVer la Sección A.11.4, página 275.

A.12.2. Representación Gráfica de Funciones Reales

Existen dos formas muy utilizadas de representar gráficamente una función real de variable real $f: X \rightarrow Y$.



- A través de diagramas de Venn.
- En un eje de coordenadas cartesianas.

Para las tradicionales funciones escalares:

$$f: \mathbb{R}^n \longrightarrow \begin{cases} \text{Si } n = 1 & \implies \text{La gráfica es una } \textit{curva} \text{ en } \mathbb{R}^2. \\ \text{Si } n = 2 & \implies \text{La gráfica es una } \textit{superficie} \text{ en } \mathbb{R}^3. \\ \text{Si } n \geq 3 & \implies \text{La gráfica } \textit{no} \text{ se puede visualizar.} \end{cases}$$

Las funciones vectoriales clásicas que se pueden dibujar son:

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \quad ; \quad f: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

Representación gráfica de funciones.

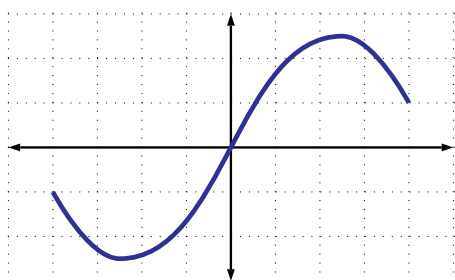


Figura A.4: $f: \mathbb{R} \mapsto \mathbb{R}$

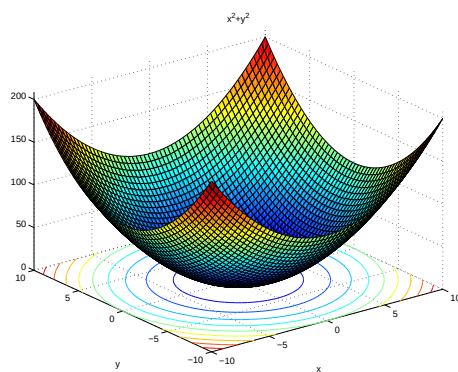


Figura A.5: $f: \mathbb{R}^2 \mapsto \mathbb{R}$

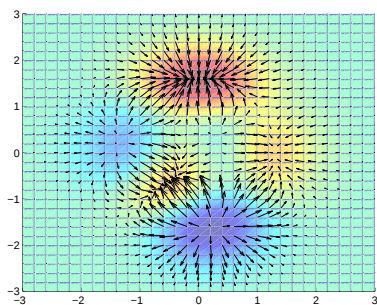


Figura A.6: $f: \mathbb{R}^2 \mapsto \mathbb{R}^2$

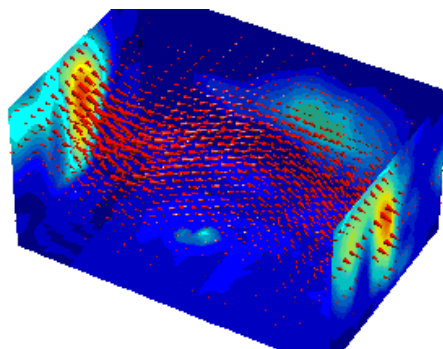


Figura A.7: $f: \mathbb{R}^3 \mapsto \mathbb{R}^3$

A.12.3. Tipos de Funciones

- **Función Inyectiva:** si a elementos distintos del dominio le corresponden elementos distintos de la imagen.

$$f : A \mapsto B \text{ es inyectiva} \iff \forall x, y \in A : x \neq y \longrightarrow f(x) \neq f(y)$$

- **Función Sobreyectiva:** si todo elemento del conjunto imagen es imagen de algún elemento del dominio.

$$f : A \mapsto B \text{ es sobreyectiva} \iff \forall y \in B : \exists x \in A : f(x) = y$$

- **Función Biyectiva:** si es inyectiva y sobreyectiva a la vez.

Tipos de funciones.

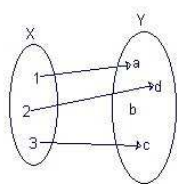


Figura A.8: Inyectiva, no sobreyectiva.

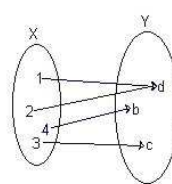


Figura A.9: Sobreyectiva, no inyectiva.

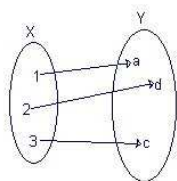


Figura A.10: Biyectiva.

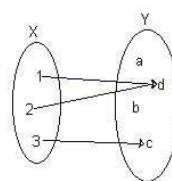


Figura A.11: No sobreyectiva, no inyectiva.

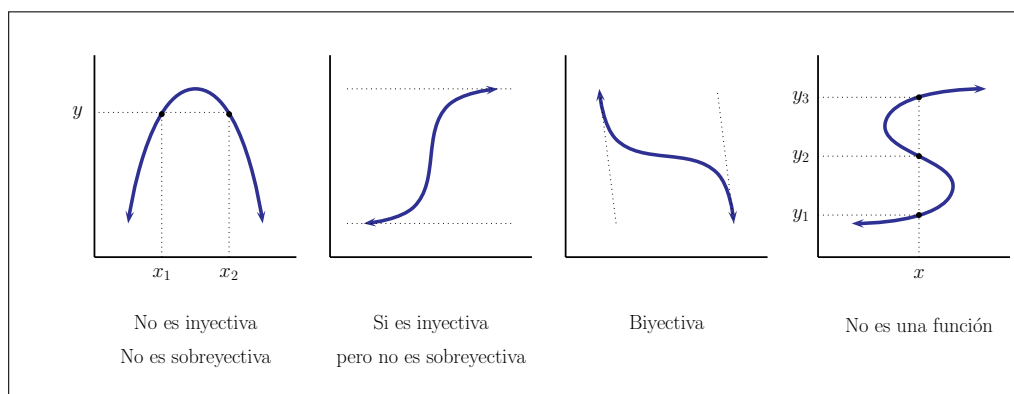


Figura A.12: Ejemplos gráficos de funciones reales de variable real.

A.12.4. Función Inversa (o Recíproca) $f^{-1}(x)$

Definición A.26. Función Inversa o Recíproca

Dada una función *biyectiva* $f : X \longrightarrow Y$, se define la **función inversa** (o **recíproca**) de f como:

$$f^{-1} : Y \longrightarrow X$$

En caso de que exista, la función inversa $f^{-1}(x)$ de una función dada $f(x)$ es aquella función que tiene por conjunto dominio el conjunto imagen de $f(x)$ y por conjunto imagen el conjunto dominio de $f(x)$.

La representación gráfica de $f^{-1}(x)$ es una reflexión de la gráfica de $f(x)$ con respecto a la recta $y = x$. La **Figura A.13** muestra gráficamente dos casos sencillos de la función inversa, en relación a la función original.

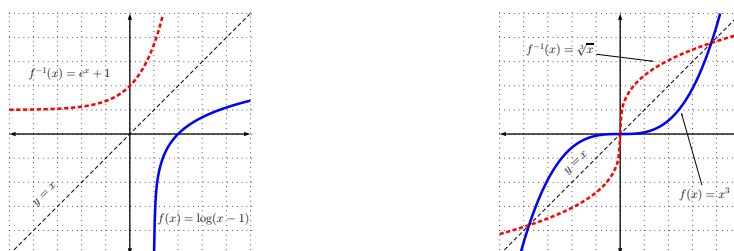


Figura A.13: Dos ejemplos de representación gráfica de una función $f(x)$ junto a su correspondiente función inversa $f^{-1}(x)$. Puede apreciarse como ambas son una reflexión respecto de la recta $y = x$.

La condición necesaria y suficiente para que una función sea *invertible* es que sea *biyectiva*¹⁴.

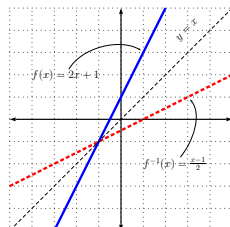
Nota A.9. La función inversa $f^{-1}(x)$ es distinta que la inversa de una función $[f(x)]^{-1}$:

$$\underbrace{[f(x)]^{-1} = \frac{1}{f(x)}}_{\text{Inversa de la función}} \neq \underbrace{f^{-1}(x)}_{\text{Función inversa}}$$

Ejercicio A.18. Calcular, en caso de que exista, la función inversa $f^{-1}(x)$ de $f(x) = 2x + 1$.

Esta función es claramente inyectiva al ser una recta no vertical $\implies$ sí existe su función inversa. Para calcularla:

1. $y = 2x + 1$
2. $x = \frac{y - 1}{2}$
3. $f^{-1}(x) = \frac{x - 1}{2}$



¹⁴Ver el Apéndice A.12.3, página 279.

A.12.5. Composición de Funciones

Definición A.27. Composición de Funciones

Dadas las funciones:

$$\begin{cases} g: X \mapsto Y \\ f: Y \mapsto Z \end{cases}$$

se define, la *composición de las funciones* g y f de la siguiente forma¹⁵:

$$(f \circ g)(x) = f[g(x)] \qquad x \xrightarrow{g} g(x) \xrightarrow{f} f[g(x)]$$

$$(f \circ g): X \mapsto Z$$

- Para poder realizar la composición $(f \circ g)(x) = f[g(x)]$ es necesario que el dominio de f coincida con la imagen de g .
- El dominio de la nueva función $(f \circ g)(x)$ es el dominio de $g(x)$ y su imagen es la imagen de $f(x)$.

Ejercicio A.19. Dadas las siguientes funciones $f(x)$ y $g(x)$ determinar, en caso de ser posible, las composiciones $(f \circ g)(x)$ y $(g \circ f)(x)$.

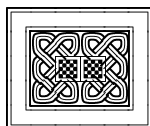
$$f(x) = \cos(x) \qquad ; \qquad g(x) = x^2 + 1$$

$$\left. \begin{array}{l} f(x) = \cos(x) \\ g(x) = x^2 + 1 \end{array} \right\} \Rightarrow \begin{cases} (f \circ g)(x) = f[g(x)] = \cos(x^2 + 1) \\ (g \circ f)(x) = g[f(x)] = \cos^2(x) + 1 \end{cases}$$

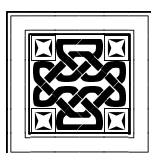
La composición múltiple de funciones es *asociativa* pero no es *conmutativa*.

$$\left. \begin{array}{l} [h \circ (g \circ f)](x) = h[(g \circ f)(x)] = h\{g[f(x)]\} \\ [(h \circ g) \circ f](x) = (h \circ g)[f(x)] = h\{g[f(x)]\} \end{array} \right\} \text{ Iguales !! } \Rightarrow \text{ Es asociativa}$$

Sin embargo generalmente: $(f \circ g)(x) \neq (g \circ f)(x)$ como se comprueba en el anterior **Ejemplo A.19**.



¹⁵La expresión $f \circ g$ se lee *composición de f y g*, o bien, *g compuesta con f*. Notar que no se nombra siguiendo el orden de escritura, sino siguiendo el orden en que se aplican en la composición: $(f \circ g)(x) = f[g(x)]$. Primero se aplica la función g y posteriormente la función f .



Aplicaciones Geométricas

Contenido

B.1	Introducción	284
B.2	Aplicaciones Geométricas de la Proyección Ortogonal	284
B.2.1	Distancia de un Punto A a una Recta en $\mathbb{R}^2$ y $\mathbb{R}^3$	284
B.2.2	Distancia entre dos Rectas paralelas en $\mathbb{R}^2$ y $\mathbb{R}^3$	285
B.2.3	Distancia de un Punto A a un Plano en $\mathbb{R}^3$	286
B.2.4	Distancia entre Dos Planos Paralelos en $\mathbb{R}^3$	286
B.3	Transformaciones Geométricas. Movimientos en $\mathbb{R}^2$ y $\mathbb{R}^3$	287
B.3.1	Traslaciones en $\mathbb{R}^2$ y $\mathbb{R}^3$	287
B.3.2	Rotaciones en $\mathbb{R}^2$	288
B.3.3	Rotaciones en $\mathbb{R}^3$	290
B.3.4	Simetría Puntual en $\mathbb{R}^2$	290
B.3.5	Reflexiones sobre Rectas en $\mathbb{R}^2$	290
B.3.6	Reflexiones sobre Rectas en $\mathbb{R}^3$	292
B.3.7	Reflexiones sobre Planos en $\mathbb{R}^3$	293
B.3.8	Homotecias en $\mathbb{R}^2$ y $\mathbb{R}^3$	293

B.1. Introducción

En este apéndice se muestran algunas aplicaciones geométricas de la *proyección ortogonal*¹ y de las *transformaciones lineales*².

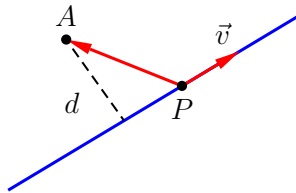
B.2. Aplicaciones Geométricas de la Proyección Ortogonal

A continuación se muestran algunas fórmulas prácticas en la determinación de *distancias* en $\mathbb{R}^2$ y $\mathbb{R}^3$, obtenidas mediante consideraciones geométricas y utilizando el concepto de *proyección ortogonal*³.

B.2.1. Distancia de un Punto A a una Recta en $\mathbb{R}^2$ y $\mathbb{R}^3$

Para calcular la distancia d de un punto A a una determinada recta, podemos utilizar la proyección ortogonal de la siguiente forma:

1. Tomamos un punto P cualquiera de la recta y formamos el vector⁴ $\overrightarrow{PA}$.
2. Calculamos la proyección ortogonal de $\overrightarrow{PA}$ sobre el vector director de la recta $\vec{v}$.
3. La distancia buscada será la norma del vector diferencia $\overrightarrow{PA} - \text{proy}_{\vec{v}}(\overrightarrow{PA})$.

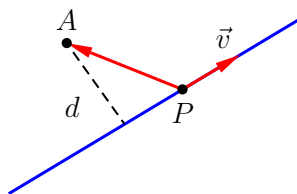


$$d = \|\overrightarrow{PA} - \text{proy}_{\vec{v}}(\overrightarrow{PA})\| \quad \left\{ \begin{array}{ll} P & : \text{Punto cualquiera de la recta.} \\ \vec{v} & : \text{Vector director de la recta.} \end{array} \right.$$

Ejercicio B.1. En $\mathbb{R}^2$ encontrar la distancia entre el punto $P = (2, 2)$ y la recta:

$$\ell: \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Tomamos un punto cualquiera de la recta, por ejemplo para $t = 1$ se tiene $P = (0, 1)$. Entonces calculamos el vector $\overrightarrow{PA}$ y su proyección ortogonal sobre el vector director de la recta $\vec{v} = [1, -1]$. Finalmente aplicamos la fórmula de la distancia punto-recta.



$$d = \|\overrightarrow{PA} - \text{proy}_{\vec{v}}(\overrightarrow{PA})\|$$

$$\overrightarrow{PA} = [2, 2] - [0, 1] = [2, 1] \quad ; \quad \text{proy}_{\vec{v}}(\overrightarrow{PA}) = \frac{[1, -1] \cdot [2, 1]}{[1, -1] \cdot [1, -1]} [1, -1] = \frac{1}{2} [1, -1]$$

$$d = \|\overrightarrow{PA} - \text{proy}_{\vec{v}}(\overrightarrow{PA})\| = \|[2, 1] - \frac{1}{2} [1, -1]\| = \left\| \left[\frac{3}{2}, \frac{3}{2} \right] \right\| = \frac{3}{2} \sqrt{2}$$

¹Ver la Sección 6.5, página 178.

