

UNIVERSIDAD COMPLUTENSE DE MADRID

Manual de Álgebra

Grado en Física

Curso 2020/2021

PROF. PIERGIULIO TEMPESTA

ACTUALIZACIÓN: 27 DE JUNIO DE 2021

Contents

1 Estructuras algebraicas	6
1.1 Notas históricas	6
1.2 Leyes de composición	7
1.3 Estructuras algebraicas	8
1.3.1 Semigrupos, monoides, grupos	8
1.3.2 Propiedades elementales de los grupos	9
1.3.3 Subestructuras	10
1.4 Anillos	10
1.5 Cuerpos	11
1.6 Números Complejos	13
1.6.1 Conjugación compleja	14
1.6.2 Módulo de un número complejo	15
1.6.3 Argumento	16
1.7 Fórmula de de Moivre	17
1.8 Raíces n-ésimas: caso general	18
1.9 Función exponencial compleja	19
1.10 Raíces n-ésimas de la unidad	20
1.11 Polinomios	20
1.11.1 Definiciones	20
1.11.2 Propiedades del grado	20
1.11.3 Raíces de polinomios y completitud algebraica	21
1.12 Fórmulas de Cardano - Viète	23
1.13 Matrices	23
1.13.1 Definiciones	23
1.13.2 Suma de matrices	25
1.13.3 Producto de una matriz por un escalar	25
1.13.4 Producto de dos matrices	25
1.13.5 Propiedades del producto de matrices	26
1.14 Matrices invertibles	27
1.15 Operaciones elementales	29
1.15.1 Definiciones	29
1.15.2 Un algoritmo para invertir matrices	29
1.15.3 Matrices por bloques	30
1.16 Reducción por filas	31

1.17	Sistemas de ecuaciones lineales	33
1.17.1	Definiciones	33
1.17.2	Método de Gauss-Jordan	34
1.18	Producto de matrices por bloques	35
2	Espacios Vectoriales	36
2.1	Espacios vectoriales: definición	36
2.1.1	Ejemplos de espacios vectoriales	37
2.1.2	Propiedades elementales de los espacios vectoriales	37
2.1.3	Combinaciones lineales	38
2.2	Subespacios vectoriales	38
2.2.1	Criterios para subespacios	39
2.2.2	Suma e intersección de subespacios vectoriales	40
2.3	Dependencia e independencia lineal de un conjunto (finito) de vectores	42
2.3.1	Propiedades elementales	42
2.4	Bases de un espacio vectorial	44
2.5	Suma directa	49
2.5.1	Definición	49
2.6	Identidad de Grassmann	50
2.6.1	Otras propiedades de la suma directa de subespacios	52
2.6.2	Bases adaptadas a una suma directa	52
2.7	Ecuaciones intrínsecas de un subespacio vectorial	52
2.8	Cambio de base	54
3	Aplicaciones Lineales	57
3.1	Definiciones	57
3.2	Criterio para aplicaciones lineales	58
3.3	Núcleo e imagen de una aplicación lineal	60
3.4	Composición de aplicaciones lineales	62
3.5	El teorema de las dimensiones	63
3.6	Teoremas de equivalencia	64
3.6.1	Primer teorema de equivalencia	64
3.6.2	Segundo teorema de equivalencia	65
3.7	Matriz asociada a una aplicación lineal	65
3.7.1	Aplicación lineal asociada a una matriz	66
3.7.2	Matriz asociada a la composición de aplicaciones	68
3.7.3	Ejercicios	68
3.8	Cambios de base y matrices asociadas a aplicaciones lineales	69
3.9	Formas lineales	71
3.9.1	Base dual	71
3.10	Aplicaciones multilineales	72
3.11	Aplicaciones alternadas	73
3.12	Grupo de permutaciones	75
3.13	Determinantes e invertibilidad	80

4	Autovalores y Autovectores	84
4.1	Definiciones y propiedades básicas	84
4.2	Subespacios invariantes	86
4.3	Ecuación característica	86
4.4	Ejemplos	89
4.5	Independencia lineal	91
4.6	Multiplicidad algebraica y geométrica	91
4.7	Diagonalización de un endomorfismo	92
4.7.1	Primer criterio de diagonalización	93
4.7.2	Segundo criterio de diagonalización	93
4.7.3	Tercer criterio de diagonalización	94
4.8	Ejercicios	96
4.9	Teorema de Caley-Hamilton y polinomio mínimo	98
5	Formas bilineales y cuadráticas. Productos escalares	100
5.1	Aplicaciones y formas bilineales	100
5.2	Formas cuadráticas	105
5.2.1	Propiedades básicas	105
5.2.2	Forma polar de una forma cuadrática	107
5.2.3	Matriz asociada a Φ	108
5.3	Vectores conjugados	109
5.3.1	Propiedades de los vectores conjugados	109
5.3.2	Núcleo de una forma bilineal simétrica	111
5.3.3	Conjugación respecto de una forma cuadrática	111
5.4	Clasificación de las formas cuadráticas reales	112
5.5	Desigualdades de Schwarz y de Minkowski	113
5.5.1	Desigualdad de Schwarz	113
5.5.2	Desigualdad de Minkowski	114
5.6	Ejercicios	115
5.7	Productos escalares	119
5.7.1	Definición	119
5.7.2	Norma asociada a un producto escalar	119
5.7.3	Distancia inducida por una norma	120
5.7.4	Ángulo entre vectores no nulos	121
5.8	Bases ortogonales y ortonormales	123
5.8.1	Teorema de la independencia lineal	124
5.8.2	Descomposición en coeficientes de Fourier	124
5.9	Método de ortonormalización de Gram-Schmidt	126
5.9.1	Procedimiento general	126
5.9.2	Observaciones sobre el método	128
5.9.3	Bases ortonormales	129
5.10	Complemento ortogonal	129
5.11	Proyección ortogonal	132
5.12	Producto vectorial en \mathbb{R}^3	133

5.13	Formas sesquilineales y cuadráticas en espacios vectoriales complejos	134
5.13.1	Definición	134
5.13.2	Teorema de los autovalores reales para matrices hermíticas	136
5.14	Productos escalares hermíticos	137
5.14.1	Identidad de polarización	137
5.14.2	Desigualdad de Cauchy-Schwarz y triangular en un espacio hermítico	138
5.14.3	Método de Gram-Schmidt en espacios hermíticos	138
5.15	Ejercicios	139
6	Aplicaciones lineales entre espacios con producto escalar	141
6.1	Aplicación adjunta	141
6.2	Núcleo de f e imagen de la aplicación adjunta	142
6.3	Representación matricial de la aplicación adjunta	143
6.4	Operadores normales, autoadjuntos y unitarios	144
6.4.1	Definiciones	144
6.4.2	Propiedades generales	146
6.5	Teoría espectral para operadores normales	148
6.6	Teoría espectral de los operadores autoadjuntos	150
6.7	Teoría espectral para operadores unitarios	150
6.8	Otras propiedades de los operadores unitarios	151
6.9	Diagonalización simultanea de operadores normales que conmutan	152
6.10	Proyectores ortogonales en espacios hermíticos	152
6.10.1	Descomposición espectral	152
6.10.2	Cálculo de proyectores ortogonales	154
6.11	Ejemplos	154
6.12	Operadores positivos	158
6.13	Operadores simétricos en espacios vectoriales reales	158
6.13.1	Descomposición espectral de un operador simétrico	159
6.14	Complejización de un espacio vectorial real	159
6.14.1	Espacio complejificado	159
6.14.2	Complejización de un operador lineal	160
6.14.3	Complejización de un producto escalar real	160
6.15	Operadores ortogonales en espacios euclídeos	160
6.16	Diagonalización por bloques de un operador ortogonal	161
6.16.1	Consideraciones generales	161
6.16.2	Caso $n = 2$	162
6.16.3	Caso $n = 3$	164
6.17	Fórmula de Rodrigues	165
7	Diagonalización de formas cuadráticas reales	167
7.1	Reducción a forma diagonal	167
7.2	Método de Gauss-Lagrange	169
7.3	Ley de inercia de Sylvester	170
7.4	Diagonalización por congruencia de una forma cuadrática	173

7.4.1	Criterio de Sylvester	176
7.5	Cónicas	177
7.5.1	Secciones cónicas	177
7.5.2	Ecuación general de una cónica	178
7.6	Ecuación reducida de una cónica	179
7.7	Invariantes métricos de las cónicas	180
7.8	Clasificación de las cónicas por invariantes	181
7.8.1	Centro de simetría de cónicas	181
7.8.2	Reducción a forma canónica. Caso $I_2 \neq 0$	181
7.8.3	Reducción a forma canónica. Caso $I_2 = 0$	184

Chapter 1

Estructuras algebraicas

En este capítulo estudiaremos la noción de estructura algebraica, que juega un papel fundamental en matemática y en muchas ramas de la ciencia. Conceptos como los de semigrupo, grupo, anillo, campo, espacio vectorial, categoría, etc., permiten enfocar el estudio del álgebra en un sentido abstracto, y deducir propiedades comunes de clases muy amplias de objetos matemáticos.

1.1 Notas históricas

La palabra álgebra es de origen árabe. El primer tratado moderno de álgebra fue escrito en el año 820 d.C. por Al-Khwarizmi (desde su nombre deriva la palabra moderna “algoritmo”).

Sin embargo, el origen de la disciplina remonta a la civilización babilónica, que ya poseía nociones de ecuaciones algebraicas de I y II grado.

Los antiguos griegos se enfocaron esencialmente en la resolución geométrica de las ecuaciones algebraicas. Diofanto de Alejandría (III siglo a.C.) escribió un tratado, “*Arithmetica*”, sobre la resolución de dichas ecuaciones.

A partir del siglo XIII, el estudio del Álgebra empieza a desarrollarse en Europa de forma sistemática. El más importante matemático de la edad media, Leonardo Fibonacci (Pisa, 1170-1250), dio una significativa contribución al desarrollo del álgebra en su “*Liber Abaci*” (1202). En particular, él introdujo varias sucesiones numéricas y resolvió las ecuaciones cúbicas.

En la edad renacentista, la matemática en general, y el álgebra en particular, tuvieron un gran desarrollo con Nicola Tartaglia, Gerolamo Cardano, etc. Así, se organizaron desafíos matemáticos y competiciones de álgebra en las principales plazas italianas.

En el siglo XVII, la escuela francesa, con François Viète, René Descartes, Pierre Fermat, dio contribuciones fundamentales al desarrollo de una visión nueva de la geometría, expresada a través de nociones algebraicas.

Sería imposible describir en pocas palabras los avances impresionantes de la matemática y en particular del álgebra obtenidos a partir del siglo XVIII, por

matemáticos como Gauss, Lagrange, Euler, etc. y posteriormente con Galois, Abel, Ruffini, Dedekind, Riemann, etc.

Cabe destacar que, a pesar de que el álgebra se haya convertido, cada vez más, en una disciplina abstracta, sobre todo en el siglo XX (con la geometría algebraica y la obra de Grothendieck y muchos más) y en la actualidad, sin embargo, paralelamente, su aplicabilidad en otras ramas de la ciencia no ha dejado de crecer.

1.2 Leyes de composición

Denotaremos los conjuntos con letras latinas mayúsculas (A, B, C, \dots, X, \dots). Los elementos de un conjunto se denotarán con letras latinas minúsculas.

Por definición, el conjunto de los números naturales incluye el cero:

$$\mathbb{N} = 0, 1, 2, \dots$$

Dados dos conjuntos X e Y , su producto cartesiano es el conjunto de los pares ordenados de elementos de X e Y :

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\} .$$

Definición 1. Sea X un conjunto no vacío (no necesariamente numérico). Una ley de composición interna en X es una aplicación

$$\perp : X \times X \rightarrow X$$

$$(a, b) \in X \times X \rightarrow a \perp b \in X \tag{1.1}$$

El símbolo \perp es utilizado para denotar una ley de composición arbitraria.

Ejemplo 1.1. La suma $+$ y la multiplicación \cdot en \mathbb{N} son leyes de composición internas: a dos números naturales asocian otro natural.

Observación Sea $I = \{1, 3, 5, \dots\} \subset \mathbb{N}$ el subconjunto de los números impares. El producto

$$\cdot : I \times I \rightarrow I \quad \text{es una ley de composición interna en } I .$$

Sin embargo, la suma

$$+ : I \times I \rightarrow \mathbb{N} \quad \text{no es una ley de composición interna en } I$$

porque la suma de dos impares es un número par.

Definición 2. Sea \perp una ley de composición interna en X . Diremos que la ley es asociativa si

$$(a \perp b) \perp c = a \perp (b \perp c), \quad \forall a, b, c \in X . \tag{1.2}$$

En este caso, utilizaremos directamente el símbolo $a \perp b \perp c$ en el caso de composición de tres elementos. Diremos que la ley es conmutativa si

$$a \perp b = b \perp a, \quad \forall a, b \in X .$$

Ejemplo 1.2. En \mathbb{R} , la suma $+$ y el producto \cdot son leyes asociativas y conmutativas.

1.3 Estructuras algebraicas

1.3.1 Semigrupos, monoides, grupos

Introducimos ahora la noción más sencilla de estructura algebraica: el magma.

Definición 3. *Un magma es un par (X, \perp) , donde X es un conjunto no vacío, llamado el **soporte** de la estructura, y \perp es una ley de composición en X .*

Esta es la estructura más general en Álgebra. Sin embargo, necesitamos que la ley de composición posea más propiedades para que sea de mayor utilidad en muchos contextos, matemáticos y físicos. La propiedad más básica que usualmente se requiere es la asociatividad.

Definición 4. *Un **semigrupo** es un par (X, \perp) , donde $\perp : X \times X \rightarrow X$ es una ley de composición interna **asociativa**. Un semigrupo se dirá **conmutativo** si la ley de composición \perp es conmutativa.*

Ejemplo 1.3. Los ejemplos posiblemente más básicos de semigrupo son $(\mathbb{N}, +)$ y (\mathbb{N}, \cdot) , ambos conmutativos.

Construiremos ahora estructuras cada vez más refinadas, pidiendo que la ley de composición satisfaga a más propiedades.

Definición 5. *Sea (X, \perp) un semigrupo. Un elemento $e \in X$ se dirá **elemento neutro** de la ley \perp si*

$$\forall x \in X : \quad x \perp e = e \perp x = x . \quad (1.3)$$

En otras palabras, el elemento neutro deja invariado el elemento $x \in X$ con el que se componga. Una propiedad interesante de e es su unicidad.

Teorema 1. *Si un semigrupo admite elemento neutro, este elemento es único.*

Demostración. Sea (X, \perp) un semigrupo y supongamos que existan e, e' elementos neutros de \perp en X . Entonces

$$\begin{cases} e \perp e' = e' \perp e = e' & \text{porque } e \text{ es elemento neutro} \\ e \perp e' = e' \perp e = e & \text{porque } e' \text{ es elemento neutro} \end{cases} \implies e = e' .$$

□

Definición 6. *Un semigrupo (X, \perp) dotado de elemento neutro de \perp se llamará **monoide**.*

Un monoide se dirá conmutativo si su ley de composición es conmutativa. Los ejemplos anteriores, es decir $(\mathbb{N}, +)$ y (\mathbb{N}, \cdot) son también monoides (conmutativos).

Definición 7. Sea (X, \perp) un monoide. Diremos que un elemento $x \in X$ admite un inverso en X , denotado con x^{-1} , si

$$x \perp x^{-1} = x^{-1} \perp x = e . \quad (1.4)$$

Observamos que los monoides $(\mathbb{N}, +)$ y (\mathbb{N}, \cdot) no admiten elementos invertibles (aparte el caso trivial de sus respectivos elementos neutros). Para que los elementos de \mathbb{N} sean invertibles respecto de la suma, debemos extender este conjunto. El conjunto $(\mathbb{Z}, +)$ es un monoide conmutativo: para cada $n \in \mathbb{Z}$ el número $-n$ es su inverso respecto de la suma.

El conjunto de los números racionales (\mathbb{Q}, \cdot) es un monoide conmutativo respecto del producto. En general, no todos los elementos de un monoide son invertibles. Cuando esta propiedad se cumple, obtenemos una estructura algebraica de gran relevancia teórica e aplicativa: la de grupo.

Definición 8. Un grupo es un monoide (X, \perp) en el que todo elemento admite un inverso.

Si la ley es conmutativa, el grupo se dirá conmutativo o abeliano (desde Niels Abel, matemático noruego (1802-1829)). En otras palabras, un grupo es un par (X, \perp) en que se cumple la propiedad asociativa, existe el elemento neutro y el inverso de cada elemento.

1.3.2 Propiedades elementales de los grupos

La teoría de grupos es una de las ramas más importantes de la matemática. Consideraremos en esta sección algunas propiedades sencillas pero fundamentales de la noción de grupo.

Teorema 2. Sea (G, \perp) un grupo. El inverso de cada elemento de G es único.

Demostración. Sea $a \in G$. Supongamos que a posea dos elementos inversos:

$$\exists a^{-1} \in G \text{ tal que } a \perp a^{-1} = a^{-1} \perp a = e ,$$

$$\exists a' \in G \text{ tal que } a \perp a' = a' \perp a = e .$$

Entonces,

$$\begin{aligned} a' \perp a \perp a^{-1} &= a' \perp (a \perp a^{-1}) = a' \perp e = a' \\ \parallel & \\ (a' \perp a) \perp a^{-1} &= e \perp a^{-1} = a^{-1} \end{aligned} \quad \Longrightarrow \quad a' = a^{-1}$$

Teorema 3. Sea (G, \perp) un grupo. Para todo $a, b \in G$:

$$(a \perp b)^{-1} = b^{-1} \perp a^{-1} . \quad (1.5)$$

Demostración. Tenemos:

$$(a \perp b) \perp (b^{-1} \perp a^{-1}) = a \perp (b \perp b^{-1}) \perp a^{-1} = a \perp e \perp a^{-1} = a \perp a^{-1} = e .$$

□

Ejemplos de grupos: $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, (\mathbb{R}^*, \cdot) , donde $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ son grupos conmutativos. Sin embargo, (\mathbb{Z}, \cdot) no es un grupo. Veremos en la sección 1.14 que el conjunto de las matrices invertibles con coeficientes en un cuerpo \mathbb{K} es un grupo.

1.3.3 Subestructuras

Una pregunta muy natural es la siguiente: si el par (X, \perp) representa una estructura algebraica (semigrupo, monoide, grupo), entonces, dado un subconjunto $A \subset X$ del soporte de la estructura, ¿ bajo qué condiciones (A, \perp) es *per se* una estructura algebraica ? En este caso, diremos que (A, \perp) es una subestructura de (X, \perp) .

Definición 9. Sea (X, \perp) un semigrupo, $A \subset X$, $A \neq \emptyset$. Diremos que A es estable respecto a la ley \perp si

$$\forall x, y \in A : \quad x \perp y \in A .$$

Es evidente que, si A es estable respecto de \perp , entonces (A, \perp) es una subestructura (precisamente, un sub-semigrupo) de X . En el caso de subconjuntos de un grupo, vale el siguiente teorema de caracterización

Proposición 1. Sea (G, \perp) un grupo y $H \subset G$, $H \neq \emptyset$. Entonces H es un subgrupo de $G \iff \forall a, b \in H: a \perp b^{-1} \in H$.

Demostración. Si H es un subgrupo y $b \in H$, entonces $b^{-1} \in H$. Además, la composición de dos elementos de H pertenece a H , en particular $a \perp b^{-1} \in H \forall a, b \in H$.

Viceversa, supongamos que $a \perp b^{-1} \in H \forall a, b \in H$. Está claro que la ley \perp , siendo asociativa en G , seguirá siendo asociativa para cada terna de elementos de H . También, eligiendo $a = b$, tenemos que $b \perp b^{-1} = e \in H$, es decir H contiene el elemento neutro (único) de G . Finalmente, eligiendo $a = e$, tenemos que $\forall b \in H, b^{-1} \in H$, es decir todo elemento de H es invertible. Por tanto, H es un subgrupo de G . □

Bajo las hipótesis de la Proposición 1, si G es un grupo conmutativo, evidentemente H es un subgrupo conmutativo.

1.4 Anillos

Dado un conjunto $X \neq \emptyset$, podemos también construir, con X como soporte, estructuras con más de una ley de composición. La más sencilla entre ellas es la de anillo.

Definición 10. Un anillo $(A, +, \cdot)$, es un conjunto $A \neq \emptyset$ dotado de dos leyes de composición internas, llamadas por convenio suma y multiplicación, tales que:

- $(A, +)$ es un grupo conmutativo
- la ley producto en A es asociativa:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c), \quad \forall a, b, c \in A$$

y distributiva respecto de la suma:

$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c, \\ (a + b) \cdot c &= a \cdot c + b \cdot c, \quad \forall a, b, c \in A \end{aligned}$$

Por definición, en un anillo la primera ley (denotada simbólicamente como " + ") es conmutativa. Si también la segunda lo es, diremos que el anillo $(A, +, \cdot)$ es **conmutativo**. Si en el anillo $(A, +, \cdot)$ el producto admite un elemento neutro, llamado por convenio **unidad**, entonces diremos que A es un **anillo unitario**¹

Ejemplos muy relevantes de anillos son:

- (1) El anillo de los enteros $(\mathbb{Z}, +, \cdot)$, que es unitario y conmutativo
- (2) El anillo de los polinomios $(\mathbb{R}[x], +, \cdot)$ en la variable x , con coeficientes en \mathbb{R} , que es conmutativo y unitario
- (3) El anillo, no conmutativo, y unitario, de las matrices $\mathcal{M}_{n \times m}(A)$ cuyos elementos pertenecen a un anillo A .

1.5 Cuerpos

Cuando en un anillo unitario todos los elementos son invertibles también respecto de la segunda ley de composición (salvo el "cero" de la primera), tenemos una estructura algebraica fundamental: la de **cuerpo**.

Definición 11. Un **cuerpo** es un conjunto $K \neq \emptyset$, dotado de dos leyes de composición internas, llamadas suma y producto, tales que:

- $(K, +)$ es un grupo conmutativo, con elemento neutro 0
- (K^*, \cdot) es un grupo, donde $K^* = K \setminus \{0\}$ y $1 (\neq 0)$ es el elemento neutro del producto
- La suma es distributiva respecto del producto:

$$(a + b) \cdot c = a \cdot c + b \cdot c, \quad a \cdot (b + c) = a \cdot b + a \cdot c, \quad \forall a, b, c \in A$$

¹Varios autores (como Bourbaki y Lang) piden la existencia de la unidad en un anillo, es decir, definen como anillo lo que para nosotros es un anillo unitario.

Si el producto es conmutativo, el cuerpo $(K, +, \cdot)$ se dirá **conmutativo**.

Por ej., $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ son cuerpos conmutativos. Los cuaterniones $(\mathbb{H}, +, \cdot)$ forman un cuerpo no conmutativo.

Resumen. Utilizando la notación $+$ y \cdot para las leyes de composición:

$(X, +)$	semigrupo :	$+$ asociativa
$(X, +)$	monoide :	$+$ asociativa, $\exists e$
$(G, +)$	grupo :	$+$ asociativa, $\exists e$, $\exists x^{-1}$
$(A, +, \cdot)$	anillo :	$(A, +)$ grupo conmutativo, “ \cdot ” asociativa y distributiva
$(A, +, \cdot)$	anillo unitario :	$(A, +)$ grupo conmutativo, “ \cdot ” asociativa y distributiva, $\exists “1”$
$(K, +, \cdot)$	cuerpo :	$(K, +)$ grupo conmutativo, (K^*, \cdot) grupo, “ \cdot ” distributiva

Las estructuras anteriores se dirán conmutativas si, en el caso de semigrupos, monoides y grupos, es conmutativa la ley $+$; en el caso de anillos y cuerpos, si lo es la ley \cdot (la $+$ lo es por definición).

1.6 Números Complejos

Presentaremos la construcción rigurosa del cuerpo de los números complejos como estructura algebraica. Con más precisión, los complejos son pares de números reales, que se componen entre sí mediante dos leyes de composición internas específicas, como se detalla a continuación.

Definición 12. *El conjunto de los números complejos es el conjunto de los pares de números reales $\mathbb{C} := (\mathbb{R}^2, +, \cdot)$, dotado de dos leyes de composición internas definidas de la forma siguiente:*

$$\begin{aligned}
 + : \mathbb{R}^2 \times \mathbb{R}^2 \\
 (x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2) \\
 \cdot : \mathbb{R}^2 \times \mathbb{R}^2 \\
 (x_1, y_1) \cdot (x_2, y_2) &= (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)
 \end{aligned}$$

El subconjunto de los pares $(x, 0) \in \mathbb{C}$ se identificará con \mathbb{R} .

Observamos que la suma de pares de reales es estándar, mientras que el producto está definido de forma tal que los complejos adquieren propiedades de cuerpo muy interesantes.

Definición 13. *Se define unidad imaginaria al par*

$$i := (0, 1) . \tag{1.6}$$

Esta definición nos permite reobtener las propiedades usuales de la unidad imaginaria.

Proposición 2. *La unidad imaginaria $i \in \mathbb{C}$ satisface a la ecuación*

$$i^2 = -1 \tag{1.7}$$

Demostración.

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1 .$$

□

Sea $z = (x, y) \in \mathbb{C}$. Valen las igualdades siguientes:

$$(x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1) \cdot (y, 0)$$

Por tanto, dado que $(x, 0) = x \forall x \in \mathbb{R}$, tenemos la representación cartesiana de un número complejo:

$$z = x + iy . \tag{1.8}$$

Definición 14. *Dado el número $z = x + iy \in \mathbb{C}$, el número real x se dirá la **parte real** de z , y el número real y se dirá la **parte imaginaria** de z .*

Utilizaremos las notaciones $x = \operatorname{Re} z$, $y = \operatorname{Im} z$. Dados dos números complejos z_1 y z_2 ,

$$z_1 = z_2 \iff \begin{cases} x_1 = x_2 \\ y_1 = y_2 \end{cases} .$$

En particular, $z = 0 \iff \begin{cases} x = 0 \\ y = 0 \end{cases} .$

Determinemos el **inverso** de un número complejo. Sea $z = a+ib$, y queremos demostrar que $\exists z^{-1} = x + iy$ tal que $zz^{-1} = 1 = z^{-1}z$. Estas relaciones son equivalentes a

$$\begin{cases} ax - by = 1 \\ ay + bx = 0 \end{cases} \iff z^{-1} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2} .$$

Utilizando la definición de \mathbb{C} , es fácil demostrar el siguiente

Teorema 4. *El conjunto \mathbb{C} es un cuerpo conmutativo. Precisamente, valen las propiedades siguientes:*

$$z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3 \quad \forall z_1, z_2, z_3 \in \mathbb{C}$$

$$z + 0 = z \quad \forall z \in \mathbb{C}$$

$$z + (-z) = 0 \quad \forall z \in \mathbb{C}$$

$$z_1 + z_2 = z_2 + z_1 \quad \forall z_1, z_2 \in \mathbb{C}$$

$$z_1(z_2 z_3) = (z_1 z_2) z_3 \quad \forall z_1, z_2, z_3 \in \mathbb{C}$$

$$z \cdot 1 = z \quad \forall z \in \mathbb{C}$$

$$z \cdot z^{-1} = 1, \quad \forall z \in \mathbb{C}, \quad z \neq 0$$

$$z_1 z_2 = z_2 z_1 \quad \forall z_1, z_2 \in \mathbb{C}$$

$$z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3 \quad \forall z_1, z_2, z_3 \in \mathbb{C}$$

$$(z_1 + z_2) z_3 = z_1 z_3 + z_2 z_3 \quad \forall z_1, z_2, z_3 \in \mathbb{C}$$

1.6.1 Conjugación compleja

Definición 15. *Sea $z = x + iy \in \mathbb{C}$. Se define complejo conjugado de z , y se denota con \bar{z} , al número complejo*

$$\bar{z} = x - iy . \tag{1.9}$$

Proposición 3. *Como consecuencia inmediata de la definición anterior, valen las propiedades siguientes:*

- (1) $\overline{(\bar{z})} = z$
- (2) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$
- (3) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$

Además,

$$\operatorname{Re} z = \frac{1}{2}(z + \bar{z}), \quad \operatorname{Im} z = \frac{1}{2i}(z - \bar{z}) \quad (1.10)$$

1.6.2 Módulo de un número complejo

Definición 16. Se define módulo del número complejo $z = x + iy$ a la función

$$\begin{aligned} |\cdot| : \mathbb{C} &\rightarrow \mathbb{R} \\ |z| &= \sqrt{x^2 + y^2}. \end{aligned} \quad (1.11)$$

Proposición 4. La función módulo satisface las propiedades siguientes:

- (1) $|z_1 z_2| = |z_1| |z_2| \quad \forall z_1, z_2 \in \mathbb{C}$
- (2) $|\bar{z}| = |z| \quad \forall z \in \mathbb{C}$
- (3) $z \cdot \bar{z} = |z|^2 \quad \forall z \in \mathbb{C}$
- (4) $|\operatorname{Re} z| \leq |z|, \quad |\operatorname{Im} z| \leq |z| \quad \forall z \in \mathbb{C}$
- (5) $|z_1 + z_2| \leq |z_1| + |z_2| \quad (\text{desigualdad triangular}) \quad \forall z_1, z_2 \in \mathbb{C}$
- (6) $||z_1| - |z_2|| \leq |z_1 - z_2|$
- (7) $z^{-1} = \frac{\bar{z}}{|z|^2}, \quad \forall z \in \mathbb{C}, z \neq 0$

Demostración.

(1)

$$\begin{aligned} |z_1 z_2| &= \sqrt{(x_1 x_2 - y_1 y_2)^2 + (x_1 y_2 + x_2 y_1)^2} = \sqrt{x_1^2 x_2^2 + x_1^2 y_2^2 + y_1^2 x_2^2 + y_1^2 y_2^2} \\ &= \sqrt{x_1^2 + y_1^2} \sqrt{x_2^2 + y_2^2} = |z_1| |z_2|. \end{aligned}$$

(2)

$$|\bar{z}| = \sqrt{x^2 + (-y)^2} = |z|$$

(3)

$$z \bar{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2$$

(4)

$$|\operatorname{Re} z| = |x| \leq \sqrt{x^2 + y^2} = |z|.$$

Análogamente se demuestra que $|\operatorname{Im} z| \leq |z|$.

(5)

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1 + z_2}) = |z_1|^2 + |z_2|^2 + (z_1 \bar{z}_2 + \bar{z}_1 z_2) = |z_1|^2 + |z_2|^2 + 2\operatorname{Re}(z_1 \bar{z}_2) \\ &\leq |z_1|^2 + |z_2|^2 + 2|z_1 \bar{z}_2| = |z_1|^2 + |z_2|^2 + 2|z_1| |z_2| = (|z_1| + |z_2|)^2. \end{aligned}$$

(6)

$$|z_1| = |(z_1 - z_2) + z_2| \leq |z_1 - z_2| + |z_2| \implies |z_1| - |z_2| \leq |z_1 - z_2|.$$

Cambiando z_1 por z_2 , obtenemos la desigualdad $|z_2| - |z_1| \leq |z_1 - z_2|$. Ambas desigualdades implican $||z_1| - |z_2|| \leq |z_1 - z_2|$.

(7)

Desde la relación $z \cdot \bar{z} = |z|^2$, deducimos que, si $z \neq 0$, $z^{-1} = \frac{\bar{z}}{|z|^2}$. En particular, si $|z| = 1$, tenemos $\bar{z} = z^{-1}$. □

1.6.3 Argumento

Sea $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, y $z \in \mathbb{C}^*$. Existe $\theta \in \mathbb{R}$ tal que podemos escribir z en la forma siguiente:

$$z = |z| (\cos \theta + i \sin \theta) \quad (1.12)$$

La fórmula (1.12) es conocida como la **representación polar o trigonométrica** de un número complejo. Observamos que θ es el ángulo que el vector z forma con el eje real positivo. Claramente, añadiendo múltiplos de 2π a dicho ángulo, obtenemos una familia de ángulos que representan en la fórmula (1.12) el mismo número z . En otras palabras, θ está definido módulo un múltiplo de 2π .

Definición 17. Sea $z \in \mathbb{C}^*$. Llamaremos *argumento* de z , y lo denotaremos con $\arg(z)$ a cualquier $\theta \in \mathbb{R}$ tal que $z = |z| (\cos \theta + i \sin \theta)$.

En base a la definición anterior, a cada $0 \neq z$ asociamos una familia infinita de argumentos. Por ejemplo

$$\begin{aligned} z = i &\implies \theta \in \left\{ \frac{\pi}{2}, \frac{\pi}{2} \pm 2\pi, \dots \right\} = \left\{ \frac{\pi}{2} + 2k\pi \mid k \in \mathbb{Z} \right\} \\ z = -1 - i &\implies \theta \in \left\{ \frac{5\pi}{4} + 2k\pi \mid k \in \mathbb{Z} \right\} = \left\{ -\frac{3\pi}{4} + 2k\pi \mid k \in \mathbb{Z} \right\}. \end{aligned}$$

Evidentemente, $\arg(z)$ no es una función. Para que lo sea, es decir para que θ sea único, es necesario elegir un cierto intervalo de valores reales, de longitud 2π , e imponer que θ pertenezca a dicho intervalo. Cuando escogemos semejante intervalo I , diremos que hemos tomado una **determinación** del argumento. De esta forma, podemos construir una función, denotada con $\arg_I : \mathbb{C}^* \rightarrow I$. Típicos ejemplos de posibles elecciones de I son representados por los intervalos $[0, 2\pi)$, $(-\pi, \pi]$, $-\left(\frac{\pi}{2}, \frac{3\pi}{2}\right]$, etc.

Definición 18. La función $\arg_I(z)$ es el único valor de $\arg(z)$ que pertenece al intervalo I . En particular, si $I = (-\pi, \pi]$, la función

$$\text{Arg } z := \arg_{(-\pi, \pi]}(z), \quad z \in \mathbb{C}^* \quad (1.13)$$

se dirá la **determinación principal** del argumento de z .

Ejercicio 1.4.

$$\begin{aligned} \text{Arg}(1) = 0, \quad \text{Arg}(1+i) = \pi/4, \quad \text{Arg}(i) = \pi/2, \quad \text{Arg}(-1) = \pi, \\ \text{Arg}(-1-i) = -3\pi/4, \quad \text{Arg}(-i) = -\pi/2, \quad \text{Arg}(1-i) = -\pi/4. \end{aligned}$$

Observamos que $Arg(z) : \mathbb{C}^* \rightarrow (-\pi, \pi]$ es una función discontinua en el semieje $\mathbb{R}^- \cup \{0\}$. La determinación $arg_{[0, 2\pi)}(z)$ es discontinua en el semieje $\mathbb{R}^+ \cup \{0\}$. Si tomamos un intervalo arbitrario, la determinación $arg_{[\alpha, \alpha+2\pi)}$ (y también $arg_{(\alpha, \alpha+2\pi]}$) es discontinua en la semirrecta cerrada que forma un ángulo α con el eje real positivo.

La representación polar de un número complejo se escribe comúnmente en la forma

$$z = r(\cos \theta + i \sin \theta), \quad r = |z|, \theta = arg(z), \quad z \in \mathbb{C}^* .$$

Tenemos que

$$z, w \in \mathbb{C}^*, z = w \iff (|z| = |w|, \quad arg z = arg w \pmod{2\pi})$$

Calculemos ahora el producto de dos complejos a través de su representación trigonométrica. Si $z_1, z_2 \in \mathbb{C}^*$, tenemos:

$$\begin{aligned} z_1 z_2 &= r_1(\cos \theta_1 + i \sin \theta_1) r_2(\cos \theta_2 + i \sin \theta_2) \\ &= r_1 r_2 [(\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2) + i(\cos \theta_1 \sin \theta_2 + \sin \theta_1 \cos \theta_2)] \\ &= r_1 r_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)] \end{aligned}$$

Deducimos inmediatamente que

$$arg(z_1 z_2) = arg z_1 + arg z_2 \pmod{2\pi}$$

Sin embargo,

$$Arg(z_1 z_2) \neq Arg(z_1) + Arg(z_2).$$

Como contraejemplo, observamos que

$$Arg(-i) = -\frac{\pi}{2} \neq Arg(-1) + Arg(i) = \frac{3\pi}{2} .$$

Otras propiedades elementales del argumento son:

$$\begin{aligned} (zz^{-1} = 1 \implies) \quad arg(z^{-1}) &= -arg z && \pmod{2\pi} \\ (z\bar{z} = |z|^2 \implies) \quad arg(\bar{z}) &= -arg z && \pmod{2\pi} \\ arg(z_1/z_2) &= arg z_1 - arg z_2 && \pmod{2\pi} \end{aligned}$$

validas para todo $z, z_1, z_2 \in \mathbb{C}^*$.

1.7 Fórmula de de Moivre

La representación polar es particularmente versatil a la hora de multiplicar entre sí dos números complejos.

Teorema 5. Sea $z \in \mathbb{C}^*$, $z = r(\cos \theta + i \sin \theta)$. Entonces

$$z^n = r^n(\cos n\theta + i \sin n\theta), \quad n \in \mathbb{Z} . \quad (1.14)$$

Este teorema se puede demostrar por inducción (nótese que es válida también para cualquier $n \in \mathbb{Z}$).

Una consecuencia interesante de la fórmula de de Moivre es la posibilidad de calcular las funciones $\cos n\theta$ y $\sin n\theta$ en términos de $\cos \theta$ y $\sin \theta$. Por ejemplo, para calcular $\cos 3\theta$ y $\sin 3\theta$, observamos que

$$\begin{aligned} (\cos \theta + i \sin \theta)^3 &= \cos 3\theta + i \sin 3\theta \iff \\ \cos^3 \theta + 3i \cos^2 \theta \sin \theta - 3 \cos \theta \sin^2 \theta - i \sin^3 \theta &= \cos 3\theta + i \sin 3\theta . \end{aligned}$$

Identificando la parte real y la parte imaginaria de primero y segundo miembro, obtenemos

$$\begin{cases} \cos 3\theta = \cos^3 \theta - 3 \cos \theta \sin^2 \theta \\ \sin 3\theta = -\sin^3 \theta + 3 \cos^2 \theta \sin \theta . \end{cases}$$

1.8 Raíces n-ésimas: caso general

Resolvemos ahora el problema de determinar todas las raíces de un número complejo.

Sea $z \in \mathbb{C}^*$, $z = r(\cos \theta + i \sin \theta)$. Determinemos los números complejos

$$w \in \mathbb{C}, \quad w^n = z . \quad (1.15)$$

Sea $w = \rho(\cos \varphi + i \sin \varphi)$. La relación (1.15) implica

$$\rho^n(\cos n\varphi + i \sin n\varphi) = r(\cos \theta + i \sin \theta) \implies \begin{cases} \rho^n = r \\ n\varphi = \theta + 2k\pi, \quad k \in \mathbb{Z} \end{cases} .$$

Evidentemente, tenemos que los n números

$$w_k = \sqrt[n]{r} \left[\cos \left(\frac{\theta}{n} + \frac{2k\pi}{n} \right) + i \sin \left(\frac{\theta}{n} + \frac{2k\pi}{n} \right) \right], \quad k = 0, 1, \dots, n-1 \quad (1.16)$$

satisfacen la ecuación (1.15). Estos números se dirán las **raíces n-ésimas** del número z . Observamos que en la fórmula (1.16) sólo aparecen las n raíces distintas. Está claro que w_k y w_{k+ln} coinciden para todo $l \in \mathbb{Z}$.

Ejemplo 1.5. Calcular las raíces cúbicas de i . Sea $z = i$, $|z| = 1$. Entonces

$$w_k = \cos \left(\frac{\pi}{6} + \frac{2k\pi}{3} \right) + i \sin \left(\frac{\pi}{6} + \frac{2k\pi}{3} \right), \quad k = 0, 1, 2$$

Tenemos: $w_0 = \frac{1}{2}(\sqrt{3} + i)$, $w_1 = \frac{1}{2}(-\sqrt{3} + i)$, $w_2 = -i$.

Ejemplo 1.6. Calcular $(-8i)^{1/3}$. Sea $z = -8i$: $|z| = 8$, $\arg z = \theta = \frac{3\pi}{2}$. Entonces,

$$(-8i)^{1/3} = 2 \left[\cos \left(\frac{\pi}{2} + \frac{2k\pi}{3} \right) + i \sin \left(\frac{\pi}{2} + \frac{2k\pi}{3} \right) \right], \quad k = 0, 1, 2 .$$

Precisamente, tenemos

$$\begin{cases} w_0 = 2i, \\ w_1 = 2 \left[\cos \frac{7\pi}{6} + i \sin \frac{7\pi}{6} \right] = -\sqrt{3} - i, \\ w_2 = 2 \left[\cos \frac{11\pi}{6} + i \sin \frac{11\pi}{6} \right] = \sqrt{3} - i. \end{cases}$$

1.9 Función exponencial compleja

Definición 19. Sea $z = x + iy \in \mathbb{C}$. Definimos la exponencial compleja como

$$e^z = e^x (\cos y + i \sin y). \quad (1.17)$$

A partir de esta definición se puede demostrar también que valen las representaciones siguientes:

$$e^z = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n, \quad e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

como se verá en cursos superiores.

Claramente, si $z \in \mathbb{R}$, entonces la exponencial compleja se reduce a la exponencial real.

Valen las propiedades siguientes:

- (i) $e^z e^w = e^{z+w}$
- (ii) $e^0 = 1$
- (iii) $\overline{e^z} = e^{\bar{z}}$

Una consecuencia inmediata es la célebre fórmula de Euler

$$e^{i\theta} = \cos \theta + i \sin \theta \quad (1.18)$$

que nos permite introducir la representación exponencial de un número complejo:

$$z = r e^{i\theta}. \quad (1.19)$$

Tenemos las fundamentales relaciones

$$e^0 = 1, \quad e^{i\pi/2} = i, \quad e^{i\pi} = -1, \quad e^{3\pi i/2} = -i, \quad e^{2\pi i} = 1.$$

También podemos escribir

$$z = |z| e^{i \arg z} \quad (1.20)$$

Si $z = r e^{i\theta}$, tenemos que $z^n = r^n e^{in\theta}$.

1.10 Raíces n-ésimas de la unidad

Estudiamos ahora la ecuación

$$z^n = 1, \quad z \in C^* . \quad (1.21)$$

Aplicando la teoría anterior, obtenemos inmediatamente

$$z_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \quad k = 0, 1, \dots, n-1 . \quad (1.22)$$

Las soluciones de la ecuación (1.21) por tanto son

$$z_0 = 1, \quad z_1 = e^{\frac{2\pi i}{n}}, \quad z_2 = e^{\frac{4\pi i}{n}}, \dots, \quad z_{n-1} = e^{\frac{2(n-1)\pi i}{n}} .$$

Por ejemplo, para $n = 3$, tenemos que la ecuación $z^3 = 1$ admite las tres soluciones

$$z_0 = 1, \quad z_1 = e^{\frac{2\pi i}{3}}, \quad z_2 = e^{\frac{4\pi i}{3}} .$$

Interpretación geométrica. Las raíces n -ésimas de 1 determinan un polígono regular de n lados inscrito en el círculo de radio 1.

1.11 Polinomios

1.11.1 Definiciones

Definición 20. Sea $(A, +, \cdot)$ un anillo conmutativo unitario. Sea

$$p_n(x) := a_0 + a_1x + a_2x^2 + \dots + a_nx^n, \quad a_0, a_1, \dots, a_n \in A, \quad a_n \neq 0 \quad (1.23)$$

Diremos que $p_n(x)$ es un polinomio en x de grado n con coeficientes en el anillo A . El conjunto de todos los polinomios con coeficientes en A se denota con $A[x]$.

Una constante $a_0 \in A \iff$ polinomio de grado cero en $A[x]$.

Si n es el grado de un polinomio $p(x)$, el término a_nx^n se llamará **término dominante** del polinomio. Si el coeficiente del término dominante es igual a 1, el polinomio se dirá **mónico**.

1.11.2 Propiedades del grado

Sean $P, Q \in A[x]$. Valen las propiedades siguientes:

- (1) $\text{grad}(P \cdot Q) = \text{grad} P + \text{grad} Q$
- (2) $\text{grad}(P + Q) \leq \max(\text{grad} P, \text{grad} Q)$.

Ejemplo 1.7. $\mathbb{Z}[x]$ es el conjunto de todos los polinomios en x con coeficientes enteros.

$$p(x) = 3x^2 - 5x + 1 \in \mathbb{Z}[x], \quad q(x) = 9x^3 + 6x^3 + 7x - 8 \in \mathbb{Z}[x]$$

Otros ejemplos son $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$, etc.

Proposición 5. Sea $(A, +, \cdot)$ un anillo. Entonces $(A[x], +, \cdot)$ es también un anillo, respecto de las operaciones:

(1) suma de polinomios:

$$p(x) + q(x) := \sum_{k=0}^{\max(n,m)} (a_k + b_k)x^k$$

(2) producto de polinomios:

$$p(x) \cdot q(x) := \sum_{k=0}^{n+m} \left(\sum_{i+j=k} a_i b_j \right) x^k$$

donde $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$, $q(x) = b_0 + b_1x + b_2x^2 + \dots + b_mx^m \in A[x]$. Además, en la expresión de la suma se asume que si $\max(n, m) = n$, entonces $b_k = 0$ para $k > m$; análogamente, si $\max(n, m) = m$, tomaremos $a_k = 0$ para $k > n$.

Explícitamente, tenemos (sea por ej. $n > m$):

$$p(x) + q(x) := (a_0 + b_0) + (a_1 + b_1)x + \dots + (a_m + b_m)x^m + a_{m+1}x^{m+1} + a_nx^n ;$$

$$p(x) \cdot q(x) = a_0b_0 + (a_1b_0 + a_0b_1)x + (a_2b_0 + a_1b_1 + a_0b_2)x^2 + \dots + a_nb_mx^{n+m} .$$

$(A[x], +, \cdot)$: el anillo de los polinomios con coeficientes en el anillo conmutativo unitario A .

1.11.3 Raíces de polinomios y completitud algebraica

Sea \mathbb{K} un cuerpo conmutativo, y $p(x) \in \mathbb{K}[x]$: $p(x) = a_0 + a_1x + \dots + a_nx^n$, $a_i \in \mathbb{K}$.

Definición 21. Diremos que el escalar $\alpha \in \mathbb{K}$ es una raíz de $p(x)$ si $p(\alpha) = 0$.

El lema a continuación nos permitirá demostrar un resultado interesante de la teoría de los polinomios.

Lema 1. Para todo par de polinomios $p(x), q(x) \in \mathbb{K}[x]$, $q(x) \neq 0$, existen dos polinomios $d(x), r(x) \in \mathbb{K}[x]$, tales que

$$p(x) = d(x)q(x) + r(x)$$

con $\text{grad } r(x) < \text{grad } q(x)$. Además, dichos polinomios son únicos.

Teorema 6. Sea $\alpha \in \mathbb{K}$ y $p(x) \in \mathbb{K}[x]$. El escalar α es una raíz de $p(x) \iff p(x) = (x - \alpha)d(x)$.

Demostración. Sea α una raíz de $p(x)$. Como consecuencia del Lema 1, dados los polinomios $p(x)$ y $(x - \alpha)$, existen sendos polinomios $d(x)$ y $r(x)$ tales que

$$p(x) = (x - \alpha)d(x) + r(x)$$

con $\text{grad } r(x) < \text{grad } q(x) = 1 \implies r(x) = \text{const.}$ Además, dado que

$$0 = p(\alpha) = 0 + r(\alpha) \implies r(\alpha) = 0$$

así que $p(x) = (x - \alpha)d(x)$. □

El resultado siguiente establece que cada polinomio con coeficientes en \mathbb{C} posee al menos una raíz en \mathbb{C} . Por su relevancia es denominado *Teorema fundamental del álgebra*.

Teorema 7. *Sea $p(z) \in \mathbb{C}[z]$. Entonces $\exists z_1 \in \mathbb{C}$ tal que $p(z_1) = 0$.*

Desde el Teorema fundamental del álgebra desciende el

Teorema 8. *Todo polinomio $p(z) \in \mathbb{C}[z]$ se factoriza como producto de factores de grado 1, es decir*

$$p(x) = a(z - z_1)(z - z_2) \cdot \dots \cdot (z - z_n) \tag{1.24}$$

donde $a \in \mathbb{C}$, y z_1, \dots, z_n son las raíces de $p(z)$.

Demostración. Sea $\text{grad } p(z) = n$. Aplicando el teorema fundamental del álgebra y el Teorema 6, obtenemos $p(z) = (z - z_1)q(z)$, con $\text{grad } q(z) = n - 1$. Utilizando el mismo razonamiento a partir del polinomio $q(z)$ así obtenido, y reiterando el procedimiento, obtenemos la factorización (1.24). □

En general, cada raíz en la factorización anterior tendrá su multiplicidad algebraica, es decir, si $p(z)$ es un polinomio de grado N , en general

$$p(z) = a(z - z_1)^{n_1} \cdot \dots \cdot (z - z_k)^{n_k}, \quad n_1 + \dots + n_k = N.$$

Diremos entonces que la raíz z_i posee *multiplicidad algebraica* $n_i \geq 1$, $i = 1, \dots, k$. En particular, si $n_1 = \dots = n_k = 1$, diremos que las raíces de $p(z)$ son **simples**.

Esta fundamental propiedad de factorización para los polinomios en $\mathbb{C}[z]$ también se expresa diciendo que el cuerpo \mathbb{C} es **algebraicamente cerrado**. Sin embargo, ni \mathbb{Q} ni \mathbb{R} son algebraicamente cerrados. Como es bien conocido, la ecuación $x^2 = 2$ no admite soluciones en \mathbb{Q} , y la ecuación $x^2 + 1 = 0$ no admite soluciones en \mathbb{R} .

Si un polinomio tiene coeficientes reales, sus eventuales raíces complejas aparecen en pares $\{z_k, \bar{z}_k\}$.

1.12 Fórmulas de Cardano - Viète

Sea

$$p_n(z) := a_n z^n + a_{n-1} z^{n-1} + \dots + a_2 z^2 + a_1 z + a_0, \quad a_0, \dots, a_n \in \mathbb{C}, a_n \neq 0.$$

Como consecuencia del Teorema fundamental del álgebra, el polinomio tiene n raíces (no necesariamente distintas) z_1, \dots, z_n .

Problema. ¿Qué relación hay entre $\{a_0, \dots, a_n\}$ y $\{z_1, \dots, z_n\}$?

Consideremos algunos casos particulares.

$$p_1(z) = a_1 z + a_0, \quad z_1 = -\frac{a_0}{a_1}$$

$$p_2(z) = a_2 z^2 + a_1 z + a_0 = a_2(z - z_1)(z - z_2) = a_2[z^2 - z(z_1 + z_2) + z_1 z_2]$$

Por tanto,

$$\begin{cases} z_1 + z_2 = -\frac{a_1}{a_2}, \\ z_1 z_2 = \frac{a_0}{a_2}. \end{cases}$$

$$p_3(z) = a_3 z^3 + a_2 z^2 + a_1 z + a_0 = a_3(z - z_1)(z - z_2)(z - z_3) \implies$$

$$\begin{cases} z_1 + z_2 + z_3 = -\frac{a_2}{a_3} \\ z_1 z_2 + z_1 z_3 + z_2 z_3 = \frac{a_1}{a_3} \\ z_1 z_2 z_3 = -\frac{a_0}{a_3} \end{cases}$$

Caso general: sea

$$p_n(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_2 z^2 + a_1 z + a_0, \quad a_n \neq 0.$$

Tenemos entonces las fórmulas generales de Cardano-Viète, que relacionan las raíces con los coeficientes del polinomio:

$$\begin{cases} z_1 + z_2 + \dots + z_n = -\frac{a_{n-1}}{a_n} \\ (z_1 z_2 + z_1 z_3 + \dots + z_1 z_n) + (z_2 z_3 + \dots + z_2 z_n) + \dots + z_{n-1} z_n = \frac{a_{n-2}}{a_n} \\ \vdots \\ z_1 z_2 \cdot \dots \cdot z_n = (-1)^n \frac{a_0}{a_n} \end{cases}$$

1.13 Matrices

1.13.1 Definiciones

Definición 22. Una matriz es una colección de objetos, dispuestos en una forma rectangular, en filas y columnas.

Consideraremos siempre matrices formadas por escalares. Estos escalares se denominan los coeficientes de la matriz. Dada una matriz A , indicaremos con $(A)_{ij} \equiv a_{ij}$ el coeficiente situado en la fila i -ésima y en la columna j -ésima. El conjunto de todas las matrices con n filas y m columnas, con coeficientes en un cuerpo \mathbb{K} se denota con el símbolo $\mathcal{M}_{n \times m}(\mathbb{K})$. Es decir

$$A \in \mathcal{M}_{n \times m}(\mathbb{K}), \quad A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}, \quad a_{ij} \in \mathbb{K}.$$

También se usan las notaciones $A = (a_{ij})$, $A = \{a_{ij}\}$, $i = 1, \dots, n$, $j = 1, \dots, m$.

Definición 23. Sea $A \in \mathcal{M}_{n \times m}(\mathbb{K})$, con coeficientes a_{ij} . La matriz transpuesta de A , denotada con A^t , es la matriz en $\mathcal{M}_{m \times n}$ con coeficientes $(A^t)_{ij}$ tales que $(A^t)_{ij} := (A)_{ji}$, $i = 1, \dots, m$, $j = 1, \dots, n$.

En otras palabras, la matriz transpuesta de A se obtiene intercambiando las filas de A con sus columnas.

Definición 24. Sea $A \in \mathcal{M}_n(\mathbb{K})$. Si $A = A^t$, A se dice **simétrica**. Si $A = -A^t$, A se dice **antisimétrica**.

Definición 25. Una matriz $A \in \mathcal{M}_{n \times n}(\mathbb{K})$ se dirá **triangular superior** si $a_{ij} = 0 \forall i > j$, y **triangular estrictamente superior** si $a_{ij} = 0 \forall i \geq j$. Una matriz $A \in \mathcal{M}_{n \times n}(\mathbb{K})$ se dirá **triangular inferior** si $a_{ij} = 0 \forall i < j$ y **triangular estrictamente inferior** si $a_{ij} = 0 \forall i \leq j$.

Ejemplo 1.8. Las matrices

$$T_1 = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & a_{22} & \dots & \dots & a_{2n} \\ \vdots & \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}, \quad T_2 = \begin{pmatrix} a_{11} & 0 & \dots & \dots & 0 \\ a_{21} & a_{22} & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & \dots & 0 \\ a_{n,1} & a_{n,2} & \dots & a_{n,n-1} & a_{nn} \end{pmatrix}$$

son, respectivamente, triangular superior y triangular inferior. Si la diagonal principal también se anula, serían triangulares en sentido estricto.

Definición 26. Una matriz en la que los elementos de las diagonales descendentes son iguales se denomina **matriz de Toeplitz**, es decir

$$a_{ij} = a_{i+1,j+1} := a_{i-j}.$$

Ejemplo 1.9. La matriz

$$A = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \dots & \dots & a_{-(n-1)} \\ a_1 & a_0 & a_{-1} & \ddots & \dots & \vdots \\ a_2 & a_1 & a_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & & \ddots & a_1 & a_0 & a_{-1} \\ a_{n-1} & \dots & \dots & a_2 & a_1 & a_0 \end{pmatrix}$$

es de Toeplitz.

1.13.2 Suma de matrices

Sean $A = (a_{ij}), B = (b_{ij}) \in \mathcal{M}_{n \times m}(\mathbb{K})$.

Definición 27. *La aplicación*

$$\begin{aligned} + : \mathcal{M}_{n \times m}(\mathbb{K}) \times \mathcal{M}_{n \times m}(\mathbb{K}) &\rightarrow \mathcal{M}_{n \times m}(\mathbb{K}) \\ (A, B) &\rightarrow A + B, \quad (A + B)_{ij} := a_{ij} + b_{ij} \end{aligned}$$

se dirá *suma de las matrices A y B*.

Dicho de otra forma, la matriz suma tiene como coeficientes la suma (en \mathbb{K}) de los coeficientes que ocupan la misma posición en las matrices sumandas.

Definición 28. *La matriz tal que $a_{ij} = 0 \forall i, j$ se dirá matriz nula.*

1.13.3 Producto de una matriz por un escalar

Sea $\lambda \in \mathbb{K}, A = (a_{ij}) \in \mathcal{M}_{n \times m}(\mathbb{K})$.

Definición 29. *La aplicación*

$$\begin{aligned} \mathbb{K} \times \mathcal{M}_{n \times m}(\mathbb{K}) &\rightarrow \mathcal{M}_{n \times m}(\mathbb{K}) \\ (\lambda, A) &\rightarrow \lambda A, \quad (\lambda A)_{ij} := \lambda a_{ij} . \end{aligned}$$

se dirá *producto del escalar λ por la matriz A*.

Evidentemente, el producto por un escalar consiste en multiplicar cada entrada de la matriz A por el escalar λ .

Nótese que la aplicación anterior es un ejemplo de ley de composición **externa**. Si X es un conjunto no vacío, y \mathbb{K} un cuerpo, una ley de composición externa es una aplicación

$$* : \mathbb{K} \times X \rightarrow X.$$

De forma más general, \mathbb{K} puede ser remplazado por un conjunto no vacío arbitrario.

1.13.4 Producto de dos matrices

Definición 30. *Sean $A = (a_{ij}) \in \mathcal{M}_{n \times m}, B = (b_{ij}) \in \mathcal{M}_{m \times k}(\mathbb{K})$. Llamaremos producto de matrices a la aplicación*

$$\begin{aligned} \cdot : \mathcal{M}_{n \times m}(\mathbb{K}) \times \mathcal{M}_{m \times k}(\mathbb{K}) &\rightarrow \mathcal{M}_{n \times k}(\mathbb{K}) \\ (A, B) &\rightarrow A \cdot B \end{aligned}$$

tal que

$$c_{ij} = (A \cdot B)_{ij}, \quad c_{ij} := \sum_{k=1}^m a_{ik} b_{kj} .$$

Habitualmente omitiremos el \cdot para denotar el producto de dos matrices.

Explícitamente, si $A = (a_{ij}) \in \mathcal{M}_{n \times m}(\mathbb{K})$, $B = (b_{kl}) \in \mathcal{M}_{m \times k}(\mathbb{K})$, tenemos

$$AB = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1k} \\ \vdots & \dots & \vdots \\ b_{m1} & \dots & b_{mk} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1k} \\ \vdots & \dots & \vdots \\ c_{n1} & \dots & c_{nk} \end{pmatrix}$$

donde $c_{jl} = a_{j1}b_{1l} + \dots + a_{jm}b_{ml}$ es el producto de la fila j -ésima por la columna l -ésima.

1.13.5 Propiedades del producto de matrices

Teorema 9. Sean $A \in \mathcal{M}_{n \times m}(\mathbb{K})$, $B \in \mathcal{M}_{m \times p}(\mathbb{K})$, $C \in \mathcal{M}_{p \times r}(\mathbb{K})$. Entonces

$$A(BC) = (AB)C . \quad (1.25)$$

Demostración.

$$\begin{aligned} [A(BC)]_{ij} &= \sum_{k=1}^m A_{ik}(BC)_{kj} = \sum_{k=1}^m A_{ik} \left(\sum_{l=1}^p B_{kl}C_{lj} \right) = \sum_{l=1}^p \left(\sum_{k=1}^m A_{ik}B_{kl} \right) C_{lj} \\ &= \sum_{l=1}^p (AB)_{il}C_{lj} = [(AB)C]_{ij} . \end{aligned}$$

□

Teorema 10. Sean $A \in \mathcal{M}_{n \times m}(\mathbb{K})$, $B, C \in \mathcal{M}_{m \times p}(\mathbb{K})$. Entonces

$$A(B + C) = AB + AC . \quad (1.26)$$

□

Demostración.

$$\begin{aligned} [A(B + C)]_{ij} &= \sum_{k=1}^m A_{ik}(B + C)_{kj} = \sum_{k=1}^m A_{ik}(B_{kj} + C_{kj}) \\ &= \sum_{k=1}^m A_{ik}B_{kj} + \sum_{k=1}^m A_{ik}C_{kj} = (AB)_{ij} + (AC)_{ij} . \end{aligned}$$

□

Teorema 11. Sean $A, B \in \mathcal{M}_n(\mathbb{K})$. Entonces

$$(AB)^t = B^t A^t \quad (1.27)$$

Demostración.

$$[(AB)^t]_{ij} = (AB)_{ji} = \sum_{k=1}^n A_{jk}B_{ki} = \sum_{k=1}^n (B^t)_{ik}(A^t)_{kj} = [B^t A^t]_{ij} .$$

□

Observación El producto de matrices no es necesariamente conmutativo, es decir, en general $AB \neq BA$.

Otras propiedades elementales. Si $A, B \in \mathcal{M}_n(\mathbb{K})$, $\lambda \in \mathbb{K}$, tenemos:

$$\begin{aligned}(A^t)^t &= A \\ (A+B)^t &= A^t + B^t \\ (\lambda A)^t &= \lambda(A^t)\end{aligned}$$

Observación La discusión anterior nos lleva a la conclusión que el conjunto $(\mathcal{M}_n(\mathbb{K}), +, \cdot)$ es un **anillo** unitario: la matriz

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \dots & \ddots & 0 \\ 0 & \dots & \dots & 1 \end{pmatrix}$$

se dirá **matriz identidad** de $\mathcal{M}_n(\mathbb{K})$. En general, este anillo no es conmutativo.

Una cuestión natural es averiguar las propiedades de invertibilidad en este anillo respecto del producto de matrices.

1.14 Matrices invertibles

Definición 31. Sea $A \in \mathcal{M}_n(\mathbb{R})$. Se dice que A es invertible si existe otra matriz $C \in \mathcal{M}_n(\mathbb{K})$ tal que

$$AC = \mathbf{1}, \quad CA = \mathbf{1} .$$

Diremos entonces que C es la matriz inversa de A , y la denotaremos con A^{-1} .

Proposición 6. La matriz inversa es única.

Demostración. Esta propiedad es una consecuencia inmediata del Teorema 2. También podemos demostrarlo directamente. Supongamos que $\exists B \in \mathcal{M}_n(K)$ tal que $BA = AB = I$. Entonces

$$B = B \mathbf{1} = B(AA^{-1}) = (BA)A^{-1} = \mathbf{1}A^{-1} = A^{-1} .$$

Por tanto, $B = A^{-1}$, es decir, la inversa es única. □

Definición 32. Una matriz no invertible se dirá singular. Una matriz invertible se dirá regular o no singular.

Consideremos ahora el caso sencillo $n = 2$.

Teorema 12. Sea $A \in M_2(\mathbb{K})$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Si $ad - bc \neq 0$, entonces A es invertible, y

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad (1.28)$$

Demostración. Es suficiente comprobar con un cálculo directo que $AA^{-1} = \mathbf{1}$. \square

El escalar $ad - bc$ es el determinante de A , como veremos más adelante.

Teorema 13. Sean $A, B \in \mathcal{M}_n(\mathbb{K})$.

(a) Si A es invertible, entonces A^{-1} es invertible, y

$$(A^{-1})^{-1} = A. \quad (1.29)$$

(b) Si A y B son invertibles, también lo es AB , y

$$(AB)^{-1} = B^{-1}A^{-1}. \quad (1.30)$$

(c) Si A es invertible, también lo es A^t , y

$$(A^t)^{-1} = (A^{-1})^t. \quad (1.31)$$

Demostración.

(a) Por definición, la matriz inversa de A^{-1} tiene que cumplir la propiedad

$$(A^{-1})^{-1}A^{-1} = A^{-1}(A^{-1})^{-1} = A.$$

Observamos que A cumple esta propiedad. Dado que la inversa de una matriz es única, deducimos inmediatamente que $(A^{-1})^{-1} = A$.

(b) Es una consecuencia del Teorema 3. También podemos observar directamente que

$$(AB) \cdot B^{-1}A^{-1} = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = \mathbf{1}.$$

(c) Sea A invertible. Tenemos:

$$\mathbf{1} = (AA^{-1})^t = (A^{-1})^t \cdot A^t \iff (A^t)^{-1} = (A^{-1})^t.$$

\square

La discusión anterior nos permite enunciar el siguiente

Teorema 14. El conjunto de todas las matrices invertibles de $\mathcal{M}_n(\mathbb{K})$ es un grupo respecto del producto de matrices.

Definición 33. El grupo de las matrices invertibles de $\mathcal{M}_n(\mathbb{K})$ se denomina **grupo lineal** y se denota con $GL(n, \mathbb{K})$.

1.15 Operaciones elementales

1.15.1 Definiciones

Definición 34. Una operación elemental de filas consiste en una de las tres transformaciones siguientes:

- (1) Cambiar entre sí dos filas.
- (2) Multiplicar una fila por un escalar (no nulo)
- (3) Sumar una fila a otra fila.

De manera análoga se definen las operaciones elementales de columnas.

Definición 35. Se dice que dos matrices son equivalentes por filas (columnas) si existe una sucesión de operaciones elementales de filas (columnas) que convierte una matriz en la otra.

Denotaremos con $e_f(A)$ a un conjunto de operaciones elementales de filas aplicadas a una matriz A y con $e_c(A)$ a un conjunto de operaciones elementales de columnas.

Observación Para indicar que dos matrices son equivalentes por filas utilizaremos la notación $A \sim_f B$. Sean $A, B, C \in \mathcal{M}_{n \times m}(\mathbb{K})$. Tenemos:

$$\begin{aligned} A &\sim_f A \\ A \sim_f B &\implies B \sim_f A \\ A \sim_f B \text{ y } B \sim_f C &\implies A \sim_f C . \end{aligned}$$

Por tanto, siendo verificadas las propiedades reflexiva, simétrica y transitiva, \sim_f es una *relación de equivalencia*.

Lema 2. Sean $A \in \mathcal{M}_{n \times m}(\mathbb{K})$, $B \in \mathcal{M}_{m \times k}(\mathbb{K})$. Tenemos

$$e_f(AB) = e_f(A) \cdot B , \quad (1.32)$$

$$e_c(AB) = A \cdot e_c(B) , \quad (1.33)$$

$$e_f(A) = (e_c(A^t))^t, \quad e_c(A) = (e_f(A^t))^t . \quad (1.34)$$

1.15.2 Un algoritmo para invertir matrices

Teorema 15. Una matriz $A \in \mathcal{M}_n(\mathbb{K})$ es invertible si y sólo si es equivalente por filas a $\mathbf{1}_n$. En este caso, cualquier sucesión de operaciones elementales de fila que reduzca A a $\mathbf{1}_n$ también transforma $\mathbf{1}_n$ en A^{-1} .

Demostración. Según el Lema 2, sabemos que $e_f(AB) = e_f(A) \cdot B$. Entonces,

$$e_f(\mathbf{1}) = e_f(AA^{-1}) = (e_f(A))A^{-1}$$

Por tanto, deducimos la igualdad $e_f(\mathbf{1}) = (e_f(A))A^{-1}$, de donde se obtiene que, si $e_f(A) = \mathbf{1}$, entonces $A^{-1} = e_f(\mathbf{1})$. \square

Dada una matriz $A \in \mathcal{M}_n(\mathbb{K})$, tenemos entonces un simple **algoritmo** para determinar A^{-1} . Introducimos la matriz

$$(A \mid \mathbb{1}) = \left(\begin{array}{ccc|ccc} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array} \right)$$

Si A es equivalente por filas a $\mathbb{1}$, entonces $(A \mid \mathbb{1})$ es equivalente por filas a $(\mathbb{1} \mid A^{-1})$. De esta forma, construimos A^{-1} . Si no conseguimos reducir A a $\mathbb{1}$ mediante operaciones elementales, A no es invertible.

Ejemplo 1.10. Sean f_1, f_2 la primera y la segunda fila de la matriz A .

$$\begin{aligned} & \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right) \xrightarrow{f_2 \rightarrow f_2 - 3f_1} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right) \xrightarrow{f_2 \rightarrow (-1/2)f_2} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right) \\ & \xrightarrow{f_1 \rightarrow f_1 - 2f_2} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right). \end{aligned}$$

1.15.3 Matrices por bloques

Cuando una matriz tiene una estructura por bloques, la discusión de sus eventuales propiedades de invertibilidad se simplifica. Consideraremos dos casos sencillos.

(a) Supongamos que una matriz $F \in \mathcal{M}_n(\mathbb{K})$ tenga una estructura por bloques de la forma

$$F = \left(\begin{array}{c|c} A & B \\ \hline O & C \end{array} \right)$$

con $A \in \mathcal{M}_k(\mathbb{K})$, $B \in \mathcal{M}_{k \times (n-k)}(\mathbb{K})$, $C \in \mathcal{M}_{(n-k) \times (n-k)}$, $O \in \mathcal{M}_{(n-k) \times k}(\mathbb{K})$. Supongamos que A y C sean invertibles. Entonces, F es invertible, y

$$F^{-1} = \left(\begin{array}{c|c} A^{-1} & -A^{-1}BC^{-1} \\ \hline O & C^{-1} \end{array} \right)$$

Ejemplo 1.11. Sea

$$F = \left(\begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ 3 & 5 & 0 & 0 \\ \hline 0 & 0 & 3 & 8 \\ 0 & 0 & 1 & 3 \end{array} \right)$$

Identificamos los bloques:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}, \quad \det A = -1 \quad C = \begin{pmatrix} 3 & 8 \\ 1 & 3 \end{pmatrix}, \quad \det C = 1.$$

Entonces,

$$A^{-1} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} 3 & -8 \\ -1 & 3 \end{pmatrix}$$

así que

$$F^{-1} = \begin{pmatrix} -5 & 2 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & -1 & 3 \end{pmatrix}.$$

(b) Sea ahora una matriz por bloques de la forma

$$H = \left(\begin{array}{c|c} A & O \\ \hline C & B \end{array} \right).$$

Si A , B son invertibles, entonces

$$H^{-1} = \left(\begin{array}{c|c} A^{-1} & O \\ \hline -B^{-1}CA^{-1} & B^{-1} \end{array} \right).$$

1.16 Reducción por filas

Definición 36. Sea $A \in \mathcal{M}_{n \times m}(\mathbb{K})$. Llamaremos **pivote** o **entrada principal** de una fila (columna) de A al primer elemento no nulo de dicha fila (columna), si es que hay alguno. Una columna que contiene a un pivote se llamará columna pivote.

Definición 37. Diremos que una matriz rectangular está en forma escalonada por filas si satisface las tres propiedades siguientes:

- (1) Todas las filas distintas de cero están situadas arriba de cualquier fila integrada sólo por ceros.
- (2) El pivote de cada fila no nula está a la derecha del de la fila anterior.
- (3) Todas las entradas de una columna situadas debajo de un pivote son ceros.

Definición 38. Diremos que una matriz en forma escalonada está en su forma escalonada **reducida** si, aparte de las tres propiedades anteriores, se cumplen también las siguientes:

- (4) El pivote de cada fila (distinta de la fila nula) es igual a 1.
- (5) Cada pivote representa la única entrada distinta de cero en su columna.

Reducir por filas significa transformar una matriz, mediante operaciones elementales de fila, en una matriz escalonada.

Ejemplo 1.12. La matriz

$$F = \begin{pmatrix} 5 & -4 & 3 & 1 \\ 0 & 1 & -4 & 7 \\ 0 & 0 & 0 & 3/4 \end{pmatrix}$$

está en forma escalonada. La matriz

$$R = \begin{pmatrix} 1 & 0 & 0 & 15 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

está en forma escalonada reducida.

Reduciendo una matriz se puede obtener más de una matriz en forma escalonada; sin embargo, la forma escalonada reducida es **única**.

Teorema 16. *Cada matriz es equivalente por filas a una única matriz escalonada reducida, llamada la forma normal de Hermite de la matriz.*

Consecuencia: Las entradas principales siempre están en las mismas posiciones en cualquier forma escalonada.

Ejercicio 1.13. Sea la matriz

$$A = \begin{pmatrix} 0 & 3 & -6 & 6 \\ 3 & -7 & 8 & -5 \\ 3 & -9 & 12 & -9 \end{pmatrix}$$

Intercambiando filas, y procediendo con otras operaciones elementales, obtenemos su forma escalonada reducida:

$$\begin{aligned} & \begin{pmatrix} 3 & -9 & 12 & -9 \\ 3 & -7 & 8 & -5 \\ 0 & 3 & -6 & 6 \end{pmatrix} \xrightarrow{f_2 \rightarrow f_2 - f_1} \begin{pmatrix} 3 & -9 & 12 & -9 \\ 0 & 2 & -4 & 4 \\ 0 & 3 & -6 & 6 \end{pmatrix} \xrightarrow{f_{1,3} \rightarrow 1/3} \xrightarrow{f_{1,3}, f_2 \rightarrow 1/2} f_2} \begin{pmatrix} 1 & -3 & 4 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 1 & -2 & 2 \end{pmatrix} \\ \Rightarrow & \begin{pmatrix} 1 & -3 & 4 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{f_1 \rightarrow f_1 + 3f_2} \begin{pmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Definición 39. *Sea $A \in \mathcal{M}_{n \times m}(\mathbb{K})$. Llamaremos rango por fila de A , y lo denotaremos con $r_f(A)$ al número de filas no nulas de su forma escalonada reducida por filas.*

Por tanto,

$$r_f(A) = n - c_f(A), \quad (1.35)$$

donde c_f es el número de filas nulas (es decir, integradas por ceros) de A . Análogamente se introduce en rango por columnas:

$$r_c(A) = m - c_c(A), \quad (1.36)$$

siendo c_c el número de columnas nulas de A .

En realidad, las dos nociones de rango son equivalentes:

Teorema 17. *Sea $A \in \mathcal{M}_{n \times m}(\mathbb{K})$. Entonces*

$$r_f(A) = r_c(A). \quad (1.37)$$

Por tanto, hablaremos del **rango de A tout court**, y lo denotaremos con $r(A)$.

Proposición 7. (a) *Sea $A \in \mathcal{M}_{n \times m}(\mathbb{K})$. Entonces*

$$r(A) \leq \min\{n, m\}. \quad (1.38)$$

1.17.2 Método de Gauss-Jordan

La teoría de las transformaciones elementales nos permite afirmar que

$$AX = B \iff e_f(A)X = e_f(B) .$$

Entonces X es solución del sistema $AX = B$ si y sólo si X es solución del sistema reducido mediante operaciones elementales. El método de Gauss-Jordan consiste en llevar A a una forma escalonada (posiblemente reducida) mediante operaciones elementales de fila, obteniendo la matriz $e_f(A)$.

Introducimos la matriz ampliada del sistema

$$(A|B) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right)$$

Entonces, para resolver un sistema, ante todo llevamos A a una forma escalonada. Distinguimos varios casos.

(1) Si $r(A) = m$ (no hay filas de ceros), el rango del sistema es máximo y evidentemente $r(A) = r(A|B)$. Sea $r = r(A)$. Introducimos la cantidad $k = n - m$: es el número de parámetros libres. Observamos que $k \geq 0$, dado que $r \leq \min(m, n)$. Tenemos entonces $r = m \leq n$.

(1a) Si $r = m = n$, el sistema es compatible determinado. En este caso, $e_f(A) = \mathbf{1}$ y $X = e_f(B) = A^{-1} \cdot B$.

(1b) Si $r = m < n$, despejemos las x_i de la forma escalonada, dejando a la izquierda las m incógnitas independientes. El sistema es compatible, y dependerá de k parámetros, es decir, es compatible indeterminado.

(2) Si $r(A) < m$ (hay $m - r$ filas de ceros), el sistema es compatible $\iff [e_f(B)]_i = 0 \forall i = r + 1, \dots, m$. Esto es equivalente a pedir $r(A) = r(A|B)$. En caso contrario, será incompatible. El número de parámetros libres es ahora $k = n - r$, y por tanto el sistema es compatible determinado si $r = n$, compatible indeterminado para $r < n$.

Observamos que en todo caso $r + k = n$.

Si $B = 0$, es decir el sistema es homogéneo, siempre será compatible, y será determinado si $k = n - r = 0$ o indeterminado si $k > 0$.

La discusión anterior es equivalente al teorema fundamental de la teoría de las ecuaciones lineales:

Teorema 18 (Rouché-Frobenius-Capelli). *Sea un sistema de m ecuaciones lineales con n incógnitas de la forma $AX = B$.*

El sistema es compatible $\iff r(A) = r(A|B) = r$

En particular, el sistema es compatible determinado $\iff r = m$, y es indeterminado $\iff r < m$.

1.18 Producto de matrices por bloques

Sea $A \in \mathcal{M}_{m \times n}(\mathbb{K})$, $B \in \mathcal{M}_{n \times p}(\mathbb{K})$. Supongamos que A y B tengan una estructura por bloques:

$$A = \begin{bmatrix} A_{[11]} & \dots & A_{[1r]} \\ \vdots & \dots & \vdots \\ A_{[q1]} & \dots & A_{[qr]} \end{bmatrix}, \quad B = \begin{bmatrix} B_{[11]} & \dots & B_{[1s]} \\ \vdots & \dots & \vdots \\ B_{[r1]} & \dots & B_{[rs]} \end{bmatrix} \quad (1.42)$$

donde el bloque $A_{[ij]}$ tiene tamaño $m_i \times n_j$, $i = 1, \dots, q$, $j = 1, \dots, r$, $B_{[jv]}$ tiene tamaño $n_j \times p_v$, $v = 1, \dots, s$. También tenemos

$$m_1 + \dots + m_q = m, \quad n_1 + \dots + n_r = n, \quad p_1 + \dots + p_s = p.$$

Definimos el producto por bloques de las matrices A y B de la forma siguiente:

$$AB = \begin{bmatrix} (AB)_{[11]} & \dots & (AB)_{[1s]} \\ \vdots & \dots & \vdots \\ (AB)_{[q1]} & \dots & (AB)_{[qs]} \end{bmatrix} \quad (1.43)$$

donde

$$(AB)_{[iv]} = \sum_{k=1}^r A_{[ik]} \cdot B_{[kv]}. \quad (1.44)$$

Observación El producto de matrices triangulares superiores (o inferiores) por bloques sigue siendo triangular superior (inferior).

Chapter 2

Espacios Vectoriales

2.1 Espacios vectoriales: definición

Definición 41. Sea V un conjunto no vacío, cuyos elementos denotaremos con $\mathbf{v}_1, \mathbf{v}_2, \dots$, y \mathbb{K} un cuerpo conmutativo. Supongamos que en V sea definida una ley de composición interna, que llamaremos suma

$$\begin{aligned} + : V \times V &\longrightarrow V \\ (\mathbf{v}_1, \mathbf{v}_2) &\longrightarrow \mathbf{v}_1 + \mathbf{v}_2 \end{aligned}$$

y de una ley de composición externa sobre \mathbb{K} , que llamaremos producto por escalares

$$\begin{aligned} \cdot : \mathbb{K} \times V &\longrightarrow V \\ (\lambda, \mathbf{v}) &\longrightarrow \lambda \mathbf{v} . \end{aligned}$$

El conjunto $(V, +, \cdot)$ se denomina espacio vectorial sobre \mathbb{K} (o \mathbb{K} -espacio vectorial) si las leyes de composición satisfacen los axiomas siguientes:

(I) $(V, +)$ es un grupo conmutativo:

$$(I1) \quad \forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V : \quad \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3) = (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 \quad (2.1)$$

$$(I2) \quad \exists \mathbf{0} \in V \quad t.q. \quad \forall \mathbf{v} \in V : \mathbf{v} + \mathbf{0} = \mathbf{v} \quad (2.2)$$

$$(I3) \quad \forall \mathbf{v} \in V \exists \mathbf{v}' \in V \quad t.q. \quad \mathbf{v} + \mathbf{v}' = \mathbf{0} \quad (2.3)$$

$$(I4) \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in V : \quad \mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1 \quad (2.4)$$

(II) La ley de composición externa satisface las propiedades siguientes:

$$(II1) \quad \forall \lambda \in \mathbb{K}, \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in V : \quad \lambda(\mathbf{v}_1 + \mathbf{v}_2) = \lambda \mathbf{v}_1 + \lambda \mathbf{v}_2 \quad (2.5)$$

$$(II2) \quad \forall \lambda, \mu \in \mathbb{K}, \quad \forall \mathbf{v} \in V : \quad (\lambda + \mu)\mathbf{v} = \lambda \mathbf{v} + \mu \mathbf{v} \quad (2.6)$$

$$(II3) \quad \forall \lambda, \mu \in \mathbb{K}, \quad \forall \mathbf{v} \in V : \quad (\lambda \mu)\mathbf{v} = \lambda(\mu \mathbf{v}) \quad (2.7)$$

$$(II4) \quad \forall \mathbf{v} \in V : \quad 1 \cdot \mathbf{v} = \mathbf{v}, \quad 1: \text{unidad de } \mathbb{K} \quad (2.8)$$

Los elementos de V se llaman **vectores**, los de \mathbb{K} **escalares**. El elemento $\mathbf{0}$ se denomina **vector nulo** de V , el elemento \mathbf{v}' es dicho el opuesto de \mathbf{v} y se denota con $\mathbf{v}' = -\mathbf{v}$.

2.1.1 Ejemplos de espacios vectoriales

Ejemplo 1. El conjunto de los números reales $(\mathbb{R}, +, \cdot)$ es un espacio vectorial sobre \mathbb{R} . Más en general, si \mathbb{K} es un cuerpo, $(\mathbb{K}, +, \cdot)$ es un espacio vectorial sobre \mathbb{K} .

Ejemplo 2. Sea $V = \mathbb{R}^2$, el conjunto de los pares de números reales (x, y) , $x, y \in \mathbb{R}$. Introducimos las leyes suma y producto por escalares

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad (2.9)$$

$$\lambda(x, y) = (\lambda x, \lambda y) \quad \lambda \in \mathbb{R} \quad (2.10)$$

Por tanto, \mathbb{R}^2 es un espacio vectorial sobre \mathbb{R} .

Ejemplo 3. Sea $V = \mathbb{R}^n$, el conjunto de las n -plas de números reales (x_1, \dots, x_n) , $x_1, \dots, x_n \in \mathbb{R}$. Análogamente al caso anterior, introducimos las leyes suma y producto por escalares

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \quad (2.11)$$

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n) \quad \lambda \in \mathbb{R} \quad (2.12)$$

Por tanto, \mathbb{R}^n es un espacio vectorial sobre \mathbb{R} . Más en general, dado un cuerpo conmutativo \mathbb{K} , resulta que $(\mathbb{K}^n, +, \cdot)$ es un espacio vectorial sobre \mathbb{K} . En particular, $(\mathbb{C}^n, +, \cdot)$ es un espacio vectorial sobre \mathbb{C} .