²Ver el Tema 4, página 119.

³En el Apéndice A.11, página 275, se encuentran algunas fórmulas alternativas.

⁴Podemos calcular indistintamente $\overrightarrow{PA}$ ó $\overrightarrow{AP}$.

Ejercicio B.2. Utilizando el concepto de proyección ortogonal, encontrar la distancia del punto $A = (1, -1, 0)$ a la recta:

$$\ell : \begin{cases} x + y - z = 1 \\ x - y - z = 0 \end{cases}$$

Vamos a utilizar la fórmula anterior. Conocido el punto A de los datos del problema, se necesita un punto de la recta P y un vector director de la misma $\vec{v}$. Para obtenerlos, vamos a expresar la ecuación general de la recta en su forma vectorial, resolviendo el SEL por Gauss.

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{array} \right] \xrightarrow{F_2 - F_1} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -2 & 0 & -1 \end{array} \right] \Rightarrow \begin{cases} x = 1 - y + z = 1 - \frac{1}{2} + t = \frac{1}{2} + t \\ y = \frac{-1}{-2} = \frac{1}{2} \\ z = t \end{cases}$$

A través de la forma vectorial se puede obtener fácilmente un punto y el vector director⁵:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \Rightarrow \begin{cases} P = (\frac{1}{2}, \frac{1}{2}, 0) \\ \vec{v} = [1, 0, 1] \end{cases}$$

Vamos ahora a calcular los vectores $\overrightarrow{PA}$ y $\text{proy}_{\vec{v}}(\overrightarrow{PA})$.

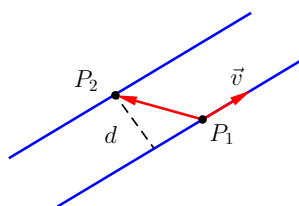
$$\overrightarrow{PA} = (1, -1, 0) - \left(\frac{1}{2}, \frac{1}{2}, 0\right) = \left[\frac{1}{2}, -\frac{3}{2}, 0\right]; \text{proy}_{\vec{v}}(\overrightarrow{PA}) = \frac{[1/2, -3/2, 0] \cdot [1, 0, 1]}{[1, 0, 1] \cdot [1, 0, 1]} [1, 0, 1] = \frac{1}{4} [1, 0, 1]$$

Finalmente:

$$d = \left\| \left[\frac{1}{2}, -\frac{3}{2}, 0\right] - \left[\frac{1}{4}, 0, \frac{1}{4}\right] \right\| = \left\| \left[\frac{1}{4}, -\frac{3}{2}, -\frac{1}{4}\right] \right\| = \sqrt{\frac{1}{16} + \frac{9}{4} + \frac{1}{16}} = \sqrt{\frac{38}{16}} = \frac{\sqrt{38}}{4}$$

B.2.2. Distancia entre dos Rectas paralelas en $\mathbb{R}^2$ y $\mathbb{R}^3$

Este caso es semejante al anterior, tomando un punto cualquiera P_2 de una de las rectas y calculando la distancia de ese punto a la otra recta.



$$d = \|\overrightarrow{P_1 P_2} - \text{proy}_{\vec{v}}(\overrightarrow{P_1 P_2})\|$$

$$\begin{cases} P_1 : & \text{Pto cualquiera de la recta 1.} \\ P_2 : & \text{Pto cualquiera de la recta 2.} \\ \vec{v} : & \text{Vector director de las rectas.} \end{cases}$$

Ejercicio B.3. En $\mathbb{R}^3$ encontrar la distancia entre las siguientes dos rectas paralelas:

$$\ell_1 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad ; \quad \ell_2 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Tomamos un punto cualquiera de la primera recta, por ejemplo para $s = 0$ se tiene $P_2 = (1, 0, -1)$. Entonces calculamos la distancia de P_2 a la recta ℓ_2 .

$$d = \|\overrightarrow{P_1 P_2} - \text{proy}_{\vec{v}}(\overrightarrow{P_1 P_2})\|$$

⁵Como punto P se puede tomar cualquiera de la recta, dando valores al parámetro t . Lo más sencillo es tomar el punto para $t = 0$, como se ha hecho en este caso. También como vector director se puede tomar cualquier vector paralelo al vector $[1, 0, 1]$.

$$\overrightarrow{P_1P_2} = [1, 0, -1] - [0, 1, 1] = [1, -1, -2] \quad ; \quad \text{proy}_{\vec{v}}(\overrightarrow{P_1P_2}) = \frac{[1, 1, 1] \cdot [1, -1, -2]}{[1, 1, 1] \cdot [1, 1, 1]} [1, 1, 1] = \frac{-2}{3} [1, 1, 1]$$

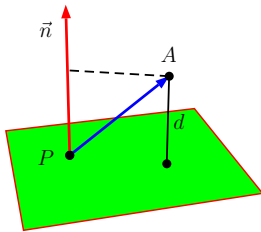
$$\overrightarrow{P_1P_2} - \text{proy}_{\vec{v}}(\overrightarrow{P_1P_2}) = [1, -1, -2] - \frac{-2}{3} [1, 1, 1] = \left[\frac{5}{3}, \frac{-1}{3}, \frac{-4}{3} \right] \quad ; \quad d = \frac{1}{3} \sqrt{5^2 + (-1)^2 + (-4)^2} = \frac{1}{3} \sqrt{42}$$

$$d = \frac{\sqrt{42}}{3}$$

B.2.3. Distancia de un Punto A a un Plano en $\mathbb{R}^3$

En este caso podemos utilizar la proyección ortogonal de la siguiente forma:

1. Tomamos un punto P cualquiera del plano y formamos el vector $\overrightarrow{PA}$.
2. Calculamos la proyección ortogonal de $\overrightarrow{PA}$ sobre el vector normal al plano $\vec{n}$.
3. La distancia buscada coincide con la norma del vector proyección $\text{proy}_{\vec{n}}(\overrightarrow{PA})$.



$$d = \|\text{proy}_{\vec{n}}(\overrightarrow{PA})\| \quad \begin{cases} P & : \text{ Punto cualquiera del plano.} \\ \vec{n} & : \text{ Vector normal al plano.} \end{cases}$$

Ejercicio B.4. En $\mathbb{R}^3$ encontrar la distancia entre el punto $A = (1, 0, 1)$ y el plano $x - y + z = 1$.

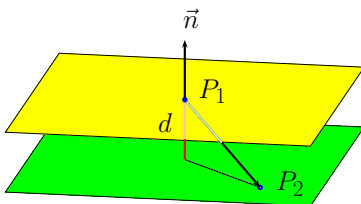
En este caso un vector normal al plano es $\vec{n} = [1, -1, 1]$ y punto del plano puede ser por ejemplo $P = (1, 0, 0)$, por tanto:

$$\overrightarrow{PA} = [1, 0, 1] - [1, 0, 0] = [0, 0, 1] \quad ; \quad \text{proy}_{\vec{n}}(\overrightarrow{PA}) = \frac{[1, -1, 1] \cdot [0, 0, 1]}{[1, -1, 1] \cdot [1, -1, 1]} [1, -1, 1] = \frac{1}{3} [1, -1, 1]$$

$$d = \frac{1}{3} \sqrt{1 + 1 + 1} = \frac{\sqrt{3}}{3}$$

B.2.4. Distancia entre Dos Planos Paralelos en $\mathbb{R}^3$

Este caso es semejante al anterior, tomando un punto cualquiera P_1 de uno de los planos y calculando la distancia de ese punto al otro plano.



$$d = \|\text{proy}_{\vec{n}}(\overrightarrow{P_1P_2})\| \quad \begin{cases} P_1 & : \text{ Pto cualquiera del plano 1.} \\ P_2 & : \text{ Pto cualquiera del plano 2.} \\ \vec{n} & : \text{ Vector normal a los planos.} \end{cases}$$

Ejercicio B.5. En $\mathbb{R}^3$ encontrar la distancia entre los dos planos paralelos:

$$\mathcal{M}_1 : x + y + z = 1 \quad ; \quad \mathcal{M}_2 : x + y + z = 3$$

Tomamos un punto perteneciente al primer plano $P_1 = (1, 0, 0)$ y calculamos la distancia de P_1 al plano $\mathcal{M}_2$. Para ello tomamos un segundo punto M_2 , $P_2 = (1, 1, 1)$ y el un vector normal del mismo $\vec{n} = [1, 1, 1]$. Aplicando la fórmula anterior se tiene:

$$\overrightarrow{P_1 P_2} = [1, 1, 1] - [1, 0, 0] = [0, 1, 1] \quad ; \quad \text{proy}_{\vec{n}}(\overrightarrow{P_1 P_2}) = \frac{[1, 1, 1] \cdot [0, 1, 1]}{[1, 1, 1] \cdot [1, 1, 1]} [1, 1, 1] = \frac{2}{3} [1, 1, 1]$$

$$d = \frac{2\sqrt{3}}{3}$$

B.3. Transformaciones Geométricas. Movimientos en $\mathbb{R}^2$ y $\mathbb{R}^3$

Una *transformación geométrica*, es una aplicación que hace corresponder a cada punto del plano (o espacio) otro punto del plano (o espacio). Como consecuencia, las figuras se transforman en otras figuras.

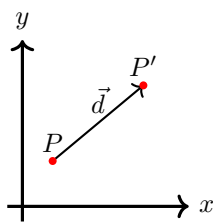
Las *transformaciones geométricas* más usuales son las *traslaciones*, *rotaciones*, *reflexiones*⁶ y las *homotecias*. Todas ellas mantienen la forma de las figuras, pero pueden cambiar su tamaño, orientación y/o posición.

Un *movimiento* de una figura es una combinación de *traslaciones*, *rotaciones* y *reflexiones*.

Cuando a una figura se le aplica un *movimiento*, produce otra figura de la misma forma y tamaño pero con distinta orientación. Por su parte, las *homotecias* cambian el tamaño de la figura, pero conservan su forma y orientación. Vamos a ver con más detalle cada una de ellas.

B.3.1. Traslaciones en $\mathbb{R}^2$ y $\mathbb{R}^3$

Desde un punto de vista matemático, trasladar un punto (o vector) P de $\mathbb{R}^n$ en una determinada dirección $\vec{d} = [d_1, \dots, d_n]$, consiste en añadir a las coordenadas de P los correspondientes desplazamientos d_i .

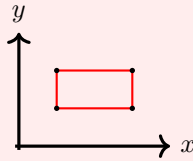


$$\vec{p}' = \vec{p} + \vec{d} \quad ; \quad \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$$

Desde un punto de vista matricial, el movimiento de *traslación* **no** corresponde con una *transformación lineal*.

⁶Estas tres correspondiente al grupo de *simetrías*.

Ejercicio B.6. Sea el rectángulo con vértices $(1, 1)$, $(3, 1)$, $(3, 2)$ y $(1, 2)$. Encontrar los valores de sus nuevos vértices, cuando se traslada la figura según el vector de desplazamiento $\vec{d} = [3, -4]$.



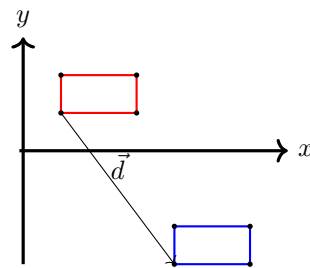
Desde un punto de vista computacional, las figuras se almacenan en memoria como matrices de puntos. En nuestro caso, colocando los vértices del rectángulo como vectores columna se obtiene la siguiente matriz que define la figura:

$$P = \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix}$$

Los vértices de la nueva figura trasladada según nuestro problema se puede obtener realizando la siguiente operación matricial:

$$P' = P + D \quad ; \quad P' = \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 3 & 3 & 3 \\ -4 & -4 & -4 & -4 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 6 & 4 \\ -3 & -3 & -2 & -2 \end{bmatrix}$$

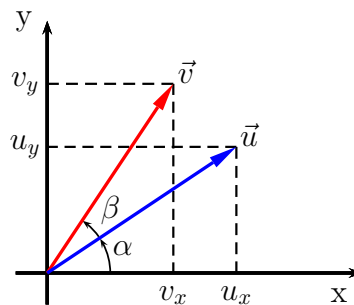
Siendo por tanto los nuevos vértices $(4, -3)$, $(6, -3)$, $(6, -2)$ y $(4, -2)$ respectivamente.



B.3.2. Rotaciones en $\mathbb{R}^2$

Primeramente vamos a estudiar el efecto de una rotación sobre un único vector de $\mathbb{R}^2$. Este proceso se puede generalizar sobre un conjunto de vectores o puntos para calcular el efecto de una rotación sobre una figura plana cualquiera.

Dado un vector $\vec{u} = [u_x, u_y]$ queremos encontrar las coordenadas que tendrá si lo giramos⁷ con centro en el origen $(0, 0)$ un determinado ángulo de β , obteniéndose un nuevo vector $\vec{v} = [v_x, v_y]$.



El vector original $\vec{u}$ tiene unas determinadas coordenadas u_x y u_y y un ángulo α . A través de estos valores queremos obtener el valor de las nuevas coordenadas v_x y v_y . Teniendo en cuenta consideraciones trigonométricas se tiene:

$$\begin{cases} u_x = \|\vec{u}\| \cos \alpha \\ u_y = \|\vec{u}\| \sin \alpha \end{cases} \quad (\text{B.1})$$

⁷En sentido positivo, es decir, en sentido contrario a las agujas del reloj.

En la operación del giro de un vector, no cambia la longitud (norma) del mismo, únicamente su dirección y sentido, por tanto $\|\vec{v}\| = \|\vec{u}\|$. Por otra parte, el ángulo γ del nuevo vector girado $\vec{v}$ será igual al original α más el valor del giro β , por tanto $\gamma = \alpha + \beta$. Se tiene entonces:

$$\begin{cases} v_x = \|\vec{v}\| \cos \gamma &= \|\vec{u}\| \cos (\alpha + \beta) \\ v_y = \|\vec{v}\| \sin \gamma &= \|\vec{u}\| \sin (\alpha + \beta) \end{cases} \quad (\text{B.2})$$

Teniendo en cuenta las relaciones trigonométricas:

$$\begin{cases} \cos (\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin (\alpha + \beta) = \cos \alpha \sin \beta + \sin \alpha \cos \beta \end{cases}$$

la **Ecuación (B.2)** se puede expresar finalmente como:

$$\begin{cases} v_x = \|\vec{u}\| (\cos \alpha \cos \beta - \sin \alpha \sin \beta) &= u_x \cos \beta - u_y \sin \beta \\ v_y = \|\vec{u}\| (\cos \alpha \sin \beta + \sin \alpha \cos \beta) &= u_x \sin \beta + u_y \cos \beta \end{cases} \quad (\text{B.3})$$

donde se ha considerado los valores de u_x y u_y dados en la **Ecuación (B.1)**.