Ejemplo 4. Sea $\mathcal{C}[0, 1]$ el conjunto de las funciones continuas en $[0, 1]$ con valores en \mathbb{R} :

$$\mathcal{C}[0, 1] := \{f : [0, 1] \rightarrow \mathbb{R}, f \text{ continua}\}$$

Observamos que si f y g son continuas, $f + g$ es continua, y λf es continua $\forall \lambda \in \mathbb{R}$. Por tanto, $(\mathcal{C}, +, \cdot)$ es un espacio vectorial sobre \mathbb{R} .

Ejemplo 5. $(\mathcal{M}_{m \times n}(\mathbb{K}), +, \cdot)$ donde $+$ denota la suma de matrices, y \cdot el producto de un escalar $\lambda \in \mathbb{K}$ por una matriz, es un espacio vectorial sobre \mathbb{K} .

Ejemplo 6. Sea $\mathbb{R}_n[x]$ el conjunto de todos los polinomios en la variable x , de grado menor o igual a $n \in \mathbb{N}$, con coeficientes en \mathbb{R} . Tenemos que $(\mathbb{R}_n[x], +, \cdot)$, donde $+$ es la suma de dos polinomios y \cdot el producto de un escalar $\lambda \in \mathbb{R}$ por un polinomio, es un espacio vectorial sobre \mathbb{R} .

Ejemplo 7. Sea S^O el conjunto de los segmentos orientados de \mathbb{R}^3 , de origen común O , dotado de la suma $+$ de dos segmentos orientados, definida por la regla del paralelograma, y el producto \cdot de un escalar $\lambda \in \mathbb{R}$ por un segmento orientado. Es inmediato comprobar que $(S^O, +, \cdot)$ es un espacio vectorial sobre \mathbb{R} .

2.1.2 Propiedades elementales de los espacios vectoriales

Teorema 19. Sea $(V, +, \cdot)$ un espacio vectorial sobre un cuerpo conmutativo \mathbb{K} . Se cumplen las propiedades siguientes:

1. $\forall \mathbf{v} \in V, \quad 0 \cdot \mathbf{v} = \mathbf{0}$
2. $\forall \lambda \in \mathbb{K}, \quad \lambda \mathbf{0} = \mathbf{0}$
3. $\forall \lambda \in \mathbb{K}, \forall \mathbf{v} \in V, \quad (-\lambda)\mathbf{v} = -(\lambda\mathbf{v})$
4. $\lambda\mathbf{v} = \mathbf{0} \implies \lambda = 0 \text{ o } \mathbf{v} = \mathbf{0}$

Proof. Consideremos cada caso por separado.

1. Tenemos:

$$\mathbf{v} + \mathbf{0} = \mathbf{v} = (1 + 0)\mathbf{v} = 1 \cdot \mathbf{v} + 0\mathbf{v} = \mathbf{v} + 0\mathbf{v}$$

así que

$$\mathbf{0} = 0\mathbf{v} .$$

2.

$$\lambda\mathbf{v} + \mathbf{0} = \lambda\mathbf{v} = \lambda(\mathbf{v} + \mathbf{0}) = \lambda\mathbf{v} + \lambda\mathbf{0} .$$

3.

$$\mathbf{0} = 0\mathbf{v} = [\lambda + (-\lambda)]\mathbf{v} = \lambda\mathbf{v} + (-\lambda)\mathbf{v} .$$

4. Si $\lambda = 0$, la propiedad 1 implica $0\mathbf{v} = \mathbf{0}$. Entonces, sea $\lambda \neq 0$. Por tanto, $\exists \lambda^{-1}$ tal que $\lambda\lambda^{-1} = 1$.

$$\lambda\mathbf{v} = \mathbf{0} \iff \lambda^{-1}\lambda\mathbf{v} = \lambda^{-1}\mathbf{0} \iff \mathbf{v} = \lambda^{-1}\mathbf{0} = \mathbf{0} .$$

□

2.1.3 Combinaciones lineales

Definición 42. Sean $\mathbf{v}_1, \dots, \mathbf{v}_n$ vectores de un espacio vectorial V sobre \mathbb{K} y $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. Se llama **combinación lineal** de los vectores $\mathbf{v}_1, \dots, \mathbf{v}_n$ con coeficientes $\lambda_1, \dots, \lambda_n$ al vector

$$\lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n .$$

2.2 Subespacios vectoriales

Un subespacio vectorial es un subconjunto no vacío de un espacio vectorial, que en sí mismo satisface a los axiomas de espacio vectorial. Con más precisión, tenemos la siguiente

Definición 43. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y $W \subseteq V$, $W \neq \emptyset$. Diremos que el conjunto W es un subespacio vectorial de V si $(W, +, \cdot)$ es un espacio vectorial sobre \mathbb{K} .

Ejemplos.

1. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Los conjuntos $\{\mathbf{0}\}$ y V son subespacios vectoriales de V (usualmente denominados subespacios triviales).
2. Sea $(\mathbb{R}, +, \cdot)$. Los subespacios vectoriales de \mathbb{R} son $\{\mathbf{0}\}$ y \mathbb{R} .
3. Sea $(\mathbb{R}^2, +, \cdot)$. Los subespacios vectoriales de \mathbb{R}^2 son $\{\mathbf{0}\}$, las rectas que pasan por el origen, y \mathbb{R}^2 .
4. Sea $(\mathbb{R}^3, +, \cdot)$. Los subespacios vectoriales de \mathbb{R}^3 son $\{\mathbf{0}\}$, las rectas que pasan por el origen, los planos de \mathbb{R}^3 que pasan por el origen y \mathbb{R}^3 .

Claramente, el vector nulo de cada subespacio vectorial coincide con el vector nulo de V . Es evidente que un subconjunto W de un espacio vectorial V que no contiene el vector nulo no es un subespacio. Más en general, es útil tener criterios que permitan determinar si un cierto subconjunto no vacío de un espacio vectorial es un subespacio.

2.2.1 Criterios para subespacios

Criterio 1. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Un subconjunto no vacío $W \subseteq V$ es un subespacio de V si y sólo si

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in W, \quad \forall \lambda \in \mathbb{K} : \quad \mathbf{v}_1 + \mathbf{v}_2 \in W, \quad \lambda \mathbf{v} \in W. \quad (2.13)$$

Demostración. Claramente, si W es un subespacio, entonces la condición (2.13) se cumple trivialmente.

Vice versa, asumimos que la condición (2.13) esté satisfecha. Como los vectores de W son también vectores de V , entonces satisfacen las propiedades asociativa y conmutativa de la suma. Además, siendo $0\mathbf{v} = \mathbf{0}$, entonces $\mathbf{0} \in W$. También, eligiendo $\lambda = -1$, tenemos que todo vector $\mathbf{v} \in W$ tiene su opuesto $-\mathbf{v} \in W$. Es inmediato comprobar que los cuatro axiomas relativos a la ley de composición externa siguen siendo válidos en W . En definitiva, los ocho axiomas de espacio vectorial están satisfechos. Por tanto $(W, +, \cdot)$ en sí mismo es un espacio vectorial sobre \mathbb{K} , es decir, es un subespacio de V . \square

Equivalentemente, podemos enunciar el

Criterio 2. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Un subconjunto no vacío $W \subseteq V$ es un subespacio de V si y sólo si

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in W, \quad \forall \lambda_1, \lambda_2 \in \mathbb{K} : \quad \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in W. \quad (2.14)$$

Demostración. Es suficiente demostrar que el criterio 2 es equivalente al criterio 1. Con este fin, observamos que si se cumple la condición (2.13), entonces $\forall \mathbf{v}_1, \mathbf{v}_2 \in W, \lambda_1, \lambda_2 \in \mathbb{K}$, tenemos que $\lambda_1 \mathbf{v}_1 \in W, \lambda_2 \mathbf{v}_2 \in W$; por tanto $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in W$. Vice versa, si la condición (2.14) está satisfecha, eligiendo $\lambda_1 = 1, \lambda_2 = 1$ tenemos que $\mathbf{v}_1 + \mathbf{v}_2 \in W$, y eligiendo $\lambda_2 = 0$, tenemos que $\lambda \mathbf{v} \in W \forall \lambda \in \mathbb{K}$. \square

Ejercicio 9. Sea V un espacio vectorial, $\mathbf{u} \in V$, $\mathbf{u} \neq \mathbf{0}$. Sea

$$W = \{\mathbf{v} = \lambda\mathbf{u}, \lambda \in \mathbb{K}\} .$$

Demostrar que W es un subespacio vectorial de V , llamado la **recta vectorial** generada por \mathbf{u} .

Resolución. Es suficiente comprobar que se cumple el Criterio 2. De hecho, sean $\mathbf{v}_1 = \lambda_1\mathbf{u}$, $\mathbf{v}_2 = \lambda_2\mathbf{u}$ dos vectores de W , y $\alpha_1, \alpha_2 \in \mathbb{K}$. Tenemos que $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 \in W$, porque

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \alpha_1\lambda_1\mathbf{u} + \alpha_2\lambda_2\mathbf{u} = (\alpha_1\lambda_1 + \alpha_2\lambda_2)\mathbf{u} = \gamma\mathbf{u}, \quad \gamma \in \mathbb{K}.$$

Teorema 20. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} y $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ un conjunto de p vectores de V . El conjunto $\text{lin } S$ de todas las posibles combinaciones lineales de los vectores de S con coeficientes en \mathbb{K} es un subespacio vectorial de V , denominado el **subespacio generado por S** .

Demostración. Sean $\mathbf{x}_1, \mathbf{x}_2$ dos vectores de $\text{lin } S$. Entonces, $\mathbf{x}_1 = \lambda_1\mathbf{v}_1 + \dots + \lambda_p\mathbf{v}_p$, $\mathbf{x}_2 = \mu_1\mathbf{v}_1 + \dots + \mu_p\mathbf{v}_p$. Es evidente que $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 \in \text{lin } S$, y por tanto, en virtud del Criterio 2, $\text{lin } S$ es un subespacio vectorial de V . \square

Definición 44. Los vectores $\mathbf{v}_1, \dots, \mathbf{v}_p$ se llamarán los **generadores** de $\text{lin } S$.

Ejemplo 10. En \mathbb{R}^3 , tenemos por ej. los subespacios

$$\text{lin}\{(1, 3, 0), (-1, 5, 7)\}, \quad \text{lin}\{(1, 2, 3)\} .$$

2.2.2 Suma e intersección de subespacios vectoriales

Definición 45. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y sean W_1, \dots, W_p p subespacios vectoriales de V . El conjunto de vectores

$$H = \{\mathbf{v} \in V \mid \exists \mathbf{v}_i \in W_i \ (i = 1, \dots, p), \mathbf{v} = \mathbf{v}_1 + \dots + \mathbf{v}_p\} \quad (2.15)$$

es dicho **suma de los subespacios** W_1, \dots, W_p y se denota con

$$H = W_1 + \dots + W_p .$$

Definición 46. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y sean W_1, \dots, W_p p subespacios vectoriales de V . El conjunto de vectores

$$G = \{\mathbf{v} \in W_i \ \forall i = 1, \dots, p\} \quad (2.16)$$

es llamado **intersección de los subespacios** W_1, \dots, W_p y se denota con

$$G = \bigcap_{i=1}^p W_i .$$

Teorema 21. *La intersección y la suma de subespacios vectoriales son subespacios vectoriales.*

Demostración. Sean p subespacios W_1, \dots, W_p de un espacio vectorial $(V, +, \cdot)$ sobre \mathbb{K} , y $G = \{\mathbf{v} \in W_i \forall i = 1, \dots, p\}$, $H = W_1 + \dots + W_p$ sean su intersección y suma, respectivamente.

(1) Sean \mathbf{v}_1 y $\mathbf{v}_2 \in G$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Claramente, $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2$ pertenece a cada uno de los subespacios W_i , $i = 1, \dots, p$ y por tanto pertenece a su intersección G . Como consecuencia del criterio 2, G es un subespacio vectorial de V .

(2) Sean $\mathbf{v} = \mathbf{v}_1 + \dots + \mathbf{v}_p$, $\mathbf{v}' = \mathbf{v}'_1 + \dots + \mathbf{v}'_p$, con $\mathbf{v}_i, \mathbf{v}'_i \in W_i$, $i = 1, \dots, p$. Por tanto, \mathbf{v} y $\mathbf{v}' \in H$. Sean $\lambda, \mu \in \mathbb{K}$. Tenemos que

$$\lambda \mathbf{v} + \mu \mathbf{v}' = \lambda(\mathbf{v}_1 + \dots + \mathbf{v}_p) + \mu(\mathbf{v}'_1 + \dots + \mathbf{v}'_p) = (\lambda \mathbf{v}_1 + \mu \mathbf{v}'_1) + \dots + (\lambda \mathbf{v}_p + \mu \mathbf{v}'_p) \in H.$$

Por tanto, H es un subespacio vectorial de V .

Definición 47. *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y sean W_1 y W_2 dos subespacios vectoriales de V . Si $W_1 \cup W_2 = \{\mathbf{0}\}$, es decir si W_1 y W_2 tienen en común sólo el vector nulo, se dice que W_1 y W_2 son **disjuntos**.*

Otro teorema interesante es el siguiente.

Teorema 22. *El subespacio suma es el subespacio generado por la unión de los subespacios W_1, \dots, W_p :*

$$H = \text{lin}(W_1 \cup \dots \cup W_p). \quad (2.17)$$

Observación 2.1. En general, la unión de dos subespacios de un espacio V no es un subespacio. Como contraejemplo, podemos considerar el caso de las dos rectas en \mathbb{R}^2 dada por el eje $X = \text{lin}\{(1, 0)\}$ y el eje $Y = \text{lin}\{(0, 1)\}$. Ahora bien, tenemos que el conjunto de vectores

$$\text{lin}\{(1, 0)\} \cup \text{lin}\{(0, 1)\}$$

no es un subespacio de \mathbb{R}^2 :

$$\mathbf{e}_1 = (1, 0) \in X, \mathbf{e}_2 = (0, 1) \in Y, \quad \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 = (\lambda_1, \lambda_2)$$

y en general el vector (λ_1, λ_2) pertenece a una recta por el origen del plano XY , pero no pertenece necesariamente ni a X , ni a Y . Sin embargo

$$X + Y = \text{lin}\{(1, 0), (0, 1)\} = \mathbb{R}^2$$

y por tanto es un subespacio vectorial.

2.3 Dependencia e independencia lineal de un conjunto (finito) de vectores

Definición 48. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . El conjunto $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ es linealmente independiente si la relación

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0} \quad (2.18)$$

implica

$$\alpha_1 = 0, \quad \alpha_2 = 0, \quad \dots, \quad \alpha_k = 0. \quad (2.19)$$

Si existe al menos un escalar $\alpha_i \neq 0$ tal que la relación (2.18) está satisfecha, los vectores $\mathbf{v}_1, \dots, \mathbf{v}_k$ se dirán linealmente dependientes.

2.3.1 Propiedades elementales

Proposición 10. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Valen las propiedades siguientes:

- (1) Un vector $\mathbf{v} \neq \mathbf{0}$ es linealmente independiente.
- (2) Dos vectores \mathbf{v}_1 y \mathbf{v}_2 son linealmente dependientes $\iff \mathbf{v}_1 = \lambda \mathbf{v}_2$
- (3) Un sistema de vectores S que contiene el vector nulo es linealmente dependiente
- (4) Si $\mathbf{v}_k = \alpha_1 \mathbf{v}_1 + \dots + \alpha_{k-1} \mathbf{v}_{k-1}$, entonces el conjunto $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ es linealmente dependiente.
- (5) Si un conjunto de vectores es linealmente dependiente, cualquier conjunto que lo contenga es también linealmente dependiente.
- (6) Si un conjunto de vectores es linealmente independiente, cualquier subconjunto de él es linealmente independiente.

Demostración.

(1) Si $\mathbf{v} \neq \mathbf{0}$, entonces la relación $\alpha \mathbf{v} = \mathbf{0}$ implica $\alpha = 0$, y por tanto \mathbf{v} es independiente.

(2) Si \mathbf{v}_1 y \mathbf{v}_2 son linealmente dependientes, entonces existe $\alpha_i \neq 0$, ($i = 1, 2$) tal que $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 = \mathbf{0}$. Supongamos $\alpha_1 \neq 0$. Entonces $\mathbf{v}_1 = \lambda \mathbf{v}_2$ con $\lambda = \frac{\alpha_2}{\alpha_1}$. Vice versa, si $\mathbf{v}_1 = \lambda \mathbf{v}_2$, entonces $\mathbf{v}_1 - \lambda \mathbf{v}_2 = \mathbf{0}$ y la relación (2.18) se cumple con $\alpha_1 = 1 \neq 0$.

(3) Si en un sistema de vectores uno de ellos, por ejemplo $\mathbf{v}_k = \mathbf{0}$, entonces la relación (2.18) se cumple con $\alpha_k \neq 0$ y entonces el sistema es dependiente.

(4) Es evidente que vale la relación $\alpha_1 \mathbf{v}_1 + \dots + \alpha_{k-1} \mathbf{v}_{k-1} - \mathbf{v}_k = \mathbf{0}$, y por tanto, siendo $\alpha_k = -1$, el sistema $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ es dependiente.

(5) Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ un sistema de vectores linealmente dependientes. Entonces por definición existe $\alpha_i \neq 0$ tal que la relación $\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$ está satisfecha. Por tanto, dado el conjunto de vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$, es evidente que la relación $\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n = \mathbf{0}$ se cumple con escalares

$$\alpha_1, \dots, \alpha_i \neq 0, \dots, \alpha_k, \alpha_{k+1} = 0, \dots, \alpha_n = 0$$

no todos nulos.

Proposición 11. Si $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ son n vectores linealmente dependientes de V , entonces por lo menos uno de estos vectores se puede expresar como combinación lineal de los demás vectores.

Demostración. Si $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ son linealmente dependientes, entonces por definición la relación

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

se cumple con algún coeficiente no nulo. Supongamos que, por ej., $\lambda_k \neq 0$. Entonces tenemos

$$\lambda_k \mathbf{v}_k = -\lambda_1 \mathbf{v}_1 - \dots - \lambda_{k-1} \mathbf{v}_{k-1} - \lambda_{k+1} \mathbf{v}_{k+1} - \dots - \lambda_n \mathbf{v}_n$$

que, siendo $\lambda_k \neq 0$, se puede escribir como

$$\mathbf{v}_k = -\frac{1}{\lambda_k} \sum_{\substack{i=1 \\ i \neq k}}^n \lambda_i \mathbf{v}_i$$

es decir, el vector \mathbf{v}_k es combinación lineal de los demás vectores. \square

Ejemplo 11. Los tres vectores $\mathbf{v}_1 = (2, 1, 2)$, $\mathbf{v}_2 = (7, -5/2, 3)$, $\mathbf{v}_3 = (-3, 1, -1)$ son linealmente independientes en \mathbb{R}^3 :

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \lambda_3 \mathbf{v}_3 = \mathbf{0} \iff \lambda_1 (2, 1, 2) + \lambda_2 (7, -5/2, 3) + \lambda_3 (-3, 1, -1) = (0, 0, 0)$$

$$\begin{cases} 2\lambda_1 + 7\lambda_2 - 3\lambda_3 = 0 \\ \lambda_1 - 5/2\lambda_2 + \lambda_3 = 0 \\ 2\lambda_1 + 3\lambda_2 - \lambda_3 = 0 \end{cases} \iff \begin{cases} 2\lambda_1 + 7\lambda_2 - 3\lambda_3 = 0 \\ 12\lambda_2 - 5\lambda_3 = 0 \\ \lambda_3 = 0 \end{cases} \iff \lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Observación 2.2. Mediante *operaciones elementales de filas* es posible determinar, dado un conjunto de generadores, los que son linealmente independientes entre sí. Por ejemplo, sea en \mathbb{R}^4 el subespacio generado por los vectores

$$W = \text{lin}\{(1, 3, 4, 1), (2, 6, 8, 2), (2, 5, 7, 2)\}.$$

Consideremos la matriz cuyas filas son las coordenadas de los vectores del sistema. Mediante operaciones elementales de fila, podemos reducir esta matriz hasta llegar a su forma escalonada reducida (o de Hermite) por filas. Tenemos:

$$\begin{pmatrix} 1 & 3 & 4 & 1 \\ 2 & 6 & 8 & 2 \\ 2 & 5 & 7 & 2 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 3 & 4 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Entonces, los vectores $(1, 3, 4, 1)$, $(2, 6, 8, 2)$ son linealmente independientes. Entonces, $W = \text{lin}\{(1, 0, 1, 1), (0, 1, 1, 0)\}$. Como veremos, los vectores $\{(1, 0, 1, 1), (0, 1, 1, 0)\}$ forman una base de W .

Definición 49. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , $S \subset V$ un conjunto no vacío de vectores de V . El rango de S es el máximo número de vectores linealmente independientes contenidos en S .

2.4 Bases de un espacio vectorial

La noción de **base** de un espacio vectorial es una de las más relevantes de toda la teoría.

Definición 50. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Una base de V es un conjunto de generadores de V linealmente independientes.

En otras palabras, \mathcal{B} es una base del espacio vectorial V si \mathcal{B} es un conjunto de vectores linealmente independientes y $\text{lin } \mathcal{B} = V$.

Teorema 23. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Un conjunto de vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ es una base de V si y sólo si cada $\mathbf{v} \in V$ se escribe de manera única como

$$\mathbf{v} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n, \quad x_1, \dots, x_n \in \mathbb{K} \quad (2.20)$$

Demostración. (\implies) Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base de V . Si la expresión (2.20) no fuese única, entonces existirían escalares y_1, \dots, y_n tales que, al mismo tiempo, $\mathbf{v} = y_1\mathbf{v}_1 + \dots + y_n\mathbf{v}_n$. Entonces

$$\mathbf{v} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n = y_1\mathbf{v}_1 + \dots + y_n\mathbf{v}_n$$

y por tanto

$$(x_1 - y_1)\mathbf{v}_1 + \dots + (x_n - y_n)\mathbf{v}_n = \mathbf{0}.$$

Como los vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ son independientes, la relación anterior implica inmediatamente $x_i = y_i$, $i = 1, \dots, n$.

(\impliedby) Supongamos que cada vector $\mathbf{v} \in V$ se escriba de forma única como $\mathbf{v} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$. Entonces, $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ son un sistema de generadores de V . Demostremos que además son linealmente independientes. Con tal fin, observamos que $0\mathbf{v}_1 + \dots + 0\mathbf{v}_n = \mathbf{0}$, y dado que esta expresión por hipótesis es única, entonces la relación $\alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n = \mathbf{0}$ implica $\alpha_1 = 0, \dots, \alpha_n = 0$ y los vectores son independientes. \square

Definición 51. La expresión $\mathbf{v} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$ se dirá *descomposición del vector \mathbf{v} respecto de la base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ de V* . Los coeficientes x_1, \dots, x_n , determinados de forma única, se llamarán *las componentes de \mathbf{v} respecto de la base dada*.

Lema 3. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Si $\mathbf{y} = c_1\mathbf{v}_1 + \dots + c_m\mathbf{v}_m$, con $c_1 \neq 0$, entonces

$$\text{lin}\{\mathbf{v}_1, \dots, \mathbf{v}_m\} = \text{lin}\{\mathbf{y}, \mathbf{v}_2, \dots, \mathbf{v}_m\}.$$

Demostración. Observamos que

$$\mathbf{v}_1 = \frac{1}{c_1}\mathbf{y} - \frac{c_2}{c_1}\mathbf{v}_2 - \dots - \frac{c_m}{c_1}\mathbf{v}_m.$$

Por tanto, cualquier combinación lineal de $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ se puede escribir como combinación lineal de $\{\mathbf{y}, \mathbf{v}_2, \dots, \mathbf{v}_m\}$. \square

Lema 4. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} y $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ un sistema de generadores de V . Si \mathbf{v}_i , $i \in \{1, \dots, p\}$, es combinación lineal de los demás vectores del sistema, entonces el conjunto de vectores que se obtiene eliminando \mathbf{v}_i de S sigue siendo un sistema de generadores de V .

Demostración. Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ un sistema de generadores de V . Por tanto, si $\mathbf{x} \in V$, se puede escribir como combinación lineal de los generadores:

$$\mathbf{x} = x_1\mathbf{v}_1 + \dots + x_i\mathbf{v}_i + \dots + x_p\mathbf{v}_p . \quad (2.21)$$

Por otro lado, tenemos que

$$\mathbf{v}_i = \alpha_1\mathbf{v}_1 + \dots + \alpha_{i-1}\mathbf{v}_{i-1} + \alpha_{i+1}\mathbf{v}_{i+1} + \dots + \alpha_p\mathbf{v}_p .$$

Por tanto, sustituyendo esta expresión de \mathbf{v}_i en la ecuación (2.21), tenemos:

$$\mathbf{x} = (x_1 + \alpha_1 x_i)\mathbf{v}_1 + \dots + (x_{i-1} + \alpha_{i-1} x_i)\mathbf{v}_{i-1} + (x_{i+1} + \alpha_{i+1} x_i)\mathbf{v}_{i+1} + \dots + (x_p + \alpha_p x_i)\mathbf{v}_p ,$$

lo que demuestra que $\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_p\}$ es un sistema de generadores de V . \square

Teorema 24 (Independencia lineal). Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Supongamos que $V = \text{lin}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$, con $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ y sea $S = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ un conjunto de vectores linealmente independientes de V . Entonces $k \leq m$.

Demostración. Cada vector de S , siendo también vector de V , es una combinación lineal de los generadores $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$. Por ejemplo,

$$\mathbf{y}_1 = a_1\mathbf{v}_1 + \dots + a_m\mathbf{v}_m .$$

Como S es un sistema de vectores independientes, necesariamente $\mathbf{y}_1 \neq \mathbf{0}$, así que no todos los coeficientes a_i pueden ser nulos. Sin pérdida de generalidad, supongamos que $a_1 \neq 0$. Aplicando el Lema 3, deducimos que

$$\text{lin}\{\mathbf{v}_1, \dots, \mathbf{v}_m\} = \text{lin}\{\mathbf{y}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}.$$

Por tanto, $\mathbf{y}_2 \in \text{lin}\{\mathbf{y}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$:

$$\mathbf{y}_2 = b_1\mathbf{y}_1 + b_2\mathbf{v}_2 + \dots + b_m\mathbf{v}_m$$

donde algunos de los b_2, \dots, b_m son distintos de cero, ya que en caso contrario $\mathbf{y}_2 = b_1\mathbf{y}_1$, y entonces \mathbf{y}_1 y \mathbf{y}_2 serían dependientes (Proposición 10, (2)). Supongamos que por ej. $b_2 \neq 0$, y aplicando de nuevo el Lema 3, obtenemos

$$\text{lin}\{\mathbf{v}_1, \dots, \mathbf{v}_m\} = \text{lin}\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{v}_m\}$$

así que $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{v}_m\}$ es un sistema de generadores de V .

Repetimos el razonamiento hasta agotar uno de los dos conjuntos. Pero si fuera $k > m$, entonces $\{\mathbf{y}_1, \dots, \mathbf{y}_m\}$ sería un sistema de generadores de V ; entonces los vectores $\mathbf{y}_{m+1}, \dots, \mathbf{y}_k$ se expresarían como combinación lineal de los vectores $\{\mathbf{y}_1, \dots, \mathbf{y}_m\}$, lo cual es imposible dado que $\{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ son linealmente independientes. \square

Teorema 25 (de la dimensión). *Si en un espacio vectorial $(V, +, \cdot)$ existe una base integrada por $n \in \mathbb{N} \setminus \{0\}$ vectores, todas las demás bases tienen el mismo número de vectores.*

Demostración. Sean $\mathcal{B}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ y $\mathcal{B}_2 = \{\mathbf{v}'_1, \dots, \mathbf{v}'_k\}$ dos bases de V . Aplicando el teorema de la independencia lineal (intercambiando el papel de \mathcal{B}_1 y \mathcal{B}_2), se obtiene

$$k \leq n \quad \text{y} \quad n \leq k \quad \implies \quad k = n .$$

□

El teorema anterior nos permite introducir la noción de dimensión de un espacio vectorial.

Definición 52. *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Si el espacio admite una base compuesta por $n \in \mathbb{N} \setminus \{0\}$ vectores, dicho número n se llamará la dimensión del espacio V . Si dicho entero n no existe, diremos que V tiene dimensión infinita.*

Por convenio, el subespacio $\{\mathbf{0}\}$ tiene dimensión nula.

Teorema 26 (Ampliación de la base). *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , $\dim V = n$, y sea $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base de V . Si $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$, $k \leq n$ es un sistema de vectores linealmente independientes de V , entonces existen vectores $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ de \mathcal{B} tales que el conjunto $\{\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ es una base de V .*

Demostración. Demostremos el teorema por recurrencia. Sea $k = 1$. El vector \mathbf{w}_1 , siendo por hipótesis linealmente independiente, es no nulo. Tenemos, respecto de la base \mathcal{B} , la descomposición

$$\mathbf{w}_1 = x_1 \mathbf{v}_1 + \dots + x_n \mathbf{v}_n , \tag{2.22}$$

donde al menos uno de los coeficientes es no nulo. Sea por ej. $x_1 \neq 0$. Entonces $\mathbf{v}_1 = \frac{1}{x_1}(\mathbf{w}_1 - x_2 \mathbf{v}_2 - \dots - x_n \mathbf{v}_n)$. Por tanto, cualquier combinación lineal de $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ es también combinación lineal de $\{\mathbf{w}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, que entonces son un sistema de generadores de V . Además, son linealmente independientes. En efecto, sea

$$\alpha_1 \mathbf{w}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_n \mathbf{v}_n = \mathbf{0} . \tag{2.23}$$

Si $\alpha_1 \neq 0$, tendríamos

$$\mathbf{w}_1 = -\frac{1}{\alpha_1}(\alpha_2 \mathbf{v}_2 + \dots + \alpha_n \mathbf{v}_n) \tag{2.24}$$

Sin embargo, esta relación está en contradicción con la ec. (2.22), porque en aquella ecuación $x_1 \neq 0$, mientras que en la (2.24) el coeficiente de \mathbf{v}_1 es nulo. Esto viola la unicidad de la descomposición de \mathbf{w}_1 en la base \mathcal{B} (Teorema (23)). Entonces, $\alpha_1 = 0$, y la ecuación (2.23) se reduce a

$$\alpha_2 \mathbf{v}_2 + \dots + \alpha_n \mathbf{v}_n = \mathbf{0} . \tag{2.25}$$

Como $\mathbf{v}_2, \dots, \mathbf{v}_n$ son independientes, entonces $\alpha_2 = 0, \dots, \alpha_n = 0$. En definitiva, hemos demostrado que en la relación (2.23) $\alpha_1 = 0, \dots, \alpha_n = 0$, por tanto los vectores $\{\mathbf{w}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ son linealmente independientes y forman una nueva base de V .

El teorema queda demostrado para $k = 1$. Ahora bien, supongamos que sea cierto para $k - 1$ vectores y demostremos que sigue siendo cierto para k vectores. Con tal fin, consideremos k vectores $\mathbf{w}_1, \dots, \mathbf{w}_k$ linealmente independientes. Los vectores $\mathbf{w}_1, \dots, \mathbf{w}_{k-1}$ también son independientes, y aplicando la hipótesis, deducimos que existen $n - k + 1$ vectores en la base \mathcal{B} , por ejemplo $\mathbf{v}_k, \dots, \mathbf{v}_n$ tales que el conjunto $\{\mathbf{w}_1, \dots, \mathbf{w}_{k-1}, \mathbf{v}_k, \dots, \mathbf{v}_n\}$ es una base de V . Por tanto, el vector \mathbf{w}_k será combinación lineal de estos vectores de base:

$$\mathbf{w}_k = \lambda_1 \mathbf{w}_1 + \dots + \lambda_{k-1} \mathbf{w}_{k-1} + \lambda_k \mathbf{v}_k + \dots + \lambda_n \mathbf{v}_n . \quad (2.26)$$

Por hipótesis, $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ son linealmente independientes; por tanto, necesariamente uno de los coeficientes $\lambda_k, \dots, \lambda_n$ es no nulo. Sea por ejemplo $\lambda_k \neq 0$. Deducimos que $\{\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ son un sistema de generadores de V (Lema 3). Además, con el mismo razonamiento visto por el caso $k = 1$ deducimos que estos vectores son independientes. Por tanto, forman una base de V y el teorema queda demostrado para k vectores. \square

En el caso $k = n$, obtenemos inmediatamente el siguiente

Corolario 1. *Si un espacio vectorial V tiene dimensión n , entonces cada sistema de n vectores linealmente independientes es una base de V .*

Teorema 27. *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , $\dim V = n$. Entonces, cada conjunto de vectores linealmente independientes posee como mucho n vectores.*

Demostración. Sean $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ p vectores linealmente independientes de V . Supongamos por absurdo $p > n$. Los vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ son independientes (siendo un subconjunto de un conjunto independiente). Por tanto, en base al Corolario 1, forman una base de V . Entonces los vectores $\mathbf{v}_{n+1}, \dots, \mathbf{v}_p$ se expresarían como combinación lineal de $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, lo que es en contradicción con la hipótesis de la independencia lineal de $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$. \square

También podemos observar que, dado un espacio vectorial $V \neq \{\mathbf{0}\}$, es posible construir una base eliminando vectores dependientes de un sistema de generadores de V .

Teorema 28. *En un espacio vectorial no nulo, de cada sistema de generadores finito se puede extraer una base.*

Demostración. Sea un sistema finito de generadores de un espacio $V \neq \{\mathbf{0}\}$. Si son independientes, ya son una base, y no hay que eliminar vectores. Si son dependientes, entonces al menos un vector del sistema será combinación lineal de los demás vectores. Eliminando este vector, el conjunto que queda sigue siendo un sistema de generadores, en virtud del Lema 4. Si el nuevo conjunto con un

vector menos es independiente, entonces es una base. Si todavía es dependiente, repetiremos el proceso hasta llegar a un sistema de generadores independientes, es decir, una base de V . \square

Teorema 29. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , $\dim V = n$, y sea $W \subset V$ un subespacio vectorial de V . Tenemos:

$$\dim W \leq \dim V . \quad (2.27)$$

Demostración. Sea $\mathbf{0} \neq \mathbf{w}_1 \in W$. Si $\text{lin}\{\mathbf{w}_1\} = W$, entonces $\dim W = 1 \leq n$. Si $\text{lin}\{\mathbf{w}_1\} \neq W$, entonces $\exists \mathbf{w}_2 \in W$, linealmente independiente de \mathbf{w}_1 . Si $\text{lin}\{\mathbf{w}_1, \mathbf{w}_2\} \neq W$, entonces $\exists \mathbf{w}_3 \in W$, linealmente independiente de $\mathbf{w}_1, \mathbf{w}_2$. Dado que la dimensión de V es finita, el proceso se acaba en un número finito de pasos. Dado que los vectores de W están generados por cualquier base de V , por el teorema de la independencia lineal, en cada paso, resulta que $\dim W \leq \dim V$. \square

Remark 1. En \mathbb{R}^n sea $\{\mathbf{e}_1 = (1, 0, \dots, 0), \mathbf{e}_2 = (0, 1, 0, \dots, 0), \dots, \mathbf{e}_n = (0, 0, \dots, 0, 1)\}$. Es inmediato demostrar que este conjunto es una base:

$$\lambda_1 \mathbf{e}_1 + \dots + \lambda_n \mathbf{e}_n = \mathbf{0} \implies (\lambda_1, \dots, \lambda_n) = (0, \dots, 0)$$

es decir, se trata de un conjunto de n vectores linealmente independientes. Además, son generadores del espacio, por que cualquier vector $\mathbf{x} \in \mathbb{R}^n$ se puede escribir como combinación lineal de ellos:

$$\mathbf{x} = (x_1, \dots, x_n) = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n .$$

Definición 53. Sea \mathbb{K} un cuerpo conmutativo, y el espacio vectorial $(\mathbb{K}^n, +, \cdot)$ sobre \mathbb{K} . Se denomina base canónica a la base siguiente:

$$\{(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)\}.$$

En la definición anterior, obviamente, 0 y 1 representan el elemento neutro de la suma $+$ y del producto \cdot definidos en el cuerpo conmutativo \mathbb{K} .

Ejemplo 12. La base canónica de \mathbb{R}^2 es $\{(1, 0), (0, 1)\}$, de \mathbb{R}^3 es $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$, etc.

Ejercicio 13. Demostrar que el conjunto

$$\left\{ A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

es una base del espacio vectorial $(M_{2 \times 2}(\mathbb{R}), +, \cdot)$ (base canónica).

Ejercicio 14. Sea \mathbb{R}^3 dotado de la base canónica $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Sean los subespacios

$$F_1 = \text{lin}\{(1, 2, 0), (0, 1, 1)\}, \quad F_2 = \text{lin}\{(0, 1, 2), (-1, 0, 1)\}$$

Determinar los subespacios $F_1 \cap F_2$ y $F_1 + F_2$.

Resolución.

Sea $\mathbf{w} \in F_1 \cap F_2$, $\mathbf{u}_1 = (1, 2, 0)$, $\mathbf{u}_2 = (0, 1, 1)$, $\mathbf{u}_3 = (0, 1, 2)$, $\mathbf{u}_4 = (-1, 0, 1)$.
Entonces

$$\begin{aligned} \mathbf{w} &= \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 = \lambda_3 \mathbf{u}_3 + \lambda_4 \mathbf{u}_4 \\ \iff \lambda_1(1, 2, 0) + \lambda_2(0, 1, 1) &= \lambda_3(0, 1, 2) + \lambda_4(-1, 0, 1) \end{aligned}$$

Tenemos el sistema

$$\begin{cases} \lambda_1 = -\lambda_4 \\ 2\lambda_1 + \lambda_2 = \lambda_3 \\ \lambda_2 = 2\lambda_3 + \lambda_4 \end{cases} \iff \begin{cases} \lambda_1 = -\lambda_4 \\ 2\lambda_1 + \lambda_2 = \lambda_3 \\ \lambda_1 + \lambda_2 = 2\lambda_3 \end{cases} \iff \begin{cases} \lambda_4 = -\lambda_3 \\ \lambda_2 = 3\lambda_3 \\ \lambda_1 = -\lambda_3 \end{cases}$$

Por tanto, llamando $\lambda_3 = \alpha$, tenemos que $\mathbf{w} = \alpha(-1, 1, 3)$, es decir

$$F_1 \cap F_2 = \text{lin}\{(-1, 1, 3)\}$$

es decir, es la recta vectorial generada por \mathbf{w} , así que $\dim F_1 \cap F_2 = 1$.
Como $F_1 + F_2 = \text{lin}(F_1 \cup F_2)$, tenemos

$$F_1 + F_2 = \text{lin}\{(1, 2, 0), (0, 1, 1), (0, 1, 2), (-1, 0, 1)\} .$$

Claramente, los cuatro vectores no pueden ser independientes. Es fácil comprobar que $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_4$ son linealmente independientes, y por tanto forman una base de $F_1 + F_2$, y que $\mathbf{u}_3 = -\mathbf{u}_1 + 3\mathbf{u}_2 - \mathbf{u}_4$. En definitiva,

$$F_1 + F_2 = \mathbb{R}^3 .$$

2.5 Suma directa

2.5.1 Definición

Definición 54. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y $W_1, \dots, W_p \subset V$ sean p subespacios vectoriales de V . Si $\mathbf{v}_i \in W_i$, y

$$\mathbf{v}_1 + \dots + \mathbf{v}_p = \mathbf{0} \implies \mathbf{v}_1 = \mathbf{0}, \mathbf{v}_2 = \mathbf{0}, \dots, \mathbf{v}_p = \mathbf{0} \quad (2.28)$$

diremos que la suma de los subespacios W_1, \dots, W_p es directa, y se denotará con

$$W_1 \oplus W_2 \dots \oplus W_p .$$

Teorema 30. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y $W_1, \dots, W_p \subset V$ sean p subespacios vectoriales de V . Si su suma es directa, entonces cualquier vector $\mathbf{v} \in W_1 \oplus \dots \oplus W_p$ se escribe de manera única de la forma

$$\mathbf{v} = \mathbf{v}_1 + \dots + \mathbf{v}_p, \quad \mathbf{v}_i \in W_i . \quad (2.29)$$

Demostración. Supongamos que la expresión (2.29) no sea única: entonces $\exists \mathbf{v}'_i \in W_i, i = 1, \dots, p$ tales que

$$\mathbf{v} = \mathbf{v}'_1 + \dots + \mathbf{v}'_p .$$

Obviamente

$$\mathbf{v}_1 + \dots + \mathbf{v}_p = \mathbf{v}'_1 + \dots + \mathbf{v}'_p ,$$

por tanto,

$$(\mathbf{v}_1 - \mathbf{v}'_1) + \dots + (\mathbf{v}_p - \mathbf{v}'_p) = \mathbf{0} .$$

Como $\mathbf{v}_i - \mathbf{v}'_i \in W_i, i = 1, \dots, p$, de la definición de suma directa se sigue $\mathbf{v}_i - \mathbf{v}'_i = \mathbf{0}$, es decir $\mathbf{v}_i = \mathbf{v}'_i$, por tanto la expresión (2.29) es única. \square

Proposición 12. *Sean W_1 y W_2 dos subespacios de V . La suma $W_1 + W_2$ es directa si y sólo si*

$$W_1 \cap W_2 = \{\mathbf{0}\} . \quad (2.30)$$

Demostración. (\implies) Supongamos que la suma sea directa. Si fuera $W_1 \cap W_2 \neq \{\mathbf{0}\}$, entonces existiría un vector $\mathbf{u} \neq \mathbf{0}$ en dicha intersección, es decir, $\mathbf{u} \in W_1$ y $\mathbf{u} \in W_2$. Sea $\mathbf{v} \in W_1 \oplus W_2$. Por tanto,

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2, \quad \mathbf{v}_1 \in W_1, \quad \mathbf{v}_2 \in W_2 . \quad (2.31)$$

Ahora bien, como

$$\mathbf{v} = (\mathbf{v}_1 + \mathbf{u}) + (\mathbf{v}_2 - \mathbf{u}),$$

con $\mathbf{v}_1 + \mathbf{u} \in W_1$ y $\mathbf{v}_2 - \mathbf{u} \in W_2$, evidentemente la expresión (2.31) no sería única.