Observando el resultado obtenido en la **Ecuación (B.3)** podemos representar la operación *giro de un vector* sobre un ángulo β como una **transformación lineal** de la siguiente forma:

$$\vec{v} = T(\vec{u}) = A\vec{u} \quad ; \quad \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} \quad \Rightarrow \quad \begin{cases} v_x = u_x \cos \beta - u_y \sin \beta \\ v_y = u_x \sin \beta + u_y \cos \beta \end{cases}$$

La matriz de esta transformación lineal se determina *matriz de rotación* y se suele representar mediante la letra R :

$$R = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}$$

Esta matriz es *ortogonal* y la transformación lineal que define es una *transformación lineal ortogonal*, cuyas propiedades se estudiarán en la **Sección 6.3.5, página 173**.

Ejercicio B.7. Sea el rectángulo con vértices $(1,1)$, $(3,1)$, $3,2)$ y $(1,2)$. Encontrar los valores de sus nuevos vértices, cuando se gira la figura un ángulo de 30° en sentido contrario a las agujas del reloj, respecto al origen.

La matriz correspondiente a este giro, teniendo en cuenta que $30^\circ = \frac{\pi}{6}$ rad es:

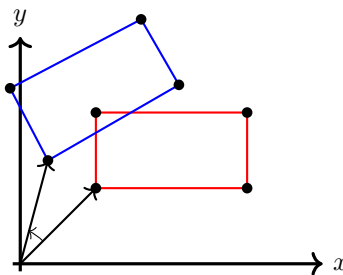
$$R_{\frac{\pi}{6}} = \begin{bmatrix} \cos \frac{\pi}{6} & -\sin \frac{\pi}{6} \\ \sin \frac{\pi}{6} & \cos \frac{\pi}{6} \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix}$$

Para calcular el movimiento de giro más fácilmente, vamos a disponer los vértices del rectángulo como vectores columna de una matriz:

$$P = \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix}$$

Los nuevos vértices transformados son:

$$T = R(P) = R_{\frac{\pi}{6}} P = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}-1}{2} & \frac{3\sqrt{3}-1}{2} & \frac{3\sqrt{3}-2}{2} & \frac{\sqrt{3}-2}{2} \\ \frac{1+\sqrt{3}}{2} & \frac{3+\sqrt{3}}{2} & \frac{3+2\sqrt{3}}{2} & \frac{1+2\sqrt{3}}{2} \end{bmatrix}$$



B.3.3. Rotaciones en $\mathbb{R}^3$

Las rotaciones básicas sobre $\mathbb{R}^3$ son aquellas que se realizan sobre cada uno de los ejes de coordenadas X , Y y Z . Cada una de estas rotaciones tiene asociada una matriz de rotación, similar a la que se obtuvo en la sección anterior.

- Rotación sobre el eje X (plano YZ) :

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & -\operatorname{sen} \beta \\ 0 & \operatorname{sen} \beta & \cos \beta \end{bmatrix}$$

- Rotación sobre el eje Y (plano XZ) :

$$R_y = \begin{bmatrix} \cos \beta & 0 & -\operatorname{sen} \beta \\ 0 & 1 & 0 \\ \operatorname{sen} \beta & 0 & \cos \beta \end{bmatrix}$$

- Rotación sobre el eje Z (plano XY) :

$$R_z = \begin{bmatrix} \cos \beta & -\operatorname{sen} \beta & 0 \\ \operatorname{sen} \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

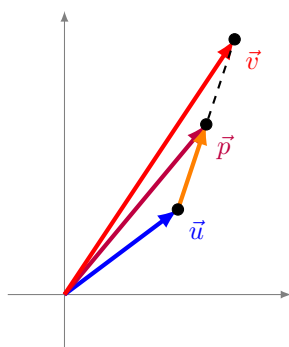
Todas estas *matrices de rotación* pertenecen al importante grupo de las *matrices ortogonales*, que tienen bastantes propiedades y que se estudiarán en detalle en la [Sección 6.3.4, página 171](#).

Cualquier *rotación* general en el espacio $\mathbb{R}^3$ se puede descomponer en una serie de *rotaciones* básicas sobre los ejes, mediante composición de transformaciones lineales.

B.3.4. Simetría Puntual en $\mathbb{R}^2$

Una simetría puntual de centro P es una transformación que hace corresponder a cada punto U otro punto V tal que P es el punto medio del segmento $\overline{UV}$.

Asignando a cada punto P , U y V sus correspondientes vectores $\vec{p}$, $\vec{u}$ y $\vec{v}$ se puede representar gráficamente de la siguiente forma:



Claramente se tiene:

$$\vec{v} = \vec{u} + 2(\vec{p} - \vec{u}) = 2\vec{p} - \vec{u}$$

que no corresponde con una TL.

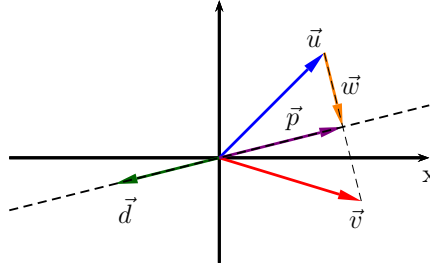
B.3.5. Reflexiones sobre Rectas en $\mathbb{R}^2$

Al igual que en las secciones anteriores, estudiaremos este movimiento inicialmente sobre un único vector. Este procedimiento se puede generalizar para todos los puntos de una figura.

Dado un vector $\vec{u} = [u_x, u_y]$ se quiere determinar las coordenadas de un nuevo vector $\vec{v}$ que geoméricamente sea la *reflexión*⁸ del vector $\vec{u}$ sobre una *recta* que pasa por el origen y tiene por dirección el vector $\vec{d} = [d_x, d_y]$ ⁹.

⁸Proceso de trasladar o copiar todos los puntos de una figura a otra posición equidistante de un objeto geométrico (punto, recta o plano).

⁹Dicha recta tiene por ecuación vectorial: $[x, y] = t[d_x, d_y]$ y general: $y = (d_y/d_x)x$.



Si siguiendo el gráfico mostrado, la solución de este problema se puede calcular fácilmente haciendo uso del vector *proyección ortogonal*¹⁰, $\vec{p}$.

$$\vec{v} = \vec{u} + 2\vec{w} = \vec{u} + 2(\vec{p} - \vec{u}) = 2\vec{p} - \vec{u} \quad (\text{B.4})$$

En este caso el vector $\vec{p}$ corresponde con la proyección ortogonal del vector $\vec{u}$ sobre el vector director de la recta $\vec{d}$:

$$\vec{v} = 2\vec{p} - \vec{u} = 2 \text{proy}_{\vec{d}}(\vec{u}) - \vec{u} = 2 \frac{\vec{d} \cdot \vec{u}}{\vec{d} \cdot \vec{d}} \vec{d} - \vec{u} = 2 \frac{\vec{d} \cdot \vec{u}}{\|\vec{d}\|^2} \vec{d} - \vec{u} \quad (\text{B.5})$$

Realizando operaciones se tiene:

$$\vec{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} = 2 \frac{d_x u_x + d_y u_y}{\|\vec{d}\|^2} \begin{bmatrix} d_x \\ d_y \end{bmatrix} - \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} 2 \frac{d_x^2}{\|\vec{d}\|^2} u_x + 2 \frac{d_x d_y}{\|\vec{d}\|^2} u_y - u_x \\ 2 \frac{d_x d_y}{\|\vec{d}\|^2} u_x + 2 \frac{d_y^2}{\|\vec{d}\|^2} u_y - u_y \end{bmatrix}$$

Esto se puede expresar matricialmente mediante una *transformación lineal*:

$$\vec{v} = T(\vec{u}) = A\vec{u} \quad ; \quad \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} 2 \frac{d_x^2}{\|\vec{d}\|^2} - 1 & 2 \frac{d_x d_y}{\|\vec{d}\|^2} \\ 2 \frac{d_x d_y}{\|\vec{d}\|^2} & 2 \frac{d_y^2}{\|\vec{d}\|^2} - 1 \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} \quad (\text{B.6})$$

La **Ecuación (B.6)** no es necesario memorizarla y desde un punto de vista práctico interesa mucho más entender y conocer el proceso realizado para su obtención, dado en la **Ecuación (B.4)** y **Ecuación (B.5)**. Este proceso es general e independiente de la dimensión de espacio en que nos encontremos.

Ejercicio B.8. Encontrar las matrices asociadas a las TL correspondientes a las reflexiones de un vector $\vec{u} = [u_x, u_y]$ de $\mathbb{R}^2$ sobre las siguientes rectas, encontrando las coordenadas del vector reflejado $\vec{v}$:

- Eje x .
- Eje y .
- Recta $y = x$.

Vamos a hacer uso de la matriz encontrada en la **Ecuación (B.6)**.

- El eje x corresponde con la recta $y = 0$ que tiene por vector director $\vec{d} = [1, 0]$. Por tanto la matriz correspondiente a la TL que define la reflexión es:

$$M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \implies \vec{v} = M\vec{x} = \begin{bmatrix} u_x \\ -u_y \end{bmatrix}$$

¹⁰Ver la **Sección 6.5**, página 178.

- b) El eje y corresponde con la recta $x = 0$ que tiene por vector director $\vec{d} = [0, 1]$. Por tanto la matriz correspondiente a la TL que define la reflexión es:

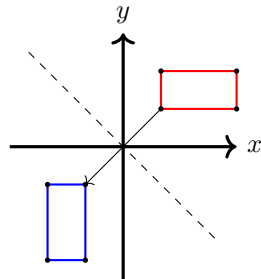
$$M = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \implies \vec{v} = M\vec{x} = \begin{bmatrix} -u_x \\ u_y \end{bmatrix}$$

- c) La recta $y = x$ tiene por vector director $\vec{d} = [1, 1]$. Por tanto la matriz correspondiente a la TL que define la reflexión es:

$$M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \implies \vec{v} = M\vec{x} = \begin{bmatrix} u_y \\ u_x \end{bmatrix}$$

Ejercicio B.9. En $\mathbb{R}^2$, encontrar las nuevas coordenadas del rectángulo con vértices $(1, 1)$, $(3, 1)$, $(3, 2)$ y $(1, 2)$, cuando se le realiza un movimiento de reflexión sobre la recta $y = -x$.

La recta $y = -x$ tiene por vector director $\vec{d} = [1, -1]$. Por tanto la matriz correspondiente a la TL que define la reflexión es:



$$M = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

$$P' = T(P) = MP = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} -1 & -1 & -2 & -2 \\ -1 & -3 & -3 & -1 \end{bmatrix}$$

B.3.6. Reflexiones sobre Rectas en $\mathbb{R}^3$

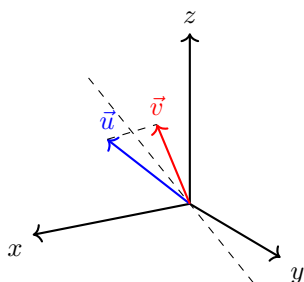
Como se ha comentado, el proceso seguido para obtener la **Ecuación (B.5)** es válido para cualquier dimensión, de forma que se puede obtener, siguiendo los mismos pasos de la sección anterior, la matriz 3×3 asociada a la TL de una reflexión de un vector $\vec{u}$ de $\mathbb{R}^3$ sobre una recta que pasa por el origen con vector director $\vec{d} = [d_x, d_y, d_z]$ y el correspondiente vector reflejado $\vec{v}$.

Desde un punto de vista operacional, resulta más interesante realizar el proceso de cálculo dado en la **Ecuación (B.5)** que calcular la matriz de reflexión y utilizarla para calcular el vector reflejado. Por otra parte, la matriz de reflexión tiene mucho interés desde un punto de vista geométrico.

Ejercicio B.10. Calcular la reflexión del vector $\vec{u} = [1, -1, 1]$ de $\mathbb{R}^3$ sobre la recta:

$$\begin{cases} x = 2t \\ y = -t \\ z = 3t \end{cases}$$

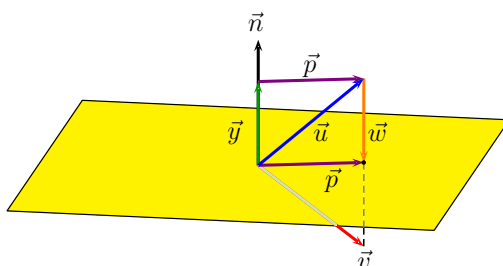
El vector director de esta recta es $\vec{d} = [2, -1, 3]$, de forma que se tiene:



$$\begin{aligned} \vec{v} &= 2 \frac{\vec{d} \cdot \vec{u}}{\|\vec{d}\|^2} \vec{d} - \vec{u} = 2 \frac{[2, -1, 3] \cdot [1, -1, 1]}{[2, -1, 3] \cdot [2, -1, 3]} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \\ &= \frac{6}{7} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 5 \\ 1 \\ 11 \end{bmatrix} \end{aligned}$$

B.3.7. Reflexiones sobre Planos en $\mathbb{R}^3$

Al igual que se hizo en las reflexiones de vectores sobre rectas, también se puede determinar en $\mathbb{R}^3$ la reflexión $\vec{v}$ de un vector $\vec{u}$ sobre un determinado plano que pasa por el origen¹¹.



Para ello sigue siendo válida la **Ecuación (B.4)**:

$$\vec{v} = \vec{u} + 2\vec{w}$$

En este caso, si se denomina $\vec{y}$ al vector *proyección ortogonal* del vector $\vec{u}$ sobre el vector normal al plano $\vec{n}$, se puede determinar geoméricamente, (ver la figura), que $\vec{w} = -\vec{y}$, por tanto:

$$\vec{v} = \vec{u} + 2\vec{w} = \vec{u} - 2\vec{y} = \vec{u} - 2\text{proy}_{\vec{n}}(\vec{u}) = \vec{u} - 2\frac{\vec{n} \cdot \vec{u}}{\vec{n} \cdot \vec{n}} \vec{n} \quad (\text{B.7})$$

Ejercicio B.11. Calcular la reflexión del vector $\vec{u} = [1, -1, 1]$ de $\mathbb{R}^3$ sobre el plano:

$$x + y - 2z = 0$$

En este caso el vector normal del plano es $\vec{n} = [1, 1, -2]$, de forma que aplicando la **Ecuación (B.7)** se tiene:

$$\vec{v} = \vec{u} - 2\frac{\vec{n} \cdot \vec{u}}{\vec{n} \cdot \vec{n}} \vec{n} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - 2\frac{[1, 1, -2] \cdot [1, -1, 1]}{[1, 1, -2] \cdot [1, 1, -2]} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 5 \\ -1 \\ -1 \end{bmatrix}$$

B.3.8. Homotecias en $\mathbb{R}^2$ y $\mathbb{R}^3$

Se denomina *homotecia* a las transformaciones que modifican las dimensiones de una figura pero sin cambiar su forma ni orientación. Las *homotecias* corresponden con *expansiones* y *compresiones* de una figura. Matricialmente, las homotecias con centro en el origen y razón d se pueden representar mediante una **transformación lineal** cuya matriz asociada es:

$$P' = T(P) = HP \quad ; \quad H = dI = d \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} d & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & d \end{bmatrix}$$

Si $d > 1$ corresponderá con una expansión (aumento de tamaño) y si $0 < d < 1$ corresponderá con una compresión (disminución de tamaño).

¹¹Y por tanto con ecuación general $ax + by + cz = 0$.

Ejercicio B.12. En $\mathbb{R}^2$, encontrar las nuevas coordenadas del rectángulo con vértices $(1, 1)$, $(3, 1)$, $(3, 2)$ y $(1, 2)$ sometido a una homotecia de centro en el origen y razón 3.

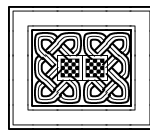
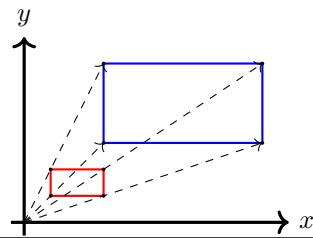
La matriz asociada a la TL de la homotecia es:

$$H = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

Aplicando esta TL a la matriz de los vértices del rectángulo se obtiene:

$$P' = T(P) = HP = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 9 & 3 \\ 3 & 3 & 6 & 6 \end{bmatrix}$$

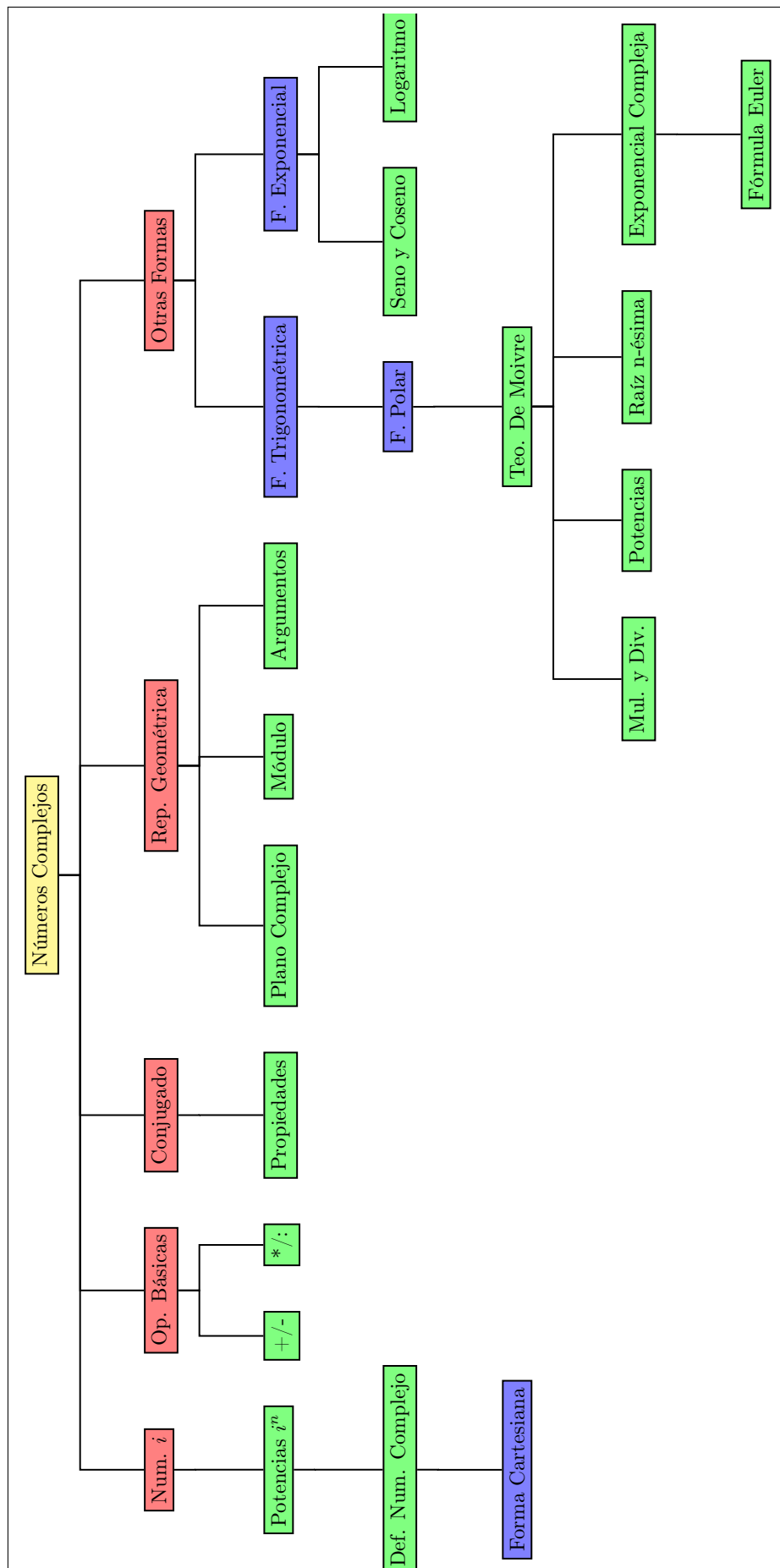
Siendo por tanto los nuevos vértices $(3, 3)$, $(9, 3)$, $(9, 6)$ y $(3, 6)$ respectivamente.



Números Complejos

Contenido

C.1	Introducción	297
C.2	Formas de los Números Complejos	297
C.2.1	Forma Cartesiana, $z = (a, b)$	297
C.2.2	Forma Binómica, $z = a + bi$	298
C.3	Operaciones Básicas en $\mathbb{C}$	298
C.3.1	Suma de Números Complejos	298
C.3.2	Producto de Números Complejos	299
C.3.3	Potencias Enteras de i	299
C.3.4	Elemento Inverso, z^{-1}	300
C.3.5	División en $\mathbb{C}$	300
C.3.6	El Cuerpo de los Números Complejo	301
C.4	Conjugado de un Número Complejo	301
C.4.1	Propiedades del Conjugado	301
C.4.2	Conjugado y División de Números Complejos	301
C.5	Representación Geométrica de $\mathbb{C}$	302
C.5.1	El Plano Complejo	302
C.5.2	Módulo de un Número Complejo	303
C.5.3	Argumento de un Número Complejo	304
C.6	Otras Formas de los Números Complejos	306
C.7	Forma Trigonométrica y Polar	306
C.7.1	Propiedades y Operaciones Básicas de la Forma Polar	307
C.7.2	Potencia n-ésima en Forma Polar	308
C.8	Forma Exponencial	309
C.8.1	Exponencial Compleja. Fórmula de Euler	309
C.8.2	Fórmula de De Moivre	310
C.8.3	Forma Exponencial	311
C.8.4	Operaciones en Forma Exponencial	311
C.9	Logaritmo Complejo	311
C.10	Potencia Compleja	313
C.11	Raíz n-ésima de un Número Complejo	314
C.11.1	Raíz Cuadrada Generalizada	316
C.12	Seno y Coseno Complejo	317



C.1. Introducción

Se puede pensar en principio que los números complejos solo están relacionados con problemas que de alguna forma no tienen solución y que por tanto su utilidad básica radica en separar lo útil (real) de lo *irreal* (imaginario). Sin embargo, los números complejos son una herramienta matemática muy apropiada y *natural* en el estudio y resolución de una gran cantidad de problemas prácticos que aparecen en la ciencia y la tecnología. En este tema se verán los principios básicos de estos números.

C.2. Formas de los Números Complejos

A los números complejos se les puede tratar como *puntos*, como *vectores*, como *pares ordenados* o como *números*, ya que comparten propiedades con varias áreas.

Vamos a ver primeramente dos *formas* distintas de expresar los números complejos junto con sus propiedades.

C.2.1. Forma Cartesiana, $z = (a, b)$

Definición C.1. Forma Cartesiana

Un *número complejo* z se define como un *par ordenado*¹ de *números reales* (a, b) , donde:

- La primera componente a se denomina **parte real** y se representa por $\text{Re}(z) = a$.
- La segunda componente b se denomina **parte imaginaria** y se representa por $\text{Im}(z) = b$.

Definición de suma y producto de números complejos en la forma cartesiana:

$$(a, b) + (c, d) = (a + c, b + d) \quad ; \quad (a, b)(c, d) = (ac - bd, ad + bc) \quad (\text{C.1})$$

Definición C.2. Cuerpo de los Números Complejos

El conjunto de todos los pares ordenados junto con las propiedades de suma y producto tiene estructura de *cuerpo*².

Definición C.3. Unidad Imaginaria

Se define la **unidad imaginaria** y se representa por³ i al número complejo:

$$i = (0, 1)$$

Teniendo en cuenta la definición de la unidad imaginaria i y el producto de números complejos en forma cartesiana (C.1), se tiene:

$$i^2 = ii = (0, 1)(0, 1) = (0 - 1, 0 + 0) = (-1, 0)$$

La forma cartesiana tiene un interés puramente teórico, siendo más interesantes desde el punto de vista práctico otras formas de los números complejos.

¹Se llama *par ordenado* a un conjunto formado por dos elementos y un criterio de ordenación que establece cuál es primer elemento y cuál el segundo.

²En Matemáticas un *cuerpo* o *campo* es una estructura algebraica en la cual las operaciones de adición y producto cumplen las propiedades: asociativa, conmutativa y distributiva del producto respecto de la adición, además de la existencia de inverso aditivo, de inverso multiplicativo y de un elemento neutro para la adición y otro para la multiplicación, los cuales permiten efectuar además las operaciones de sustracción y división.

³En algunos textos y aplicaciones se acostumbra a utilizar la letra j en lugar de la i para identificar la unidad imaginaria.

C.2.2. Forma Binómica, $z = a + bi$

Definición C.4. Forma Binómica

Se denomina **forma binómica** de un número complejo $z = (a, b)$ al siguiente modo de expresar sus partes real y compleja:

$$z = a + bi \quad \text{donde} \quad \begin{cases} a \in \mathbb{R} & \text{Parte Real : } \operatorname{Re}(z) = a \\ b \in \mathbb{R} & \text{Parte Imaginaria : } \operatorname{Im}(z) = b \end{cases}$$

Definición C.5. Imaginario Puro

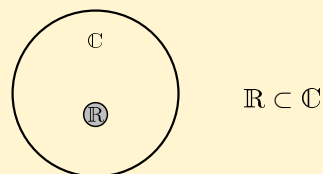
Se denominan **imaginarios puros** a los números complejos $z = bi$ con parte real $a = 0$ y parte imaginaria $b \neq 0$.

Definición C.6. Conjunto de los Números Complejos, $\mathbb{C}$

El conjunto de todos los *números complejos* se representa por $\mathbb{C}$ y se define como:

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$$

El conjunto de los números reales $\mathbb{R}$ es un (pequeño) subconjunto de $\mathbb{C}$ y puede definirse como aquellos números complejos que tienen su *parte imaginaria* nula.



C.3. Operaciones Básicas en $\mathbb{C}$

La forma binómica es la forma preferida para realizar las operaciones básicas entre números complejos.

C.3.1. Suma de Números Complejos

Definición C.7. Suma en $\mathbb{C}$

Dados dos números complejos $z_1 = a_1 + b_1i$ y $z_2 = a_2 + b_2i$ se define la *suma* (+) de z_1 y z_2 como el número complejo dado por:

$$z_1 + z_2 = (a_1 + b_1i) + (a_2 + b_2i) = (a_1 + a_2) + (b_1 + b_2)i$$

Propiedades de la Suma

- **Conmutativa:** $z_1 + z_2 = z_2 + z_1$
- **Asociativa:** $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$
- **Elemento Neutro** (0): $z_1 + 0 = 0 + z_1 = z_1$
- **Elemento Opuesto** ($-z$): $\forall z = a + bi \quad \exists -z = -a - bi : z + (-z) = 0$

⁴Por tanto la resta no se considera una operación distinta, sino como una suma particular.

Definición C.8. Resta o Diferencia en $\mathbb{C}$

Se definen la *resta* $(-)$ de dos *números complejos* $z_1 - z_2$ como la *suma* del primero más el *opuesto* del segundo⁴.

$$z_1 - z_2 = z_1 + (-z_2)$$

C.3.2. Producto de Números Complejos**Definición C.9. Producto en $\mathbb{C}$**

Dados dos números complejos $z_1 = a_1 + b_1i$ y $z_2 = a_2 + b_2i$ se define la *producto*⁵ de z_1 y z_2 como el número complejo dado por:

$$z_1 z_2 = (a_1 + b_1i)(a_2 + b_2i) = a_1 a_2 + i a_1 a_2 + i b_1 a_2 + i^2 b_1 b_2 = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1)$$

Ejercicio C.1. Dados los números complejos $z_1 = 1 + i$ y $z_2 = 2 - 2i$ calcular su suma, resta y producto.

$$\begin{aligned} z_1 + z_2 &= (1 + i) + (2 - 2i) = 3 - i \\ z_1 - z_2 &= (1 + i) - (2 - 2i) = -1 + 3i \\ z_1 z_2 &= (1 + i)(2 - 2i) = (2 + 2) + (-2 + 2)i = 4 \end{aligned}$$

C.3.3. Potencias Enteras de i

Las primeras potencias enteras no negativas de la unidad imaginaria i tienen los siguientes valores⁶:

$$i^0 = 1 \quad i^1 = i \quad i^2 = -1 \quad i^3 = i^2 i = -i \quad i^4 = i^2 i^2 = 1$$

Las sucesivas potencias se pueden calcular fácilmente mediante la siguiente fórmula:

$$i^n = i^{n\%4} = i^r$$

donde $\%$ es el operador *módulo*, que da como resultado el resto de una división entera⁷. El valor de $r = n\%4$ solo puede ser algún natural del conjunto $\{0, 1, 2, 3\}$.

La siguiente tabla muestra los posibles casos agrupados en los cuatro posibles resultados.

$r = 0$	$r = 1$	$r = 2$	$r = 3$
$i^0 = 1$	$i = \sqrt{-1}$	$i^2 = (\sqrt{-1})^2 = -1$	$i^3 = i \cdot i^2 = -i$
$i^4 = i \cdot i^3 = i^0 = 1$	$i^5 = i \cdot i^4 = i$	$i^6 = i \cdot i^5 = i^2 = -1$	$i^7 = i \cdot i^6 = i^3 = -i$
$i^8 = i \cdot i^7 = i^0 = 1$	$i^9 = i \cdot i^8 = i$	$i^{10} = i \cdot i^9 = i^2 = -1$	$i^{11} = i \cdot i^{10} = i^3 = -i$
...	...	...	...