(\impliedby) Supongamos $W_1 \cap W_2 = \{\mathbf{0}\}$ y sea

$$\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}, \quad \mathbf{v}_1 \in W_1, \mathbf{v}_2 \in W_2 .$$

Si $\mathbf{v}_1 \neq \mathbf{0}$, entonces $\mathbf{v}_2 = -\mathbf{v}_1$, entonces existiría un vector no nulo común a ambos subespacios, lo cual es absurdo. Por tanto, $\mathbf{v}_1 = \mathbf{0}$, y entonces $\mathbf{v}_2 = \mathbf{0}$. Por tanto la suma $W_1 + W_2$ es directa. \square

Definición 55. *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} . Dos subespacios W_1 y W_2 tales que*

$$V = W_1 \oplus W_2 \quad (2.32)$$

se dirán suplementarios.

2.6 Identidad de Grassmann

El teorema de Grassmann relaciona entre sí las dimensiones de la suma y de la intersección de dos subespacios.

Teorema 31 (Grassmann). *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y W_1, W_2 dos subespacios de V . Entonces*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) . \quad (2.33)$$

Demostración. Sea $r = \dim W_1$, $s = \dim W_2$, $t = \dim (W_1 \cap W_2)$. Sea $\mathcal{B}_{12} = \{\mathbf{v}_1, \dots, \mathbf{v}_t\}$ una base de $W_1 \cap W_2$. Por el teorema de ampliación de la base, a esta base se pueden añadir $r - t$ vectores $\mathbf{x}_1, \dots, \mathbf{x}_{r-t}$ obteniendo una base de W_1 :

$$\mathcal{B}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_t, \mathbf{x}_1, \dots, \mathbf{x}_{r-t}\}.$$

También, la base \mathcal{B}_{12} se puede ampliar con $s - t$ vectores $\mathbf{y}_1, \dots, \mathbf{y}_{s-t}$, para obtener una base de W_2 :

$$\mathcal{B}_2 = \{\mathbf{v}_1, \dots, \mathbf{v}_t, \mathbf{y}_1, \dots, \mathbf{y}_{s-t}\}.$$

Demostremos que los vectores

$$\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_t, \mathbf{x}_1, \dots, \mathbf{x}_{r-t}, \mathbf{y}_1, \dots, \mathbf{y}_{s-t}\}$$

son linealmente independientes. Por ello, introducimos los vectores

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_t \mathbf{v}_t, \quad \mathbf{x} = \beta_1 \mathbf{x}_1 + \dots + \beta_{r-t} \mathbf{x}_{r-t}, \quad \mathbf{y} = \gamma_1 \mathbf{y}_1 + \dots + \gamma_{s-t} \mathbf{y}_{s-t}.$$

Los vectores de \mathcal{B} son linealmente independientes si y sólo si

$$\mathbf{v} + \mathbf{x} + \mathbf{y} = \mathbf{0}. \tag{2.34}$$

Ahora bien, observamos que $\mathbf{v} \in W_1 \cap W_2$, $\mathbf{x} \in W_1$ pero $\mathbf{x} \notin W_2$ (salvo $\mathbf{x} = \mathbf{0}$), y $\mathbf{y} \in W_2$ pero $\mathbf{y} \notin W_1$ (salvo $\mathbf{y} = \mathbf{0}$).

Ahora bien, si

$$\mathbf{y} \neq \mathbf{0}, \quad \mathbf{y} = -\mathbf{v} - \mathbf{x} \in W_1$$

lo que sería en contradicción con el hecho que $\mathbf{y} \notin W_1$. Por tanto, deducimos que $\mathbf{y} = \mathbf{0}$. Del mismo modo se obtiene $\mathbf{x} = \mathbf{0}$ y entonces desde la relación (2.34) deducimos $\mathbf{v} = \mathbf{0}$.

Observamos que

$$\mathbf{v} = \mathbf{0} \implies \alpha_1 = \dots = \alpha_t = 0$$

porque los vectores \mathbf{v}_i son linealmente independientes. Del mismo modo deducimos

$$\mathbf{x} = \mathbf{0} \implies \beta_1 = \dots = \beta_{r-t} = 0$$

$$\mathbf{y} = \mathbf{0} \implies \gamma_1 = \dots = \gamma_{s-t} = 0$$

Entonces hemos demostrado que la relación (2.34), es decir

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_t \mathbf{v}_t + \beta_1 \mathbf{x}_1 + \dots + \beta_{r-t} \mathbf{x}_{r-t} + \gamma_1 \mathbf{y}_1 + \dots + \gamma_{s-t} \mathbf{y}_{s-t} = \mathbf{0}$$

implica

$$\alpha_1 = \dots = \alpha_t = \beta_1 = \dots = \beta_{r-t} = \gamma_1 = \dots = \gamma_{s-t} = 0.$$

Por tanto, \mathcal{B} es una base de $W_1 + W_2$ y

$$\dim(W_1 + W_2) = t + r - t + s - t = r + s - t = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

□

2.6.1 Otras propiedades de la suma directa de subespacios

Desde la identidad de Grassmann se obtiene directamente el siguiente

Corolario 2. *Sea el subespacio $W = W_1 \oplus W_2$. Entonces*

$$\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2 . \quad (2.35)$$

Demostración. La suma de dos subespacios es directa si y sólo si y sólo si $W_1 \cap W_2 = \{\mathbf{0}\}$. Como $\dim\{\mathbf{0}\} = 0$, el resultado sigue de la identidad de Grassmann. \square

Además, tenemos otra propiedad interesante.

Proposición 13. *Sea $(V, +, \cdot)$ un espacio vectorial sobre K . Entonces V es suma directa de sus subespacios W_1, \dots, W_k , y escribimos $V = W_1 \oplus \dots \oplus W_k$, si*

$$\begin{cases} V_1 + \dots + V_k = V \\ (V_2 + \dots + V_k) \cap V_1 = \{\mathbf{0}\} \\ (V_1 + V_3 \dots + V_k) \cap V_2 = \{\mathbf{0}\} \\ \vdots \\ (V_1 + \dots + V_{k-1}) \cap V_k = \{\mathbf{0}\} \end{cases}$$

Si $V = W_1 \oplus \dots \oplus W_k$, entonces $\dim V = \dim W_1 + \dots + \dim W_k$.

2.6.2 Bases adaptadas a una suma directa

Definición 56. *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , $\dim V = n$, y sea la descomposición*

$$V = W_1 \oplus \dots \oplus W_p \quad (2.36)$$

donde W_1, \dots, W_p son p subespacios vectoriales de V , con $n_1 = \dim W_1, \dots, n_p = \dim W_p$, y

$$n = n_1 + \dots + n_p .$$

Diremos que una base \mathcal{B} de V está adaptada a la descomposición (2.36) si los primeros n_1 vectores de \mathcal{B} forman una base de W_1 , los siguientes n_2 vectores forman una base de W_2 y así sucesivamente, hasta los últimos n_p vectores, que formarán una base de W_p .

2.7 Ecuaciones intrínsecas de un subespacio vectorial

Consideremos un sistema homogéneo $AX = 0$ de m ecuaciones con coeficientes en un cuerpo \mathbb{K} y n incógnitas. Claramente, cada solución es un vector de \mathbb{K}^n . Demostraremos que el conjunto de las soluciones del sistema es un subespacio vectorial.

Teorema 32. Sea el espacio vectorial \mathbb{K}^n y sea $AX = 0$ un sistema homogéneo de m ecuaciones y n incógnitas con coeficientes en \mathbb{K} . El conjunto formado por **todas** las soluciones del sistema es un subespacio vectorial de \mathbb{K}^n .

Demostración. Sean (y_1, \dots, y_n) y (z_1, \dots, z_n) dos soluciones del sistema considerado, y sean Y y Z las matrices columna integradas por los coeficientes de dichas soluciones. Claramente, tenemos que $AY = 0$ y $AZ = 0$. Por tanto $A(Y + Z) = AY + AZ = 0$, as que $(y_1 + z_1, \dots, y_n + z_n)$ también es solución. Por otro lado, $(\lambda y_1, \dots, \lambda y_n)$ es solución, dado que $A(\lambda Y) = \lambda AY = 0$. Por tanto, utilizando el Criterio 1, obtenemos que el conjunto de todas las soluciones del sistema $AX = 0$ es un subespacio vectorial de \mathbb{K}^n . \square

Se puede mostrar que la dimensión del subespacio de las soluciones es igual a $n - \text{rang}(A)$, es decir, se obtiene restando a la dimensión del espacio el número de “ligaduras” independientes entre las variables.

También es cierto que cada subespacio de un espacio vectorial V de dimensión finita puede interpretarse como el conjunto de las soluciones de un sendo sistema de ecuaciones cartesianas, llamadas ecuaciones **intrínsecas** (o implícitas) del subespacio.

La relación entre la dimensión del subespacio y el número de ecuaciones intrínsecas es muy directa. Precisamente, si V es un espacio vectorial de dimensión finita y U es un subespacio de V , el número m de ecuaciones cartesianas independientes necesarias para representar U (es decir, de ecuaciones intrínsecas) es

$$m = \dim V - \dim U . \quad (2.37)$$

Ejemplo 15. Sea el subespacio de R^3 determinado por las soluciones del sistema

$$x_1 + x_2 + x_3 = 0 .$$

Despejando una de las variables, tenemos

$$x_1 = -x_2 - x_3 .$$

Por tanto, poniendo $x_2 = \alpha$, $x_3 = \beta$, tenemos que $x_1 = -\alpha - \beta$. Las tres ecuaciones son llamadas **ecuaciones paramétricas** del subespacio. El conjunto de las soluciones tiene entonces la forma

$$U = \{(-\alpha - \beta, \alpha, \beta) \in \mathbb{R}^3, \quad \alpha, \beta \in \mathbb{R}\} .$$

Una base de U se obtiene inmediatamente:

$$(-\alpha - \beta, \alpha, \beta) = \alpha(-1, 1, 0) + \beta(-1, 0, 1)$$

Por tanto $U = \text{lin}\{(-1, 1, 0), (-1, 0, 1)\}$. Siendo estos vectores independientes, forman evidentemente una base de U .

Ejercicio 16. Sea el espacio vectorial \mathbb{R}^3 y el subespacio

$$W = \text{lin}\{(1, -1, 0), (1, 1, 0)\} .$$

Determinar sus ecuaciones intrínsecas.

Resolución. Como $\dim W = 2$, necesitamos una sola ecuación intrínseca. Si $\mathbf{w} = (x, y, z)$ es un vector genérico de W , entonces

$$(x, y, z) = \lambda(1, -1, 0) + \mu(1, 1, 0)$$

de donde obtenemos las ecuaciones paramétricas

$$\begin{cases} x = \lambda + \mu \\ y = -\lambda + \mu \\ z = 0 . \end{cases}$$

La ecuación intrínseca buscada es por tanto $z = 0$.

Ejercicio 17. Sea el espacio vectorial \mathbb{R}^4 y el subespacio

$$U = \text{lin}\{(1, 0, 1, 1), (0, 1, 1, 0)\} .$$

Determinar sus ecuaciones intrínsecas.

Resolución. Sea $\mathbf{x} = (x_1, x_2, x_3, x_4) \in U$. Tenemos

$$(x_1, x_2, x_3, x_4) = \lambda(1, 0, 1, 1) + \mu(0, 1, 1, 0) .$$

Obtenemos las ecuaciones paramétricas

$$\begin{cases} x_1 = \lambda \\ x_2 = \mu \\ x_3 = \lambda + \mu \\ x_4 = \lambda \end{cases} \iff \begin{cases} x_1 = x_4 \\ x_2 = \mu \\ x_3 = x_1 + x_2 . \end{cases}$$

Por tanto, las ecuaciones intrínsecas del subespacio son

$$\begin{cases} x_1 + x_2 - x_3 = 0 \\ x_1 - x_4 = 0 . \end{cases}$$

2.8 Cambio de base

Definición 57. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ y $\hat{\mathcal{B}} = \{\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_n\}$ dos bases distintas de V . La matriz P de cambio de la base antigua \mathcal{B} a la nueva base $\hat{\mathcal{B}}$ es la matriz cuyas columnas son integradas por las coordenadas de los vectores de la base nueva calculadas respecto de la base antigua.

Ejemplo 18. Sea el espacio vectorial $(\mathbb{R}^2, +, \cdot)$ sobre \mathbb{R} . Consideremos las bases

$$\mathcal{B} = \{\mathbf{v}_1 = (1, 0), \quad \mathbf{v}_2 = (0, 1)\}$$

y

$$\hat{\mathcal{B}} = \{\hat{\mathbf{v}}_1 = (2, 1), \quad \hat{\mathbf{v}}_2 = (3, 4)\}.$$

Tenemos por tanto

$$\begin{cases} \hat{\mathbf{v}}_1 = (2, 1) = 2\mathbf{v}_1 + \mathbf{v}_2 \\ \hat{\mathbf{v}}_2 = (3, 4) = 3\mathbf{v}_1 + 4\mathbf{v}_2 \end{cases} \quad \begin{cases} \mathbf{v}_1 = (1, 0) = 4/5\hat{\mathbf{v}}_1 - 1/5\hat{\mathbf{v}}_2 \\ \mathbf{v}_2 = (0, 1) = -3/5\hat{\mathbf{v}}_1 + 2/5\hat{\mathbf{v}}_2. \end{cases}$$

Por tanto, la matriz de cambio de la base \mathcal{B} a la base $\hat{\mathcal{B}}$ es

$$P = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}.$$

La matriz P es invertible, y su inversa

$$P^{-1} = \begin{pmatrix} 4/5 & -3/5 \\ -1/5 & 2/5 \end{pmatrix}$$

representa la matriz de cambio de la base $\hat{\mathcal{B}}$ a la matriz \mathcal{B} . En general, vale la siguiente

Proposición 14. *Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ y $\hat{\mathcal{B}} = \{\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_n\}$ dos bases distintas de V . La matriz P de cambio de la base antigua \mathcal{B} a la nueva base $\hat{\mathcal{B}}$ es invertible.*

Demostración. Es suficiente observar que las columnas de la matriz P son linealmente independientes, dado que representan las componentes de los vectores de una base (respecto de otra); por tanto, el rango de la matriz es máximo, es decir, igual a n y la matriz es invertible. \square

Supongamos que V sea un espacio vectorial sobre \mathbb{K} , de dimensión n . Si P es la matriz de cambio de la base \mathcal{B} a la base $\hat{\mathcal{B}}$, valen las siguientes **ecuaciones de cambio de base**:

$$\hat{\mathbf{v}}_j = \sum_{i=1}^n P_{ij} \mathbf{v}_i, \quad j = 1, \dots, n. \quad (2.38)$$

Nos preguntamos ahora: ¿cómo cambian las componentes de un vector $\mathbf{x} \in V$, cambiando base en el espacio? Representemos el vector \mathbf{x} en dos bases distintas:

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{v}_i, \quad \mathbf{x} = \sum_{j=1}^n \hat{x}_j \hat{\mathbf{v}}_j$$

Determinemos ahora la relación entre las componentes $\{\hat{x}_1, \dots, \hat{x}_n\}$ y $\{x_1, \dots, x_n\}$. Tenemos

$$\mathbf{x} = \sum_{j=1}^n \hat{x}_j \hat{\mathbf{v}}_j = \sum_{j=1}^n \hat{x}_j \left(\sum_{i=1}^n P_{ij} \mathbf{v}_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n P_{ij} \hat{x}_j \right) \mathbf{v}_i = \sum_{i=1}^n x_i \mathbf{v}_i$$

Por tanto, deducimos

$$\hat{x}_j = \left(\sum_{i=1}^n P_{ij} \hat{x}_j \right)$$

En notación matricial, sean $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\hat{X} = \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_n \end{pmatrix}$ los vectores columna de las componentes de X en las dos bases. Evidentemente,

$$X = P\hat{X} \tag{2.39}$$

o, equivalentemente

$$\hat{X} = P^{-1}X . \tag{2.40}$$

Chapter 3

Aplicaciones Lineales

3.1 Definiciones

Definición 58. Sean V y W dos espacios vectoriales sobre el mismo cuerpo conmutativo \mathbb{K} . Una aplicación

$$f : V \rightarrow W$$

que satisface las condiciones

$$(I) \quad \forall \mathbf{v}, \mathbf{v}' \in V; \quad f(\mathbf{v} + \mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}') \quad (3.1)$$

$$(II) \quad \forall \mathbf{v} \in V, \quad \forall \lambda \in \mathbb{K}, \quad f(\lambda \mathbf{v}) = \lambda f(\mathbf{v}) \quad (3.2)$$

se dirá que es una **aplicación lineal** o un **homomorfismo** de V en W .

Sea $f : V \rightarrow W$ una aplicación lineal. En la literatura matemática se usa comúnmente la terminología siguiente:

Si f es inyectiva, f se dirá **monomorfismo**.

Si f es sobreyectiva, f se dirá **epimorfismo**.

Si f es biyectiva, f se dirá **isomorfismo**.

Además, un homomorfismo $f : V \rightarrow V$ se dirá **endomorfismo** de V ; en particular, un endomorfismo que es también biyectivo (isomorfismo) es un **automorfismo** de V .

Definición 59. El conjunto de todas las aplicaciones lineales de V en W se denota con el símbolo $L(V, W)$ o también $\text{Hom}(V, W)$. El conjunto de los endomorfismos de V se denota con $L(V)$ o con $\text{End}(V)$.

Ejemplo 3.1. Sea $(V, +, \cdot)$ un espacio vectorial sobre \mathbb{K} , y $a \in \mathbb{K}$. Introducimos la aplicación $h_a : V \rightarrow V$, así definida:

$$h_a(\mathbf{v}) = a\mathbf{v} \quad \forall \mathbf{v} \in V.$$

Verifiquemos que h_a es una aplicación lineal:

$$\forall \mathbf{v}, \mathbf{v}' \in V; \quad h_a(\mathbf{v} + \mathbf{v}') = a(\mathbf{v} + \mathbf{v}') = a\mathbf{v} + a\mathbf{v}' = h_a(\mathbf{v}) + h_a(\mathbf{v}')$$

$$\forall \lambda \in \mathbb{K} : h_a(\lambda \mathbf{v}) = a\lambda \mathbf{v} = \lambda a \mathbf{v} = \lambda h_a(\mathbf{v})$$

Entonces h_a es lineal. Además, si $a \neq 0$, h_a es biyectiva: $\forall \mathbf{v} \in V$, el vector $\mathbf{u} = a^{-1}\mathbf{v}$ es el único vector de V tal que $h_a(\mathbf{u}) = \mathbf{v}$.

Entonces, si $a \neq 0$, h_a es un automorfismo de V , llamado **homotecia** de V de constante a .

Ejercicio 3.2. Mostrar que la aplicación $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definida por

$$f(x, y, z) = (x + y, y + z)$$

es lineal.

Resolución. Utilicemos la definición de aplicación lineal.

I) Sean $\mathbf{v}_1 = (x_1, y_1, z_1)$, $\mathbf{v}_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$. Observamos que

$$\begin{aligned} f(\mathbf{v}_1 + \mathbf{v}_2) &= f[(x_1, y_1, z_1) + (x_2, y_2, z_2)] = f(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (x_1 + x_2 + y_1 + y_2, y_1 + y_2 + z_1 + z_2) = (x_1 + y_1, y_1 + z_1) + (x_2 + y_2, y_2 + z_2) \\ &= f(x_1, y_1, z_1) + f(x_2, y_2, z_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) . \end{aligned}$$

II) Sea $\lambda \in \mathbb{K}$.

$$\begin{aligned} f(\lambda \mathbf{v}) &= f[\lambda(x_1, y_1, z_1)] = f(\lambda x_1, \lambda y_1, \lambda z_1) = (\lambda x_1 + \lambda y_1, \lambda y_1 + \lambda z_1) \\ &= \lambda(x_1 + y_1, y_1 + z_1) = \lambda f(x_1, y_1, z_1) = \lambda f(\mathbf{v}) . \end{aligned}$$

3.2 Criterio para aplicaciones lineales

Teorema 33. Sean V, W espacios vectoriales sobre \mathbb{K} . La aplicación $f : V \rightarrow W$ es lineal si y sólo si

$$\forall \mathbf{v}, \mathbf{v}' \in V, \quad \forall \lambda, \lambda' \in \mathbb{K} : \quad f(\lambda \mathbf{v} + \lambda' \mathbf{v}') = \lambda f(\mathbf{v}) + \lambda' f(\mathbf{v}')$$

Proof. \implies) El vector $\lambda \mathbf{v} + \lambda' \mathbf{v}' \in V$. Entonces

$$f(\lambda \mathbf{v} + \lambda' \mathbf{v}') \stackrel{(I)}{=} f(\lambda \mathbf{v}) + f(\lambda' \mathbf{v}') \stackrel{(II)}{=} \lambda f(\mathbf{v}) + \lambda' f(\mathbf{v}')$$

\implies Elegimos $\lambda = \lambda' = 1$; entonces se cumple la condición (I). Si elegimos $\lambda = 1$, $\lambda' = 0$, se cumple la condición (II). \square

Proposición 15. Sean V, W dos espacios vectoriales sobre \mathbb{K} , $f \in L(V, W)$. Entonces

$$f(\mathbf{0}_V) = \mathbf{0}_W . \tag{3.3}$$

Proof. Si $\mathbf{v} \in V$, tenemos

$$f(\mathbf{v} + \mathbf{0}_V) = f(\mathbf{v}) = f(\mathbf{v}) + f(\mathbf{0}_V) \iff f(\mathbf{v}) - f(\mathbf{v}) = \mathbf{0}_W = f(\mathbf{0}_V) .$$

\square

Ejercicio 3.3. Sea $V = \mathcal{M}_n(\mathbb{K})$. Demostrar que la aplicación

$$f : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathcal{M}_n(\mathbb{K})$$

definida como

$$f(A) = A^T$$

es lineal.

Resolución. Sean $A_1, A_2 \in \mathcal{M}_n(\mathbb{K})$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Tenemos

$$f(\lambda_1 A_1 + \lambda_2 A_2) = (\lambda_1 A_1 + \lambda_2 A_2)^T = (\lambda_1 A_1)^T + (\lambda_2 A_2)^T = \lambda_1 (A_1)^T + \lambda_2 (A_2)^T,$$

utilizando la definición de transposición de una matriz. Por tanto, en virtud del Teorema 33, deducimos que la aplicación $f \in L(\mathcal{M}_n(\mathbb{K}))$.

Ejercicio 3.4. Sea $V = \mathbb{K}[x]$, el espacio vectorial de los polinomios en la variable x , con coeficientes en el cuerpo conmutativo \mathbb{K} . Sea

$$D : \mathbb{K}[x] \rightarrow \mathbb{K}[x], \quad Dp(x) = p'(x)$$

la aplicación que a un polinomio $p(x) \in \mathbb{K}[x]$ asocia su derivada primera $p'(x)$ respecto de la variable x . Demostrar que D es una aplicación lineal.

Resolución. Como es bien conocido, si $p_1(x)$ y $p_2(x)$ son dos polinomios,

$$D(p_1(x) + p_2(x)) = Dp_1(x) + Dp_2(x)$$

Además, si $\lambda \in \mathbb{K}$, $D(\lambda p(x)) = \lambda Dp(x)$, $\forall p(x) \in \mathbb{K}[x]$. Por tanto, $D \in L(\mathbb{K}[x])$.

Definición 60. Sean $f, g \in L(V, W)$ y $k \in \mathbb{K}$. Llamaremos suma de las aplicaciones lineales f y g a la aplicación, denotada con $f + g$, definida por

$$(f + g)(\mathbf{v}) := f(\mathbf{v}) + g(\mathbf{v}), \quad \forall \mathbf{v} \in V.$$

y producto del escalar k por la aplicación f a la aplicación

$$(kf)(\mathbf{v}) := kf(\mathbf{v}), \quad \forall \mathbf{v} \in V.$$

Teorema 34. El conjunto $L(V, W)$, dotado de las operaciones de suma de aplicaciones lineales y producto de un escalar por una aplicación lineal, es un espacio vectorial sobre \mathbb{K} .

Proof. Demostremos ante todo que la suma de dos aplicaciones lineales y el producto de un escalar por una aplicación lineal son también aplicaciones lineales. Pongamos $s = f + g$. Entonces, $\forall \mathbf{v}, \mathbf{v}' \in V, \lambda, \lambda' \in \mathbb{K}$:

$$\begin{aligned} s(\lambda \mathbf{v} + \lambda' \mathbf{v}') &= (f + g)(\lambda \mathbf{v} + \lambda' \mathbf{v}') = f(\lambda \mathbf{v} + \lambda' \mathbf{v}') + g(\lambda \mathbf{v} + \lambda' \mathbf{v}') \\ &= \lambda f(\mathbf{v}) + \lambda' f(\mathbf{v}') + \lambda g(\mathbf{v}) + \lambda' g(\mathbf{v}') \\ &= \lambda(f + g)(\mathbf{v}) + \lambda'(f + g)(\mathbf{v}') = \lambda s(\mathbf{v}) + \lambda' s(\mathbf{v}'). \end{aligned}$$

Entonces, $f + g \in L(V, W)$. Análogamente, si $p = kf$, tenemos

$$\begin{aligned} p(\lambda \mathbf{v} + \lambda' \mathbf{v}') &= k(\lambda \mathbf{v} + \lambda' \mathbf{v}') = k\lambda f(\mathbf{v}) + k\lambda' f(\mathbf{v}') \\ &= \lambda kf(\mathbf{v}) + \lambda' kf(\mathbf{v}') = \lambda p(\mathbf{v}) + \lambda' p(\mathbf{v}') . \end{aligned}$$

Por tanto, $kf \in L(V, W)$. Ahora bien, la suma de aplicaciones $f + g$ es evidentemente asociativa y conmutativa. La aplicación 0 , definida por la relación

$$0(\mathbf{v}) = \mathbf{0}_W \quad \forall \mathbf{v} \in V$$

es el elemento neutro de la suma en $L(V, W)$. El simétrico (o opuesto) de f es $(-1)f$. Tenemos:

$$\forall \mathbf{v} \in V : f(\mathbf{v}) + (-1)f(\mathbf{v}) = f(\mathbf{v} - \mathbf{v}) = f(\mathbf{0}_V) = \mathbf{0}_W = 0(\mathbf{v}) .$$

Entonces, $(-1)f = -f$. Además, se verifica fácilmente que el producto kf cumple con los axiomas 5 – 8 de la definición de espacio vectorial. \square

El resultado siguiente establece que la imagen de un subespacio de V es un subespacio de W .

Proposición 16. *Sea $f \in L(V, W)$. Si $E \subseteq V$ es un subespacio de V , entonces $f(E)$ es un subespacio de W .*

Proof. Sean $\mathbf{y}_1 = f(\mathbf{v}_1)$ e $\mathbf{y}_2 = f(\mathbf{v}_2)$ dos vectores de $f(E)$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Entonces $\mathbf{v}_1, \mathbf{v}_2 \in E$, y siendo E un subespacio, $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in E$. Tenemos que

$$\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2 = \lambda_1 f(\mathbf{v}_1) + \lambda_2 f(\mathbf{v}_2) = f(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2) \in f(E)$$

Por tanto, utilizando el criterio de subespacios, $f(E)$ es un subespacio vectorial de W . \square

3.3 Núcleo e imagen de una aplicación lineal

Definición 61. *Sea $f \in L(V, W)$. El conjunto de vectores de V cuya imagen es el vector nulo de W se llama **núcleo** de f y se denota con $\ker f$:*

$$\ker f = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}_W\} .$$

Observación 3.1. El conjunto $\ker f \neq \emptyset$, porque por lo menos el vector nulo pertenece a él:

$$\forall f \in L(V, W) \quad f(\mathbf{0}_V) = \mathbf{0}_W \implies \mathbf{0}_V \in \ker f .$$

Definición 62. *Sea $f \in L(V, W)$. Se define imagen de f , y se denota con $\text{im } f$ al conjunto*

$$\text{im } f = \{\mathbf{y} \in W \mid \exists \mathbf{v} \in V : f(\mathbf{v}) = \mathbf{y}\} .$$

Tanto $\ker f$ como $\text{im } f$ son subespacios vectoriales, como veremos a continuación.

Teorema 35. *Sia $f \in L(V, W)$. Entonces $\ker f$ un subespacio vectorial de V .*

Proof. Sean $\mathbf{v}, \mathbf{v}' \in \ker f$, $\lambda, \lambda' \in \mathbb{K}$. Entonces, por definición,

$$f(\mathbf{v}) = \mathbf{0}_W, \quad f(\mathbf{v}') = \mathbf{0}_W .$$

Además,

$$f(\lambda\mathbf{v} + \lambda'\mathbf{v}') = \lambda f(\mathbf{v}) + \lambda' f(\mathbf{v}') = \lambda\mathbf{0}_W + \lambda'\mathbf{0}_W = \mathbf{0}_W \implies \lambda\mathbf{v} + \lambda'\mathbf{v}' \in \ker f .$$

□

Teorema 36. *Sia $f \in L(V, W)$. Entonces $\operatorname{im} f$ es un subespacio vectorial de W .*

Proof. Por definición, $\operatorname{im} f = f(V)$. Por tanto, es suficiente aplicar la Proposición 16 al caso $E = V$. □

Definición 63. *Llamaremos rango de la aplicación lineal f a la dimensión del subespacio $\operatorname{im} f$:*

$$\operatorname{rang} f = \dim(\operatorname{im} f) .$$

Proposición 17. *Sea $f \in L(V, W)$. Si $W_1 \subset W$ es un subespacio de W , entonces $f^{-1}(W_1)$ es un subespacio de V .*

Proof. Sean \mathbf{x}_1 y $\mathbf{x}_2 \in f^{-1}(W_1)$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Entonces, existen \mathbf{y}_1 e \mathbf{y}_2 tales que

$$\mathbf{x}_1 = f^{-1}(\mathbf{y}_1), \quad \mathbf{x}_2 = f^{-1}(\mathbf{y}_2), \quad \mathbf{y}_1, \mathbf{y}_2 \in W_1 .$$

Aplicando el criterio de subespacios, demostramos que $\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 \in f^{-1}(W_1)$. Efectivamente, resulta que

$$\begin{aligned} \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 \in f^{-1}(W_1) &\iff \exists \mathbf{y} \in W_1 \text{ t.q. } \mathbf{y} = f(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2) = \lambda_1f(\mathbf{x}_1) + \lambda_2f(\mathbf{x}_2) \\ &= \lambda_1(f(f^{-1}(\mathbf{y}_1))) + \lambda_2(f(f^{-1}(\mathbf{y}_2))) = \lambda_1\mathbf{y}_1 + \lambda_2\mathbf{y}_2 \end{aligned}$$

lo que es cierto, siendo W_1 un subespacio de W . □

Observación 3.2. Sean V, W espacios vectoriales sobre \mathbb{K} , $\dim V = n$, y $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base de V . Sea $f \in L(V, W)$. La aplicación lineal f es unívocamente determinada cuando sean conocidos los vectores $f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)$. En efecto, si $\mathbf{v} \in V$, $\mathbf{v} = \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n$. Por tanto,

$$f(\mathbf{v}) = \alpha_1f(\mathbf{v}_1) + \dots + \alpha_nf(\mathbf{v}_n) .$$

En otras palabras, conocemos la imagen de cualquier vector $\mathbf{v} \in V$ si conocemos las imágenes de los vectores de una base de V .

Proposición 18. *Sea $f \in L(V, W)$ y $V_1 \subseteq V$ un subespacio de V . Entonces*

$$\dim f(V_1) \leq \dim V_1 . \tag{3.4}$$

En particular, si $V_1 = V$, siendo $f(V) = \operatorname{im} f$, entonces

$$\dim(\operatorname{im} f) \leq \dim V . \tag{3.5}$$

Proof. Sea $\dim V_1 = k$ y sea $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ una base de V_1 . Entonces, si $\mathbf{v} \in V_1$, tenemos que $\mathbf{v} = \alpha_1 \mathbf{e}_1 + \dots + \alpha_k \mathbf{e}_k$, y $f(\mathbf{v}) = \alpha_1 f(\mathbf{e}_1) + \dots + \alpha_k f(\mathbf{e}_k)$. Entonces $\{f(\mathbf{e}_1), \dots, f(\mathbf{e}_k)\}$ es un conjunto de generadores de $f(V_1)$, y

$$\dim f(V_1) \leq k = \dim V_1 .$$

Aplicando la desigualdad al caso $V_1 = V$, obtenemos la relación (3.5). \square

3.4 Composición de aplicaciones lineales

Teorema 37. Sean V, W y U tres espacios vectoriales sobre \mathbb{K} . Si

$$f : V \rightarrow W, \quad g : W \rightarrow U$$

son aplicaciones lineales, entonces $g \circ f : V \rightarrow U$ es lineal.

Proof. Sean $\mathbf{x}_1, \mathbf{x}_2 \in V$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Resulta que

$$\begin{aligned} (g \circ f)(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) &= g(f(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2)) = g(\lambda_1 f(\mathbf{x}_1) + \lambda_2 f(\mathbf{x}_2)) \\ &= \lambda_1 g(f(\mathbf{x}_1)) + \lambda_2 g(f(\mathbf{x}_2)) = \lambda_1 (g \circ f)(\mathbf{x}_1) + \lambda_2 (g \circ f)(\mathbf{x}_2) . \end{aligned}$$

Por tanto, $g \circ f \in L(V, U)$. \square

Como consecuencia inmediata del teorema anterior, deducimos que las derivadas de orden superior $D^2 = D \circ D, \dots, D^n = \underbrace{D \circ \dots \circ D}_n$ son también aplicaciones

lineales en los espacios vectoriales en que estén definidas (por ejemplo, el espacio $\mathbb{K}[x]$ del ejemplo anterior, o el espacio $C^n(\mathbb{R})$ de las funciones reales derivables n veces con continuidad).

Entonces

$$L = \lambda_1 D^n + \lambda_2 D^{n-1} + \dots + \lambda_n D \tag{3.6}$$

es una aplicación lineal en $C^n(\mathbb{R})$. Este objeto se denomina *operador lineal* en $C^n(\mathbb{R})$.

Proposición 19. La composición de un isomorfismo de V en W y de un isomorfismo de W en U es un isomorfismo de V en U . La composición de dos endomorfismos de V es otro endomorfismo de V . La composición de dos automorfismos de V es otro automorfismo de V .

Proof. Es suficiente tener en cuenta el Teorema 37 y el hecho que la composición de aplicaciones biyectivas es biyectiva. \square

Las proposiciones siguientes son simples consecuencias de la discusión anterior.

Proposición 20. Sean V, W, U tres espacios vectoriales sobre \mathbb{K} .

1) Si $f, g \in L(V, W)$ y $h \in L(W, U)$, entonces

$$h \circ (f + g) = h \circ f + h \circ g .$$

2) si $h \in L(V, W)$, y $f, g \in L(W, U)$, entonces

$$(f + g) \circ h = f \circ h + g \circ h .$$

Proposición 21. Sea $f : V \rightarrow W$ un isomorfismo. Entonces $f^{-1} : W \rightarrow V$ es también un isomorfismo.

Demostración. Sean $\mathbf{y}_1, \mathbf{y}_2 \in W$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Sea $\mathbf{x}_1 = f^{-1}(\mathbf{y}_1)$, $\mathbf{x}_2 = f^{-1}(\mathbf{y}_2)$; tenemos obviamente

$$\mathbf{y}_1 = f(\mathbf{x}_1), \quad \mathbf{y}_2 = f(\mathbf{x}_2) .$$

Por tanto,

$$\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2 = \lambda_1 f(\mathbf{x}_1) + \lambda_2 f(\mathbf{x}_2) = f(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2)$$

es decir,

$$f^{-1}(\lambda_1 \mathbf{y}_1 + \lambda_2 \mathbf{y}_2) = \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 = \lambda_1 f^{-1}(\mathbf{y}_1) + \lambda_2 f^{-1}(\mathbf{y}_2)$$

lo que demuestra que f^{-1} es lineal y biyectiva, o sea un isomorfismo. \square

3.5 El teorema de las dimensiones

El siguiente teorema, conocido como teorema de las dimensiones, representa uno de los resultados más relevantes de la teoría de las aplicaciones lineales.

Teorema 38. Sean V, W dos espacios vectoriales sobre \mathbb{K} y sea $f \in L(V, W)$. Entonces

$$\dim(\ker f) + \dim(\operatorname{im} f) = \dim V . \quad (3.7)$$

Demostración. Sea $\dim V = n$, y sea $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ una base de $\ker f$. Añadimos a esta base sendos vectores $\mathbf{v}_{k+1}, \dots, \mathbf{v}_{k+r}$, con $k+r = n$, hasta completar una base de V . Demostremos que $\{f(\mathbf{v}_{k+1}), \dots, f(\mathbf{v}_{k+r})\}$ es una base de $\operatorname{im} f$. Efectivamente, los vectores $\{f(\mathbf{v}_{k+1}), \dots, f(\mathbf{v}_{k+r})\}$ son un conjunto de generadores de $\operatorname{im} f$. En efecto, observamos que si $\mathbf{x} \in V$, tenemos

$$\mathbf{x} = c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k + c_{k+1} \mathbf{v}_{k+1} + \dots + c_{k+r} \mathbf{v}_{k+r} .$$

Entonces

$$f(\mathbf{x}) = c_1 f(\mathbf{v}_1) + \dots + c_k f(\mathbf{v}_k) + c_{k+1} f(\mathbf{v}_{k+1}) + \dots + c_{k+r} f(\mathbf{v}_{k+r}) ,$$

y teniendo en cuenta que los vectores $\mathbf{v}_1, \dots, \mathbf{v}_k \in \ker f$, la expresión anterior se reduce a

$$f(\mathbf{x}) = c_{k+1} f(\mathbf{v}_{k+1}) + \dots + c_{k+r} f(\mathbf{v}_{k+r}) .$$

Además, los vectores $\{f(\mathbf{v}_{k+1}), \dots, f(\mathbf{v}_{k+r})\}$ son linealmente independientes. Para demostrarlo, observamos que la relación

$$\alpha_{k+1} f(\mathbf{v}_{k+1}) + \dots + \alpha_{k+r} f(\mathbf{v}_{k+r}) = f(\alpha_{k+1} \mathbf{v}_{k+1} + \dots + \alpha_{k+r} \mathbf{v}_{k+r}) = \mathbf{0}$$

implica que $\alpha_{k+1}\mathbf{v}_{k+1} + \dots + \alpha_{k+r}\mathbf{v}_{k+r} \in \ker f$. Por tanto, este último vector se puede escribir como combinación lineal de los vectores de la base $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ de $\ker f$:

$$\alpha_{k+1}\mathbf{v}_{k+1} + \dots + \alpha_{k+r}\mathbf{v}_{k+r} = \alpha_1\mathbf{v}_1 + \dots + \alpha_k\mathbf{v}_k.$$

Como los vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_{k+r}\}$ son una base de V , sigue inmediatamente que $\alpha_1 = \dots = \alpha_{k+r} = 0$. En definitiva, hemos demostrado que los vectores $f(\mathbf{v}_{k+1}), \dots, f(\mathbf{v}_{k+r})$ son un conjunto de generadores linealmente independientes de $\text{im } f$, es decir son una base de este subespacio. Y como $\dim(\text{im } f) = r = n - k$, siendo $\dim(\ker f) = k$, obtenemos

$$\dim(\ker f) + \dim(\text{im } f) = \dim V .$$

□

3.6 Teoremas de equivalencia

3.6.1 Primer teorema de equivalencia

Teorema 39. Sean V y W dos espacios vectoriales sobre \mathbb{K} . Si $f \in L(V, W)$, las propiedades siguientes son equivalentes:

- (1) $\ker f = \{\mathbf{0}_V\}$.
- (2) La aplicación f es inyectiva.
- (3) La imagen de un sistema de vectores linealmente independientes de V es un sistema de vectores linealmente independientes de W .
- (4) Si $V' \subseteq V$ es un subespacio, $\dim f(V') = \dim V'$.

Demostración. (1) \implies (2). Sea $\ker f = \{\mathbf{0}_V\}$. Tenemos que demostrar que

$$f(\mathbf{v}) = f(\mathbf{v}') \implies \mathbf{v} = \mathbf{v}'.$$

En efecto, si $f(\mathbf{v}) = f(\mathbf{v}')$, entonces $f(\mathbf{v}) - f(\mathbf{v}') = f(\mathbf{v} - \mathbf{v}') = \mathbf{0}_W$. Por tanto $\mathbf{v} - \mathbf{v}' \in \ker f$. Como $\ker f = \{\mathbf{0}_V\}$, tenemos $\mathbf{v} - \mathbf{v}' = \mathbf{0} \iff \mathbf{v} = \mathbf{v}'$.

(2) \implies (1). Sea f inyectiva, y $\mathbf{v} \in \ker f$: $f(\mathbf{v}) = \mathbf{0}_W$. También tenemos que $f(\mathbf{0}_V) = \mathbf{0}_W$; entonces $f(\mathbf{v}) = f(\mathbf{0}_V) \implies \mathbf{v} = \mathbf{0}_V$, es decir $\ker f = \{\mathbf{0}_V\}$.

(1) \implies (3). Supongamos $\ker f = \{\mathbf{0}_V\}$, y sea $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ un sistema de vectores linealmente independientes de V . La relación

$$\begin{aligned} \alpha_1 f(\mathbf{v}_1) + \dots + \alpha_k f(\mathbf{v}_k) = \mathbf{0}_W &\iff f(\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k) = \mathbf{0}_W \\ &\iff \alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k \in \ker f . \end{aligned}$$

Entonces, $\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}_V \implies \alpha_1 = 0, \dots, \alpha_k = 0$ por la independencia lineal de $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, así que los vectores $\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_k)\}$ también son linealmente independientes.

(3) \implies (1). Si $\mathbf{v} \neq \mathbf{0}_V$, es decir es lin. independiente, entonces $f(\mathbf{v})$ es lin. independiente en W , y por tanto $f(\mathbf{v}) \neq \mathbf{0}_W$. Equivalentemente, $f(\mathbf{v}) = \mathbf{0}_W \implies \mathbf{v} = \mathbf{0}_V$, es decir $\ker f = \{\mathbf{0}_V\}$. La demostración de la equivalencia entre las propiedades (1) y (4) es dejada por ejercicio. □

3.6.2 Segundo teorema de equivalencia

Enunciamos también otro resultado interesante, válido bajo la hipótesis más restrictiva que ambos espacios vectoriales V y W tienen la misma dimensión.

Teorema 40. Sean V y W dos espacios vectoriales sobre \mathbb{K} , de dimensión n y sea $f \in L(V, W)$. Las propiedades siguientes son equivalentes:

- (1) $\ker f = \{\mathbf{0}_V\}$
- (2) f es inyectiva
- (3) La imagen de una base de V es una base de W
- (4) f es un epimorfismo (sobreyectiva)
- (5) f es un isomorfismo (biyectiva)

La demostración del teorema anterior es parecida a la del primer teorema de equivalencia. Otro resultado útil es la siguiente

Proposición 22. Sea V un espacio vectorial sobre \mathbb{K} , $\dim V = n$ y $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de V . Sea $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ un sistema de vectores linealmente independientes de otro espacio vectorial W sobre \mathbb{K} . Entonces, existe una única aplicación lineal $f \in L(V, W)$ tal que

$$f(\mathbf{e}_i) = \mathbf{w}_i, \quad i = 1, \dots, n$$

La noción de espacios isomorfos es crucial en el prosieguro de la teoría.

Definición 64. Diremos que dos espacios vectoriales V y W sobre \mathbb{K} son **isomorfos** si existe un isomorfismo entre ellos.

Como consecuencia de la discusión anterior, podemos deducir el

Teorema 41. Dos espacios vectoriales V y W sobre \mathbb{K} de la misma dimensión finita son isomorfos.

Proof. Es suficiente observar que, siendo $\dim V = \dim W = n$, entonces podemos aplicar la Proposición 22 a dos bases, una de V y la otra de W y construir una aplicación lineal que transforma una base en la otra. Por tanto, como consecuencia del segundo teorema de equivalencia, la aplicación f es un isomorfismo, y por tanto los dos espacios son isomorfos. \square

3.7 Matriz asociada a una aplicación lineal

Sean V y W dos espacios vectoriales sobre \mathbb{K} , $\dim V = n$, $\dim W = m$. Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base de V y $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ una base de W , y $f \in L(V, W)$. Calculando las imágenes de los vectores de la base de V , tenemos

$$\begin{aligned} f(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \dots + a_{m1}\mathbf{w}_m \\ &\vdots \\ f(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \dots + a_{mn}\mathbf{w}_m \end{aligned}$$

donde los $a_{ij} \in \mathbb{K}$ son sendos coeficientes. Podemos escribir las relaciones anteriores en la forma

$$f(\mathbf{v}_j) = \sum_{i=1}^m a_{ij} \mathbf{w}_i, \quad j = 1, \dots, n. \quad (3.8)$$

Por tanto, hemos definido de esta forma una matriz $A = (a_{ij}) \in \mathcal{M}_{m \times n}$. Dicha matriz se denomina **matriz asociada a la aplicación lineal** f respecto de las bases asignadas $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ y $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$.

La matriz asociada a una aplicación lineal f , respecto de dos bases asignadas, permite calcular la imagen de cualquier vector del espacio que representa el dominio de f . Con más precisión, tenemos

$$\mathbf{x} \in V, \quad \mathbf{x} = \sum_{j=1}^n x_j \mathbf{v}_j, \quad f(\mathbf{x}) \in W, \quad f(\mathbf{x}) = \sum_{i=1}^m x'_i \mathbf{w}_i$$

Entonces

$$\begin{aligned} \sum_{i=1}^m x'_i \mathbf{w}_i &= f(\mathbf{x}) = f\left(\sum_{j=1}^n x_j \mathbf{v}_j\right) = \sum_{j=1}^n x_j f(\mathbf{v}_j) \\ &= \sum_{j=1}^n x_j \left(\sum_{i=1}^m a_{ij} \mathbf{w}_i\right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j\right) \mathbf{w}_i \end{aligned}$$

En definitiva, tenemos las ecuaciones

$$x'_i = \sum_{j=1}^n a_{ij} x_j, \quad i = 1, \dots, m \quad (3.9)$$

conocidas como ecuaciones de la aplicación lineal. Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ es el vector columna de las componentes de un vector $\mathbf{x} \in V$, entonces podemos escribir las ecuaciones (3.9) en la forma matricial equivalente:

$$X' = AX. \quad (3.10)$$

3.7.1 Aplicación lineal asociada a una matriz

Sean V, W dos espacios vectoriales sobre \mathbb{K} . Para asignar una aplicación lineal $f: V \rightarrow W$, dada una base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ de V es suficiente asignar los vectores

$$f(\mathbf{v}_1), \dots, f(\mathbf{v}_n).$$

De esta forma, conoceremos la imagen de todo vector $\mathbf{x} \in V$.

Definición 65. Sean $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base de V y $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ una base de W , y $A = (a_{ij}) \in \mathcal{M}_{m \times n}(\mathbb{K})$, $i = 1, \dots, m$, $j = 1, \dots, n$. Se define aplicación lineal asociada a la matriz A , respecto de las bases dadas, a la aplicación

$$f(\mathbf{v}_j) = \sum_{i=1}^m a_{ij} \mathbf{w}_i .$$

Observación 3.3. Sea

$$\phi : L(V, W) \rightarrow \mathcal{M}_{m \times n}(\mathbb{K})$$

la aplicación que a cada f lineal asocia su matriz (respecto de dos bases dadas):

$$\phi(f) = A \in \mathcal{M}_{m \times n}(\mathbb{K}) .$$

Se puede demostrar que la aplicación ϕ es lineal; además, siendo biyectiva, establece un isomorfismo entre los dos espacios vectoriales $L(V, W)$ y $\mathcal{M}_{m \times n}(\mathbb{K})$. Por tanto, para determinar la dimensión del espacio vectorial $L(V, W)$, es suficiente determinar la dimensión del espacio isomorfo $\mathcal{M}_{m \times n}(\mathbb{K})$, lo que es mucho más asequible.

Teorema 42. La dimensión del espacio vectorial $\mathcal{M}_{m \times n}(\mathbb{K})$ es $m \cdot n$.

Proof. Sea

$$\Omega_{ij} = \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

la matriz que contiene un 1 en la fila i -ésima y en la columna j -ésima. Es fácil comprobar que el conjunto $\{\Omega_{11}, \Omega_{12}, \dots, \Omega_{mn}\}$ es una base del espacio $\mathcal{M}_{m \times n}(\mathbb{K})$. Efectivamente:

$$A = (a_{ij}) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \Omega_{ij}$$

lo que demuestra que $\{\Omega_{11}, \Omega_{12}, \dots, \Omega_{mn}\}$ es un sistema de generadores del espacio. Es inmediato comprobar que estas matrices son linealmente independientes. Por tanto, forman una base de $\mathcal{M}_{m \times n}(\mathbb{K})$. Dado que el número de estas matrices es $m \cdot n$, tenemos que $\dim \mathcal{M}_{m \times n}(\mathbb{K}) = m \cdot n$. \square

Corolario 3. Sean V, W dos espacios vectoriales sobre \mathbb{K} , $\dim V = n$, $\dim W = m$. La dimensión del espacio vectorial $L(V, W)$ es $m \cdot n$.

3.7.2 Matriz asociada a la composición de aplicaciones

Sean V, W, U , tres espacios vectoriales sobre \mathbb{K} , $\dim V = n$, $\dim W = m$, $\dim U = p$. Sean $f \in L(V, W)$, $g \in L(W, U)$ $h = g \circ f \in L(V, U)$ y sean $\mathcal{B}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base de V , $\mathcal{B}_2 = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ una base de W y $\mathcal{B}_3 = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ una base de U . Calculemos las imágenes de los vectores de las bases \mathcal{B}_1 bajo f y \mathcal{B}_2 bajo g . Tenemos

$$f(\mathbf{v}_j) = \sum_{i=1}^m a_{ij} \mathbf{w}_i, \quad g(\mathbf{w}_i) = \sum_{k=1}^p b_{ki} \mathbf{u}_k.$$

Determinemos las imágenes de los vectores de \mathcal{B}_1 bajo la acción de h :

$$\begin{aligned} h(\mathbf{v}_j) &= g(f(\mathbf{v}_j)) = g\left(\sum_{i=1}^m a_{ij} \mathbf{w}_i\right) = \sum_{i=1}^m a_{ij} g(\mathbf{w}_i) \\ &= \sum_{i=1}^m a_{ij} \left(\sum_{k=1}^p b_{ki} \mathbf{u}_k\right) = \sum_{k=1}^p \left(\sum_{i=1}^m b_{ki} a_{ij}\right) \mathbf{u}_k = \sum_{k=1}^p c_{kj} \mathbf{u}_k. \end{aligned}$$

Por tanto, hemos demostrado que, si

$$A \in \mathcal{M}_{m \times n}(\mathbb{K}), \quad B \in \mathcal{M}_{p \times m}(\mathbb{K}), \quad C \in \mathcal{M}_{p \times n}(\mathbb{K})$$

son las matrices asociadas a f, g, h respectivamente, en las bases dadas, entonces

$$C = B \cdot A. \quad (3.11)$$

En particular, deducimos que si f es biyectiva, la matriz asociada a f^{-1} es A^{-1} .

3.7.3 Ejercicios

Ejercicio 3.5. Calcular la matriz asociada a la aplicación $f \in L(\mathbb{R}^3, \mathbb{R}^2)$ definida por

$$f(x, y, z) = (x + y, y + z)$$

respecto de las bases canónicas de \mathbb{R}^2 y \mathbb{R}^3 .

Resolución. Calculemos las imágenes de los vectores de la base canónica de \mathbb{R}^3 , en la base canónica de \mathbb{R}^2 :

$$f(1, 0, 0) = (1, 0)$$

$$f(0, 1, 0) = (1, 1)$$

$$f(0, 0, 1) = (0, 1)$$

Por tanto, escribiendo en columna estas imágenes, obtenemos la matriz A asociada a f en las bases dadas:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Ejercicio 3.6. Sea $\eta : \mathcal{M}_{2 \times 2} \rightarrow \mathcal{M}_{2 \times 2}$ la aplicación definida así:

$$\eta(A) = A + A^T.$$

(1) Demostrar que η es lineal.

(2) Hallar la matriz asociada a η respecto de la base canónica de $\mathcal{M}_{2 \times 2}$.

Resolución. El punto (1) es dejado al lector.

(2) La base canónica de $\mathcal{M}_{2 \times 2}$ es

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Tenemos:

$$\eta \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\eta \left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$\eta \left[\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right] = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\eta \left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Por tanto, la matriz asociada a η respecto de la base canónica de \mathbb{R}^2 es

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

3.8 Cambios de base y matrices asociadas a aplicaciones lineales

El teorema del cambio de base es una de las herramientas más relevantes en álgebra lineal, porque permite relacionar la descripción de los vectores de un espacio vectorial en distintas bases. Recordamos que, si V es un espacio vectorial, $\dim V = n$, $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ y $\mathcal{B}' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ son dos bases de V , y $\mathbf{x} \in V$, entonces si

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

son los vectores columna de las componentes del vector \mathbf{x} en las dos bases \mathcal{B} y \mathcal{B}' , tenemos

$$X = PX'$$

donde P es la matriz (invertible) de cambio de base.

Sean ahora V, W dos espacios vectoriales sobre \mathbb{K} y $f \in L(V, W)$. Una pregunta natural es: ¿Cómo cambia la matrix asociada a f bajo un cambio de base en V y en W ? La respuesta es dada por el siguiente

Teorema 43. Sean V y W dos espacios vectoriales sobre \mathbb{K} , $\dim V = n$, $\dim W = m$. Sean \mathcal{B}_V y \mathcal{B}'_V dos bases de V , \mathcal{B}_W y \mathcal{B}'_W dos bases de W y $f \in L(V, W)$. Sea $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ la matrix asociada a f respecto de \mathcal{B}_V y \mathcal{B}_W , $A' \in \mathcal{M}_{m \times n}(\mathbb{K})$ la matrix asociada a f respecto de \mathcal{B}'_V y \mathcal{B}'_W , P la matrix de cambio de \mathcal{B}_V a \mathcal{B}'_V , Q la matrix de cambio de \mathcal{B}_W a \mathcal{B}'_W . Entonces:

$$A' = Q^{-1} \cdot A \cdot P \quad (3.12)$$

Proof. Sea $\mathbf{x} \in V$, e $\mathbf{y} = f(\mathbf{x}) \in W$. Sabemos que, si $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ es la matrix asociada a f respecto de las bases \mathcal{B}_V y \mathcal{B}_W , tenemos que

$$Y = AX . \quad (3.13)$$

Análogamente, si $A' \in \mathcal{M}_{m \times n}(\mathbb{K})$ es la matrix asociada a f respecto de las bases \mathcal{B}'_V y \mathcal{B}'_W , denotando con X' e Y' las componentes de $\mathbf{x} \in V$ y $f(\mathbf{x}) \in W$, tenemos claramente

$$Y' = A'X' . \quad (3.14)$$

Además, valen las relaciones siguientes:

$$X = PX', \quad Y = QY' .$$

Entonces, sustituyendo estas relaciones en la ec. (3.13), deducimos

$$Y = AX \iff QY' = APX' \iff Y' = Q^{-1}APX'$$

Si ahora comparamos la relación obtenida con la ec. (3.14), obtenemos

$$A' = Q^{-1}AP . \quad (3.15)$$

□

Definición 66. Una matrix $M \in \mathcal{M}_n(\mathbb{K})$ se dirá semejante a la matrix $M' \in \mathcal{M}_n(\mathbb{K})$ si existe una matrix $N \in GL_n(\mathbb{K})$ tal que

$$M' = N^{-1}MN \quad (3.16)$$

Se puede facilmente demostrar la

Proposición 23. La relación de semejanza es una relación de equivalencia.

Desde la discusión anterior, deducimos inmediatamente el

Corolario 4. Sea $f \in L(V)$. Las matrices asociadas a f en dos bases distintas de V son semejantes.

Proof. Dado que $V = W$, desde el Teorema 43 obtenemos inmediatamente la ecuación de cambio de base

$$A' = P^{-1}AP .$$

lo que demuestra que A y A' son semejantes. □

3.9 Formas lineales

Como sabemos, un cuerpo conmutativo \mathbb{K} es también un espacio vectorial, sobre \mathbb{K} mismo.

Estudiaremos ahora una familia interesante de aplicaciones lineales, precisamente las $f : V \rightarrow \mathbb{K}$, que a un vector asocian un escalar.

Definición 67. Sea V un espacio vectorial sobre \mathbb{K} , $\dim V = n$. Una aplicación $f \in L(V, \mathbb{K})$ se llamará **forma lineal** de V .

Hemos demostrado que $L(V, W)$ es un espacio vectorial por toda elección de V y W . Por tanto, $L(V, \mathbb{K})$ es obviamente un espacio vectorial.

Definición 68. El espacio vectorial $L(V, \mathbb{K})$ se llamará **espacio vectorial dual de V** , y se denota con el símbolo $V^* := L(V, \mathbb{K})$.

Ejercicio 3.7.

Sea $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ y $\omega : \mathbb{R}^3 \rightarrow \mathbb{R}$ la aplicación así definida:

$$\omega(x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3 .$$

Es fácil comprobar que es lineal: $\omega(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) = \lambda_1 \omega(\mathbf{x}_1) + \lambda_2 \omega(\mathbf{x}_2) \forall \lambda_1, \lambda_2 \in \mathbb{R}$. Por tanto $\omega \in L(\mathbb{R}^3, \mathbb{R})$, es decir, es una forma lineal de \mathbb{R}^3 . Otra manera equivalente de escribir: $\omega \in \mathbb{R}^{3*}$.

3.9.1 Base dual

Dado un espacio vectorial V sobre \mathbb{K} , de dimensión n , es posible construir de forma natural una base en el espacio dual V^* . Precisamente, sea $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de V . Introducimos las formas lineales $\{f_1, \dots, f_n\}$

$$f_i(\mathbf{e}_j) := \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (3.17)$$

Teorema 44. Las formas (3.17) son una base de V^* .

Dicha base es conocida como la **base dual** de la base dada $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

Como consecuencia de la discusión anterior, vale el siguiente

Teorema 45. Sea V un espacio vectorial sobre \mathbb{K} de dimensión finita y V^* su espacio dual. Entonces

$$\dim V = \dim V^*$$

Como V^* es también, en sí mismo, un espacio vectorial sobre \mathbb{K} , podemos reiterar la construcción anterior y construir el dual del dual, y así siguiendo.

Definición 69. Si V es un espacio vectorial sobre \mathbb{K} y V^* su dual, el dual de V^* se denomina **bidual** de V y se denota con V^{**} :

$$V^{**} := L(V^*, \mathbb{K}) .$$

3.10 Aplicaciones multilineales

Introducimos las aplicaciones multilineales, como natural generalización de las aplicaciones lineales en el caso de funciones de más variables. Como veremos, una clase particularmente relevante de formas multilineales es representada por los determinantes. Empezamos con la

Definición 70. Sean V y W dos espacios vectoriales sobre \mathbb{K} , $p \in \mathbb{N}$, $p \geq 2$. Una aplicación

$$f : \underbrace{V \times \dots \times V}_{p \text{ veces}} \rightarrow W$$

$$(\mathbf{x}_1, \dots, \mathbf{x}_p) \rightarrow f(\mathbf{x}_1, \dots, \mathbf{x}_p) \in W$$

se dirá p -lineal si es lineal respecto a cada uno de los vectores $\mathbf{x}_1, \dots, \mathbf{x}_p$, es decir, si $\forall i = 1, \dots, p$ resulta:

$$(I) \quad f(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{y}_i + \mathbf{z}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p) = f(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{y}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p) + f(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{z}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p)$$

$$(II) \quad f(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \lambda \mathbf{x}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p) = \lambda f(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p)$$

Entonces, una aplicación multilineal no es nada más que una aplicación que es lineal respecto de cada una de sus variables (vectoriales) considerada por separado.

Teorema 46. Si f es una aplicación p -lineal, entonces

$$f\left(\sum_{i_1=1}^n a_{1i_1} \mathbf{v}_{i_1}, \sum_{i_2=1}^n a_{2i_2} \mathbf{v}_{i_2}, \dots, \sum_{i_p=1}^n a_{pi_p} \mathbf{v}_{i_p}\right) = \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_p=1}^n a_{1i_1} a_{2i_2} \dots a_{pi_p} f(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_p}). \quad (3.18)$$

Demostración.

$$f\left(\sum_{i_1=1}^n a_{1i_1} \mathbf{v}_{i_1}, \dots, \sum_{i_p=1}^n a_{pi_p} \mathbf{v}_{i_p}\right) = \sum_{i_1=1}^n a_{1i_1} f\left(\mathbf{v}_{i_1}, \sum_{i_2=1}^n a_{2i_2} \mathbf{v}_{i_2}, \dots, \sum_{i_p=1}^n a_{pi_p} \mathbf{v}_{i_p}\right)$$

$$= \sum_{i_1=1}^n a_{1i_1} \sum_{i_2=1}^n a_{2i_2} f\left(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \sum_{i_p=1}^n a_{pi_p} \mathbf{v}_{i_p}\right) = \dots = \sum_{i_1=1}^n a_{1i_1} \sum_{i_2=1}^n a_{2i_2} \dots \sum_{i_p=1}^n a_{pi_p} f(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_p})$$

$$= \sum_{i_1, i_2, \dots, i_p=1}^n a_{1i_1} a_{2i_2} \dots a_{pi_p} f(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_p}).$$

□

Ejercicio 3.8. Sea $\omega : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ la aplicación

$$\omega(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Esta aplicación es bilineal. Comprobémoslo:

$$\begin{aligned}\omega(\mathbf{z} + \mathbf{w}, \mathbf{y}) &= (z_1 + w_1)y_1 + (z_2 + w_2)y_2 + (z_3 + w_3)y_3 = \\ &= z_1y_1 + z_2y_2 + z_3y_3 + w_1y_1 + w_2y_2 + w_3y_3 = \\ &= \omega(\mathbf{z}, \mathbf{y}) + \omega(\mathbf{w}, \mathbf{y}) .\end{aligned}$$

$$\omega(\lambda\mathbf{x}, \mathbf{y}) = (\lambda x_1y_1 + \lambda x_2y_2 + \lambda x_3y_3) = \lambda(x_1y_1 + x_2y_2 + x_3y_3) = \lambda\omega(\mathbf{x}, \mathbf{y}) \quad (3.19)$$

De forma totalmente análoga se comprueban las propiedades de linealidad respecto de la segunda variable. La forma bilineal $\omega(\mathbf{x}, \mathbf{y})$ definida así es conocida como el **producto escalar canónico** de \mathbb{R}^3 y en contextos físicos usualmente se denota preferentemente con $\mathbf{x} \cdot \mathbf{y}$.

Teorema 47. *Sea V un espacio vectorial sobre \mathbb{K} , y $(\mathbf{x}_1, \dots, \mathbf{x}_p) \in V^p$. Supongamos que $\exists i$ tal que $\mathbf{x}_i = \mathbf{0}_V$. Entonces, para toda aplicación $f : V^p \rightarrow W$ lineal, tenemos*

$$f(\mathbf{x}_1, \dots, \mathbf{x}_p) = \mathbf{0}_W . \quad (3.20)$$

Demostración. Observamos que, siendo $0\mathbf{v} = \mathbf{0}$ para todo $\mathbf{v} \in V$, tenemos

$$f(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{0}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p) = 0 \cdot f(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{v}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p) = \mathbf{0}_W .$$

□

3.11 Aplicaciones alternadas

Definición 71. *Una aplicación $f : V^p \rightarrow W$ p -lineal se dirá alternada si*

$$f(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_p) = -f(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_p) .$$

Por tanto, intercambiando entre sí dos vectores, una aplicación alternada cambia de signo.

Definición 72. *Una aplicación $f : V^p \rightarrow \mathbb{K}$ p -lineal se dirá forma p -lineal. Si además es alternada, se dirá forma p -lineal alternada.*

Ejercicio 3.9. Sea $\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$\alpha(\mathbf{x}, \mathbf{y}) = x_1y_2 - x_2y_1$$

Esta aplicación es una forma 2-lineal alternada. Desde un punto de vista físico, si

$$\mathbf{L} = \mathbf{x} \times \mathbf{p}, \quad \mathbf{L} \cdot \mathbf{z} = x_1p_2 - x_2p_1$$

la aplicación $\alpha(\mathbf{x}, \mathbf{p})$ se puede interpretar como la aplicación en el plano euclídeo que define la tercera componente del momento angular de un punto material de posición \mathbf{x} y momento lineal \mathbf{p} .

Teorema 48. Sea $f : V^p \rightarrow W$ una aplicación p -lineal alternada. Entonces

$$f(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_i, \dots, \mathbf{x}_p) = \mathbf{0}_W . \quad (3.21)$$

Demostración. Poniendo $\mathbf{x}_j = \mathbf{x}_i$ en la definición de aplicación alternada, obtenemos inmediatamente

$$f(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_i, \dots, \mathbf{x}_p) = -f(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_p)$$

es decir

$$2f(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_p) = \mathbf{0}_W$$

valida si el cuerpo conmutativo \mathbb{K} es de característica distinta de 2. \square

El teorema siguiente es de fundamental importancia en las aplicaciones: establece que la imagen de un conjunto de p vectores dependientes bajo cualquier aplicación p -lineal alternada es nula. Como veremos, se trata de una propiedad crucial de los determinantes.

Teorema 49. Sea $f : V^p \rightarrow W$ una aplicación p -lineal alternada. Si los vectores $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ son linealmente dependientes, entonces

$$f(\mathbf{x}_1, \dots, \mathbf{x}_p) = \mathbf{0}_W .$$

Demostración. Si $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ es un conjunto de vectores dependientes, entonces por lo menos uno entre ellos es combinación lineal de los demás. Supongamos, sin pérdida de generalidad, que

$$\mathbf{x}_1 = \alpha_2 \mathbf{x}_2 + \dots + \alpha_p \mathbf{x}_p .$$

Entonces,

$$\begin{aligned} f(\mathbf{x}_1, \dots, \mathbf{x}_p) &= f(\alpha_2 \mathbf{x}_2 + \dots + \alpha_p \mathbf{x}_p, \dots, \mathbf{x}_p) \\ &= \alpha_2 f(\mathbf{x}_2, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_p) + \alpha_3 f(\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_p) + \dots \\ &+ \alpha_p f(\mathbf{x}_p, \mathbf{x}_2, \dots, \mathbf{x}_p) = \alpha_2 \mathbf{0}_W + \alpha_3 \mathbf{0}_W + \dots + \alpha_p \mathbf{0}_W = \mathbf{0}_W . \end{aligned}$$

\square

Desde la discusión anterior se deducen fácilmente los resultados siguientes.

Corolario 5. Una aplicación p -lineal alternada $f : V^p \rightarrow W$ no cambia si a uno de los vectores $\mathbf{x}_1, \dots, \mathbf{x}_p$ añadimos una combinación lineal de los demás $p - 1$ vectores.

Corolario 6. Si $\dim V < p$, entonces toda aplicación p -lineal alternada $f : V^p \rightarrow W$ es nula, es decir

$$f(\mathbf{x}_1, \dots, \mathbf{x}_p) = \mathbf{0}_W .$$

3.12 Grupo de permutaciones

Definición 73. Sea $\mathcal{S} = \{1, 2, \dots, n\}$ un conjunto finito de números naturales, o bien de objetos identificados con esos números. Una permutación φ es una biyección de \mathcal{S} :

$$\varphi : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\} .$$

Ejemplo 3.10. Sea $\mathcal{S} = \{1, 2, 3, 4\}$. Una permutación de \mathcal{S} es la aplicación

$$\varphi : \mathcal{S} \rightarrow \mathcal{S}, \quad \varphi(1) = 3, \quad \varphi(2) = 1, \quad \varphi(3) = 2, \quad \varphi(4) = 4 .$$

Es inmediato comprobar que es biyectiva.

La permutación anterior se puede también representar con una matriz:

$$\varphi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

Dado que la primera fila sólo representa la sucesión de los primeros n números naturales (en este caso con $n = 4$), en realidad la permutación está fijada si conocemos la segunda fila de la matriz. Por tanto, una representación más compacta es

$$\varphi = 3124 .$$

Puesto que se trata de aplicaciones biyectivas, podemos obviamente definir la permutación inversa:

$$\varphi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

o bien $\varphi^{-1} = 2314$. Por supuesto,

$$\varphi \circ \varphi^{-1} = e = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} .$$

Dado un conjunto \mathcal{S} de n objetos, el número total de las posibles permutaciones de \mathcal{S} es $n!$.

Observación 3.4. El conjunto de todas las permutaciones de n objetos, dotado de la composición de permutaciones es un grupo, denotado con S_n . Efectivamente, tenemos las propiedades siguientes:

- (1) $\forall \varphi, \chi, \psi \in S_n : (\varphi \circ \chi) \circ \psi = \varphi \circ (\chi \circ \psi)$
- (2) $\exists e \in S_n$ t.q. $\forall \varphi \in S_n : e \circ \varphi = \varphi \circ e = \varphi$
- (3) $\forall \varphi \in S_n \exists \varphi^{-1} \in S_n$ t.q. $\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = e$.

Definición 74. Llamaremos *transposición* una permutación que intercambia entre sí dos elementos de \mathcal{S} dejando invariados los demás elementos.

Ejemplo 3.11. La transposición que intercambia los elementos i y j , $i \neq j$, es

$$\begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ 1 & 2 & \dots & j & \dots & i & \dots & n \end{pmatrix} .$$

Definición 75. Diremos que en una permutación $\varphi \in S_n$ los elementos $\varphi(i)$ y $\varphi(j)$ presentan una inversión si

$$i < j \quad \text{y} \quad \varphi(i) > \varphi(j) .$$

Definición 76. Una permutación se dirá **par** si el número total de inversiones que presenta es par, e **impar** si el número total de inversiones es impar.

Definición 77. Se define paridad de una permutación $\varphi \in S_n$ al número

$$\sigma_\varphi = (-1)^I$$

siendo I el número de inversiones de la permutación φ .

Entonces,

$$\sigma_\varphi = \begin{cases} +1 & \text{si la permutación es par} \\ -1 & \text{si la permutación es impar} \end{cases}$$

Se puede demostrar que

$$\sigma_\varphi = \sigma_{\varphi^{-1}}$$

Equivalentemente, podemos definir la paridad de una permutación como

$$\sigma_\varphi = \begin{cases} +1 & \text{si el número de transposiciones necesario para ordenar la permutación es par} \\ -1 & \text{si el número de transposiciones necesario para ordenar la permutación es impar} \end{cases}$$

Ejemplo 3.12. Consideremos la permutación $\varphi = 3124$. Tenemos dos inversiones: 31, 32. Por tanto $\sigma_\varphi = (-1)^2 = 1$. Equivalentemente,

$$\varphi = 3124 \xrightarrow{\text{transp}} 1324 \xrightarrow{\text{transp}} 1234 ,$$

es decir, con dos transposiciones podemos ordenar φ . Por tanto $\sigma_\varphi = 1$.

Lema 5. Sea $f : V^p \rightarrow W$ una aplicación p -lineal alternada y sea

$$\varphi = \begin{pmatrix} 1 & 2 & \dots & p \\ \varphi(1) & \varphi(2) & \dots & \varphi(p) \end{pmatrix}$$

una permutación de S_p . Entonces

$$f(\mathbf{x}_{\varphi(1)}, \mathbf{x}_{\varphi(2)}, \dots, \mathbf{x}_{\varphi(p)}) = \sigma_\varphi f(\mathbf{x}_1, \dots, \mathbf{x}_p) .$$

Demostración. Sea I el número de inversiones presentes en la permutación φ . Entonces, efectuando I transposiciones, φ se convierte en la permutación idéntica. Pero, bajo cada transposición, la aplicación f siendo alternada cambia de signo, entonces:

$$f(\mathbf{x}_{\varphi(1)}, \mathbf{x}_{\varphi(2)}, \dots, \mathbf{x}_{\varphi(p)}) = (-1)^I f(\mathbf{x}_1, \dots, \mathbf{x}_p) ,$$

es decir

$$f(\mathbf{x}_{\varphi(1)}, \mathbf{x}_{\varphi(2)}, \dots, \mathbf{x}_{\varphi(p)}) = \sigma_\varphi f(\mathbf{x}_1, \dots, \mathbf{x}_p) .$$

□

Teorema 50. Sea $f : V^p \rightarrow W$ una aplicación p -lineal alternada, y sea

$$\mathbf{x}_i = \sum_{j=1}^p a_{ij} \mathbf{v}_j \quad i = 1, \dots, p .$$

Entonces

$$f(\mathbf{x}_1, \dots, \mathbf{x}_p) = \sum_{\varphi \in S_p} \sigma_\varphi a_{1\varphi(1)} \cdots a_{p\varphi(p)} f(\mathbf{v}_1, \dots, \mathbf{v}_p) . \quad (3.22)$$

Demostración.

Ya hemos demostrado (Teorema 46) que

$$\begin{aligned} f(\mathbf{x}_1, \dots, \mathbf{x}_p) &= f\left(\sum_{i_1=1}^p a_{1i_1} \mathbf{v}_{i_1}, \sum_{i_2=1}^p a_{2i_2} \mathbf{v}_{i_2}, \dots, \sum_{i_p=1}^p a_{pi_p} \mathbf{v}_{i_p}\right) = \\ &= \sum_{i_1=1}^p \sum_{i_2=1}^p \cdots \sum_{i_p=1}^p a_{1i_1} a_{2i_2} \cdots a_{pi_p} f(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_p}) \end{aligned} \quad (3.23)$$

donde la suma involucra los p^p conjuntos de índices $\{j_1, \dots, j_p\}$, siendo $j_i = 1, \dots, p \forall i \in \{1, \dots, p\}$. Para cada conjunto de índices en que aparecen dos índices iguales, evidentemente

$$f(\mathbf{v}_{j_1}, \dots, \mathbf{v}_{j_p}) = \mathbf{0}_W .$$

Entonces, en la fórmula 3.23 sólo quedan los términos correspondientes a los sistemas de índices distintos, es decir los términos correspondientes a las permutaciones φ del conjunto $\{1, \dots, p\}$. Consecuentemente:

$$f(\mathbf{x}_1, \dots, \mathbf{x}_p) = \sum_{\varphi \in S_p} a_{1\varphi(1)} \cdots a_{p\varphi(p)} f(\mathbf{v}_{\varphi(1)}, \dots, \mathbf{v}_{\varphi(p)}) \quad (3.24)$$

$$= \sum_{\varphi \in S_p} \sigma_\varphi a_{1\varphi(1)} \cdots a_{p\varphi(p)} f(\mathbf{v}_1, \dots, \mathbf{v}_p) \quad (3.25)$$

□

El siguiente teorema establece que, dados dos espacios vectoriales V y W , siempre podemos construir una aplicación lineal alternada de V^p en W que asocia a una base de V un vector asignado de W . Además, dicha aplicación es única. Con más precisión, tenemos el

Teorema 51. Sean V y W dos espacios vectoriales sobre \mathbb{K} , $\dim V = p$. Para cada base $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ de V y $\forall \mathbf{w} \in W$, existe una única aplicación p -lineal alternada $f : V^p \rightarrow W$ tal que

$$f(\mathbf{v}_1, \dots, \mathbf{v}_p) = \mathbf{w} \quad (3.26)$$

Si, en particular, elegimos $W = \mathbb{K}$, y entonces elegimos $\mathbf{w} = 1$, vale el

Corolario 7. Sea V un espacio vectorial \mathbb{K} , $\dim V = p$. Para toda base $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ de V existe una única forma p -lineal alternada $f : V^p \rightarrow \mathbb{K}$, tal que

$$f(\mathbf{v}_1, \dots, \mathbf{v}_p) = 1 . \quad (3.27)$$

Definición 78. Sea V un espacio vectorial sobre \mathbb{K} , $\dim V = p$, $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ una base de V y $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ un conjunto de vectores de V . Llamaremos **determinante** del conjunto $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ respecto de la base $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ a la única forma p -lineal alternada $D : V^p \rightarrow \mathbb{K}$ tal que

$$D(\mathbf{v}_1, \dots, \mathbf{v}_p) = 1 .$$

Evidentemente, siendo para cada forma p -lineal

$$D(\mathbf{x}_1, \dots, \mathbf{x}_p) = \sum_{\varphi \in S_p} \sigma_\varphi a_{1\varphi(1)} \cdot \dots \cdot a_{p\varphi(p)} D(\mathbf{v}_1, \dots, \mathbf{v}_p) \quad (3.28)$$

en el caso de un determinante deducimos la fórmula fundamental

$$D(\mathbf{x}_1, \dots, \mathbf{x}_p) = \sum_{\varphi \in S_p} \sigma_\varphi a_{1\varphi(1)} \cdot \dots \cdot a_{p\varphi(p)} . \quad (3.29)$$

A menudo utilizaremos la notación abreviada

$$\varphi(i) =: \varphi_i, \quad i = 1, \dots, p .$$

Mostraremos ahora que la expresión (3.29) nos devuelve el algoritmo familiar para el cálculo de los determinantes.

Definición 79. Sea $A = (a_{ij}) \in \mathcal{M}_p(\mathbb{K})$, $i, j = 1, \dots, p$. Llamaremos **determinante** de A de orden p y lo denotaremos con el símbolo

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & \dots & \dots & a_{2p} \\ \vdots & \dots & \dots & \vdots \\ a_{p1} & \dots & \dots & a_{pp} \end{vmatrix}$$

al escalar así definido:

$$\det(A) = \sum_{\varphi \in S_p} \sigma_\varphi a_{1\varphi_1} \cdot \dots \cdot a_{p\varphi_p} . \quad (3.30)$$

Ejemplo 3.13. Sea $p = 2$. Calculemos el determinante de orden 2. Tenemos

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \sum_{\varphi \in S_2} \sigma_\varphi a_{1\varphi_1} a_{2\varphi_2} .$$

Estudiemos ahora las permutaciones de S_2 . Tenemos $2! = 2$ permutaciones:

$$\varphi_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{y} \quad \varphi_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \sigma_{\varphi_1} = 1, \quad \sigma_{\varphi_2} = -1$$

que en notación abreviada escribiremos como $\varphi_1 = 12$ y $\varphi = 21$. Entonces,

$$\det(A) = \sigma_{\varphi_1} a_{11} a_{22} + \sigma_{\varphi_2} a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21} .$$

Ejemplo 3.14. Sea $p = 3$ y calculemos el determinante de orden 3.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \det(A) = \sum_{\varphi \in S_3} \sigma_{\varphi} a_{1\varphi_1} a_{2\varphi_2} a_{3\varphi_3} .$$

El grupo S_3 contiene $3! = 6$ permutaciones, de las que 3 tienen paridad par y tres impar:

$$\begin{array}{lll} \varphi_1 = 123 & \varphi_2 = 231 & \varphi_3 = 312 \\ \varphi_4 = 132 & \varphi_5 = 213 & \varphi_6 = 321 \end{array}$$

Por tanto:

$$\begin{aligned} \det(A) &= \sigma_{\varphi_1} a_{11} a_{22} a_{33} + \sigma_{\varphi_2} a_{12} a_{23} a_{31} + \sigma_{\varphi_3} a_{13} a_{21} a_{32} \\ &+ \sigma_{\varphi_4} a_{11} a_{23} a_{32} + \sigma_{\varphi_5} a_{12} a_{21} a_{33} + \sigma_{\varphi_6} a_{13} a_{22} a_{31} \\ &= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31} \end{aligned}$$

es decir

$$\det(A) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) . \quad (3.31)$$

La fórmula (3.31), que es una consecuencia directa de la Definición 79, es conocida como el desarrollo de Laplace del determinante (según la primera fila).

El resultado siguiente es de gran relevancia teórica.

Teorema 52. Sea $A \in \mathcal{M}_p(\mathbb{K})$. Entonces

$$\det(A^T) = \det(A) . \quad (3.32)$$

Como consecuencia, cualquier propiedad que concierne las filas del determinante de una matriz cuadrada es válida también para sus columnas.

Una de las aplicaciones más relevantes de los determinantes es su uso para determinar si un conjunto de vectores es linealmente dependiente, o no.

Teorema 53. Condición necesaria y suficiente para que los vectores $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ de un espacio vectorial V de dimensión p sean linealmente dependientes es que su determinante, respecto de una base $\{\mathbf{e}_1, \dots, \mathbf{e}_p\}$ cualesquiera de V sea nulo.