Ejercicio C.2. Calcular i^{2010} .

$$\text{Como } 2010 = 502 \cdot 4 + 2, \text{ por tanto } 2010\%4 = 2 \quad \implies \quad i^{2010} = i^2 = -1$$

⁵Es habitual representar el *producto* de números (complejos y reales) sin utilizar ningún símbolo entre ellos.

⁶Notar que según se vio en el **Apéndice C.2, página 297** $i^2 = -1$, lo cual implica que $i = \pm\sqrt{-1}$. Esto no puede servir para la definición de i , ya que éste es único.

⁷En este caso una división entre 4.

C.3.4. Elemento Inverso, z^{-1} Definición C.10. Elemento Inverso en $\mathbb{C}$

Dado un $z \in \mathbb{C}$, se define (en caso de que exista) el elemento inverso de z y se representa por z^{-1} , al número complejo que verifica: $zz^{-1} = 1$.

$$z = a + bi \quad (a, b \neq 0) \quad \Rightarrow \quad z^{-1} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}i$$

Ejercicio C.3. Demostrar la fórmula de z^{-1} para $z = a + bi$.

Sea $z^{-1} = x + yi$, donde $x, y \in \mathbb{R}$ son los valores a determinar.

$$zz^{-1} = 1 \quad \Rightarrow \quad (a + bi)(x + yi) = (ax - by) + (ay + bx)i = 1 \quad \Rightarrow \quad \begin{cases} ax - by = 1 \\ ay + bx = 0 \end{cases}$$

Este es un sistema de dos ecuaciones con dos incógnitas (x, y) . Vamos a resolverlo por el método de igualación.

$$\left. \begin{array}{l} x = \frac{1 + by}{a} \\ x = \frac{-ay}{b} \end{array} \right\} \Rightarrow \frac{1 + by}{a} = \frac{-ay}{b} \quad ; \quad b + b^2y = -a^2y \quad \Rightarrow \quad y = \frac{-b}{a^2 + b^2}$$

$$\Rightarrow \quad x = \frac{-ay}{b} = \frac{a}{a^2 + b^2}$$

$$z = a + bi \quad (a, b \neq 0) \quad \Rightarrow \quad z^{-1} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}i$$

C.3.5. División en $\mathbb{C}$ Definición C.11. División en $\mathbb{C}$

Dados $z_1, z_2 \in \mathbb{C}$, se define la operación división z_1/z_2 como el producto⁸ $z_1z_2^{-1}$.

$$\frac{z_1}{z_2} = z_1z_2^{-1} = (a_1 + b_1i) \left(\frac{a_2}{a_2^2 + b_2^2} + \frac{-b_2}{a_2^2 + b_2^2}i \right) = \frac{a_1a_2 + b_1b_2}{a_2^2 + b_2^2} + \frac{-a_1b_2 + a_2b_1}{a_2^2 + b_2^2}i$$

Nota C.1. En el [Apéndice C.4.2](#) y [Apéndice C.5.2](#) se verá una forma más práctica de calcular la división entre números complejos sin necesidad de utilizar ninguna fórmula.

Ejercicio C.4. Dados los números complejos $z = 1 + 2i$ y $w = 2 - 2i$, calcular z/w y w/z , comprobando que el producto de ambos es igual a la unidad.

Aplicando la fórmula de la [Definición C.11](#) se tiene:

$$\begin{cases} \frac{z}{w} = \frac{1(2) + 2(-2)}{2^2 + 2^2} + \frac{-1(-2) + 2(2)}{2^2 + 2^2}i = \frac{-1}{4} + \frac{3}{4}i \\ \frac{w}{z} = \frac{2(1) + (-2)2}{1^2 + 2^2} + \frac{-2(2) + 1(-2)}{1^2 + 2^2}i = \frac{-2}{5} - \frac{6}{5}i \end{cases}$$

$$\left(\frac{z}{w} \right) \left(\frac{w}{z} \right) = \left(-\frac{1}{4} + \frac{3}{4}i \right) \left(-\frac{2}{5} - \frac{6}{5}i \right) = \frac{2}{20} + \frac{6}{20}i - \frac{6}{20}i + \frac{18}{20} = \boxed{1}$$

⁸Por tanto la división no se considera una operación nueva, sino un producto particular.

Propiedades del Producto

- **Conmutativa:** $zw = wz$, $\forall z, w \in \mathbb{C}$
- **Asociativa:** $(zw)s = z(ws)$, $\forall z, w, s \in \mathbb{C}$
- **Elemento Neutro:** $1 = 1 + 0i$: $z1 = 1z = z$, $\forall z \in \mathbb{C}$
- **Elemento Inverso:** $\forall z \in \mathbb{C} - \{0\}$, $\exists z^{-1}$: $zz^{-1} = 1$
- **Distributiva respecto a la suma:** $z(z + s) = zw + zs$, $\forall z, w, s \in \mathbb{C}$

C.3.6. El Cuerpo de los Números Complejo

Debido a las propiedades de la suma y producto de números complejos, se dice que $\mathbb{C}$ junto con las operaciones suma y producto, $(\mathbb{C}, +, *)$, es un **cuerpo**⁹

C.4. Conjugado de un Número Complejo

Definición C.12. Complejo Conjugado

Dado $z = a + bi$, se define el *conjugado* de z , y se escribe $\bar{z}$, al número complejo $\bar{z} = a - bi$.

C.4.1. Propiedades del Conjugado

- $\overline{z + w} = \bar{z} + \bar{w}$, $\forall z, w \in \mathbb{C}$
- $\overline{zw} = \bar{z}\bar{w}$, $\forall z, w \in \mathbb{C}$
- $\bar{\bar{z}} = z$, $\forall z \in \mathbb{C}$
- Si $\bar{z} = z \implies z \in \mathbb{R}$
- Si $\bar{z} = -z \implies z$ es *imaginario puro*

Ejercicio C.5. Dados $z = 5 - 3i$ y $w = -4 + 8i$,

- Calcular: $\bar{z}$, $\bar{w}$, $z + w$, $\bar{z} + \bar{w}$ y $\overline{z + w}$.
- Comprobar que $\overline{zw} = \bar{z}\bar{w}$.

$$\text{a) } \bar{z} = \boxed{5 + 3i} \quad ; \quad \bar{w} = \boxed{-4 - 8i} \quad ; \quad z + w = \boxed{1 + 5i} \quad ; \quad \bar{z} + \bar{w} = \boxed{1 + 5i} \quad ; \quad \overline{z + w} = \boxed{1 - 5i}$$

$$\text{b) } \begin{cases} \overline{zw} = \overline{(5 - 3i)(-4 + 8i)} = \overline{-20 + (40 + 12)i + 24} = \overline{4 + 52i} = \boxed{4 - 52i} \\ \bar{z}\bar{w} = (5 + 3i)(-4 - 8i) = -20 + (-40 - 12)i + 24 = \boxed{4 - 52i} \end{cases}$$

C.4.2. Conjugado y División de Números Complejos

La fórmula para realizar la división de números complejos en la forma binómica¹⁰ se puede simplificar utilizando el concepto del conjugado complejo de la siguiente forma:

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}} \quad ; \quad w\bar{w} \in \mathbb{R}$$

⁹El conjunto $\mathbb{C}$ no es un *cuerpo ordenado*, ya que no tiene sentido por ejemplo $0 < i$ ó $0 < -i$.

¹⁰Ver la **Definición C.11**, página 300.

Demostración

$$w\bar{w} = (a+bi)(a-bi) = a^2 - abi + abi + b^2 = a^2 + b^2 \implies w^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i = \frac{a-bi}{a^2 + b^2} = \frac{\bar{w}}{w\bar{w}}$$

$$\frac{z}{w} = zw^{-1} = \frac{z\bar{w}}{w\bar{w}}$$

Ejercicio C.6. Calcular las divisiones del **Ejemplo C.4**, página 300 utilizando la fórmula del **Apéndice C.4.2**, página 301.

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}} = \frac{(1+2i)(2+2i)}{(2-2i)(2+2i)} = \frac{-2+6i}{8} = \boxed{-\frac{1}{4} + \frac{3}{4}i}$$

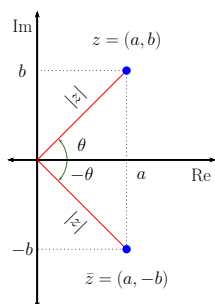
$$\frac{w}{z} = \frac{w\bar{z}}{z\bar{z}} = \frac{(2-2i)(1-2i)}{(1+2i)(1-2i)} = \frac{-2-6i}{5} = \boxed{-\frac{2}{5} - \frac{6}{5}i}$$

Debido a su simplicidad, la *forma binómica* es muy utilizada sobre todo en operaciones sencillas. Sin embargo en cálculos más avanzados esta forma no resulta práctica, utilizándose las formas **polar**, **trigonométrica** y/o **exponencial**.

C.5. Representación Geométrica de $\mathbb{C}$

C.5.1. El Plano Complejo

- Un número complejo $z = a + bi$ se puede representar gráficamente a través de un punto en un sistema de ejes coordenados, donde en el eje X , denominado **eje real** (Re), se coloca la parte real $Re(z) = a$ y en el eje Y , denominado **eje imaginario** (Im), la parte imaginaria $Im(z) = b$.
- El plano formado por este sistema de ejes cartesianos se denomina **plano complejo**.
- Cada uno de los puntos del *plano complejo* se denominan **afijos** y sus coordenadas corresponden con sus partes real y compleja (a, b) .
- Se denomina **módulo** de un número complejo a la distancia del segmento que une el origen del *plano complejo* con su afijo.
- Se denomina **argumento** de un número complejo al ángulo que forma el segmento que define su módulo y el *semieje real* positivo.



- Re : Eje real
- Im : Eje imaginario
- Afijo: punto del plano complejo con coordenadas (a, b)
- Módulo $|z|$: distancia del origen al afijo.
- Argumento θ : ángulo del segmento del módulo.

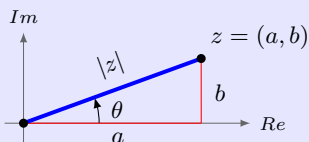
Figura C.1: Representación gráfica en el plano complejo y conceptos asociados.

Nota C.2. El conjunto de los números reales $\mathbb{R}$ se representa a través de la recta real y ésta aparece en el plano complejo representada por el eje horizontal.

C.5.2. Módulo de un Número Complejo

Definición C.13. Módulo

El **módulo** de un número complejo $z = a + bi$ se representa por $|z|$ o por ρ y corresponde con el número real no negativo tal que: $|z| = \rho = \sqrt{a^2 + b^2}$.



$$z = a + bi \quad \begin{cases} \operatorname{Re}(z) = a = |z| \cos \theta = \rho \cos \theta \\ \operatorname{Im}(z) = b = |z| \operatorname{sen} \theta = \rho \operatorname{sen} \theta \end{cases}$$

Propiedades del Módulo, $|z|$

Sean $z, w \in \mathbb{C}$, entonces se cumplen las siguientes propiedades:

- $|z| > 0 \quad \forall z \neq 0 \quad ; \quad |0| = 0$
- $z\bar{z} = a^2 + b^2 = |z|^2 \implies |z| = \sqrt{z\bar{z}}$
- $|-z| = |z|$
- $|\bar{z}| = |z|$
- $|zw| = |z| |w|$
- $z^{-1} = \frac{\bar{z}}{|z|^2} \quad ; \quad |z^{-1}| = \frac{1}{|z|} \quad , \quad (z \neq 0)$
- $|z + w| \leq |z| + |w|$ (Desigualdad Triangular)
- El número complejo $\frac{z}{|z|}$ tiene módulo unidad, $(z \neq 0)$.

Ejercicio C.7. Dado $z = 1 + i\sqrt{3}$, calcular su módulo, conjugado e inverso.

$$|z| = \sqrt{1^2 + (\sqrt{3})^2} = \sqrt{4} = \boxed{2} \quad ; \quad \bar{z} = \boxed{1 - i\sqrt{3}} \quad ; \quad z^{-1} = \frac{\bar{z}}{|z|^2} = \frac{1 - i\sqrt{3}}{4} = \boxed{\frac{1}{4} - \frac{\sqrt{3}}{4}i}$$

Comprobación del inverso:

$$zz^{-1} = (1 + i\sqrt{3}) \left(\frac{1}{4} - \frac{\sqrt{3}}{4}i \right) = \frac{1}{4} - \frac{\sqrt{3}}{4}i + \frac{\sqrt{3}}{4}i + \frac{3}{4} = 1$$

Módulo y División de Números Complejos

Atendiendo a las propiedades del producto¹¹, conjugado¹² y módulo en $\mathbb{C}$ podemos llegar a las siguientes relaciones prácticas:

$$\frac{z}{w} = zw^{-1} = \frac{z\bar{w}}{w\bar{w}} = \frac{z\bar{w}}{|w|^2}$$

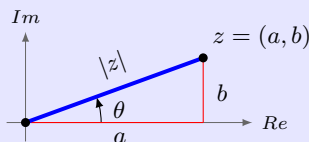
¹¹Ver el Apéndice C.3.2, página 299.

¹²Ver el Apéndice C.4.1, página 301.

C.5.3. Argumento de un Número Complejo

Definición C.14. Argumento

Si las partes real e imaginaria de un número complejo $z = a + bi$ son ambas positivas, el **argumento**¹³ θ de z expresado en radianes verifica la siguiente expresión¹⁴:



$$\tan \theta = \frac{b}{a} \quad ; \quad a, b > 0$$

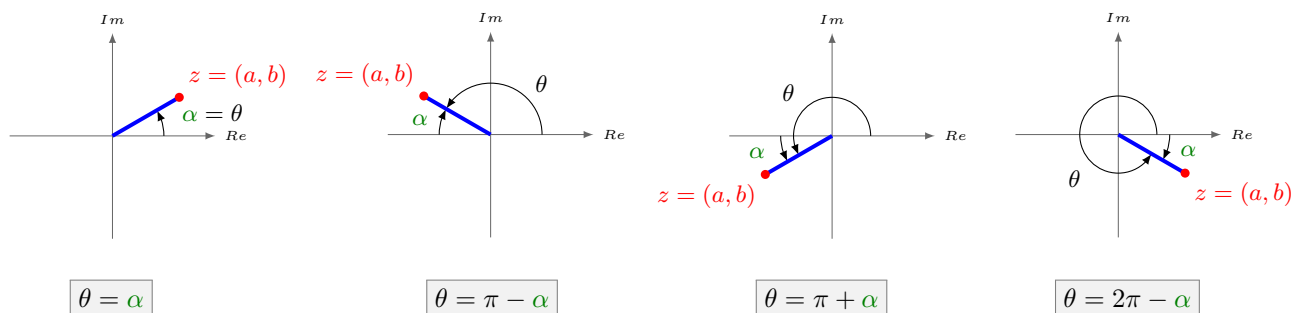
Determinación General del Argumento

La forma general de calcular el argumento θ de cualquier número complejo $z = a + bi$ depende del *cuadrante*¹⁵ del *plano complejo*¹³ en el que esté situado z . Para ello primero determinamos el siguiente ángulo¹⁶:

$$\alpha = \arctan \frac{|b|}{|a|}$$

que corresponde con el ángulo (argumento) que tendría si z estuviera en el primer cuadrante (valores positivos de a y b). En función del cuadrante en que se encuentre z , que dependerá de la posición del punto (a, b) en el plano complejo, así tendremos que modificar el valor de α para determinar el argumento.

En las siguientes figuras se ha representado los cuatro casos posibles en los que se puede encontrar z (primer cuadrante, segundo, tercero o cuarto) y la modificación que hay que hacer para obtener el valor del correspondiente argumento θ .



Notar que en todos estos casos se calcula un valor positivo del argumento $\theta \in [0, 2\pi]$, de los infinitos valores que tiene z .

En los siguientes ejercicios se muestran un par de ejemplos de este cálculo.

¹³Ver el Apéndice C.5.1, página 302

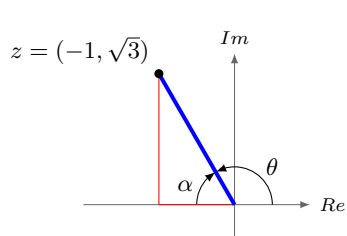
¹⁴Ver el Apéndice A.1.2, página 251.

¹⁵Ver el Apéndice A.1.2, página 253.

¹⁶En los casos prácticos, sin calculadora, es bueno recordar los valores de la tabla dada en el Apéndice A.1.2, página 253, para los valores trigonométricos de ángulos sencillos del primer cuadrante.

Ejercicio C.8. Dado $z = -1 + i\sqrt{3}$, determinar $\arg(z)$.

En este caso el afijo de z se encuentra en el segundo cuadrante del plano complejo. Teniendo en cuenta los valores trigonométricos dados en el [Apéndice A.1.2, página 252](#) se tiene:



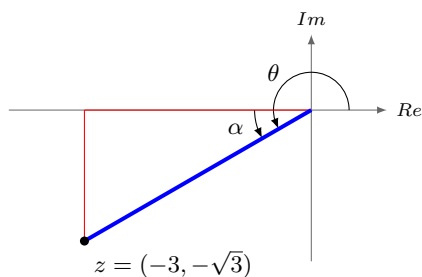
$$\tan \alpha = \frac{|b|}{|a|} = \frac{\sqrt{3}}{1} \implies \alpha = \frac{\pi}{3} \quad (60^\circ)$$

$$\theta = \pi - \alpha = \pi - \frac{\pi}{3} = \frac{2\pi}{3} \quad (120^\circ)$$

$$\arg(z) = \frac{2\pi}{3} + 2k\pi \quad ; \quad k \in \mathbb{Z}$$

Ejercicio C.9. Dado $z = -3 - i\sqrt{3}$, determinar $\arg(z)$.

En este caso el afijo de z se encuentra en el tercer cuadrante del plano complejo. Teniendo en cuenta los valores trigonométricos dados en el [Apéndice A.1.2, página 252](#) se tiene:



$$\tan \alpha = \frac{|b|}{|a|} = \frac{\sqrt{3}}{3} \implies \alpha = \frac{\pi}{6} \quad (30^\circ)$$

$$\theta = \pi + \alpha = \pi + \frac{\pi}{6} = \frac{7\pi}{6} \quad (210^\circ)$$

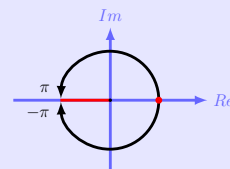
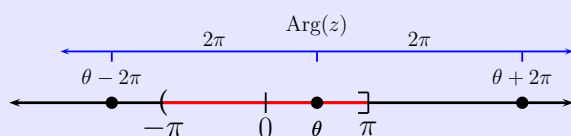
$$\arg(z) = \frac{7\pi}{6} + 2k\pi \quad ; \quad k \in \mathbb{Z}$$

Si θ es el argumento de un número complejo z , entonces $\theta + 2k\pi$ para cualquier entero k también lo será¹⁷, por tanto existen infinitos argumentos asociados a todo número complejo z y el conjunto de todos ellos se denota por $\arg(z)$.

$$\arg(z) = \theta + 2k\pi \quad ; \quad k \in \mathbb{Z}$$

Definición C.15. Argumento Principal

Se denomina **argumento principal**¹⁸ de un número complejo z , al único ángulo $\theta \in \arg(z)$ tal que $\theta \in (-\pi, \pi]$. Se representa por $\text{Arg}(z)$.



El argumento principal es un elemento clave en la identificación de un número complejo, y es utilizado de forma habitual para identificar un número complejo y hacer operaciones con él.

¹⁷Tener en cuenta que $2k\pi$ radianes corresponde con k vueltas completas de 360° , en el sentido positivo (contrario a las agujas del reloj) si $k > 0$ y en el sentido negativo (a favor de las agujas del reloj) si $k < 0$.

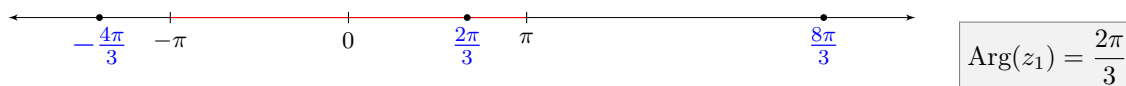
¹⁸A veces, por abuso del lenguaje, se denomina simplemente *argumento* al *argumento principal*.

Ejercicio C.10. Basándose en lo realizado en el **Ejemplo C.8** y **Ejemplo C.9** determinar el argumento principal de los números complejos $z_1 = -1 + i\sqrt{3}$ y $z_2 = -3 - i\sqrt{3}$.

Teniendo en cuenta que el argumento principal debe cumplir: $-\pi < \text{Arg}(z) \leq \pi$, se tiene:

a) $z_1 = -1 + i\sqrt{3}$

Según el resultado del **Ejemplo C.8, página 305** $\implies \arg(z_1) = \frac{2\pi}{3} + 2k\pi$



b) $z_2 = -3 - i\sqrt{3}$

Según el resultado del **Ejemplo C.9, página 305** $\implies \arg(z_2) = \frac{7\pi}{6} + 2k\pi$



C.6. Otras Formas de los Números Complejos

Además de las formas¹⁹ cartesiana y binómica, existen otras formas de representar los números complejos, que son muy utilizadas en la práctica, en función de la aplicación u operaciones que se realicen. Estas formas son:

- Forma trigonométrica y polar.
- Forma exponencial.

Vamos a ver cada una de ellas, explicando sus aplicaciones, ventajas e inconvenientes.

C.7. Forma Trigonométrica y Polar

Definición C.16. Forma Trigonométrica o Polar

Teniendo en cuenta la geometría de un número complejo en el *plano complejo*²⁰ se puede definir un número complejo a través de su *forma trigonométrica* y/o *polar* como sigue:

$$z = a + bi = \rho \cos \theta + i \rho \sin \theta = \rho (\cos \theta + i \sin \theta) = \rho_{\theta}$$

$$z = a + bi \quad \begin{cases} \text{Forma Trigonométrica de } z & : \rho (\cos \theta + i \sin \theta) \\ \text{Forma Polar de } z & : \rho_{\theta} \quad [\text{Nota: } \theta \text{ está en el subíndice}] \end{cases}$$

La razón fundamental de expresar un número complejo z a través de su *forma trigonométrica* y/o *polar* es la sencillez de operaciones: *multiplicación*, *división*, *potenciación* y *raíz n-ésima*, como se verá en el **Apéndice C.7.2** y **Apéndice C.11**. Sin embargo, las operaciones de suma y resta de números complejos es mucho más sencillo de realizar en la forma binómica o cartesiana.

¹⁹Ver el **Apéndice C.2.1, página 297**.

²⁰Ver el **Apéndice C.5.1, página 302**.

Ejercicio C.11. Pasar a la forma trigonométrica y polar los siguientes números complejos: $z_1 = 1 - i$; $z_2 = -1 + i\sqrt{3}$.

a)

$$z_1 = 1 - i \left\{ \begin{array}{l} |z_1| = \sqrt{(1)^2 + (-1)^2} = \sqrt{2} \\ \theta = \arctan \frac{|-1|}{|1|} = \arctan(1) = \frac{\pi}{4} \end{array} \right\}$$

Como el afijo de z_1 cae en el cuarto cuadrante, el valor del argumento será: $\theta_1 = -\theta = -\frac{\pi}{4}$

$$z_1 = 1 - i = \sqrt{2} \left(\cos \frac{-\pi}{4} + i \operatorname{sen} \frac{-\pi}{4} \right) = \sqrt{2} \left(\cos \frac{\pi}{4} - i \operatorname{sen} \frac{\pi}{4} \right) = \sqrt{2}_{-\frac{\pi}{4}}$$

b)

$$z_2 = -1 + i\sqrt{3} \left\{ \begin{array}{l} |z_2| = \sqrt{(-1)^2 + (\sqrt{3})^2} = \sqrt{4} = 2 \\ \theta = \arctan \frac{|\sqrt{3}|}{|-1|} = \arctan \sqrt{3} = \frac{\pi}{3} \end{array} \right\}$$

En este caso el afijo de z_2 cae en el segundo cuadrante y por tanto su argumento será: $\theta_2 = \pi - \theta = \frac{2\pi}{3}$

$$z_2 = -1 + i\sqrt{3} = 2 \left(\cos \frac{2\pi}{3} + i \operatorname{sen} \frac{2\pi}{3} \right) = 2_{\frac{2\pi}{3}}$$

Ejercicio C.12. Dar la forma polar de un número real y de un imaginario puro, tanto positivo como negativo.

- $a + 0i = a_0$; $-a + 0i = a_\pi$
- $0 + bi = b_{\frac{\pi}{2}}$; $0 - bi = b_{-\frac{\pi}{2}}$

C.7.1. Propiedades y Operaciones Básicas de la Forma Polar

- $-z = \rho_{\theta+\pi}$
- $iz = \rho_{\theta+\frac{\pi}{2}}$
- $\bar{z} = \rho_{-\theta}$

Multiplicación y División en Forma Polar

Sean $z = a + bi = \rho_\theta$; $z' = a' + b'i = \rho'_{\theta'}$

- $zz' = (\rho\rho')_{(\theta+\theta')}$
- $\frac{z}{z'} = \left(\frac{\rho}{\rho'} \right)_{(\theta-\theta')}$

Demostración

Teniendo en cuenta que:

$$\begin{aligned}\cos(a+b) &= \cos a \cos b - \sin a \sin b & ; & \quad \cos(a-b) = \cos a \cos b + \sin a \sin b \\ \sin(a+b) &= \sin a \cos b + \cos a \sin b & ; & \quad \sin(a-b) = \sin a \cos b - \cos a \sin b\end{aligned}$$

Tenemos que:

$$\begin{aligned}z_1 z_2 &= (\rho_\theta)(\rho_{\theta'}) = \rho(\cos \theta + i \sin \theta) \rho'(\cos \theta' + i \sin \theta') = \\ &= \rho \rho' [\cos \theta \cos \theta' - \sin \theta \sin \theta' + (\cos \theta \sin \theta' + \sin \theta \cos \theta')i] = \\ &= \rho \rho' [\cos(\theta + \theta') + i \sin(\theta + \theta')] = (\rho \rho')_{(\theta + \theta')}.\end{aligned}$$

Siguiendo el mismo procedimiento se puede demostrar la fórmula para el cociente de números complejos en forma polar, utilizando en este caso las fórmulas de coseno y seno de diferencia de ángulos.

Ejercicio C.13. Dados los números complejos $z_1 = -1 - i$ y $z_2 = 1 + i\sqrt{3}$, determinar su producto $z_1 z_2$ y su cociente z_1/z_2 en forma polar.

Primero pasamos los complejos a su forma polar:

$$\begin{aligned}z_1 = -1 - i & \left\{ \begin{array}{l} |z_1| = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} \\ \theta_1 = \pi + \arctan \frac{1}{1} = \pi + \frac{\pi}{4} = \frac{5\pi}{4} \end{array} \right\} z_1 = \sqrt{2} \frac{5\pi}{4} \\ z_2 = 1 + i\sqrt{3} & \left\{ \begin{array}{l} |z_2| = \sqrt{(1)^2 + (\sqrt{3})^2} = 2 \\ \theta_2 = \arctan \frac{\sqrt{3}}{1} = \frac{\pi}{3} \end{array} \right\} z_2 = 2 \frac{\pi}{3}\end{aligned}$$

$$\bullet \quad z_1 z_2 = (\sqrt{2})(2) \left(\frac{5\pi}{4} + \frac{\pi}{3} \right) = 2\sqrt{2} \frac{19\pi}{12}$$

$$\bullet \quad \frac{z_1}{z_2} = \frac{\sqrt{2}}{2} \left(\frac{5\pi}{4} - \frac{\pi}{3} \right) = \frac{\sqrt{2}}{2} \frac{11\pi}{12}$$

C.7.2. Potencia n-ésima en Forma Polar

Definición C.17. Potencia n-ésima de un Complejo

Dado $z = a + bi$ se define la potencia n-ésima de z como:

$$z^n = (a + bi)^n = (a + bi)(a + bi) \dots (a + bi) \quad ; \quad n \in \mathbb{N}$$

Teniendo en cuenta la *fórmula de De Moivre*²¹ es fácil comprobar que:

$$z^n = [\rho_\theta]^n = \rho_{n\theta}^n$$

Ejercicio C.14. Demostrar la fórmula de potenciación en forma polar.

$$z^n = [\rho_\theta]^n = [\rho(\cos \theta + i \sin \theta)]^n = \rho^n (\cos \theta + i \sin \theta)^n = \rho^n [\cos(n\theta) + i \sin(n\theta)] = \rho_{n\theta}^n$$

²¹Ver el Teorema C.2, página 310.

Ejercicio C.15. Calcular z^6 siendo $z = 1 + i$, dando el resultado en forma binómica.

A través de la forma tradicional:

$$z^6 = (1 + i)^6 = (1 + i)(1 + i)(1 + i)(1 + i)(1 + i)(1 + i) = \dots$$

En forma polar:

$$z = 1 + i = \left\{ \begin{array}{l} \rho = \sqrt{1^2 + 1^2} = \sqrt{2} \\ \theta = \arctan \frac{1}{1} = \frac{\pi}{4} \end{array} \right\} = \sqrt{2} \frac{\pi}{4} \quad ; \quad z^6 = (\sqrt{2})^6 \frac{6\pi}{4} = (2^3) \frac{3\pi}{2}$$

$$z^6 = 8 \left(\cos \frac{3\pi}{2} + i \operatorname{sen} \frac{3\pi}{2} \right) = 8(0 - i) = \boxed{-8i}$$

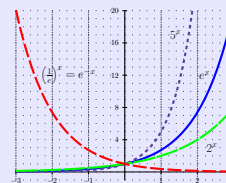
C.8. Forma Exponencial

C.8.1. Exponencial Compleja. Fórmula de Euler

Definición C.18. Función Exponencial en $\mathbb{R}$

La **función exponencial** es la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por:

$$f(x) = e^x \quad \left\{ \begin{array}{l} x \in \mathbb{R} \\ e = \text{número } e = 2,718281828459\dots \end{array} \right.$$



Por claridad en la notación, a veces la función exponencial suele expresarse mediante el término *exp* de la siguiente forma:

$$e^x = \exp(x)$$

Teorema C.1. Exponencial Compleja: Fórmula de Euler

Para todo $x \in \mathbb{R}$ se cumple²²:

$$e^{xi} = \cos x + i \operatorname{sen} x \quad ; \quad x \in \mathbb{R}$$

denominada **exponencial compleja** o **fórmula de Euler**²³.