Demostración. La condición es necesaria: en virtud del Teorema 49, ya sabemos que si $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ son linealmente dependientes, entonces cualquier forma p -lineal alternada $f : V^p \rightarrow W$ se anula: $f(\mathbf{x}_1, \dots, \mathbf{x}_p) = \mathbf{0}_W$. Entonces, en particular

$$D(\mathbf{x}_1, \dots, \mathbf{x}_p) = 0 .$$

Demostremos que la condición es suficiente, es decir:

$$D(\mathbf{x}_1, \dots, \mathbf{x}_p) = 0 \implies \{\mathbf{x}_1, \dots, \mathbf{x}_p\} \text{ lin. dependientes.}$$

Demostremoslo por absurdo. Supongamos que $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ sean linealmente independientes; entonces, forman una base de V . Por tanto, los vectores $\{\mathbf{e}_1, \dots, \mathbf{e}_p\}$ se pueden expresar como combinación lineal de ellos:

$$\mathbf{e}_i = \sum_{j=1}^p \alpha_{ij} \mathbf{x}_j, \quad i = 1, \dots, p.$$

Tenemos:

$$D(\mathbf{e}_1, \dots, \mathbf{e}_p) = \left(\sum_{\varphi} \sigma_{\varphi} \alpha_{1\varphi_1} \dots \alpha_{p\varphi_p} \right) D(\mathbf{x}_1, \dots, \mathbf{x}_p) = 0$$

lo que implica $D(\mathbf{e}_1, \dots, \mathbf{e}_p) = 0$, pero esto es absurdo porque, siendo $\{\mathbf{e}_1, \dots, \mathbf{e}_p\}$ una base de V , debe ser $D(\mathbf{e}_1, \dots, \mathbf{e}_p) = 1$. \square

3.13 Determinantes e invertibilidad

Teorema 54 (Criterio para matrices invertibles). *Una matriz $A \in M_n(\mathbb{K})$ es invertible si y sólo si $\det(A) \neq 0$.*

Demostración. Es suficiente observar que las afirmaciones siguientes son equivalentes:

A invertible $\iff A$ es la matriz asociada a una aplicación lineal biyectiva (automorfismo) \iff La imagen de una base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ de V es otra base $\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)\}$ de $V \iff$ Los vectores columna de A son linealmente independientes $\iff \det(A) \neq 0$.

Corolario 8. *Si una fila (columna) de un determinante es combinación lineal de las otras filas (columnas), el determinante es nulo.*

Corolario 9. *Si una fila (columna) de un determinante tiene todas sus entradas nulas, o si dos filas (columnas) son proporcionales, el determinante es nulo.*

Proposición 24. *Si añadimos a una fila (columna) una combinación lineal de otras filas (columnas) el determinante queda invariado.*

Demostración. Sea V un espacio vectorial de dimensión p y sea un sistema de vectores $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$. En el determinante $D(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_p)$ vamos a añadir al vector \mathbf{x}_i una combinación lineal de los demás $(p-1)$ vectores:

$$\begin{aligned} & D(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_i + \alpha_1 \mathbf{x}_1 + \dots + \alpha_{i-1} \mathbf{x}_{i-1} + \alpha_{i+1} \mathbf{x}_{i+1} + \dots + \alpha_p \mathbf{x}_p, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p) = \\ & = D(\mathbf{x}_1, \dots, \mathbf{x}_p) + D(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \alpha_1 \mathbf{x}_1 + \dots + \alpha_{i-1} \mathbf{x}_{i-1} + \alpha_{i+1} \mathbf{x}_{i+1} + \dots + \alpha_p \mathbf{x}_p, \mathbf{x}_{i+1}, \dots, \mathbf{x}_p) = \\ & = D(\mathbf{x}_1, \dots, \mathbf{x}_p), \end{aligned}$$

dado que el segundo determinante se puede repartir en la suma de determinantes que contienen, cada uno de ellos, un par de vectores iguales.

Corolario 10. Si se multiplican todas las entradas de una fila (columna) de un determinante por un escalar λ , el determinante queda multiplicado por λ .

Corolario 11. Si se intercambian entre sí dos filas (columnas) de un determinante, el determinante cambia signo.

El siguiente resultado, que presentamos sin demostración, es conocido como desarrollo de Laplace de un determinante según la i -ésima fila.

Teorema 55. Sea la matriz $A = (a_{ij})$, $(i, j = 1, \dots, p)$. Si denotamos con A_{ij} la matriz obtenida desde A eliminando en ella la fila i -ésima y la columna j -ésima, tenemos:

$$\det(A) = \sum_{j=1}^p (-1)^{i+j} a_{ij} \det(A_{ij}), \quad i \in \{1, \dots, p\} .$$

Existe también la fórmula de desarrollo de un determinante según la j -ésima columna:

$$\det(A) = \sum_{i=1}^p (-1)^{i+j} a_{ij} \det(A_{ij}), \quad j \in \{1, \dots, p\} .$$

Definición 80. El determinante $\det(A_{ij})$ se denomina menor del elemento a_{ij} de A . La cantidad

$$\Delta_{ij} := (-1)^{i+j} \det(A_{ij})$$

es dicha cofactor del elemento a_{ij} de la matriz A .

Una consecuencia obvia, pero muy útil, es la siguiente

Observación 3.5. El determinante de una matriz diagonal es igual al producto de los elementos de su diagonal principal. Efectivamente, si

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & & \dots & \lambda_n \end{pmatrix}$$

entonces evidentemente

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n . \tag{3.33}$$

En particular, para el determinante de la matriz identidad I_n obtenemos inmediatamente $\det(I_n) = 1$.

Enunciamos también un resultado, bien conocido, que proporciona un algoritmo para calcular la inversa de una matriz utilizando la noción de determinante.

Teorema 56. Sea una matriz cuadrada $A = (a_{ij})$, $(i, j = 1, \dots, p)$ invertible. Entonces la matriz inversa de A es dada por

$$A^{-1} = \frac{1}{\det(A)} \|\Delta_{ji}\|.$$

Por tanto, para calcular la inversa de una matriz A invertible, podemos operar así:

- Se escribe A^T , la matriz transpuesta de A
- Se substituye cada elemento de A^T con su cofactor
- Se divide la matriz así determinada por $\det(A) \neq 0$.

Un resultado crucial de la teoría de los determinantes es debido al matemático francés Binet:

Teorema 57. Sean $A, B \in M_p(\mathbb{K})$. Entonces

$$\det(AB) = (\det A)(\det B). \quad (3.34)$$

Demostración. Consideremos un espacio vectorial V de dimensión p sobre el cuerpo conmutativo \mathbb{K} , y sea $\{\mathbf{e}_1, \dots, \mathbf{e}_p\}$ una base de V . Sea en V un sistema de vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$, definido por:

$$\mathbf{v}_j := \sum_{k=1}^p b_{jk} \mathbf{e}_k, \quad (j = 1, \dots, p) \quad (3.35)$$

y otro sistema $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$, definido por

$$\mathbf{u}_i := \sum_{j=1}^p a_{ij} \mathbf{v}_j, \quad (i = 1, \dots, p). \quad (3.36)$$

Introducimos las matrices $A = (a_{ij})$, $B = (b_{jk})$. Además, sea D la única forma p -lineal alternada en V tal que

$$D(\mathbf{e}_1, \dots, \mathbf{e}_p) = 1, \quad (3.37)$$

es decir, el determinante. Tenemos entonces

$$D(\mathbf{u}_1, \dots, \mathbf{u}_p) = (\det A)D(\mathbf{v}_1, \dots, \mathbf{v}_p) \quad (3.38)$$

y

$$D(\mathbf{v}_1, \dots, \mathbf{v}_p) = (\det B)D(\mathbf{e}_1, \dots, \mathbf{e}_p) \quad (3.39)$$

Sustituyendo la ecuación (3.39) en la (3.38), y teniendo en cuenta la ecuación (3.37), obtenemos

$$D(\mathbf{u}_1, \dots, \mathbf{u}_p) = (\det A)(\det B). \quad (3.40)$$

Por otro lado, observamos que, siendo los vectores $\mathbf{u}_i \in V$, se podrán escribir, cada uno de ellos, como combinación lineal de los vectores de la base $\{\mathbf{e}_1, \dots, \mathbf{e}_p\}$. Por tanto,

$$\mathbf{u}_i = \sum_{k=1}^p c_{ik} \mathbf{e}_k \quad (i = 1, \dots, p) . \quad (3.41)$$

Utilizando las relaciones (3.35) y (3.36), obtenemos entonces

$$\mathbf{u}_i = \sum_{j=1}^p a_{ij} \mathbf{v}_j = \sum_{j=1}^p a_{ij} \left(\sum_{k=1}^p b_{jk} \mathbf{e}_k \right) = \sum_{k=1}^p \left(\sum_{j=1}^p a_{ij} b_{jk} \right) \mathbf{e}_k ,$$

de donde se deduce que

$$c_{ik} = \sum_{j=1}^p a_{ij} b_{jk} .$$

Entonces, si $C := (c_{ik})$, $i, k = 1, \dots, p$, tenemos

$$C = A \cdot B .$$

Ahora bien, desde la eq. (3.41) obtenemos

$$D(\mathbf{u}_1, \dots, \mathbf{u}_p) = (\det C) D(\mathbf{e}_1, \dots, \mathbf{e}_p) = \det(A \cdot B) . \quad (3.42)$$

En definitiva, comparando las ecuaciones (3.40) y (3.42), deducimos que

$$\det(A \cdot B) = (\det A)(\det B) .$$

□

Chapter 4

Autovalores y Autovectores

4.1 Definiciones y propiedades básicas

Sea V un espacio vectorial sobre un cuerpo conmutativo \mathbb{K} , y sea $f \in L(V)$ una aplicación lineal de V en V (es decir, un *endomorfismo* de V).

Definición 81. Sea $\mathbf{v} \neq \mathbf{0}$ de V . El vector \mathbf{v} se dirá *autovector* (o *vector propio*) del endomorfismo f si existe un escalar $\lambda \in \mathbb{K}$ tal que se cumple la condición

$$f(\mathbf{v}) = \lambda \mathbf{v} . \quad (4.1)$$

El escalar λ se dirá *autovalor* (o *valor propio*) de f correspondiente al autovector \mathbf{v} .

Definición 82. El conjunto de los autovalores de un endomorfismo se denomina *espectro* de los autovalores de f y se denota con $\text{Spec}(f)$.

Teorema 58. Sea $f \in L(V)$, y sea λ un autovalor de f . El conjunto V_λ formado por los autovectores correspondientes a λ y por el vector nulo $\mathbf{0}$ es un subespacio vectorial de V .

Demostración. Sean \mathbf{v}_1 y \mathbf{v}_2 dos autovectores correspondientes al mismo autovalor λ y sea $\alpha_1, \alpha_2 \in \mathbb{K}$. Entonces,

$$f(\mathbf{v}_1) = \lambda \mathbf{v}_1, \quad f(\mathbf{v}_2) = \lambda \mathbf{v}_2 .$$

Por tanto tenemos

$$f(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) = \alpha_1 \lambda \mathbf{v}_1 + \alpha_2 \lambda \mathbf{v}_2 = \lambda(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) ,$$

Esto implica que $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$ es un autovector correspondiente al autovalor λ . \square

Definición 83. El subespacio V_λ se llamará *autoespacio* (o *subespacio propio*) correspondiente al autovalor λ .

Proposición 25. *Los subespacios correspondientes a dos autovalores distintos λ_1 y λ_2 son disjuntos, es decir*

$$V_{\lambda_1} \cap V_{\lambda_2} = \{\mathbf{0}\} .$$

Demostración. Sea $\lambda_1 \neq \lambda_2$ y sean $V_{\lambda_1}, V_{\lambda_2}$ los autoespacios correspondientes. Supongamos que $\mathbf{v} \in V_{\lambda_1} \cap V_{\lambda_2}$. Entonces se cumple

$$f(\mathbf{v}) = \lambda_1 \mathbf{v}, \quad f(\mathbf{v}) = \lambda_2 \mathbf{v}$$

lo que implica

$$(\lambda_1 - \lambda_2) \mathbf{v} = \mathbf{0} .$$

Dado que, por hipótesis, $\lambda_1 \neq \lambda_2$, entonces $\mathbf{v} = \mathbf{0}$, y por tanto $V_{\lambda_1} \cap V_{\lambda_2} = \{\mathbf{0}\}$. \square

Teorema 59. *Si $V_{\lambda_1}, \dots, V_{\lambda_p}$ son autoespacios correspondientes a p autovalores distintos $\lambda_1, \dots, \lambda_p$, entonces su suma $V_{\lambda_1} + \dots + V_{\lambda_p}$ es directa.*

Demostración. Utilizaremos el procedimiento por inducción. Sea $p = 2$. Si $\lambda_1 \neq \lambda_2$, entonces $V_{\lambda_1} \cap V_{\lambda_2} = \{\mathbf{0}\}$ y la suma $V_{\lambda_1} + V_{\lambda_2}$ es directa.

Supondremos ahora la propiedad cierta para $p - 1$ subespacios propios, y demostraremos que es cierta para p subespacios propios.

Sea $\mathbf{v}_i \in V_{\lambda_i}$, $i = 1, \dots, p$. Ahora bien, sabemos que la suma $V_{\lambda_1} + \dots + V_{\lambda_p}$ es directa si y sólo si la relación

$$\sum_{i=1}^p \mathbf{v}_i = \mathbf{0} \implies \mathbf{v}_i = \mathbf{0} \quad (i = 1, \dots, p). \quad (4.2)$$

Entonces tenemos

$$f\left(\sum_{i=1}^p \mathbf{v}_i\right) = \sum_{i=1}^p f(\mathbf{v}_i) = f(\mathbf{0}) = \mathbf{0} .$$

Al mismo tiempo, dado que

$$f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i ,$$

obtenemos

$$\sum_{i=1}^p \lambda_i \mathbf{v}_i = \mathbf{0} . \quad (4.3)$$

Para demostrar que la suma es directa, utilicemos la propiedad (4.2). Claramente, si $\sum_{i=1}^p \mathbf{v}_i = \mathbf{0}$, también resulta (trivialmente) que

$$\lambda_1 \sum_{i=1}^p \mathbf{v}_i = \mathbf{0}. \quad (4.4)$$

Por tanto, restando las ecuaciones (4.3) y (4.4), deducimos que

$$(\lambda_2 - \lambda_1) \mathbf{v}_2 + \dots + (\lambda_p - \lambda_1) \mathbf{v}_p = \mathbf{0} .$$

Está claro que $\mathbf{w}_k := (\lambda_k - \lambda_1)\mathbf{v}_k \in V_{\lambda_k}$, $k = 2, \dots, p$. Entonces, utilizando la hipótesis de inducción, obtenemos

$$\mathbf{w}_2 + \dots + \mathbf{w}_p = \mathbf{0} \implies \mathbf{w}_2 = \mathbf{0}, \dots, \mathbf{w}_p = \mathbf{0},$$

es decir $(\lambda_k - \lambda_1)\mathbf{v}_k = \mathbf{0}$, $k = 2, \dots, p$. Esto implica $\mathbf{v}_k = \mathbf{0}$, $k = 2, \dots, p$ dado que $\lambda_k \neq \lambda_1$. Por tanto, en la suma $\sum_{i=1}^p \mathbf{v}_i = \mathbf{0}$, siendo $\mathbf{v}_2 = \dots = \mathbf{v}_p = \mathbf{0}$, resulta que también $\mathbf{v}_1 = \mathbf{0}$. En definitiva, hemos demostrado que la condición (4.2) se cumple, y que la suma es directa. □

4.2 Subespacios invariantes

Introducimos la noción de subespacio invariante por una aplicación lineal f .

Definición 84. Sea V un espacio vectorial sobre el cuerpo conmutativo \mathbb{K} , $f \in L(V)$. Diremos que un subespacio $W \subset V$ es invariante bajo f o es f -invariante si

$$f(W) \subseteq W$$

es decir, f transforma vectores de W en vectores de W .

Supongamos ahora que V_λ sea un autoespacio de f relativo al autovalor λ . Por definición, tenemos que

$$f(\mathbf{v}) = \lambda\mathbf{v} \quad \forall \mathbf{v} \in V_\lambda$$

y como V_λ es un subespacio vectorial, obviamente $\lambda\mathbf{v} \in V_\lambda$. Por tanto, según la definición anterior, los autoespacios de f son subespacios invariantes por f .

4.3 Ecuación característica

El teorema siguiente posee un papel crucial en la teoría.

Teorema 60. Sea V un espacio vectorial sobre \mathbb{K} , $\dim V = n$, y sea $f \in L(V)$. Sea $A \in M_n(\mathbb{K})$ la matriz asociada a f respecto de la base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ de V . El escalar $\lambda \in \mathbb{K}$ es autovalor de f si y sólo si

$$\det(A - \lambda I_n) = 0, \tag{4.5}$$

donde I_n es la matriz identidad $n \times n$.

Demostración. Sea $\lambda \in \text{Spec}(f)$. Entonces $\exists \mathbf{v} \neq \mathbf{0}$ tal que $f(\mathbf{v}) = \lambda\mathbf{v}$, es decir

$$f(\mathbf{v}) - \lambda\mathbf{v} = \mathbf{0} \iff (f - \lambda\mathcal{I})(\mathbf{v}) = \mathbf{0}. \tag{4.6}$$

donde \mathcal{I} es la aplicación identidad:

$$\mathcal{I}(\mathbf{v}) = \mathbf{v}, \quad \forall \mathbf{v} \in V.$$

Sea A la matriz asociada a f , e I_n la matriz identidad (que es la asociada a \mathcal{I}). Entonces la relación (4.6), pasando a las matrices asociadas, implica

$$(A - \lambda I_n) X = 0 \quad (4.7)$$

Esta ecuación matricial representa un sistema lineal homogéneo en que el vector columna X representa las componentes (x_1, \dots, x_n) de \mathbf{v} en la base asignada. Dado que $\mathbf{v} \neq \mathbf{0}$, el sistema obviamente admite soluciones no triviales. Esto se cumple si y sólo si el determinante del sistema es nulo, es decir

$$\det(A - \lambda I_n) = 0 .$$

□

El teorema anterior nos proporciona un algoritmo para determinar los autovalores de un endomorfismo f . Sea $A = (a_{ij})$, $i, j = 1, \dots, n$ la matriz asociada a f respecto de una cierta base. La ecuación (4.5) entonces se escribe en la forma

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \cdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 . \quad (4.8)$$

Expandiendo el determinante obtenemos una ecuación algebraica de grado n en la variable λ , con coeficientes en \mathbb{K} , llamada *ecuación característica del endomorfismo f o de la matriz A* (o bien “*ecuación secular*”).

Definición 85. *El polinomio*

$$p(\lambda) = \det(A - \lambda I_n) \quad (4.9)$$

se dirá polinomio característico de la aplicación f o de la matriz asociada A .

Se puede demostrar que el polinomio característico asume la forma general

$$p(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \text{tr}(A) \lambda^{n-1} + \dots + \det(A). \quad (4.10)$$

donde $\text{tr}(A) = a_{11} + \dots + a_{nn}$ es la traza de A .

Definición 86. *Si λ es una raíz simple de $p(\lambda)$, se dirá que λ es un autovalor simple de f ; si λ es una raíz de multiplicidad h , diremos que λ es un autovalor de f de **multiplicidad algebraica** h .*

Por ejemplo, si

$$p(\lambda) = \lambda^2(\lambda - 4)^3(\lambda - 1)^4(\lambda - 5) ,$$

los autovalores son $\lambda_1 = 0, \lambda_2 = 4, \lambda_3 = 1, \lambda_4 = 5$. Además, tenemos que $\lambda_1 = 0$ es autovalor de multiplicidad 2, $\lambda_2 = 4$ de multiplicidad 3, $\lambda_3 = 1$ de multiplicidad 4 y $\lambda_4 = 5$ es autovalor simple.

Hemos definido la multiplicidad algebraica de un autovalor como su multiplicidad como raíz del polinomio característico. Existe también la multiplicidad geométrica de un autovalor, definida como la dimensión del subespacio propio asociado a λ .

Definición 87. Sea λ un autovalor de $f \in L(V)$. Definimos multiplicidad geométrica de λ a la dimensión del autoespacio asociado V_λ .

Ejemplo 4.1. Sea la matriz

$$A = \begin{pmatrix} 1 & -1 \\ -4 & 1 \end{pmatrix}.$$

La ecuación característica se escribe

$$\begin{vmatrix} 1-\lambda & -1 \\ -4 & 1-\lambda \end{vmatrix} = 0 \iff (1-\lambda)^2 - 4 = 0 \iff \lambda^2 - 2\lambda - 3 = 0.$$

Tenemos los dos autovalores $\lambda_1 = 3$, $\lambda_2 = -1$, ambos simples. Vamos a hallar los autoespacios asociados.

i) Sea $\lambda_1 = 3$. Los vectores del autoespacio V_3 satisfacen la relación

$(A - \lambda I)X = O$ para $\lambda = 3$ (véase la ec.(4.7)) :

$$\left(\begin{pmatrix} 1 & -1 \\ -4 & 1 \end{pmatrix} - 3 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} -2 & -1 \\ -4 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Obtenemos la ecuación $2x_1 + x_2 = 0$, así que

$$V_3 = \{(\alpha, -2\alpha) \mid \alpha \in \mathbb{R}\} \text{ es decir } V_3 = \text{lin}\{(1, -2)\}.$$

ii) Sea ahora $\lambda_2 = -1$. Análogamente al caso anterior, los vectores del autoespacio V_{-1} satisfacen la relación

$$\begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} 2x_1 - x_2 = 0 \\ -4x_1 + 2x_2 = 0 \end{cases}$$

Por tanto

$$V_{-1} = \{(\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\} \text{ es decir } V_{-1} = \text{lin}\{(1, 2)\}.$$

Una propiedad interesante de las matrices semejantes es la siguiente.

Teorema 61. Dos matrices semejantes poseen los mismos autovalores.

Demostración. Sean A y B dos matrices semejantes: es decir, $\exists P$ matriz invertible tal que

$$B = P^{-1} A P.$$

Consecuentemente,

$$B - \lambda I_n = P^{-1} A P - \lambda I_n = P^{-1} A P - P^{-1}(\lambda I_n) P = P^{-1}(A - \lambda I_n)P$$

Entonces, utilizando el teorema de Binet (es decir, $\det(AB) = \det(A)\det(B)$), tenemos

$$\det(B - \lambda I_n) = (\det P^{-1})[\det(A - \lambda I_n)]\det(P) =$$

$$\det(P^{-1}P)\det(A - \lambda I_n) = \det(I_n)\det(A - \lambda I_n) = \det(A - \lambda I_n).$$

Por tanto las ecuaciones

$$\det(A - \lambda I_n) = 0 \quad \text{y} \quad \det(B - \lambda I_n) = 0$$

tienen las mismas raíces.

4.4 Ejemplos

Ejemplo 4.2. Sea $f \in L(\mathbb{R}^2)$, y sea A su matriz asociada respecto de la base canónica:

$$A = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}.$$

La ecuación característica es

$$\det(A - \lambda I) = 0 \iff \begin{vmatrix} 1 - \lambda & -1 \\ 2 & -1 - \lambda \end{vmatrix} = 0 \iff (1 - \lambda)(-1 - \lambda) + 2 = 0 \iff \lambda^2 + 1 = 0.$$

Dado que $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, buscaremos soluciones reales de la ecuación característica. Sin embargo, $\lambda^2 + 1 = 0$ no admite soluciones reales. Por tanto, $\text{Spec}(f) = \emptyset$.

Sin embargo, si consideramos una aplicación lineal $f \in L(\mathbb{C}^2)$, entonces los autovalores complejos son admisibles. Tenemos:

$$\lambda^2 + 1 = 0 \iff (\lambda - i)(\lambda + i) = 0, \quad \text{Spec}(f) = \{i, -i\}.$$

Los autoespacios se determinan como antes:

i) Consideremos $\lambda_1 = i$.

$$(A - iI)X = O \iff \begin{pmatrix} 1 - i & -1 \\ 2 & -1 - i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} (1 - i)x_1 - x_2 = 0 \\ 2x_1 - (1 + i)x_2 = 0 \end{cases}$$

$$\begin{cases} x_1 = \left(\frac{1+i}{2}\right)x_2 \\ (1-i)\left(\frac{1+i}{2}\right)x_2 = x_2 \end{cases} \iff V_i = \left\{ \left(\left(\frac{1+i}{2} \right) \alpha, \alpha \right) \mid \alpha \in \mathbb{C} \right\} \quad \text{es decir} \quad V_i = \text{lin}\{(1, 1-i)\}$$

ii) Análogamente, para $\lambda_2 = -i$, resolviendo el sistema de ecuaciones lineales correspondientes obtenemos

$$V_{-i} = \left\{ \left(\left(\frac{1-i}{2} \right) \alpha, \alpha \right) \mid \alpha \in \mathbb{C} \right\} \quad \text{es decir} \quad V_{-i} = \text{lin}\{(1, 1+i)\}.$$

Ejemplo 4.3. Sea el espacio vectorial \mathbb{R}^3 , dotado de la base canónica, y sea $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la aplicación lineal definida por

$$f(x_1, x_2, x_3) = (x_1 + 2x_2 + 10x_3, 2x_1 + x_2 + 10x_3, -x_1 - x_2 - 6x_3).$$

Determinemos sus autovalores y autovectores. La matriz A de f respecto de la base canónica es

$$A = \begin{pmatrix} 1 & 2 & 10 \\ 2 & 1 & 10 \\ -1 & -1 & -6 \end{pmatrix}.$$

La ecuación característica es

$$\begin{vmatrix} 1-\lambda & 2 & 10 \\ 2 & 1-\lambda & 10 \\ -1 & -1 & -6-\lambda \end{vmatrix} = 0$$

es decir

$$(1-\lambda)\{(1-\lambda)(-6-\lambda)+10\}-2\{-2(6+\lambda)+10\}+10\{-2+1-\lambda\}=0$$

y simplificando

$$-(\lambda+1)^2(\lambda+2)=0.$$

Por tanto, tenemos los autovalores $\lambda_1 = -1$ (de multiplicidad algebraica 2) y $\lambda_2 = -2$ (autovalor simple). Vamos ahora a determinar los autoespacios correspondientes.

i) Hallemos V_{-1} . La ecuación matricial para los autovectores es

$$(A+I)X=O \iff \begin{pmatrix} 2 & 2 & 10 \\ 2 & 2 & 10 \\ -1 & -1 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

que nos proporciona la ecuación independiente $x_1+x_2+5x_3=0$. Esta es la ecuación intrínseca del autoespacio V_{-1} . Resolviéndola, obtenemos una base de él

$$V_{-1} = \text{lin}\{(1, -1, 0), (5, 0, -1)\}.$$

i) Hallemos V_{-2} . La ecuación matricial para los autovectores es

$$(A+2I)X=O \iff \begin{pmatrix} 3 & 2 & 10 \\ 2 & 3 & 10 \\ -1 & -1 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Obtenemos el sistema

$$\begin{cases} 3x_1+2x_2+10x_3=0, \\ 2x_1+3x_2+10x_3=0, \\ -x_1-x_2-4x_3=0 \end{cases} \iff \begin{cases} x_1+x_2+4x_3=0, \\ x_2+2x_3=0, \end{cases} \iff \begin{cases} x_1=-2x_3 \\ x_2=-2x_3 \end{cases}$$

$$V_{-2} = \{(-2\alpha, -2\alpha, \alpha) \mid \alpha \in \mathbb{R}\} = \text{lin}\{(2, 2, -1)\}.$$

Observación 4.1. Si $\lambda = 0 \in \text{Spec}(f)$, evidentemente la ecuación $f(\mathbf{v}) = \mathbf{0}$ admite soluciones no triviales, es decir su núcleo es no nulo. Por tanto, la aplicación no es invertible y consecuentemente el determinante de la matriz asociada a f (en una base cualesquiera) es nulo:

$$\lambda = 0 \in \text{Spec}(f) \iff \ker(f) \neq \{\mathbf{0}\} \iff f \text{ no inyectiva} \iff \det(A) = 0.$$

4.5 Independencia lineal

Otro resultado fundamental es el siguiente, conocido como **teorema de la independencia lineal**. Esencialmente, establece que un conjunto de vectores no nulos, cada uno de ellos extraído desde un autoespacio distinto, es linealmente independiente.

Teorema 62. *Sea V un espacio vectorial sobre \mathbb{K} , $f \in L(V)$, y sean $\lambda_1, \dots, \lambda_p \in \text{Spec}(f)$ p autovalores distintos. Sean $\mathbf{v}_1 \in V_{\lambda_1}, \dots, \mathbf{v}_p \in V_{\lambda_p}$ p autovectores no nulos de f . Entonces $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ es un conjunto de vectores linealmente independientes.*

Demostración. El resultado es una consecuencia directa del Teorema 59: la suma de autoespacios correspondientes a autovalores distintos es directa.

Sea $\mathbf{v}_1 \in V_{\lambda_1}$, $i = 1, \dots, p$. Si por absurdo $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ fueran linealmente dependientes, entonces al menos uno de ellos sería combinación lineal de los demás, por ejemplo:

$$\mathbf{v}_1 = \sum_{i=2}^p \alpha_i \mathbf{v}_i$$

Pero en este caso, evidentemente, \mathbf{v}_1 pertenecería tanto al autoespacio V_1 como al subespacio suma $V_{\lambda_2} + \dots + V_{\lambda_p}$. Por tanto

$$V_1 \cap (V_{\lambda_2} + \dots + V_{\lambda_p}) \neq \{\mathbf{0}\}$$

y esto es absurdo porque la suma $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_p}$ es directa y por tanto (como hemos visto) la intersección de cada subespacio con la suma de los demás debe contener sólo al vector nulo. \square

4.6 Multiplicidad algebraica y geométrica

El siguiente teorema aclara la relación entre las nociones de multiplicidad algebraica y multiplicidad geométrica.

Teorema 63 (Teorema de la multiplicidad). *Sea V un espacio vectorial sobre \mathbb{K} y $f \in L(V)$. Si λ es un autovalor de f de multiplicidad algebraica h , y V_λ el autoespacio correspondiente, entonces*

$$1 \leq \dim V_\lambda \leq h .$$

Demostración. Sea λ un autovalor de f . Entonces, por definición $\exists \mathbf{v} \neq \mathbf{0}$, y por tanto $\dim V_\lambda \geq 1$. Para demostrar que vale también la desigualdad

$$\dim V_\lambda \leq h$$

procederemos por absurdo. Supondremos entonces que

$$s = \dim V_\lambda \geq h + 1, \quad h + 1 \leq s \leq n .$$

Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_s\}$ una base de V_λ . El subespacio V_λ admite por lo menos un subespacio suplementario $W \subset V$, con $\dim W = n - s$. Sea $\{\mathbf{u}_1, \dots, \mathbf{u}_{n-s}\}$ una base de W . Entonces,

$$\{\mathbf{v}_1, \dots, \mathbf{v}_s, \mathbf{u}_1, \dots, \mathbf{u}_{n-s}\} \quad (4.11)$$

es una base de V . Tenemos que

$$f(\mathbf{v}_i) = \lambda \mathbf{v}_i, \quad i = 1, \dots, s$$

Por otro lado, los vectores $f(\mathbf{u}_k)$ serán combinaciones lineales de los vectores de la base (4.11), es decir

$$f(\mathbf{u}_k) = \sum_{i=1}^s a_{ik} \mathbf{v}_i + \sum_{j=1}^{n-s} b_{jk} \mathbf{u}_j, \quad k = 1, \dots, n-s$$

con sendos coeficientes $a_{ik}, b_{jk} \in \mathbb{K}$. Entonces la matriz asociada a f en la base (4.11) es

$$\begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1,n-s} \\ 0 & \lambda & & & \vdots & a_{21} & a_{22} & \dots & a_{2,n-s} \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ 0 & \dots & \dots & \dots & \lambda & a_{s1} & a_{s2} & \dots & a_{s,n-s} \\ 0 & \dots & \dots & \dots & 0 & b_{11} & b_{12} & \dots & b_{1,n-s} \\ \vdots & & & & \vdots & \vdots & & \dots & \vdots \\ 0 & \dots & \dots & \dots & 0 & b_{n-s,1} & b_{n-s,2} & \dots & b_{n-s,n-s} \end{pmatrix}$$

Claramente, desde el polinomio característico de f deducimos que λ es un autovalor múltiple de orden s con $s > h$, lo que es absurdo. En definitiva,

$$\dim V_\lambda \leq h.$$

□

4.7 Diagonalización de un endomorfismo

Definición 88. Sea V un espacio vectorial sobre \mathbb{K} , $\dim V = n$ y $f \in L(V)$. Diremos que el endomorfismo f es diagonalizable si existe una base de V tal que la matriz asociada a f , respecto de esta base, es diagonal.

Definición 89. Diremos que la matriz $A \in M_n(\mathbb{K})$ es diagonalizable si existe una matriz cuadrada $P \in M_n(\mathbb{K})$ invertible tal que la matriz $P^{-1}AP$ es diagonal.

Claramente, las dos condiciones

1. $f \in L(V)$ es diagonalizable
2. La matriz A asociada a f es diagonalizable

son equivalentes.

4.7.1 Primer criterio de diagonalización

Teorema 64. *Sea V un espacio vectorial sobre \mathbb{K} y $f \in L(V)$. El endomorfismo f es diagonalizable si y sólo si es posible hallar una base de V formada por autovectores de f .*

Demostración

Supongamos que exista en V una base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ formada por autovectores de f , es decir

$$f(\mathbf{e}_i) = \lambda_i \mathbf{e}_i, \quad i = 1, \dots, n \quad (4.12)$$

Entonces, la matriz de f respecto de la base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ es la matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_n \end{pmatrix} \quad (4.13)$$

y por tanto f es diagonalizable.

Vice versa, si respecto de la base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ la matriz asociada a f es (4.13), entonces las ecuaciones de f , es decir las imágenes de los vectores de una base, son exactamente las relaciones (4.12); consecuentemente, los vectores de base son autovectores.

4.7.2 Segundo criterio de diagonalización

Teorema 65. *Sea V un espacio vectorial sobre \mathbb{K} y $f \in L(V)$, y $\{\lambda_1, \dots, \lambda_m\}$ sus autovalores distintos. El endomorfismo f es diagonalizable si y sólo si la suma directa de sus autoespacios coincide con el espacio vectorial V :*

$$V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_m} \quad (4.14)$$

Demostración. \implies) Supongamos que f sea diagonalizable. Entonces, según el criterio anterior, existe una base $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ formada por autovectores de f . Sean V_{λ_i} , $i = 1, \dots, m$ los autoespacios correspondientes a los autovalores λ_i . Claramente, dado que los vectores \mathbf{e}_i pertenecen cada uno a un cierto subespacio, tenemos que

$$\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subset V_{\lambda_1} \cup \dots \cup V_{\lambda_m} .$$

Entonces, considerando los conjuntos de las combinaciones lineales asociadas, tenemos

$$L(\{\mathbf{e}_1, \dots, \mathbf{e}_n\}) \subset L(V_{\lambda_1} \cup \dots \cup V_{\lambda_m}) .$$

Esto equivale a la relación

$$V \subseteq V_{\lambda_1} \oplus \dots \oplus V_{\lambda_m}$$

dado que $V = L(\{\mathbf{e}_1, \dots, \mathbf{e}_n\})$, y por definición de suma de subespacios, tenemos $L(V_{\lambda_1} \cup \dots \cup V_{\lambda_m}) = V_{\lambda_1} + \dots + V_{\lambda_m}$; además la suma es directa como consecuencia del Teorema 59.

Por otro lado, siendo V_{λ_i} subespacios del mismo espacio vectorial V , obviamente

$$V \supseteq V_{\lambda_1} \oplus \dots \oplus V_{\lambda_m}$$

así que, en definitiva

$$V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_m} . \quad (4.15)$$

\Leftarrow) Supongamos ahora que f posee m autoespacios $V_{\lambda_1}, \dots, V_{\lambda_m}$, y que

$$V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_m} . \quad (4.16)$$

Sean entonces:

$\{\mathbf{e}_1, \dots, \mathbf{e}_{h_1}\}$ base de V_{λ_1} formada por autovectores correspondientes a λ_1

$\{\mathbf{e}_{h_1+1}, \dots, \mathbf{e}_{h_1+h_2}\}$ base de V_{λ_2} formada por autovectores correspondientes a λ_2 , etc.

Dado que vale la relación (4.16), entonces la unión de todas estas bases es una base de V , claramente formada por autovectores. Desde el primer criterio de diagonalización deducimos que f es diagonalizable. \square

Vamos a enunciar (sin demostración) el

4.7.3 Tercer criterio de diagonalización

Teorema 66. *Sea V un espacio vectorial sobre \mathbb{K} y $f \in L(V)$. El endomorfismo f es diagonalizable si y sólo si se cumplen las propiedades siguientes.*

1. *El polinomio característico se puede descomponer completamente en el cuerpo conmutativo \mathbb{K} .*

2. *Para cada autovalor λ_i , de multiplicidad algebraica h_i , resulta que*

$$\dim V_{\lambda_i} = h_i, \quad i = 1, \dots, m. \quad (4.17)$$

Una interesante consecuencia del tercer criterio es el siguiente

Teorema 67. *Sea V un espacio vectorial sobre \mathbb{K} , $\dim V = n$ y $f \in L(V)$. Si f tiene n autovalores distintos, entonces f es diagonalizable.*

Demostración. Dado que f tiene n autovalores distintos, su polinomio característico es completamente factorizable (descomponible) sobre \mathbb{K} . Además, $\dim V_{\lambda_i} = 1 = h_i$. Entonces se cumplen las hipótesis del tercer criterio de diagonalización.

Teorema 68. *Sea V un espacio vectorial sobre \mathbb{K} , $\dim V = n$. Sea $f \in L(V)$ un endomorfismo diagonalizable de V y A la matriz asociada a f respecto de la base $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ de V . Denotemos con $\lambda_1, \dots, \lambda_n$ los autovalores de f (algunos de los cuales posiblemente coincidentes) y con $\mathbf{v}_1, \dots, \mathbf{v}_n$ n autovectores asociados (si $\lambda_1 = \lambda_2$, se supondrá que \mathbf{v}_1 y \mathbf{v}_2 sean elegidos linealmente independientes en V_{λ_1}). Si P es la matriz cuyas columnas son las componentes*

de los vectores $\mathbf{v}_1, \dots, \mathbf{v}_n$ respecto de la base \mathcal{B} , y D es la matriz diagonal de los autovalores $\lambda_1, \dots, \lambda_n$, entonces

$$P^{-1}AP = D. \quad (4.18)$$

Demostración. Sean X_1, \dots, X_n los vectores columna de las componentes de $\mathbf{v}_1, \dots, \mathbf{v}_n$ respecto de la base \mathcal{B} . Resulta que

$$AX_1 = \lambda_1 X_1 \quad (4.19)$$

$$AX_2 = \lambda_2 X_2 \quad (4.20)$$

$$\vdots \quad (4.21)$$

$$AX_n = \lambda_n X_n \quad (4.22)$$

Entonces, si denotamos con P la matriz cuyas columnas son integradas por los vectores X_1, \dots, X_n , es decir las componentes de los vectores $\mathbf{v}_1, \dots, \mathbf{v}_n$ respecto de la base \mathcal{B} , obtenemos

$$AP = PD$$

Como la suma de los autoespacios es directa, entonces los autovectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ son independientes, y P es invertible. Desde la relación anterior, obtenemos

$$P^{-1}AP = D.$$

□

Proposición 26. Si A es una matriz diagonalizable y tiene autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ (distintos o no), entonces la matriz A^m es diagonalizable y tiene como autovalores

$$\lambda_1^m, \lambda_2^m, \dots, \lambda_n^m.$$

Demostración

$$\begin{aligned} A^m &= (PDP^{-1}) \cdot (PDP^{-1}) \dots (PDP^{-1}) = \\ &= PD(P^{-1} \cdot P)D(P^{-1} \cdot P)D \dots DP^{-1} = PD^m P^{-1} \end{aligned}$$

Por otro lado

$$D^m = \begin{pmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & 0 & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_n^m \end{pmatrix} \quad (4.23)$$

Entonces A^m es diagonalizable y tiene como autovalores $\lambda_1^m, \dots, \lambda_n^m$. □

4.8 Ejercicios

Ejercicio 4.4. Sea la matriz

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 3 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

Determinar sus autovalores y autovectores, y establecer si es diagonalizable.

Resolución. Los autovalores de A son las raíces de la ecuación característica:

$$\begin{vmatrix} 1 - \lambda & 0 & 3 \\ 0 & 3 - \lambda & 0 \\ 3 & 0 & 1 - \lambda \end{vmatrix} = 0 \iff -(\lambda + 2)(\lambda - 3)(\lambda - 4) = 0$$

Deducimos que los autovalores son

$$\lambda_1 = -2, \quad \lambda_2 = 3, \quad \lambda_3 = 4.$$

Entonces A es diagonalizable, porque posee tres autovalores distintos (véase el Teorema (67)). Vamos a construir los autoespacios asociados.

i) $\lambda_1 = -2$.

$$\begin{pmatrix} 3 & 0 & 3 \\ 0 & 5 & 0 \\ 3 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Obtenemos el sistema $\begin{cases} 3x + 3z = 0 \\ y = 0 \end{cases} \iff \begin{cases} x = -z \\ y = 0 \end{cases}$.

Por tanto,

$$V_{-2} = \text{lin}\{(1, 0, -1)\}.$$

ii) $\lambda_2 = 3$.

$$\begin{pmatrix} -2 & 0 & 3 \\ 0 & 0 & 0 \\ 3 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Resolviendo este sistema, obtenemos

$$V_3 = \text{lin}\{(0, 1, 0)\}.$$

iii) $\lambda_3 = 4$.

$$\begin{pmatrix} -3 & 0 & 3 \\ 0 & -1 & 0 \\ 3 & 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Obtenemos

$$V_4 = \text{lin}\{(1, 0, 1)\} .$$

La matriz de paso es

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

Entonces $A = PDP^{-1}$, con

$$D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} .$$

Ejercicio 4.5. Determinar el valor del parámetro a tal que la matriz

$$A = \begin{pmatrix} 1 & a & a \\ -1 & 1 & -1 \\ 1 & 0 & 2 \end{pmatrix}$$

sea diagonalizable.