²²La expresión $\cos x + i \operatorname{sen} x$ a veces se denota por $\operatorname{cis} x$.

²³Con este nombre es también conocida la ecuación:

$$e^{i\pi} + 1 = 0$$

que corresponde al caso particular $x = \pi$ y que es considerada *la ecuación más bella del mundo*.

Demostración C.1.

La función exponencial puede definirse a través de su *desarrollo de McLaurin*²⁴ de la siguiente forma:

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad x \in \mathbb{R}$$

Por tanto:

$$\begin{aligned} e^{xi} &= \sum_{n=0}^{\infty} \frac{i^n}{n!} x^n = 1 + \frac{i}{1!}x - \frac{1}{2!}x^2 - \frac{i}{3!}x^3 + \frac{1}{4!}x^4 - \frac{i}{5!}x^5 - \frac{1}{6!}x^6 - \frac{i}{7!}x^7 + \cdots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots\right) + i \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\right) \end{aligned}$$

Teniendo en cuenta los desarrollos de McLaurin de las funciones seno y coseno:

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \quad x \in \mathbb{R}$$

$$\operatorname{sen} x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \quad x \in \mathbb{R}$$

se obtiene la fórmula de Euler:

$$e^{xi} = \cos x + i \operatorname{sen} x$$

Definición C.19. Exponencial de un Complejo

Si $z \in \mathbb{C}$ se tiene que la **exponencial de un número complejo** es:

$$e^z = e^{a+bi} = e^a e^{bi} = e^a (\cos b + i \operatorname{sen} b)$$

En la forma polar/trigonométrica:

$$e^z = \exp(\rho_\theta) = \exp[\rho(\cos \theta + i \operatorname{sen} \theta)] = \exp(\rho \cos \theta) \exp(i \rho \operatorname{sen} \theta) = e^{\rho \cos \theta} [\cos(\rho \operatorname{sen} \theta) + i \operatorname{sen}(\rho \operatorname{sen} \theta)]$$

C.8.2. Fórmula de De Moivre**Teorema C.2. Fórmula de De Moivre**

$$(\cos \theta + i \operatorname{sen} \theta)^n = \cos(n\theta) + i \operatorname{sen}(n\theta)$$

Demostración C.2.

Aplicando la Fórmula de Euler y teniendo en cuenta que $(a^x)^y = a^{xy}$ tenemos :

$$(\cos \theta + i \operatorname{sen} \theta)^n \stackrel{\text{F. Euler}}{=} (e^{i\theta})^n = e^{in\theta} \stackrel{\text{F. Euler}}{=} \cos(n\theta) + i \operatorname{sen}(n\theta)$$

²⁴Correspondiente al *desarrollo de Taylor* en torno al punto $x = 0$. Ver el concepto en algún texto de Cálculo Infinitesimal.

C.8.3. Forma Exponencial

Definición C.20. Forma Exponencial

Un número $z \in \mathbb{C}$ se puede expresar en la *forma exponencial* como sigue:

$$z = a + bi = \rho_{\theta} = \rho (\cos \theta + i \operatorname{sen} \theta) = \boxed{\rho e^{i\theta}}$$

Las utilidades de la forma exponencial son las mismas que las de la forma polar.

C.8.4. Operaciones en Forma Exponencial

Sean $z = a + bi = \rho_{\theta} = \rho e^{i\theta}$; $z' = a' + b'i = \rho'_{\theta'} = \rho' e^{i\theta'}$

- $zz' = (\rho\rho')e^{(\theta+\theta')i}$
- $\frac{z}{z'} = \left(\frac{\rho}{\rho'}\right)e^{(\theta-\theta')i}$
- $z^n = \rho^n e^{n\theta i}$

Nota C.3. La demostración es inmediata en base a el [Apéndice C.7.1, página 307](#) y [Apéndice C.7.2, página 308](#).

Ejercicio C.16. Dados los números $z_1 = 1 + i$ y $z_2 = 1 - i$ calcular $z_1 z_2$, z_1/z_2 y z_1^{100} utilizando la forma exponencial.

$$z = 1 + i = \sqrt{2}e^{i\frac{\pi}{4}} \quad ; \quad z = 1 - i = \sqrt{2}e^{-i\frac{\pi}{4}}$$

$$z_1 z_2 = (\sqrt{2})(\sqrt{2})e^{(\frac{\pi}{4}-\frac{\pi}{4})i} = 2e^0 = \boxed{2}$$

$$\frac{z_1}{z_2} = \frac{\sqrt{2}}{\sqrt{2}}e^{(\frac{\pi}{4}+\frac{\pi}{4})i} = e^{i\frac{\pi}{2}} = \boxed{i}$$

$$z_1^{100} = (\sqrt{2})^{100}e^{i\frac{100\pi}{4}} = 2^{50}e^{i25\pi} = 2^{50}e^{i[12(2\pi)+\pi]} = 2^{50}e^{i\pi} = \boxed{-2^{50}}$$

C.9. Logaritmo Complejo

Teorema C.3. Logaritmo Complejo

Dado un número complejo $z = a + bi = \rho_{\theta} = \rho e^{i\theta}$, con $z \neq 0$, se define el *logaritmo (natural o neperiano) complejo* de la siguiente forma:

$$\log z = \log(\rho_{\theta}) = \log \rho + (\theta + 2k\pi)i \quad ; \quad k \in \mathbb{Z}$$

Demostración C.3.

Sea $z = \rho_{\theta}$ y queremos calcular $\log z = x + yi$. Despejando de esta última expresión se tiene:

$$z = e^{x+yi} = e^x(\cos y + i \operatorname{sen} y) = e_y^x \implies \begin{cases} \rho = e^x & \implies x = \log \rho \\ \theta = y + 2k\pi & \implies y = \theta - 2k\pi \end{cases}$$

Notar que $k \in \mathbb{Z}$ y por tanto es indiferente que aparezca como positivo o negativo.

Existen infinitos valores del logaritmo de un $z \neq 0$, que se obtienen a través de la fórmula del **Teorema C.3**, dando valores a k .

El cambio de *base*²⁵ en los logaritmos complejos se hace de forma análoga al de los logaritmos reales:

$$\log_a z = \log_b z \cdot \log_a b \quad (\text{Fórmula de Cambio de Base})$$

Definición C.21. Logaritmo Principal

El **logaritmo principal** de un número complejo z se define como:

$$\log z = \log |z| + i \arg z \quad ; \quad z \neq 0$$

donde $\arg z$ es el *argumento principal*²⁶ de z .

Ejercicio C.17. Calcular el logaritmo principal del número $z = \sqrt{3} + i$.

$$z = \sqrt{3} + i = 2 \frac{\pi}{6}$$

$$\log(z) = \log 2 + \left(\frac{\pi}{6} + 2k\pi\right)i \quad ; \quad k \in \mathbb{Z}$$

Tenemos que encontrar el valor de k que hace $(\theta + 2k\pi)$, es positiva y lo más pequeña posible.

$$\vdots$$

$$k = -1 \quad ; \quad (\theta - 2\pi) = -\frac{11\pi}{6}$$

$$k = 0 \quad ; \quad (\theta) = \frac{\pi}{6}$$

$$k = 1 \quad ; \quad (\theta + 2\pi) = \frac{13\pi}{6}$$

$$k = 2 \quad ; \quad (\theta + 4\pi) = \frac{25\pi}{6}$$

$$\vdots$$

Por tanto, en este caso el logaritmo principal es:

$$\log(\sqrt{3} + i) = \log 2 + \frac{\pi}{6}i$$

Una forma mucho más efectiva de calcular el valor principal de un número complejo es expresar el número complejo en forma exponencial y aplicar las propiedades de las potencias y logaritmos, de la siguiente forma:

$$\log(\sqrt{3} + i) = \log 2e^{i\pi/6} = \log 2 + \log e^{i\pi/6} = \log 2 + \frac{\pi}{6}i \log e = \log 2 + \frac{\pi}{6}i$$

²⁵El logaritmo en *base* b de un número x se define como:

$$\log_b x = y \quad \implies \quad b^y = x$$

²⁶Ver la **Definición C.15**, página 305.

Ejercicio C.18. Comprobar la veracidad del resultado del ejemplo anterior.

Como por definición de logaritmo:

$$\log_b x = y \implies b^y = x$$

En nuestro caso la base es el número e y por tanto:

$$\log(\sqrt{3} + i) = \log 2 + \frac{\pi}{6}i \implies e^{\log 2 + \frac{\pi}{6}i} = \sqrt{3} + i$$

Vamos a comprobarlo:

$$e^{\log 2 + \frac{\pi}{6}i} = e^{\log 2} e^{\frac{\pi}{6}i} = 2e^{\frac{\pi}{6}i} = 2e^{\frac{\pi}{6}} = 2\left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6}\right) = 2\left(\frac{\sqrt{3}}{2} + \frac{1}{2}i\right) = \sqrt{3} + i$$

Por tanto, el resultado es correcto.

C.10. Potencia Compleja

Definición C.22. Potencia Compleja

Dados dos números complejos $z \neq 0$ y w , se define la *potencia compleja* z^w de la siguiente forma:

$$z^w = e^{w \log z} \quad ; \quad z \neq 0$$

Ya que existen infinitos valores del logaritmo de un número complejo, también existen infinitos valores asociados a una potencia compleja.

Definición C.23. Valor Principal de una Potencia Compleja

Se denomina *valor principal*²⁷ de una potencia compleja z^w al único valor resultante de tomar el *logaritmo principal*²⁸ de z en la definición de potencia compleja.

Ejercicio C.19. Calcular el valor principal de la siguiente potencia compleja $(1 + i\sqrt{3})^{1-i}$.

Aplicando la definición de potencia de un número complejo, se tiene:

$$\begin{aligned} (1 + i\sqrt{3})^{1-i} &= e^{(1-i) \log(1+i\sqrt{3})} = e^{(1-i) \log(2e^{i\pi/3})} = e^{(1-i)(\log 2 + \log e^{i\pi/3})} = e^{(1-i)(\log 2 + \frac{i\pi}{3} \log e)} = \\ &= e^{(1-i)(\log 2 + \frac{i\pi}{3})} = e^{(\log 2 + \frac{\pi}{3}) + i(\frac{\pi}{3} - \log 2)} = e^{\log 2} e^{\pi/3} e^{i(\pi/3 - \log 2)} = 2e^{\pi/3} e^{i(\pi/3 - \log 2)} \end{aligned}$$

Notar que esta forma de calcular el logaritmo, siguiendo las reglas habituales de los números reales, produce un único resultado correspondiente al valor principal del logaritmo.

²⁷También se le denomina *rama principal*.

²⁸Ver la **Definición C.21**, página 312.

C.11. Raíz n-ésima de un Número Complejo

Definición C.24. Raíz n-ésima de un Número Complejo

Un número complejo no nulo $z = \rho_\theta = \rho e^{i\theta}$ tiene n raíces n-ésimas, que pueden ser expresadas a través la siguiente fórmula²⁹:

$$z^{1/n} = \sqrt[n]{z} = r e^{i\phi} \implies \begin{cases} r = \sqrt[n]{\rho} > 0 & \text{(siempre existe y es único)} \\ \phi = \frac{\theta + 2k\pi}{n} & ; k = 0, 1, 2, \dots, n-1 \end{cases} ; n \in \mathbb{N} \quad (\text{C.2})$$

En forma exponencial³⁰:

$$\sqrt[n]{z} = \sqrt[n]{\rho e^{i\theta}} = \sqrt[n]{\rho} e^{\frac{\theta + 2k\pi}{n} i} ; k = 0, 1, \dots, n-1$$

Demostración C.3.

$$\text{Si } \sqrt[n]{\rho_\theta} = r_\varphi \implies (r_\varphi)^n = \rho_\theta \implies \begin{cases} r^n = \rho \implies r = \sqrt[n]{\rho} \\ n\varphi = \theta + 2k\pi \implies \varphi = \frac{\theta + 2k\pi}{n} \end{cases}$$

El **valor principal** de la raíz n-ésima de $z = \rho_\theta$ es $\sqrt[n]{z} = (\sqrt[n]{\rho})_{\theta/n}$ donde θ es el argumento principal de z y corresponde con tomar el valor $k = 0$ en la Ecuación (C.2).

Ejercicio C.20. Calcular $\sqrt{1+i}$

Como es una raíz cuadrada tendrá dos soluciones z_1 y z_2 .

$$1 + i = \sqrt{2}_{\frac{\pi}{4}} \implies \sqrt{1+i} \left\{ \begin{array}{l} \rho = \sqrt{\sqrt{2}} = \sqrt[4]{2} \\ \theta = \frac{\pi/4 + 2k\pi}{2} \left\{ \begin{array}{ll} k=0 & \theta_1 = \pi/8 \\ k=1 & \theta_2 = 9\pi/8 \end{array} \right. \end{array} \right\} \quad \boxed{\sqrt{1+i} = \{ \sqrt[4]{2}_{\frac{\pi}{8}}, \sqrt[4]{2}_{\frac{9\pi}{8}} \}}$$

En este caso, las dos raíces solución son números complejos, cuyo valor puede dejarse en forma polar. Notar que ambas soluciones son complejos opuestos³¹, ya que $\theta_2 = \theta_1 + \pi$ y por tanto³² $z_1 = -z_2$.

Ejercicio C.21. Calcular $\sqrt[3]{-1}$

En este caso al ser una raíz cúbica tendrá tres soluciones z_1, z_2 y z_3 .

$$-1 + 0i = 1_\pi \implies \sqrt[3]{-1} \left\{ \begin{array}{l} \rho = \sqrt[3]{1} = 1 \\ \theta = \frac{\pi + 2k\pi}{3} \left\{ \begin{array}{ll} k=0 & \theta_1 = \frac{\pi}{3} \\ k=1 & \theta_2 = \pi \\ k=2 & \theta_3 = \frac{5\pi}{3} \end{array} \right. \end{array} \right\} \quad \boxed{\sqrt[3]{-1} = \{ 1_{\frac{\pi}{3}}, 1_\pi, 1_{\frac{5\pi}{3}} \}}$$

Notar que 1_π es el número real -1 , por tanto $\sqrt[3]{-1}$ tiene dos soluciones complejas ($1_{\frac{\pi}{3}}, 1_{\frac{5\pi}{3}}$) y una real (-1) .

²⁹Notar que θ representa el argumento principal de z , $\theta = \text{Arg}(z)$.

³⁰Ver la Definición C.20, página 311.

³¹Ver el Apéndice C.3.1, página 298.

³²Ver el Apéndice C.7.1, página 307.

Ejercicio C.22. Calcular $\sqrt{(1+i)^2}$

Primeramente calculamos el valor interior de la raíz:

$$1+i = \sqrt{2}_{\pi/4} \implies (1+i)^2 = 2_{\pi/2}$$

$$\sqrt{(1+i)^2} \left\{ \begin{array}{l} \rho = \sqrt{2} \\ \theta = \frac{\pi/2+2k\pi}{2} \left\{ \begin{array}{ll} k=0 & \theta_1 = \pi/4 \\ k=1 & \theta_2 = 5\pi/4 \end{array} \right. \end{array} \right\} \quad \boxed{\sqrt{(1+i)^2} = \{\sqrt{2}_{\pi/4}, \sqrt{2}_{5\pi/4}\}}$$

Notar que este resultado es igual al obtenido mediante la operación tradicional con números reales:

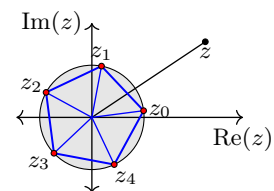
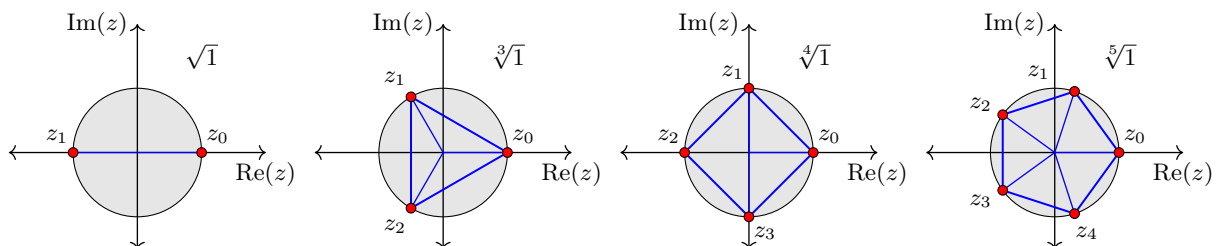
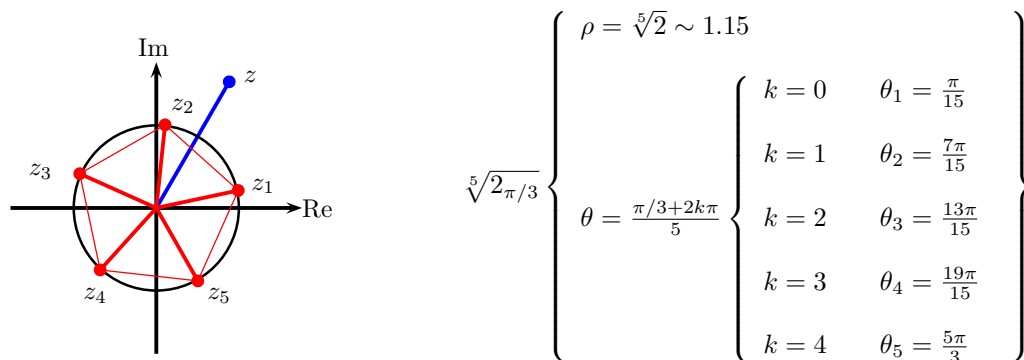
$$\sqrt{(1+i)^2} = \pm(1+i) = \left\{ \begin{array}{ll} 1+i & = \sqrt{2}_{\pi/4} \\ -1-i & = \sqrt{2}_{5\pi/4} \end{array} \right.$$

Geometría de las Raíces n-ésimas

Si $n > 2$, las n raíces n-ésimas de un complejo z forman al representarlas en el plano complejo un *polígono regular*³³ de n lados inscrito en una circunferencia de radio $\sqrt[n]{|z|}$.

En la siguiente figura se muestra gráficamente las cinco raíces z_k con $k = \{0, 1, 2, 3, 4\}$ de $\sqrt[5]{z}$, siendo $z = 4 + 3i$. Vemos que las raíces forman en el plano complejo un pentágono inscrito en una circunferencia de radio $\sqrt[5]{|z|}$ y centro el origen.

Notar que el argumento de z_0 es $\theta/5$, siendo θ el argumento de z , y el argumento del resto z_k es $(\theta + 2k\pi)/5$.

**Ejercicio C.23.** Representar gráficamente en el plano complejo las raíces segunda, tercera, cuarta y quinta de la unidad.**Ejercicio C.24.** Calcular $\sqrt[5]{z}$ con $z = 2_{\pi/3}$ y representar el resultado en el plano complejo.

³³Si $n = 2$ se obtiene una recta.

C.11.1. Raíz Cuadrada Generalizada

La fórmula general **Ecuación (C.2)** no siempre resulta una forma práctica para determinar las raíces de un número complejo, principalmente cuando los argumentos de las raíces no son ángulos simples conocidos. Vamos a ver una forma más general de determinar las raíces cuadradas de un número complejo.

Sea:

$$\sqrt{a+bi} = x+iy$$

donde $a+bi$ es un número complejo conocido y $x+iy$ los valores desconocidos de su correspondiente raíz cuadrada. Entonces despejando:

$$a+bi = (x+iy)^2 = x^2 - y^2 + 2xyi \implies \begin{cases} x^2 - y^2 = a \\ 2xy = b \end{cases}$$

dando lugar a un sistema no lineal de dos ecuaciones con dos incógnitas, cuya resolución:

$$y = \frac{b}{2x} \implies x^2 - y^2 = x^2 - \frac{b^2}{4x^2} = a \implies x^4 - b^2 = 4ax^2 \implies x^4 - 4ax^2 - b^2 = 0$$

da lugar a una ecuación *bicuatrálica*. Para resolverla podemos hacer un cambio de variable:

$$t = x^2 \implies t^2 - 4at - b^2 = 0 \implies t = \frac{4a \pm \sqrt{16a^2 + 4b^2}}{2} = 2a \pm \sqrt{4a^2 + b^2}$$

De estos dos valores de t , solo el positivo tiene sentido al deshacer el cambio de variable, obteniéndose los valores $x, y \in \mathbb{R}$:

$$x_{\pm} = \pm\sqrt{t} \implies y = \frac{b}{2x_{\pm}} \implies \text{Raíces: } \boxed{x_{\pm} + iy_{\pm}}$$

Ejercicio C.25. Encontrar las soluciones de la siguiente ecuación cuadrática compleja:

$$z^2 + 4iz - 7 + 4i = 0$$

Aplicando la conocida fórmula para la resolución de ecuaciones cuadráticas reales a este caso se tiene³⁴:

$$z = \frac{-4i + \sqrt{-16 - 4(-7 + 4i)}}{2} = \frac{-4i + \sqrt{-16 + 28 - 16i}}{2} = \frac{-4i + \sqrt{12 - 16i}}{2} = -2i + \sqrt{3 - 4i}$$

Notar que en este caso al pasar el número complejo $3 - 4i$ a su forma polar, el argumento no es uno de los valores conocidos sencillos. Por tanto vamos a calcular la raíz cuadrada a través de la forma generalizada.

$$\sqrt{3 - 4i} = x + iy \implies 3 - 4i = (x + iy)^2 = x^2 - y^2 + 2xyi \implies \begin{cases} x^2 - y^2 = 3 \\ 2xy = -4 \end{cases} \implies y = \frac{-2}{x}$$

$$x^2 - \frac{4}{x^2} = 3 \quad ; \quad x^4 - 4 = 3x^2 \quad ; \quad x^4 - 3x^2 - 4 = 0 \quad \xrightarrow{\text{C.V.}} \quad t = x^2$$

$$t^2 - 3t - 4 = 0 \implies t = \frac{3 \pm \sqrt{9 + 16}}{2} = \frac{3 \pm 5}{2} = \{-1, 4\}$$

Deshaciendo el cambio de variable:

$$x = \sqrt{t} \implies \begin{cases} x = \sqrt{-1} & \rightarrow \text{Sin sentido ya que } x \in \mathbb{R} \\ x = \sqrt{4} = \pm 2 & \rightarrow y = \frac{-2}{x} = \{-1, 1\} \end{cases}$$

Por tanto, los valores de la raíz cuadrada buscados son:

$$\sqrt{3 - 4i} = \{2 - i, -2 + i\}$$

Por tanto las soluciones buscadas de la ecuación cuadrática compleja original son:

$$z = -2i + \sqrt{3 - 4i} = \begin{cases} z_1 = -2i + 2 - i = 2 - 3i \\ z_2 = -2i - 2 + i = -2 - i \end{cases}$$

C.12. Seno y Coseno Complejo

Teniendo en cuenta la *fórmula de Euler*³⁵, las funciones trigonométricas *seno* y *coseno* de un número real θ se pueden definir como:

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{\cos \theta + i \operatorname{sen} \theta + \cos \theta - i \operatorname{sen} \theta}{2} = \frac{2 \cos \theta}{2} = \cos \theta$$

$$\operatorname{sen} \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} = \frac{\cos \theta + i \operatorname{sen} \theta - (\cos \theta - i \operatorname{sen} \theta)}{2i} = \frac{2i \operatorname{sen} \theta}{2i} = \operatorname{sen} \theta$$

Siguiendo estas fórmulas se pueden definir las funciones trigonométricas *seno* y *coseno* en el campo de los números complejos de la siguiente forma.

Definición C.25. Seno y Coseno Complejo

Las funciones trigonométricas *seno*, *coseno* y *tangente* se definen en el plano complejo de la siguiente forma:

$$\left. \begin{aligned} \cos z &= \frac{e^{zi} + e^{-zi}}{2} \\ \operatorname{sen} z &= \frac{e^{zi} - e^{-zi}}{2i} \end{aligned} \right\} \quad \tan z = \frac{\operatorname{sen} z}{\cos z} = \frac{e^{zi} - e^{-zi}}{i(e^{zi} + e^{-zi})} = -i \left(\frac{e^{zi} - e^{-zi}}{e^{zi} + e^{-zi}} \right)$$

Demostración C.3.

Comprobamos las fórmulas del coseno y seno complejo usando la fórmula de Euler y teniendo en cuenta que $\cos(-\alpha) = \cos(\alpha)$ y que $\operatorname{sen}(-\alpha) = -\operatorname{sen}(\alpha)$.

$$\frac{e^{zi} + e^{-zi}}{2} = \frac{\cos(z) + i \operatorname{sen}(z) + \cos(-z) + i \operatorname{sen}(-z)}{2} = \frac{\cos(z) + i \operatorname{sen}(z) + \cos(z) - i \operatorname{sen}(z)}{2} = \frac{2 \cos(z)}{2} = \cos(z)$$

$$\frac{e^{zi} - e^{-zi}}{2i} = \frac{\cos(z) + i \operatorname{sen}(z) - \cos(-z) - i \operatorname{sen}(-z)}{2i} = \frac{\cos(z) + i \operatorname{sen}(z) - \cos(z) + i \operatorname{sen}(z)}{2i} = \frac{2i \operatorname{sen}(z)}{2i} = \operatorname{sen}(z)$$

Ejercicio C.26. Expresar en forma binómica el valor de $\cos(1+i)$.

Aplicando la definición del coseno de un número complejo se tiene:

$$\begin{aligned} \cos(1+i) &= \frac{e^{(1+i)i} + e^{-(1+i)i}}{2} = \frac{e^{-1+i} + e^{1-i}}{2} = \frac{e^{-1}e^i + ee^{-i}}{2} = \frac{e^{-1}(\cos 1 + i \operatorname{sen} 1) + e(\cos(-1) + i \operatorname{sen}(-1))}{2} \\ &= \frac{e^{-1}e^i + ee^{-i}}{2} = \frac{e^{-1}(\cos 1 + i \operatorname{sen} 1) + e(\cos 1 - i \operatorname{sen} 1)}{2} = \frac{e + e^{-1}}{2} \cos 1 + i \frac{e^{-1} - e}{2} \operatorname{sen} 1 \end{aligned}$$

³⁴Notar que en la fórmula no aparece el doble signo $\pm$ delante de la raíz, ya que en este caso se obtendrán dos soluciones complejas cualesquiera, no necesariamente la misma con signos opuestos.

³⁵Ver el Teorema C.1, página 309.

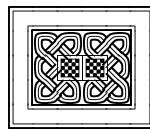
Nota C.4. Basadas en las fórmulas anteriores, se definen las *Funciones Hiperbólicas*: seno hiperbólico, coseno hiperbólico y tangente hiperbólica, de un número $x \in \mathbb{R}$, de la siguiente forma:

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \quad ; \quad \cosh(x) = \frac{e^x + e^{-x}}{2} \quad ; \quad \tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Verificándose, entre otras, la siguiente propiedad básica: $\cosh^2(x) - \sinh^2(x) = 1$.

En base a ellas se definen también las siguientes funciones hiperbólicas complejas para cualquier número $z \in \mathbb{C}$:

$$\sinh(z) = \frac{e^z - e^{-z}}{2} \quad ; \quad \cosh(z) = \frac{e^z + e^{-z}}{2} \quad ; \quad \tanh(z) = \frac{\sinh(z)}{\cosh(z)} = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$



Bibliografía

- O. BRETSCHER, **Linear algebra with applications**, Prentice Hall. 5^a Edition. 2019.
- J. BURGOS, **Álgebra lineal** (Ed. McGraw-Hill, 2000)
- M. CORRAL, **Trigonometry**, <http://www.mecmath.net/trig>. 2009.
- P. J. HERNANDO, **Clases de Cálculo I para Ingeniería**, Versión 3.8, (Septiembre 2025).
- P. J. HERNANDO, **Clases de Cálculo II para Ingeniería**, Versión 4.2, (Septiembre 2025).
- P. J. HERNANDO, **Clases de Cálculo III para Ingeniería**, Versión 2.2, (Septiembre 2025).
- P. J. HERNANDO, **Clases de Ampliación de Matemáticas para Ingeniería**, Versión 6.0, (Septiembre 2025).
- B. KOLMAN, **Álgebra lineal**, Octava Edición, (Prentice Hall, 2006).
- D. C. LAY, **Álgebra lineal y sus aplicaciones**, Cuarta Edición, (Addison-Wesley, 2012).
- B. NOBLE y J. W. DANIEL, **Álgebra lineal aplicada**, Tercera Edición, (Prentice Hall Hispanoamericana, 1989).
- D. POOLE, **Álgebra lineal, una introducción moderna**, Tercera Edición, Thomson. 2011.
- G. STRANG, **Álgebra lineal y sus aplicaciones**, Cuarta Edición, Thomson. 2007.
- G. STRANG, **Introduction to linear algebra**, Fifth Edition, Wellesley-Cambridge Press. 2016.
- D. G. ZILL y M. CULLEN, **Ecuaciones diferenciales**, Séptima Edición, (Cengage Learning Editores, 2009).