Resolución. Escribimos la ecuación característica:

$$\begin{vmatrix} 1 - \lambda & a & a \\ -1 & 1 - \lambda & -1 \\ 1 & 0 & 2 - \lambda \end{vmatrix} = 0 \iff -(\lambda - 1)^2(\lambda - 2) = 0 ,$$

después de simplificaciones algebraicas elementales. Obtenemos los autovalores

$$\begin{aligned} \lambda_1 &= 1, & h_1 &= 2 \\ \lambda_2 &= 2, & h_2 &= 1 \end{aligned}$$

donde h_1 y h_2 son las multiplicidades algebraicas de λ_1 y λ_2 , respectivamente.

i) $\lambda_1 = 1$.

$$\begin{pmatrix} 0 & a & a \\ -1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Obtenemos el sistema $\begin{cases} a(y + z) = 0 \\ x + z = 0 \end{cases}$.

Tenemos dos casos.

i₁) $a = 0 \implies x = -z$. Por tanto,

$$V_1 = \text{lin}\{(1, 0, -1), (0, 1, 0)\} \quad \text{y} \quad \dim V_1 = 2 = h_1 .$$

$$i_2) a \neq 0 \implies \begin{cases} x = -z \\ y = -z \end{cases} .$$

Entonces

$$V_1 = \text{lin}\{(1, 1, -1)\} \quad y \quad \dim V_1 = 1 < h_1 .$$

i) $\lambda_2 = 2$.

$$\begin{pmatrix} -1 & a & a \\ -1 & -1 & -1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} -x + a(y+z) = 0 \\ x + y + z = 0 \\ x = 0 \end{cases} \iff \begin{cases} y = -z \\ x = 0 \end{cases} .$$

Entonces

$$V_2 = \text{lin}\{(0, 1, -1)\} \quad y \quad \dim V_2 = 1 = h_2 .$$

En definitiva, la matriz A es diagonalizable si y solo si $a = 0$. La matriz que diagonaliza A es la matriz

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} .$$

Como sabemos, P es invertible y satisface la relación $P^{-1}AP = D$, con

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

4.9 Teorema de Caley-Hamilton y polinomio mínimo

El siguiente resultado nos permite construir identidades polinomiales que involucran potencias de un endomorfismo f .

Teorema 69 (Caley-Hamilton). *Sea V un espacio vectorial sobre \mathbb{R} o \mathbb{C} , $\dim V = n$ y sea $f \in L(V)$. Entonces, si $p(x)$ es el polinomio característico de f , tenemos*

$$p(f) = 0 . \tag{4.24}$$

Como consecuencia inmediata de este teorema, si A es la matriz asociada a f en una cierta base, tenemos que

$$p(A) = 0 \tag{4.25}$$

En otras palabras, f (o equiv. A) anula a su polinomio característico.

Ejemplo 4.6. Sea V espacio vectorial sobre \mathbb{R} , $\dim V = 2$. Sea $f \in L(V)$, y

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

su matriz asociada respecto de una base de V . El polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 2 \\ 3 & 4 - \lambda \end{vmatrix} \iff \lambda^2 - 5\lambda - 2.$$

Por tanto, el teorema de Caley-Hamilton nos asegura que

$$p(A) = A^2 - 5A - 2I = 0$$

como podemos fácilmente comprobar.

Observación. El polinomio característico no es el único polinomio que anula f , y en general tampoco es él de grado menor.

Definición 90. Llamaremos *polinomio mínimo* de $f \in L(V)$ al polinomio mónico (es decir, el coeficiente del término dominante es 1) de grado menor entre todos los que satisfacen la propiedad

$$p(f) = 0.$$

Este polinomio se denotará con $m(\lambda)$.

Se puede demostrar que

Proposición 27. El polinomio característico es un múltiplo del polinomio mínimo.

Entonces, dado que el polinomio mínimo divide el polinomio característico, cada factor lineal del polinomio característico tiene que aparecer en el polinomio mínimo. En general las raíces del polinomio mínimo también son autovalores.

La noción de polinomio mínimo nos proporciona un criterio ulterior de diagonalización.

Teorema 70. Sea V un espacio vectorial sobre \mathbb{C} , $\dim V = n$, y $f \in L(V)$. Entonces f es diagonalizable si y sólo si las raíces del polinomio mínimo tienen multiplicidad igual a 1:

$$f \text{ diagonal} \iff m(\lambda) = (\lambda - \lambda_1) \cdot \dots \cdot (\lambda - \lambda_p).$$

Si f es diagonalizable, entonces podemos construir inmediatamente el polinomio mínimo utilizando el resultado anterior: si $\lambda_1, \dots, \lambda_p$ son los autovalores distintos de f , $m(\lambda) = (\lambda - \lambda_1) \cdot \dots \cdot (\lambda - \lambda_p)$.

Chapter 5

Formas bilineales y cuadráticas. Productos escalares

5.1 Aplicaciones y formas bilineales

Definición 91. Sean E, F, G tres espacios vectoriales sobre el mismo cuerpo \mathbb{K} . Llamaremos *aplicación bilineal de $E \times F$ en G* a una aplicación

$$\varphi : E \times F \rightarrow G$$

tal que

1°

$$\forall \mathbf{x}, \mathbf{x}' \in E, \quad \forall \mathbf{y} \in F, \quad \varphi(\mathbf{x} + \mathbf{x}', \mathbf{y}) = \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{x}', \mathbf{y})$$

$$\forall \mathbf{x} \in E, \quad \forall \mathbf{y}, \mathbf{y}' \in F, \quad \varphi(\mathbf{x}, \mathbf{y} + \mathbf{y}') = \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{x}, \mathbf{y}')$$

2°

$$\forall \mathbf{x} \in E, \quad \forall \mathbf{y} \in F, \quad \forall \lambda \in \mathbb{K}, \quad \varphi(\lambda \mathbf{x}, \mathbf{y}) = \varphi(\mathbf{x}, \lambda \mathbf{y}) = \lambda \varphi(\mathbf{x}, \mathbf{y}) .$$

En el caso particular $E = F$, obtenemos la definición de aplicación 2-lineal de $E \times E$ en G que ya conocemos.

Denotaremos con $\mathcal{B}(E \times F, G)$ el conjunto de todas las aplicaciones bilineales de $E \times F$ en G , y con $\mathcal{B}(E, G)$ (o bien $\mathcal{B}(E \times E, G)$) el conjunto de todas las aplicaciones bilineales de $E \times E$ en G .

Podemos introducir dos leyes de composición en $\mathcal{B}(E \times F, G)$. La suma de dos aplicaciones $\varphi_1, \varphi_2 \in \mathcal{B}(E \times F, G)$ es la aplicación

$$\varphi_1 + \varphi_2 : E \times F \longrightarrow G$$

definida por

$$\forall \mathbf{x} \in E, \quad \forall \mathbf{y} \in F, \quad (\varphi_1 + \varphi_2)(\mathbf{x}, \mathbf{y}) := \varphi_1(\mathbf{x}, \mathbf{y}) + \varphi_2(\mathbf{x}, \mathbf{y}) .$$

El producto del escalar $\alpha \in \mathbb{K}$ por la aplicación φ es la aplicación $\alpha\varphi : E \times F \rightarrow G$, definida por

$$\forall \mathbf{x} \in E, \quad \forall \mathbf{y} \in F, \quad (\alpha\varphi)(\mathbf{x}, \mathbf{y}) := \alpha\varphi(\mathbf{x}, \mathbf{y}) .$$

Es inmediato comprobar la siguiente

Proposición 28. Si $\varphi_1, \varphi_2 \in \mathcal{B}(E \times F, G)$ y $\alpha \in \mathbb{K}$, entonces

$$\varphi_1 + \varphi_2 \in \mathcal{B}(E \times F, G), \quad \alpha\varphi \in \mathcal{B}(E \times F, G) .$$

También es inmediata la demostración del

Teorema 71. El conjunto $(\mathcal{B}(E \times F, G), +, \cdot)$ es un espacio vectorial sobre \mathbb{K} .

Como siempre, cuando el espacio vectorial codominio es \mathbb{K} , hablaremos de *formas* en lugar de aplicaciones.

Definición 92. Una aplicación bilineal de $E \times F \rightarrow \mathbb{K}$ se dirá *forma bilineal de $E \times F$* . Una aplicación bilineal $\varphi \in \mathcal{B}(E, \mathbb{K})$ se dirá *forma bilineal de E* .

Definición 93. Una forma bilineal $\varphi \in \mathcal{B}(E, \mathbb{K})$ se dirá *simétrica* si

$$\forall \mathbf{x}, \mathbf{y} \in E, \quad \varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x}) .$$

Ejemplo 5.1. En \mathbb{R}^3 :

$$\varphi(\mathbf{x}, \mathbf{y}) = 2x_1y_1 + 3x_2y_2 + x_3y_3$$

es una forma bilineal simétrica.

Proposición 29. El conjunto $\mathcal{B}_S(E, \mathbb{K})$ de todas las formas bilineales simétricas de E es un subespacio vectorial de $\mathcal{B}(E, \mathbb{K})$.

Demostración. Sean $\varphi_1, \varphi_2 \in \mathcal{B}_S(E, \mathbb{K})$, $\alpha_1, \alpha_2 \in \mathbb{K}$. Entonces

$$\begin{aligned} \forall \mathbf{x}, \mathbf{y} \in E : (\alpha_1\varphi_1 + \alpha_2\varphi_2)(\mathbf{x}, \mathbf{y}) &= \alpha_1\varphi_1(\mathbf{x}, \mathbf{y}) + \alpha_2\varphi_2(\mathbf{x}, \mathbf{y}) = \\ &= \alpha_1\varphi_1(\mathbf{y}, \mathbf{x}) + \alpha_2\varphi_2(\mathbf{y}, \mathbf{x}) = (\alpha_1\varphi_1 + \alpha_2\varphi_2)(\mathbf{y}, \mathbf{x}) . \end{aligned}$$

□

Sea $\varphi \in \mathcal{B}(E \times F, \mathbb{K})$. Supongamos que $\dim E = n$, $\dim F = p$, y sea $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de E , $\{\mathbf{f}_1, \dots, \mathbf{f}_p\}$ una base de F . Entonces todo vector

de E y F se puede descomponer como combinación lineal de la base correspondiente: $\forall \mathbf{x} \in E, \forall \mathbf{y} \in F$, tenemos

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{j=1}^p y_j \mathbf{f}_j .$$

Por tanto

$$\varphi(\mathbf{x}, \mathbf{y}) = \varphi\left(\sum_{i=1}^n x_i \mathbf{e}_i, \sum_{j=1}^p y_j \mathbf{f}_j\right) = \sum_{i=1}^n \sum_{j=1}^p x_i y_j \varphi(\mathbf{e}_i, \mathbf{f}_j) .$$

Introducimos la matriz

$$a_{ij} = \varphi(\mathbf{e}_i, \mathbf{f}_j), \quad i = 1, \dots, n, \quad j = 1, \dots, p .$$

Obtenemos entonces

$$\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^p x_i y_j a_{ij} .$$

De esta manera hemos asociado a la forma bilineal φ la matriz $A = (a_{ij}) \in M_{n,p}(\mathbb{K})$, respecto de las bases $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ de E y $\{\mathbf{f}_1, \dots, \mathbf{f}_p\}$ de F .

Observación 5.1. Se puede notar que la correspondencia

$$\varphi \in \mathcal{B}(E \times F, \mathbb{K}) \longrightarrow A \in M_{n,p}(\mathbb{K})$$

es un isomorfismo de espacios vectoriales. Como consecuencia, tenemos que la dimensión del espacio vectorial $\mathcal{B}(E \times F, \mathbb{K})$ es np .

Teorema 72. Sean E y F dos espacios vectoriales sobre \mathbb{K} , $\dim E = n$, $\dim F = p$ y sea A la matriz asociada a $\varphi \in \mathcal{B}(E \times F, \mathbb{K})$ respecto de las bases $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ de E y $\{\mathbf{f}_1, \dots, \mathbf{f}_p\}$ de F . Denotemos por

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}$$

las matrices columnas de las componentes de los vectores $\mathbf{x} \in E$ y $\mathbf{y} \in F$. Entonces

$$\varphi(\mathbf{x}, \mathbf{y}) = X^t A Y \quad \text{o bien} \quad \varphi(\mathbf{x}, \mathbf{y}) = Y^t A^t X .$$

Demostración. Es suficiente observar que la expresión

$$\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^p x_i y_j a_{ij}$$

se puede reescribir utilizando la relación

$$X^t AY = (x_1, \dots, x_n) \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^p x_i y_j a_{ij}$$

que representa un escalar en \mathbb{K} , es decir una matriz 1×1 . Ahora bien, dado que la transpuesta de esta matriz es la matriz misma, concluimos que

$$\varphi(\mathbf{x}, \mathbf{y}) = (X^t AY)^t = Y^t A^t X .$$

□

Teorema 73. Sean E y F dos espacios vectoriales sobre \mathbb{K} , $\dim E = n$, $\dim F = p$, y sea A la matriz asociada a la forma bilineal $\varphi \in \mathcal{B}(E \times F, \mathbb{K})$ respecto de las bases \mathcal{B} de E y \mathcal{C} de F . Sean \mathcal{B}' y \mathcal{C}' dos nuevas bases de E y F respectivamente, y P la matriz de cambio de base de \mathcal{B} a \mathcal{B}' , Q la matriz de cambio de \mathcal{C} a \mathcal{C}' . Entonces la matriz de φ respecto de \mathcal{B}' y \mathcal{C}' es

$$A' = P^t A Q . \quad (5.1)$$

Demostración. Sabemos que las matrices de cambio de base verifican las relaciones

$$X = P X' , \quad Y = Q Y' .$$

Entonces

$$\begin{aligned} \varphi(\mathbf{x}, \mathbf{y}) &= X^t AY = (P X')^t A (Q Y') = \\ &= (X')^t P^t A Q Y' = (X')^t (P^t A Q) Y' \\ &= (X')^t A' Y' \end{aligned}$$

es decir

$$A' = P^t A Q .$$

□

Corolario 12. Si $\varphi \in \mathcal{B}(E, \mathbb{K})$, entonces

$$A' = P^t A P \quad (5.2)$$

Definición 94. Dos matrices cuadradas A y A' se dirán congruentes si existe otra matriz P invertible tal que

$$A' = P^t A P .$$

Como $\det P = \det(P^t) \neq 0$, tenemos que

$$\det A' = (\det P)^2 \det A$$

y por tanto $\det A$ y $\det A'$ tienen el mismo signo.

Proposición 30. La matriz A de una forma bilineal simétrica $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ es simétrica.

Demostración. Es suficiente observar que

$$a_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j) = \varphi(\mathbf{e}_j, \mathbf{e}_i) = a_{ji} .$$

□

Definición 95. Se dice que $\varphi \in \mathcal{B}(E, \mathbb{K})$ es una forma bilineal antisimétrica si

$$\varphi(\mathbf{x}, \mathbf{y}) = -\varphi(\mathbf{y}, \mathbf{x})$$

Ejemplo 5.2. En \mathbb{R}^3 :

$$\varphi(\mathbf{x}, \mathbf{y}) = a(x_1y_2 - x_2y_1) + b(x_1y_3 - x_3y_1) + c(x_2y_3 - x_3y_2), \quad a, b, c \in \mathbb{R}$$

es una forma bilineal antisimétrica.

Se demuestra fácilmente que la matriz asociada a una forma bilineal antisimétrica es antisimétrica:

$$a_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j) = -\varphi(\mathbf{e}_j, \mathbf{e}_i) = -a_{ji} .$$

Por tanto $A = -A^t$. Denotaremos con $\mathcal{B}_A(E, \mathbb{K})$ el conjunto de todas las formas bilineales antisimétricas de E .

Lema 6. El conjunto $\mathcal{B}_A(E, \mathbb{K})$ es un subespacio vectorial de $\mathcal{B}(E, \mathbb{K})$.

Demostración. Sean $\varphi_1, \varphi_2 \in \mathcal{B}_A(E, \mathbb{K})$. Tenemos

$$\begin{aligned} (\alpha_1\varphi_1 + \alpha_2\varphi_2)(\mathbf{x}, \mathbf{y}) &= \alpha_1\varphi_1(\mathbf{x}, \mathbf{y}) + \alpha_2\varphi_2(\mathbf{x}, \mathbf{y}) = -\alpha_1\varphi_1(\mathbf{y}, \mathbf{x}) - \alpha_2\varphi_2(\mathbf{y}, \mathbf{x}) = \\ &= -(\alpha_1\varphi_1 + \alpha_2\varphi_2)(\mathbf{y}, \mathbf{x}) \end{aligned}$$

y por tanto $\mathcal{B}_A(E, \mathbb{K})$ es un subespacio de $\mathcal{B}(E, \mathbb{K})$. □

Proposición 31. El espacio $\mathcal{B}(E, \mathbb{K})$ se descompone en suma directa de formas bilineales simétricas y antisimétricas:

$$\mathcal{B}(E, \mathbb{K}) = \mathcal{B}_S(E, \mathbb{K}) \oplus \mathcal{B}_A(E, \mathbb{K})$$

y

$$\dim \mathcal{B}_S(E, \mathbb{K}) = \frac{n(n+1)}{2}, \quad \dim \mathcal{B}_A(E, \mathbb{K}) = \frac{n(n-1)}{2} .$$

Demostración. Tenemos que

$$\varphi(\mathbf{x}, \mathbf{y}) = \frac{1}{2} [\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})] - \frac{1}{2} [\varphi(\mathbf{x}, \mathbf{y}) - \varphi(\mathbf{y}, \mathbf{x})] = \varphi_s(\mathbf{x}, \mathbf{y}) + \varphi_a(\mathbf{x}, \mathbf{y}).$$

Entonces una forma arbitraria se descompone en la suma de una forma simétrica y una antisimétrica. También observamos que si $\varphi \in \mathcal{B}_S(E, \mathbb{K}) \cap \mathcal{B}_A(E, \mathbb{K})$ (es decir φ es una forma al mismo tiempo simétrica e antisimétrica), evidentemente

$$\varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x}) \quad \text{y} \quad \varphi(\mathbf{x}, \mathbf{y}) = -\varphi(\mathbf{y}, \mathbf{x}) \implies \varphi = 0 .$$

La demostración de la segunda propiedad descende del hecho que el subespacio vectorial de las matrices simétricas $A \in M_n(\mathbb{K})$ tiene dimensión $\frac{n(n+1)}{2}$; análogamente para el caso antisimétrico. \square

Proposición 32. *Sea E espacio vectorial sobre $\mathbb{K} = \mathbb{R}$ o \mathbb{C} , y $\varphi \in \mathcal{B}(E, \mathbb{K})$. Entonces*

$$\varphi \text{ es antisimétrica} \iff \varphi(\mathbf{x}, \mathbf{x}) = 0 \quad \forall \mathbf{x} \in E .$$

Demostración. Si φ es antisimétrica,

$$\varphi(\mathbf{x}, \mathbf{x}) = -\varphi(\mathbf{x}, \mathbf{x}) \iff 2\varphi(\mathbf{x}, \mathbf{x}) = 0 \iff \varphi(\mathbf{x}, \mathbf{x}) = 0 .$$

Supongamos ahora $\varphi(\mathbf{x}, \mathbf{x}) = 0 \quad \forall \mathbf{x} \in \mathbb{R}$. Puesto que

$$\varphi(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}, \mathbf{x}) + \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x}) + \varphi(\mathbf{y}, \mathbf{y})$$

se tiene

$$0 = \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})$$

es decir $\varphi(\mathbf{x}, \mathbf{y}) = -\varphi(\mathbf{y}, \mathbf{x})$. \square

Teorema 74. *Sea E un espacio vectorial sobre \mathbb{K} , $\varphi \in \mathcal{B}(E, \mathbb{K})$ y sea A la matriz asociada a φ respecto de una base asignada de E . Entonces:*

1. φ es simétrica $\iff A$ es una matriz simétrica.
2. φ es antisimétrica $\iff A$ es una matriz antisimétrica.

Demostración. Las implicaciones \implies ya han sido demostradas.

\Leftarrow) Sea $A = A^t$. Entonces

$$\varphi(\mathbf{x}, \mathbf{y}) = X^t A Y = (X^t A Y)^t = Y^t A^t X = Y^t A X = \varphi(\mathbf{y}, \mathbf{x}) .$$

De forma completamente análoga se demuestra el caso antisimétrico. \square

5.2 Formas cuadráticas

5.2.1 Propiedades básicas

Definición 96. *Sea $\varphi \in \mathcal{B}(E, \mathbb{K})$. La aplicación*

$$\Phi : E \longrightarrow \mathbb{K}$$

tal que

$$\forall \mathbf{x} \in E : \Phi(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x})$$

es llamada la forma cuadrática asociada a la forma bilineal φ . Si $\mathbb{K} = \mathbb{R}$, la forma Φ se dirá forma cuadrática real.

Denotaremos con $Q(E, \mathbb{K})$ el conjunto de las formas cuadráticas de E en \mathbb{K} .

N. B. En algunos textos se prefiere definir una forma cuadrática a partir de una forma bilineal *simétrica*.

Proposición 33. Si $\Phi \in Q(E, \mathbb{K})$, entonces

$$1^\circ \quad \forall \mathbf{x} \in E, \quad \forall \lambda \in \mathbb{K}, \quad \Phi(\lambda \mathbf{x}) = \lambda^2 \Phi(\mathbf{x}) \quad (5.3)$$

$$2^\circ \quad \Phi(\mathbf{0}) = 0 \quad (5.4)$$

Demostración. Sea $\Phi(\mathbf{x})$ la forma cuadrática asociada a φ . Tenemos

$$\Phi(\lambda \mathbf{x}) = \varphi(\lambda \mathbf{x}, \lambda \mathbf{x}) = \lambda^2 \varphi(\mathbf{x}, \mathbf{x}) = \lambda^2 \Phi(\mathbf{x})$$

En particular, cuando $\lambda = 0$, tenemos

$$\Phi(0\mathbf{x}) = 0 \cdot \Phi(\mathbf{x}) = 0$$

es decir

$$\Phi(\mathbf{0}) = 0 .$$

Observamos que, si $\psi \in \mathcal{B}_A(E, \mathbb{K})$ es una forma bilineal antisimétrica, $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$, entonces la forma cuadrática

$$\Phi(\mathbf{x}) = \psi(\mathbf{x}, \mathbf{x}) = 0 \quad \forall \mathbf{x} \in E$$

es decir, la forma cuadrática asociada a una forma antisimétrica es idénticamente nula.

Proposición 34. Sea $\Phi \in Q(E, \mathbb{K})$ la forma cuadrática asociada a una forma bilineal φ . Entonces

$$\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y}) + \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in E . \quad (5.5)$$

Demostración. Es una consecuencia directa de la definición de Φ . Si $\mathbf{x}, \mathbf{y} \in E$,

$$\begin{aligned} \Phi(\mathbf{x} + \mathbf{y}) &= \varphi(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}, \mathbf{x}) + \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x}) + \varphi(\mathbf{y}, \mathbf{y}) \\ &= \Phi(\mathbf{x}) + \Phi(\mathbf{y}) + \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x}) . \end{aligned}$$

Observación 5.2. Formas bilineales distintas pueden dar lugar a la misma forma cuadrática. Por ejemplo, sea $\varphi \in \mathcal{B}(\mathbb{R}^2, \mathbb{R})$ definida por

$$\varphi[(x_1, x_2), (y_1, y_2)] = x_1 y_2 + x_2 y_1 .$$

La forma cuadrática asociada es $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$\Phi(x_1, x_2) = 2x_1x_2 .$$

Sea ahora $\chi \in \mathcal{B}(\mathbb{R}^2, \mathbb{R})$ definida por

$$\chi[(x_1, x_2), (y_1, y_2)] = 2x_1y_2 .$$

La forma cuadrática asociada es claramente la misma:

$$\Phi(x_1, x_2) = 2x_1x_2 .$$

Además, si $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ y Φ es la forma cuadrática asociada, entonces para cada forma bilineal antisimétrica $\psi \in \mathcal{B}_A(E, \mathbb{K})$ resulta que

$$\Phi(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}) + \psi(\mathbf{x}, \mathbf{x})$$

ya que $\psi(\mathbf{x}, \mathbf{x}) = 0$. En otras palabras, Φ es la forma cuadrática asociada también a $\varphi + \psi$: la forma cuadrática asociada a una forma bilineal solo depende de la parte simétrica de esta.

De entre todas las infinitas formas bilineales que dan lugar a la misma forma cuadrática, solo existe una que sea simétrica. Con más precisión, vale el siguiente

Teorema 75. *Sea $\Phi \in Q(E, \mathbb{K})$. Existe una única forma bilineal simétrica $\varphi_P \in \mathcal{B}_S(E, \mathbb{K})$ cuya forma cuadrática asociada es Φ .*

Demostración. Sea φ una forma bilineal, que tiene a Φ como su forma cuadrática asociada. Entonces desde la ec. (5.5) obtenemos

$$\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x}) = \Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x}) - \Phi(\mathbf{y}) .$$

Impongamos que φ sea simétrica; entonces queda

$$2\varphi(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x}) - \Phi(\mathbf{y})$$

con lo cual φ resulta unívocamente determinada por la fórmula

$$\varphi(\mathbf{x}, \mathbf{y}) = \frac{1}{2} [\Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x}) - \Phi(\mathbf{y})] . \quad (5.6)$$

□

5.2.2 Forma polar de una forma cuadrática

Definición 97. *La forma bilineal simétrica definida por*

$$\varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{2} [\Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x}) - \Phi(\mathbf{y})] . \quad (5.7)$$

se dirá la forma polar de la forma cuadrática Φ .

Corolario 13. Sea $\varphi \in \mathcal{B}(E, \mathbb{K})$ y $\Phi \in Q(E, \mathbb{K})$ la forma cuadrática asociada. La forma polar φ_P de Φ se puede obtener como:

$$1) \quad \varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[\Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x}) - \Phi(\mathbf{y})] \quad (5.8)$$

$$2) \quad \varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{4}[\Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x} - \mathbf{y})] \quad (5.9)$$

$$3) \quad \varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})] . \quad (5.10)$$

Demostración.

1) Obvia.

2) Tenemos que

$$\Phi(\mathbf{x} - \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y}) - 2\varphi_P(\mathbf{x}, \mathbf{y})$$

y dado que

$$\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y}) + 2\varphi_P(\mathbf{x}, \mathbf{y})$$

obtenemos

$$4\varphi_P(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x} - \mathbf{y}) \iff \\ \varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{4}[\Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x} - \mathbf{y})] .$$

3) Si Φ es la forma cuadrática asociada a φ , entonces

$$\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y}) + \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})$$

Dado que

$$\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y}) + 2\varphi_P(\mathbf{x}, \mathbf{y})$$

deducimos inmediatamente que

$$\varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})] .$$

□

5.2.3 Matriz asociada a Φ

Sea E un espacio vectorial sobre \mathbb{K} , $\dim E = n$, $\mathbf{x} \in E$ y sean x_1, \dots, x_n las componentes de \mathbf{x} respecto de una base $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ de E . Sea $\Phi \in Q(E, \mathbb{K})$.

Definición 98. Llamaremos matriz asociada a la forma cuadrática Φ respecto de la base B a la matriz asociada (respecto de la misma base) a la forma polar de Φ .

Claramente, si $\varphi \in \mathcal{B}_S(E, \mathbb{K})$, la matriz asociada a φ y la asociada a Φ coinciden. En general, tenemos

$$\Phi(\mathbf{x}) = \sum_{i,j=1}^n a_{ij}x_i x_j$$

Si X es el vector columna de las componentes de \mathbf{x} en la base asignada, podemos escribir equivalentemente

$$\Phi(\mathbf{x}) = X^t A X .$$

Observación 5.3. La matriz A asociada a Φ por definición es *simétrica*. Entonces, dado que $a_{ij} = a_{ji}$, resulta que

$$\Phi(\mathbf{x}) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{i < j} a_{ij} x_i x_j .$$

Si $\dim E = 2$, se obtiene la forma general

$$\Phi(x_1, x_2) = a_{11} x_1^2 + a_{22} x_2^2 + 2a_{12} x_1 x_2 . \quad (5.11)$$

Si $\dim E = 3$,

$$\Phi(x_1, x_2, x_3) = a_{11} x_1^2 + a_{22} x_2^2 + a_{33} x_3^2 + 2a_{12} x_1 x_2 + 2a_{13} x_1 x_3 + 2a_{23} x_2 x_3 . \quad (5.12)$$

La matriz asociada a una forma cuadrática se deduce inmediatamente comparando su expresión explícita con estas fórmulas.

Ejemplo 5.3. Consideremos la forma cuadrática $\Phi \in Q(\mathbb{R}^3, \mathbb{R})$ definida por

$$\Phi(x, y, z) = x^2 + z^2 - 2xy - 4xz .$$

Entonces

$$A = \begin{pmatrix} 1 & -1 & -2 \\ -1 & 0 & 0 \\ -2 & 0 & 1 \end{pmatrix} .$$

Definición 99. Denominaremos *rango de la forma cuadrática Φ* al rango de la matriz A . Llamaremos *discriminante de Φ* al determinante de A .

5.3 Vectores conjugados

Introducimos ahora la noción de vectores conjugados respecto de una forma bilineal simétrica.

Definición 100. Dos vectores \mathbf{x} e \mathbf{y} de E se dirán *conjugados respecto de la forma bilineal simétrica $\varphi \in \mathcal{B}_S(E, \mathbb{K})$* si

$$\varphi(\mathbf{x}, \mathbf{y}) = 0 .$$

5.3.1 Propiedades de los vectores conjugados

Proposición 35. El vector nulo es conjugado con cualquier vector $\mathbf{x} \in E$ respecto de cualquier forma $\varphi \in \mathcal{B}_S(E, \mathbb{K})$.

Demostración. $\forall \mathbf{x}, \mathbf{y} \in E, \forall \lambda \in \mathbb{K}$,

$$\varphi(\mathbf{x}, \lambda \mathbf{y}) = \lambda \varphi(\mathbf{x}, \mathbf{y}) .$$

Para $\lambda = 0$, obtenemos

$$\forall \mathbf{x} \in E, \quad \varphi(\mathbf{x}, \mathbf{0}) = 0 .$$

□

Teorema 76. *Sea E un espacio vectorial sobre \mathbb{K} y $W \subset E$ un subespacio generado por los vectores $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$. Si el vector $\mathbf{x} \in E$ es conjugado, respecto de una forma $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ con todos los vectores $\mathbf{u}_1, \dots, \mathbf{u}_p$, entonces \mathbf{x} es conjugado con todo vector de W .*

Demostración. Por hipótesis,

$$\varphi(\mathbf{x}, \mathbf{u}_i) = 0, \quad i = 1, \dots, p.$$

Dado que para todo $\mathbf{y} \in W$ resulta que $\mathbf{y} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_p \mathbf{u}_p$, entonces

$$\varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{x}, \alpha_1 \mathbf{u}_1 + \dots + \alpha_p \mathbf{u}_p) = \alpha_1 \varphi(\mathbf{x}, \mathbf{u}_1) + \dots + \alpha_p \varphi(\mathbf{x}, \mathbf{u}_p) = 0$$

es decir, \mathbf{x} es conjugado con todo $\mathbf{y} \in W$ respecto de la forma bilineal φ . □

Definición 101. *Sea E un espacio vectorial sobre \mathbb{K} y $\varphi \in \mathcal{B}_S(E, \mathbb{K})$. Dos conjuntos no vacíos $A \subset E$, $B \subset E$ se dirán conjugados respecto de φ si cada vector $\mathbf{x} \in A$ es conjugado con todo vector $\mathbf{y} \in B$.*

Teorema 77. *Si $A \subset E$, $A \neq \emptyset$, el conjunto A' de los vectores de E conjugados con A respecto de una forma bilineal simétrica $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ es un subespacio vectorial de E .*

Demostración. Tenemos que

$$A' = \{\mathbf{x} \in E \mid \forall \mathbf{y} \in A, \quad \varphi(\mathbf{x}, \mathbf{y}) = 0\}$$

Si $\mathbf{x}_1, \mathbf{x}_2 \in A'$, $\alpha_1, \alpha_2 \in \mathbb{K}$,

$$\varphi(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2, \mathbf{y}) = \alpha_1 \varphi(\mathbf{x}_1, \mathbf{y}) + \alpha_2 \varphi(\mathbf{x}_2, \mathbf{y}) = 0 \quad \forall \mathbf{y} \in A$$

entonces $\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 \in A'$. □

Proposición 36. *Sea E un espacio vectorial sobre \mathbb{K} , $\varphi \in \mathcal{B}_S(E, \mathbb{K})$. Si A es un subespacio de E , A' el subespacio conjugado con A respecto de la forma bilineal simétrica φ y A'' el subespacio conjugado con A' respecto de la misma forma φ , entonces A es un subespacio de A'' .*

Demostración. Por definición,

$$A' = \{\mathbf{x} \in E \mid \forall \mathbf{y} \in A, \quad \varphi(\mathbf{x}, \mathbf{y}) = 0\}$$

y

$$A'' = \{\mathbf{z} \in E \mid \forall \mathbf{x} \in A', \quad \varphi(\mathbf{z}, \mathbf{x}) = 0\}$$

Por tanto, si $\mathbf{y} \in A$, tenemos que $\varphi(\mathbf{x}, \mathbf{y}) = 0 \quad \forall \mathbf{x} \in A'$, es decir $\varphi(\mathbf{y}, \mathbf{x}) = 0$. Entonces $\mathbf{y} \in A''$ y A es un subespacio de A'' . \square

Definición 102. Si $A'' = A$, diremos que A y A' son totalmente conjugados, o que A es un subespacio totalmente reflexivo.

5.3.2 Núcleo de una forma bilineal simétrica

Definición 103. Se llama núcleo (o radical) de una forma bilineal simétrica $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ al subespacio E' conjugado con E respecto de φ , es decir

$$\ker \varphi = E' = \{\mathbf{y} \in E \mid \forall \mathbf{x} \in E, \quad \varphi(\mathbf{x}, \mathbf{y}) = 0\}.$$

Definición 104. Diremos que la forma $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ es regular o no degenerada si $\ker \varphi = \{\mathbf{0}\}$. Si $\ker \varphi \neq \{\mathbf{0}\}$ diremos que la forma φ es degenerada (o no regular).

5.3.3 Conjugación respecto de una forma cuadrática

La noción de conjugación se extiende de manera natural al caso de formas cuadráticas.

Definición 105. Sea $\Phi \in Q(E, \mathbb{K})$ una forma cuadrática, y $\varphi_P \in \mathcal{B}_S(E, \mathbb{K})$ su forma polar. Diremos que dos vectores \mathbf{x} e \mathbf{y} son conjugados respecto de Φ si son conjugados respecto de su forma polar φ_P , es decir $\varphi_P(\mathbf{x}, \mathbf{y}) = 0$. Diremos que $\mathbf{x} \in E$ es un vector autoconjugado (o isótropo) si es conjugado consigo mismo, esto es, si $\Phi(\mathbf{x}) = 0$.

Definición 106. Diremos que la forma cuadrática $\Phi \in Q(E, \mathbb{K})$ es regular o no degenerada si lo es su forma polar φ_P . En caso contrario, diremos que la forma Φ es degenerada (o no regular).

Lema 7. Los vectores del núcleo de una forma bilineal simétrica $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ son autoconjugados.

Demostración. Si $\mathbf{x} \in \ker \varphi$, entonces $\varphi(\mathbf{x}, \mathbf{y}) = 0 \quad \forall \mathbf{y} \in E$. En particular, para $\mathbf{y} = \mathbf{x}$ se tiene

$$\varphi(\mathbf{x}, \mathbf{x}) = \Phi(\mathbf{x}) = 0.$$

\square

Teorema 78. Sea E un espacio vectorial sobre \mathbb{K} . Una forma cuadrática $\Phi \in Q(E, \mathbb{K})$ es degenerada si y sólo si su discriminante es nulo.

Demostración. Supongamos que $\dim E = n$ y sea $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de E y $\mathbf{x} \in \ker \varphi$. La condición

$$\varphi(\mathbf{x}, \mathbf{y}) = 0 \quad \forall \mathbf{y} \in E$$

claramente es equivalente a las condiciones sobre los vectores de base

$$\varphi(\mathbf{x}, \mathbf{e}_j) = 0 \quad j = 1, \dots, n.$$

Ahora bien, como $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$, tenemos

$$\sum_{i=1}^n x_i \varphi(\mathbf{e}_i, \mathbf{e}_j) = 0 \quad j = 1, \dots, n$$

que se puede escribir como

$$\sum_{i=1}^n a_{ij} x_i = 0, \quad j = 1, \dots, n \quad (5.13)$$

donde $a_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j)$ es la entrada (i, j) de la matriz asociada a φ . El sistema de ecuaciones lineales (5.13) admite soluciones no triviales (x_1, \dots, x_n) si y sólo si $\text{rg } A < n$, es decir el discriminante de la forma cuadrática es nulo:

$$\det A = 0.$$

□

5.4 Clasificación de las formas cuadráticas reales

Las formas cuadráticas se clasifican en tres familias: definidas, semidefinidas e indefinidas.

Definición 107. Sea E un espacio vectorial sobre \mathbb{K} , $\Phi \in Q(E, \mathbb{K})$.

Se dice que Φ es **definida positiva** (negativa) si

$$\forall \mathbf{x} \in E, \quad \Phi(\mathbf{x}) \geq 0 \quad (\leq 0) \quad \text{y} \quad \Phi(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}.$$

Se dice que Φ es **semidefinida positiva** (negativa) si

$$\forall \mathbf{x} \in E, \quad \Phi(\mathbf{x}) \geq 0 \quad (\leq 0)$$

pero existen vectores $\mathbf{x} \neq \mathbf{0}$ tales que $\Phi(\mathbf{x}) = 0$.

La forma Φ se dice **indefinida** si es negativa para ciertos vectores de E y positiva para otros vectores.

Definición 108. Una forma bilineal simétrica $\varphi \in \mathcal{B}_S(E, \mathbb{K})$ se dirá definida, semidefinida o indefinida si lo es su forma cuadrática asociada $\Phi(\mathbf{x})$.

Estudiaremos ahora algunas propiedades elementales de las formas definidas.

Proposición 37. Si una forma cuadrática es definida (positiva o negativa) entonces es no degenerada.

Demostración. Sea φ_P la forma polar de Φ . Por absurdo, si Φ fuera degenerada, entonces $\ker \varphi_P \neq \{\mathbf{0}\}$; por tanto $\exists \mathbf{x} \in E, \mathbf{x} \neq \mathbf{0}$ tal que $\varphi_P(\mathbf{x}, \mathbf{y}) = 0$ para todo $\mathbf{y} \in E$. En particular, si $\mathbf{y} = \mathbf{x}$, obtendríamos

$$\Phi(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}) = 0$$

y Φ no sería definida. □

5.5 Desigualdades de Schwarz y de Minkowski

5.5.1 Desigualdad de Schwarz

Teorema 79. Sean $\Phi \in Q(E, \mathbb{R})$ una forma cuadrática real y $\varphi \in \mathcal{B}_S(E, \mathbb{R})$ su forma polar asociada. Si Φ es positiva o negativa (definida o semidefinida), entonces

$$|\varphi(\mathbf{x}, \mathbf{y})| \leq \sqrt{\Phi(\mathbf{x})\Phi(\mathbf{y})} \quad \forall \mathbf{x}, \mathbf{y} \in E. \quad (5.14)$$

Demostración. Supongamos Φ positiva, y $\lambda \in \mathbb{R}$. Tenemos

$$\Phi(\lambda\mathbf{x} + \mathbf{y}) = \varphi(\lambda\mathbf{x} + \mathbf{y}, \lambda\mathbf{x} + \mathbf{y}) = \varphi(\lambda\mathbf{x}, \lambda\mathbf{x}) + 2\varphi(\lambda\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{y})$$

es decir

$$\Phi(\lambda\mathbf{x} + \mathbf{y}) = \lambda^2\Phi(\mathbf{x}) + 2\lambda\varphi(\mathbf{x}, \mathbf{y}) + \Phi(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in E, \quad \forall \lambda \in \mathbb{R}. \quad (5.15)$$

Distinguimos dos casos.

a) Si $\Phi(\mathbf{x}) \neq 0$, el segundo miembro de la eq. (5.15) es un trinomio de segundo grado en λ . Como Φ es positiva, entonces

$$\lambda^2\Phi(\mathbf{x}) + 2\lambda\varphi(\mathbf{x}, \mathbf{y}) + \Phi(\mathbf{y}) \geq 0$$

Esta desigualdad está satisfecha si y solo si el discriminante del trinomio es negativo o nulo:

$$[\varphi(\mathbf{x}, \mathbf{y})]^2 - \Phi(\mathbf{x})\Phi(\mathbf{y}) \leq 0.$$

Dado que $\Phi(\mathbf{x}) \geq 0, \Phi(\mathbf{y}) \geq 0$ obtenemos la (5.14).

b) Si $\Phi(\mathbf{x}) = 0$, obtenemos

$$\Phi(\lambda\mathbf{x} + \mathbf{y}) = 2\lambda\varphi(\mathbf{x}, \mathbf{y}) + \Phi(\mathbf{y})$$

El segundo miembro es positivo $\forall \lambda \in \mathbb{R}$ si y solo si $\varphi(\mathbf{x}, \mathbf{y}) = 0$. En este caso la desigualdad (5.14) se cumple trivialmente. De forma análoga se puede demostrar el caso en que Φ sea negativa. \square

Una interesante consecuencia de la desigualdad de Schwarz es la siguiente

Proposición 38. *El núcleo de una forma bilineal simétrica real φ , positiva o negativa, definida o semidefinida, coincide con el conjunto de los vectores isotropos de la forma cuadrática asociada Φ .*

Demostración. \implies) coincide con el enunciado del Lema 7. Precisamente, si $\mathbf{x} \in \ker \varphi$, entonces $\varphi(\mathbf{x}, \mathbf{y}) = 0 \forall \mathbf{y} \in E$. En particular, para $\mathbf{y} = \mathbf{x}$, $\varphi(\mathbf{x}, \mathbf{x}) = \Phi(\mathbf{x}) = 0$ y por tanto $\mathbf{x} \in \ker \Phi$.

\impliedby) Si $\mathbf{x} \in \ker \Phi$, $\Phi(\mathbf{x}) = 0$ y la desigualdad de Schwarz (5.14) implica $\varphi(\mathbf{x}, \mathbf{y}) = 0$, es decir $\mathbf{x} \in \ker \varphi$. \square

5.5.2 Desigualdad de Minkowski

Teorema 80. *Sea $\Phi \in Q(E, \mathbb{R})$ una forma cuadrática real y positiva (definida o semidefinida). Entonces*

$$\sqrt{\Phi(\mathbf{x} + \mathbf{y})} \leq \sqrt{\Phi(\mathbf{x})} + \sqrt{\Phi(\mathbf{y})}, \quad \forall \mathbf{x}, \mathbf{y} \in E. \quad (5.16)$$

Demostración. Si $\varphi \in \mathcal{B}_S(E, \mathbb{R})$ es la forma polar asociada a Φ , tenemos que

$$\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + 2\varphi(\mathbf{x}, \mathbf{y}) + \Phi(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in E. \quad (5.17)$$

Utilizando la desigualdad de Schwarz (5.14), tenemos

$$\varphi(\mathbf{x}, \mathbf{y}) \leq |\varphi(\mathbf{x}, \mathbf{y})| \leq \sqrt{\Phi(\mathbf{x})\Phi(\mathbf{y})}.$$

Entonces

$$\begin{aligned} \Phi(\mathbf{x} + \mathbf{y}) &\leq \Phi(\mathbf{x}) + 2\sqrt{\Phi(\mathbf{x})\Phi(\mathbf{y})} + \Phi(\mathbf{y}) = \\ &= [\sqrt{\Phi(\mathbf{x})} + \sqrt{\Phi(\mathbf{y})}]^2 \end{aligned}$$

y siendo $\Phi(\mathbf{x} + \mathbf{y}) \geq 0$, obtenemos la desigualdad (5.16).

5.6 Ejercicios

Ejercicio 5.4. Sea V un espacio vectorial sobre \mathbb{K} , y sean $\phi, \chi : V \rightarrow \mathbb{K}$ dos formas lineales. Demostrar que la aplicación $\omega : V \times V \rightarrow \mathbb{K}$ dada por

$$\omega(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})\chi(\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in V$$

es una forma bilineal.

Resolución. Demostremos la linealidad respecto de la primera variable:

a)

$$\begin{aligned} \omega(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) &= \phi(\mathbf{x}_1 + \mathbf{x}_2)\chi(\mathbf{y}) = [\phi(\mathbf{x}_1) + \phi(\mathbf{x}_2)]\chi(\mathbf{y}) = \\ &= \phi(\mathbf{x}_1)\chi(\mathbf{y}) + \phi(\mathbf{x}_2)\chi(\mathbf{y}) = \omega(\mathbf{x}_1, \mathbf{y}) + \omega(\mathbf{x}_2, \mathbf{y}) \end{aligned}$$

b)

$$\omega(\lambda\mathbf{x}, \mathbf{y}) = \phi(\lambda\mathbf{x})\chi(\mathbf{y}) = \lambda\phi(\mathbf{x})\chi(\mathbf{y}) = \lambda\omega(\mathbf{x}, \mathbf{y})$$

De forma análoga se demuestra la linealidad respecto de la segunda variable.

Ejercicio 5.5. Sea $\varphi \in \mathcal{B}(\mathbb{R}^3, \mathbb{R})$ la forma bilineal definida por

$$\varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 + 2x_2y_2 + x_2y_3 .$$

Determinar su forma cuadrática asociada y la forma polar de ésta, y la matriz asociada a φ respecto de la base canónica de \mathbb{R}^3 .

Resolución. La forma cuadrática Φ asociada a φ es

$$\Phi(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}) = x_1^2 + 2x_2^2 + x_2x_3$$

La forma polar se puede obtener utilizando una de las expresiones (5.8)–(5.10). Por ejemplo

$$\begin{aligned} \varphi_P(\mathbf{x}, \mathbf{y}) &= \frac{1}{2}[\Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x}) - \Phi(\mathbf{y})] = \\ &= \frac{1}{2}[(x_1 + y_1)^2 + 2(x_2 + y_2)^2 + (x_2 + y_2)(x_3 + y_3) - x_1^2 - 2x_2^2 - x_2x_3 \\ &\quad - y_1^2 - 2y_2^2 - y_2y_3] = \dots = \\ &= \frac{1}{2}[2x_1y_1 + 4x_2y_2 + x_2y_3 + x_3y_2] = x_1y_1 + 2x_2y_2 + \frac{1}{2}(x_2y_3 + x_3y_2) \end{aligned}$$

Equivalentemente:

$$\varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})] = x_1y_1 + 2x_2y_2 + \frac{1}{2}(x_2y_3 + x_3y_2)$$

Hallemos ahora la matriz asociada a φ . Tenemos

$$\begin{aligned} \varphi(\mathbf{e}_1, \mathbf{e}_1) &= \varphi[(1, 0, 0), (1, 0, 0)] = 1 & \varphi(\mathbf{e}_1, \mathbf{e}_2) &= 0, & \varphi(\mathbf{e}_1, \mathbf{e}_3) &= 0 \\ \varphi(\mathbf{e}_2, \mathbf{e}_1) &= 0 & \varphi(\mathbf{e}_2, \mathbf{e}_2) &= 2 & \varphi(\mathbf{e}_2, \mathbf{e}_3) &= 1 \end{aligned}$$

$$\varphi(\mathbf{e}_3, \mathbf{e}_1) = 0 \quad \varphi(\mathbf{e}_3, \mathbf{e}_2) = 0 \quad \varphi(\mathbf{e}_2, \mathbf{e}_3) = 0 .$$

La matriz asociada a φ es dada por

$$A_\varphi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

La matriz asociada a la forma cuadrática Φ se obtiene inmediatamente utilizando la fórmula (5.12) y es

$$A_\Phi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

También, recordamos que $A_\Phi = A_{\varphi_P}$, y coherentemente con la fórmula (5.10), tenemos que

$$A_\Phi = \frac{1}{2}[A_\varphi + (A_\varphi)^t]$$

Ejercicio 5.6. Sea $\varphi \in \mathcal{B}_S(\mathbb{R}^3, \mathbb{R})$ la forma bilineal simétrica dada $\forall \mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$ por

$$\varphi(\mathbf{x}, \mathbf{y}) = 2x_1y_1 + 3x_2y_2 + \frac{28}{5}x_3y_3 - x_1y_2 - x_2y_1 - 2x_1y_3 - 2x_3y_1 + 4x_2y_3 + 4x_3y_2$$

y sea Φ la forma cuadrática asociada a φ .

- 1) Escribir la matriz de la forma φ respecto de la base canónica de \mathbb{R}^3 y calcular su rango.
- 2) Hallar $\ker \varphi$.
- 3) Demostrar que $\{(1, 1, 1), (2, -1, 2), (1, 3, -3)\}$ es una base de \mathbb{R}^3 y determinar la matriz de φ respecto de esta base.

Resolución. 1) La matriz de φ respecto de la base canónica de \mathbb{R}^3 es

$$A = \begin{pmatrix} 2 & -1 & -2 \\ -1 & 3 & 4 \\ -2 & 4 & \frac{28}{5} \end{pmatrix}$$

Tenemos que $\det(A) = 0$ y por ej. $\begin{vmatrix} 2 & -1 \\ -1 & 3 \end{vmatrix} = 5 \neq 0$ y por tanto $\text{rang}(\varphi) = \text{rang}(A) = 2$.

2)

$$\ker \varphi = \{\mathbf{y} \in \mathbb{R}^3 \mid \varphi(\mathbf{x}, \mathbf{y}) = 0 \quad \forall \mathbf{x} \in \mathbb{R}^3\} .$$

Sean $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$. Entonces $\mathbf{y} \in \ker \varphi$ si $X^t A Y = 0 \quad \forall X \in \mathbb{R}^3$, es

decir si $AY = 0$. Por tanto obtenemos el sistema

$$\begin{cases} 2y_1 - y_2 - 2y_3 = 0 \\ -y_1 + 3y_2 + 4y_3 = 0 \\ -2y_1 + 4y_2 + \frac{28}{5}y_3 = 0 \end{cases} \iff \begin{cases} y_1 = \frac{2}{5}y_3 \\ y_2 = -\frac{6}{5}y_3 \end{cases}$$

En definitiva $\ker \varphi = \{(2\alpha, -6\alpha, 5\alpha) \mid \alpha \in \mathbb{R}\}$ y una base de $\ker \varphi = \{(2, -6, 5)\}$.

3) Se observa que

$$\begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 2 \\ 1 & 3 & -3 \end{vmatrix} = 12 \neq 0$$

Por tanto los vectores $B' = \{(1, 1, 1), (2, -1, 2), (1, 3, -3)\}$ son linealmente independientes y forman una base de \mathbb{R}^3 . La matriz de φ respecto de esta nueva base es

$$A' = P^t A P$$

donde P es la matriz de cambio de la base canónica a la nueva base:

$$P = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 3 \\ 1 & 2 & -3 \end{pmatrix}.$$

En definitiva, tenemos

$$A' = \frac{1}{5} \begin{pmatrix} 63 & 36 & -29 \\ 36 & 27 & 2 \\ -29 & 2 & -67 \end{pmatrix}$$

Ejercicio 5.7. Sea \mathbb{R}^3 dotado de la base canónica $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ y sea φ una forma bilineal simétrica de \mathbb{R}^3 dada $\forall \mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3)$ por

$$\varphi(\mathbf{x}, \mathbf{y}) = x_1 y_1 + 6 x_2 y_2 + 56 x_3 y_3 - 2(x_1 y_2 + x_2 y_1) + 7(x_1 y_3 + x_3 y_1) - 18(x_2 y_3 + x_3 y_2).$$

- 1) Escribir la matriz de la forma φ respecto de la base canónica de \mathbb{R}^3 .
- 2) Mostrar que los vectores $\mathbf{e}'_1 = \mathbf{e}_1, \mathbf{e}'_2 = 2\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}'_3 = -3\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3$ forman una base de \mathbb{R}^3 .
- 3) Determinar la matriz de φ respecto de la base $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.
- 4) Determinar la forma cuadrática Φ asociada a φ respecto de las bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ y $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.

Resolución. 1) Como $a_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j)$, tenemos

$$A = \begin{pmatrix} 1 & -2 & 7 \\ -2 & 6 & -18 \\ 7 & -18 & 56 \end{pmatrix}.$$

2) Escribiendo en columna las componentes de los vectores $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ en la base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ obtenemos la matriz

$$P = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \det P = 1 \neq 0$$

Por tanto $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ es otra base de \mathbb{R}^3 y en particular P es la matriz de cambio de la base canónica a esta nueva base.

3) La matriz de φ en la nueva base es

$$A' = P^t A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

4) La forma cuadrática Φ respecto de la base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ es

$$\Phi(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}) = x_1^2 + 6x_2^2 + 56x_3^2 - 4x_1x_2 + 14x_1x_3 - 36x_2x_3$$

Respecto de la base $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ se escribe como

$$\Phi(\mathbf{x}) = (x'_1)^2 + 2(x'_2)^2 - (x'_3)^2.$$

5.7 Productos escalares

5.7.1 Definición

Una clase muy importante de formas bilineal simétricas son las definidas positivas.

Definición 109. Sea E un espacio vectorial sobre \mathbb{R} . Un **producto escalar** en E es una forma bilineal simétrica definida positiva:

$$\langle \cdot, \cdot \rangle : E \times E \longrightarrow \mathbb{R}_+ \cup \{0\}$$

Un espacio vectorial E sobre \mathbb{R} dotado de un producto escalar se llamará **espacio vectorial euclídeo**.

Propiedades. Un producto escalar, siendo una forma bilineal simétrica definida positiva, posee las siguientes propiedades:

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{y}, \mathbf{x} \rangle && \forall \mathbf{x}, \mathbf{y} \in E \\ \langle \lambda \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{x}, \lambda \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle && \forall \mathbf{x} \in E, \forall \lambda \in \mathbb{R} \\ \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle &= \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle && \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in E \\ \langle \mathbf{x}, \mathbf{x} \rangle &\geq 0 \text{ y } \langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0} \end{aligned}$$

Ejemplo 5.8. En \mathbb{R}^n podemos introducir el **producto escalar estándar o canónico**

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n .$$

También se usa comúnmente para el producto escalar estándar la notación alternativa $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$.

5.7.2 Norma asociada a un producto escalar

Definición 110. Sea E un espacio vectorial euclídeo y $\mathbf{x} \in E$. Se llama **norma** de \mathbf{x} al número real no negativo

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} . \quad (5.18)$$

Definición 111. Un vector de norma igual a 1 se dirá **vector unitario**.

Ejemplo 5.9. En \mathbb{R}^n podemos introducir la norma asociada al producto escalar canónico, llamada **norma euclídea**:

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} .$$

Desde las propiedades generales del producto escalar deducimos que una norma satisface a las propiedades siguientes:

$$\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\| \quad \forall \mathbf{x} \in E, \quad \forall \alpha \in \mathbb{R}$$

y

$$\| \mathbf{x} \| = 0 \iff \mathbf{x} = \mathbf{0}$$

Observación 5.4. Podemos escribir las desigualdades de Schwarz y Minkowski en el lenguaje de las normas. Por ello, es suficiente observar que si $\varphi(\mathbf{x}, \mathbf{y}) \equiv \langle \mathbf{x}, \mathbf{y} \rangle$ es un producto escalar, entonces $\Phi(\mathbf{x}) = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$.

Por tanto, la desigualdad de Schwarz

$$|\varphi(\mathbf{x}, \mathbf{y})| \leq \sqrt{\Phi(\mathbf{x})\Phi(\mathbf{y})} \quad \forall \mathbf{x}, \mathbf{y} \in E$$

se convierte en

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \| \mathbf{x} \| \| \mathbf{y} \| \quad \forall \mathbf{x}, \mathbf{y} \in E$$

y la desigualdad de Minkowski (5.16) nos da la desigualdad

$$\| \mathbf{x} + \mathbf{y} \| \leq \| \mathbf{x} \| + \| \mathbf{y} \| \quad \forall \mathbf{x}, \mathbf{y} \in E$$

que es también llamada **desigualdad triangular**.

5.7.3 Distancia inducida por una norma

Dado un producto escalar, podemos definir la norma asociada (5.18). Al mismo tiempo, dada una norma podemos introducir la noción de distancia.

Definición 112. Sea E un espacio vectorial euclídeo. La aplicación

$$d : E \times E \longrightarrow \mathbb{R}$$

definida por

$$d(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|, \quad \forall \mathbf{x}, \mathbf{y} \in E$$

se dirá la *distancia o métrica inducida por la norma*.

Se verifica fácilmente que una distancia satisface a las propiedades siguientes:

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) &\geq 0 \quad \text{y} \quad d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y} && \mathbf{x}, \mathbf{y} \in E \\ d(\mathbf{x}, \mathbf{y}) &= d(\mathbf{y}, \mathbf{x}) && \mathbf{x}, \mathbf{y} \in E \\ d(\mathbf{x}, \mathbf{z}) &\leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) && \mathbf{x}, \mathbf{y}, \mathbf{z} \in E \end{aligned}$$

Teorema 81 (Identidad de polarización). Si E es un espacio euclídeo, dotado del producto escalar $\langle \cdot, \cdot \rangle$ y norma $\| \cdot \|$, entonces

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4} (\| \mathbf{x} + \mathbf{y} \|^2 - \| \mathbf{x} - \mathbf{y} \|^2) \quad (5.19)$$

Demostración. Es una consecuencia inmediata de la fórmula (5.9):

$$\varphi_P(\mathbf{x}, \mathbf{y}) = \frac{1}{4}(\Phi(\mathbf{x} + \mathbf{y}) - \Phi(\mathbf{x} - \mathbf{y})) .$$

Es suficiente observar que, la norma $\|\mathbf{x}\| = \sqrt{\Phi(\mathbf{x})}$ así que la fórmula anterior se escribe en la forma (5.19). \square

Identidad del paralelogramo.

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) \quad \forall \mathbf{x}, \mathbf{y} \in E \quad (5.20)$$

Demostración. Es una consecuencia inmediata de las relaciones

$$\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + 2\varphi_P(\mathbf{x}, \mathbf{y}) + \Phi(\mathbf{y})$$

y

$$\Phi(\mathbf{x} - \mathbf{y}) = \Phi(\mathbf{x}) - 2\varphi_P(\mathbf{x}, \mathbf{y}) + \Phi(\mathbf{y})$$

que nos proporcionan la fórmula :

$$\Phi(\mathbf{x} + \mathbf{y}) + \Phi(\mathbf{x} - \mathbf{y}) = 2(\Phi(\mathbf{x}) + \Phi(\mathbf{y})) .$$

\square

5.7.4 Ángulo entre vectores no nulos

Definición 113. Sea E un espacio vectorial euclídeo y \mathbf{x}, \mathbf{y} vectores no nulos de E . El ángulo α ($0 \leq \alpha \leq \pi$) definido por

$$\cos \alpha = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \iff \alpha = \arccos \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \right)$$

se dirá ángulo entre los vectores \mathbf{x} e \mathbf{y} .

Observación 5.5. Como

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in E$$

resulta que

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1$$

y entonces la noción anterior de ángulo entre vectores está bien definida.

También tenemos obviamente la relación

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos \alpha$$

Si en \mathbb{R}^n consideramos el producto escalar estándar, la desigualdad de Schwarz se escribe como

$$|x_1 y_1 + \dots + x_n y_n| \leq \left(\sqrt{x_1^2 + \dots + x_n^2} \right) \left(\sqrt{y_1^2 + \dots + y_n^2} \right)$$

mientras que la de Minkowski se convierte en

$$\sqrt{\sum_{i=1}^n (x_i + y_i)^2} \leq \sqrt{\sum_{i=1}^n x_i^2} + \sqrt{\sum_{i=1}^n y_i^2}$$

Ejemplo 5.10. Sea \mathbb{R}^2 dotado del producto escalar estándar, $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$ vectores no nulos. El ángulo entre \mathbf{x} e \mathbf{y} es

$$\alpha = \arccos \left(\frac{x_1 y_1 + x_2 y_2}{\sqrt{x_1^2 + x_2^2} \sqrt{y_1^2 + y_2^2}} \right).$$

Problema 5.11. Verificar que en \mathbb{R}^2 la aplicación definida por

$$\langle (x_1, x_2), (y_1, y_2) \rangle = 5x_1 y_1 + x_2 y_2$$

es un producto escalar.

Ejemplo 5.12. Sea $P_n(\mathbb{R})$ el espacio de los polinomios de grado menor o igual a n , con coeficientes reales. Demostrar que la aplicación

$$\langle p(x), q(x) \rangle = \int_0^1 p(x)q(x)dx, \quad p(x), q(x) \in P_n(\mathbb{R})$$

es un producto escalar.

Ejemplo 5.13. En el espacio vectorial $M_n(\mathbb{R})$ se define el producto escalar

$$\langle A, B \rangle = \text{tr}(AB^t) = \sum_{i,j=1}^n a_{ij}b_{ij}$$

Definición 114. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo. Diremos que dos vectores $\mathbf{x}, \mathbf{y} \in E$ son **ortogonales** respecto del producto escalar $\langle \cdot, \cdot \rangle$ si

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0$$

y escribiremos $\mathbf{x} \perp \mathbf{y}$.

Definición 115. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo. Diremos que $W \subset E$ es un subconjunto ortonormal de vectores de E si

$$\forall \mathbf{x}_i, \mathbf{x}_j \in W : \quad \langle \mathbf{x}_i, \mathbf{x}_j \rangle = \delta_{ij}, \quad \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}.$$

Definición 116. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo y $W \subset E$ un conjunto de vectores. Se llama subespacio ortogonal a W el subespacio

$$W^\perp = \{ \mathbf{x} \in E \mid \forall \mathbf{y} \in W : \langle \mathbf{x}, \mathbf{y} \rangle = 0 \}.$$

En otras palabras, W^\perp es el conjugado de W respecto de la forma bilineal simétrica definida positiva $\langle \cdot, \cdot \rangle$.

Definición 117. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\dim E = n$ y sea $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de E . Se llamará **matriz de Gram** respecto de la base B a la matriz asociada al producto escalar respecto de la base B :

$$a_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle .$$

Desde la teoría de las formas bilineales simétricas sabemos que la matriz A es simétrica: $A = A^t$. Además, si $\mathbf{x}, \mathbf{y} \in E$, y X e Y son los vectores columna de sus componentes respecto de una base de E ,

$$\langle \mathbf{x}, \mathbf{y} \rangle = X^t A Y$$

fórmula que nos proporciona la expresión matricial del producto escalar.

Si B' es una nueva base de E , la matriz A' asociada al producto escalar respecto de B' es

$$A' = P^t A P$$

donde P la matriz de cambio de la base B a la base B' . En otras palabras, *las matrices de Gram de un producto escalar respecto de bases distintas son congruentes.*

5.8 Bases ortogonales y ortonormales

Definición 118. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\dim E = n$. Se dice que una base $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ es una base **ortogonal** si los vectores que la forman son ortogonales entre sí:

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0 \quad \forall i \neq j .$$

Se dice que \mathcal{B} es una base **ortonormal** si \mathcal{B} es ortogonal y además $\|\mathbf{e}_i\| = 1$, $\forall i = 1, \dots, n$.

Corolario 14. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de E y A sea la matriz de Gram del producto escalar respecto de la base B . Entonces:

- 1) La base \mathcal{B} es ortogonal si y sólo si A es una matriz **diagonal**.
- 2) La base \mathcal{B} es ortonormal si y sólo si A es la matriz **identidad**.

Demostración.

$$\begin{aligned} \mathcal{B} \text{ ortogonal} &\iff a_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0 \quad \forall i \neq j \iff A \text{ diagonal} \\ \mathcal{B} \text{ ortonormal} &\iff \delta_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle \quad \forall i, j = 1, \dots, n \iff A = \mathbf{1} \end{aligned}$$

□

Ejemplos 5.14. *i)* Sea $E = \mathbb{R}^n$, dotado del producto escalar estándar $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + \dots + x_ny_n$. La base canónica de \mathbb{R}^n es ortonormal.

ii) Sea $E = \mathbb{R}^2$. La base $\{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})\}$ es ortonormal.

Definición 119. Una matriz $A \in M_n$ invertible se dice **ortogonal** si

$$A^t = A^{-1} . \quad (5.21)$$

Proposición 39. Sea E un espacio euclídeo. La matriz de cambio de base entre dos bases ortonormales de E es ortogonal.

Demostración. Sean \mathcal{B} , \mathcal{B}' dos bases de E y P la matriz de cambio de base. Entonces, la matriz de Gram del producto escalar verifica

$$A' = P^t A P .$$

Al mismo tiempo, como \mathcal{B} y \mathcal{B}' son ortonormales, resulta que $A = A' = \mathbf{1}$. Por tanto

$$\mathbf{1} = P^t \mathbf{1} P \iff P^t = P^{-1} .$$

□

5.8.1 Teorema de la independencia lineal

Teorema 82. Sea E un espacio euclídeo. Si $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ son un conjunto de vectores no nulos y ortonormales, entonces son linealmente independientes.

Demostración. Consideremos la relación

$$\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k = \mathbf{0} . \quad (5.22)$$

Calculemos el producto escalar de I miembro y II miembro de la relación (5.22) con el vector \mathbf{x}_j :

$$\lambda_1 \langle \mathbf{x}_j, \mathbf{x}_1 \rangle + \dots + \lambda_j \langle \mathbf{x}_j, \mathbf{x}_j \rangle + \dots + \lambda_k \langle \mathbf{x}_j, \mathbf{x}_k \rangle = \langle \mathbf{x}_j, \mathbf{0} \rangle = 0 \iff \lambda_j = 0$$

donde hemos utilizado la propiedad de ortonormalidad $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = \delta_{ij}$. Entonces, repitiendo el mismo razonamiento para todo $j \in \{1, \dots, k\}$, deducimos que los escalares $\lambda_1, \dots, \lambda_k$ en la (5.22) son necesariamente nulos, y los vectores $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ son linealmente independientes. □

5.8.2 Descomposición en coeficientes de Fourier

Teorema 83. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\dim E = n$, y sea $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base ortogonal de E . Todo vector $\mathbf{x} \in E$ admite la descomposición siguiente:

$$\mathbf{x} = \frac{\langle \mathbf{e}_1, \mathbf{x} \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 + \dots + \frac{\langle \mathbf{e}_n, \mathbf{x} \rangle}{\|\mathbf{e}_n\|^2} \mathbf{e}_n \quad (5.23)$$

Demostración. Sea $\mathbf{x} \in E$; entonces $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$. Si tomamos el producto escalar de ambos miembros con el vector \mathbf{e}_i , tenemos

$$\begin{aligned}\langle \mathbf{e}_i, \mathbf{x} \rangle &= x_1\langle \mathbf{e}_i, \mathbf{e}_1 \rangle + \dots + x_i\langle \mathbf{e}_i, \mathbf{e}_i \rangle + \dots + x_n\langle \mathbf{e}_i, \mathbf{e}_n \rangle = \\ &= x_i\langle \mathbf{e}_i, \mathbf{e}_i \rangle = x_i \|\mathbf{e}_i\|^2\end{aligned}$$

es decir

$$x_i = \frac{\langle \mathbf{e}_i, \mathbf{x} \rangle}{\|\mathbf{e}_i\|^2} = \frac{\langle \mathbf{x}, \mathbf{e}_i \rangle}{\|\mathbf{e}_i\|^2}$$

□

Definición 120. Los coeficientes $x_i = \frac{\langle \mathbf{e}_i, \mathbf{x} \rangle}{\|\mathbf{e}_i\|^2}$ se denominan *coeficientes de Fourier de \mathbf{x} respecto de la base $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$* .

Ejemplo 5.15. Sea $E = \mathbb{R}^3$ dotado del producto escalar estándar, con base ortogonal $\{(2, 2, 2), (2, -2, 0), (\frac{1}{4}, \frac{1}{4}, -\frac{1}{2})\}$ y sea $\mathbf{x} = (2, 0, 3)$. Sus coeficientes de Fourier son

$$\begin{aligned}x_1 &= \frac{\langle (2, 2, 2), (2, 0, 3) \rangle}{\|(2, 2, 2)\|^2} = \frac{10}{12} = \frac{5}{6} \\ x_2 &= \frac{\langle (2, -2, 0), (2, 0, 3) \rangle}{\|(2, -2, 0)\|^2} = \frac{4}{8} = \frac{1}{2} \\ x_3 &= \frac{\langle (\frac{1}{4}, \frac{1}{4}, -\frac{1}{2}), (2, 0, 3) \rangle}{\|(\frac{1}{4}, \frac{1}{4}, -\frac{1}{2})\|^2} = -\frac{1}{3/8} = -\frac{8}{3}\end{aligned}$$

Lema 8. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo. Si $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ es una base ortogonal de E , entonces

$$\mathcal{B}' = \left\{ \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|}, \dots, \frac{\mathbf{e}_n}{\|\mathbf{e}_n\|} \right\}$$

es una base ortonormal de E .

Demostración. Los vectores de la base son normalizados por construcción:

$$\left\| \frac{\mathbf{e}_i}{\|\mathbf{e}_i\|} \right\| = \frac{\|\mathbf{e}_i\|}{\|\mathbf{e}_i\|} = 1.$$

Sea $i \neq j$. Tenemos

$$\left\langle \frac{\mathbf{e}_i}{\|\mathbf{e}_i\|}, \frac{\mathbf{e}_j}{\|\mathbf{e}_j\|} \right\rangle = \frac{1}{\|\mathbf{e}_i\| \|\mathbf{e}_j\|} \langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$$

Si $i = j$,

$$\left\langle \frac{\mathbf{e}_i}{\|\mathbf{e}_i\|}, \frac{\mathbf{e}_i}{\|\mathbf{e}_i\|} \right\rangle = \frac{\|\mathbf{e}_i\|^2}{\|\mathbf{e}_i\|^2} = 1$$

El célebre teorema de Pitágoras se puede reformular y generalizar de forma natural en el contexto de los espacios euclídeos.

Teorema 84 (Pitágoras). Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, y $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ un conjunto de vectores ortogonales de E . Entonces

$$\|\mathbf{x}_1 + \dots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \dots + \|\mathbf{x}_k\|^2 \quad (5.24)$$

Demostración.

$$\begin{aligned} \|\mathbf{x}_1 + \dots + \mathbf{x}_k\|^2 &= \langle \mathbf{x}_1 + \dots + \mathbf{x}_k, \mathbf{x}_1 + \dots + \mathbf{x}_k \rangle = \sum_{i=1}^k \sum_{j=1}^k \langle \mathbf{x}_i, \mathbf{x}_j \rangle = \\ &= \sum_{i=1}^k \langle \mathbf{x}_i, \mathbf{x}_i \rangle = \sum_{i=1}^k \|\mathbf{x}_i\|^2 . \end{aligned}$$

5.9 Método de ortonormalización de Gram-Schmidt

5.9.1 Procedimiento general

Teorema 85. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, y $B = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ base de E . Entonces existe una base ortonormal $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ de E tal que resulta

$$L(\mathbf{x}_1, \dots, \mathbf{x}_k) = L(\mathbf{u}_1, \dots, \mathbf{u}_k), \quad \forall k = 1, \dots, n$$

Demostración.

Sea

$$\begin{aligned} \mathbf{e}_1 &= \mathbf{x}_1 \\ \mathbf{e}_2 &= \mathbf{x}_2 - \frac{\langle \mathbf{e}_1, \mathbf{x}_2 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 \\ \mathbf{e}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{e}_1, \mathbf{x}_3 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 - \frac{\langle \mathbf{e}_2, \mathbf{x}_3 \rangle}{\|\mathbf{e}_2\|^2} \mathbf{e}_2 \\ &\vdots \\ \mathbf{e}_n &= \mathbf{x}_n - \frac{\langle \mathbf{e}_1, \mathbf{x}_n \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 - \frac{\langle \mathbf{e}_2, \mathbf{x}_n \rangle}{\|\mathbf{e}_2\|^2} \mathbf{e}_2 - \dots - \frac{\langle \mathbf{e}_{n-1}, \mathbf{x}_n \rangle}{\|\mathbf{e}_{n-1}\|^2} \mathbf{e}_{n-1} . \end{aligned} \quad (5.25)$$

Demostremos por inducción que:

- (1) $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ es un conjunto de vectores ortogonales dos a dos
- (2) $L(\mathbf{x}_1, \dots, \mathbf{x}_k) = L(\mathbf{e}_1, \dots, \mathbf{e}_k), \forall k = 1, \dots, n$.

Consideremos ante todo el caso $k = 2$. Tenemos

$$\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = \langle \mathbf{e}_1, \mathbf{x}_2 \rangle - \frac{\langle \mathbf{e}_1, \mathbf{x}_2 \rangle}{\|\mathbf{e}_1\|^2} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle = \langle \mathbf{e}_1, \mathbf{x}_2 \rangle - \langle \mathbf{e}_1, \mathbf{x}_2 \rangle = 0 .$$

Además, claramente $L(\mathbf{e}_1, \mathbf{e}_2) = L(\mathbf{x}_1, \mathbf{x}_2)$.

Entonces supondremos $k > 2$.

(1) Supongamos $\{\mathbf{e}_1, \dots, \mathbf{e}_{k-1}\}$ ortogonales dos a dos. Ponemos

$$\alpha_{ij} = \frac{\langle \mathbf{e}_j, \mathbf{x}_i \rangle}{\|\mathbf{e}_j\|^2}, \quad i = 1, \dots, k, \quad j = 1, \dots, k-1.$$

Entonces tenemos

$$\begin{aligned} \langle \mathbf{e}_k, \mathbf{e}_j \rangle &= \langle \mathbf{x}_k - \alpha_{k,1}\mathbf{e}_1 - \dots - \alpha_{k,k-1}\mathbf{e}_{k-1}, \mathbf{e}_j \rangle = \\ &= \langle \mathbf{x}_k, \mathbf{e}_j \rangle - \alpha_{k,1}\langle \mathbf{e}_1, \mathbf{e}_j \rangle - \dots - \alpha_{k,k-1}\langle \mathbf{e}_{k-1}, \mathbf{e}_j \rangle = \\ &= \langle \mathbf{x}_k, \mathbf{e}_j \rangle - \alpha_{k,j}\langle \mathbf{e}_j, \mathbf{e}_j \rangle = \\ &= \langle \mathbf{x}_k, \mathbf{e}_j \rangle - \frac{\langle \mathbf{e}_j, \mathbf{x}_k \rangle}{\|\mathbf{e}_j\|^2} \|\mathbf{e}_j\|^2 = 0. \end{aligned}$$

(2) Supongamos que

$$L(\mathbf{x}_1, \dots, \mathbf{x}_{k-1}) = L(\mathbf{e}_1, \dots, \mathbf{e}_{k-1}).$$

Dado que

$$\mathbf{e}_k = \mathbf{x}_k - \alpha_{k,1}\mathbf{e}_1 - \dots - \alpha_{k,k-1}\mathbf{e}_{k-1},$$

entonces $\mathbf{e}_k \in L(\mathbf{e}_1, \dots, \mathbf{e}_{k-1}, \mathbf{x}_k) = L(\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \mathbf{x}_k)$. Sea $S = L(\mathbf{x}_1, \dots, \mathbf{x}_k)$; $S \subset E$ es un subespacio, con base $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. Dado que $\{\mathbf{e}_1, \dots, \mathbf{e}_k\} \in S$, y como $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ son ortogonales dos a dos, entonces son linealmente independientes y forman otra base de S . Por tanto

$$S = L(\mathbf{x}_1, \dots, \mathbf{x}_k) = L(\mathbf{e}_1, \dots, \mathbf{e}_k).$$

En definitiva hemos demostrado por inducción que $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ es una base ortogonal de E , y $\left\{ \mathbf{u}_1 = \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|}, \dots, \mathbf{u}_n = \frac{\mathbf{e}_n}{\|\mathbf{e}_n\|} \right\}$ es una base ortonormal. \square

Ejercicio 5.16. Encontrar una base ortonormal del subespacio de \mathbb{R}^4

$$S = \text{lin}\{(1, 1, -1, -2), (-2, 1, 5, 11), (0, 3, 3, 7), (3, -3, -3, -9)\}.$$

Resolución. Los 4 vectores son dependientes: $2\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_3$; al mismo tiempo $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4$ son linealmente independientes. Entonces:

$$\begin{aligned} \mathbf{e}_1 &= \mathbf{x}_1 = (1, 1, -1, -2) & \|\mathbf{e}_1\|^2 &= 7 \\ \mathbf{e}_2 &= \mathbf{x}_2 - \frac{\langle \mathbf{e}_1, \mathbf{x}_2 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 = (-2, 1, 5, 11) - \frac{-2 + 1 - 5 - 22}{7} (1, 1, -1, -2) = \\ &= (-2, 1, 5, 11) + (4, 4, -4, -8) = (2, 5, 1, 3) & \|\mathbf{e}_2\|^2 &= 39. \\ \mathbf{e}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{e}_1, \mathbf{x}_3 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 - \frac{\langle \mathbf{e}_2, \mathbf{x}_3 \rangle}{\|\mathbf{e}_2\|^2} \mathbf{e}_2 = \\ &= (0, 3, 3, 7) - \frac{3 - 3 - 14}{7} \mathbf{e}_1 - \frac{15 + 3 + 21}{39} \mathbf{e}_2 = \\ &= (0, 3, 3, 7) + 2(1, 1, -1, -2) - (2, 5, 1, 3) = \\ &= (2, 5, 1, 3) - (2, 5, 1, 3) = (0, 0, 0, 0). \end{aligned}$$

$$\begin{aligned}
\mathbf{e}_4 &= \mathbf{x}_4 - \frac{\langle \mathbf{e}_1, \mathbf{x}_4 \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 - \frac{\langle \mathbf{e}_2, \mathbf{x}_4 \rangle}{\|\mathbf{e}_2\|^2} \mathbf{e}_2 = \\
&= (3, -3, -3, -9) - \frac{3 - 3 + 3 + 18}{7} (1, 1, -1, -2) - \frac{6 - 15 - 3 - 27}{39} (2, 5, 1, 3) = \\
&= (3, -3, -3, -9) - (3, 3, -3, -6) + (2, 5, 1, 3) = (2, -1, 1, 0) .
\end{aligned}$$

Hemos obtenido la base ortogonal de S $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$. Normalizando estos vectores llegamos a la base ortonormal deseada $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ con:

$$\begin{aligned}
\mathbf{u}_1 &= \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|} = \frac{1}{\sqrt{7}}(1, 1, -1, -2) \\
\mathbf{u}_2 &= \frac{\mathbf{e}_2}{\|\mathbf{e}_2\|} = \frac{1}{\sqrt{39}}(2, 5, 1, 3) \\
\mathbf{u}_3 &= \frac{\mathbf{e}_4}{\|\mathbf{e}_4\|} = \frac{1}{\sqrt{6}}(2, -1, 1, 0) .
\end{aligned}$$

5.9.2 Observaciones sobre el método

(1) Las ecuaciones de Gram-Schmidt se pueden también escribir recursivamente de la forma siguiente:

$$\begin{aligned}
\mathbf{u}_1 &= \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} \\
\mathbf{e}_2 &= \mathbf{x}_2 - \langle \mathbf{u}_1, \mathbf{x}_2 \rangle \mathbf{u}_1, \quad \mathbf{u}_2 = \frac{\mathbf{e}_2}{\|\mathbf{e}_2\|}, \\
\mathbf{e}_3 &= \mathbf{x}_3 - \langle \mathbf{u}_1, \mathbf{x}_3 \rangle \mathbf{u}_1 - \langle \mathbf{u}_2, \mathbf{x}_3 \rangle \mathbf{u}_2, \quad \mathbf{u}_3 = \frac{\mathbf{e}_3}{\|\mathbf{e}_3\|}, \\
&\vdots \\
\mathbf{e}_n &= \mathbf{x}_n - \langle \mathbf{u}_1, \mathbf{x}_n \rangle \mathbf{u}_1 - \langle \mathbf{u}_2, \mathbf{x}_n \rangle \mathbf{u}_2 - \dots - \langle \mathbf{u}_{n-1}, \mathbf{x}_n \rangle \mathbf{u}_{n-1}, \quad \mathbf{u}_n = \frac{\mathbf{e}_n}{\|\mathbf{e}_n\|} .
\end{aligned}$$

(2) A la hora de aplicar el método de Gram-Schmidt, a menudo es muy útil transformar una base, o bien un sistema de generadores de un espacio vectorial en otra más “sencilla” mediante **operaciones elementales de filas**. Por ejemplo, en \mathbb{R}^4 , dotado de la base canónica, sea el subespacio

$$W = \text{lin}\{(1, 3, 4, 1), (2, 6, 8, 2), (2, 5, 7, 2)\}$$

Consideremos la matriz cuyas filas son las coordenadas de los vectores del sistema de generadores considerado. Mediante operaciones elementales de fila, podemos reducir esta matriz hasta llegar a su forma escalonada reducida (o de Hermite) por filas. Tenemos:

$$\begin{pmatrix} 1 & 3 & 4 & 1 \\ 2 & 6 & 8 & 2 \\ 2 & 5 & 7 & 2 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 3 & 4 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Las primeras dos filas de la forma escalonada reducida de la matriz nos proporcionan una base de W . Entonces, $W = \text{lin}\{(1, 0, 1, 1), (0, 1, 1, 0)\}$.

En el caso del ejercicio anterior, aplicando este método llegamos al sistema de generadores

$$\{\mathbf{y}_1 = (1, 1, -1, -2), \mathbf{y}_2 = (0, 3, 3, 7), \mathbf{y}_3 = (0, 0, 6, 11)\}.$$

La aplicación del método de Gram-Schmidt a esta base es más directa.

5.9.3 Bases ortonormales

Teorema 86. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\dim E = n$. Todo conjunto de vectores no nulos $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ ortogonales dos a dos se puede ampliar a una base ortogonal de E .

Demostración. Es suficiente observar que, por el teorema de completación de una base, podemos encontrar vectores independientes $\mathbf{y}_{k+1}, \dots, \mathbf{y}_n$ tales que el conjunto $\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_{k+1}, \dots, \mathbf{y}_n\}$ sea una base de E . Aplicando el procedimiento de Gram-Schmidt a esta base, los vectores $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ quedan invariados (siendo ortogonales), mientras que los demás son transformados en sendos vectores $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ de modo que el nuevo conjunto $\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ es una base ortogonal de E . \square

5.10 Complemento ortogonal

Definición 121. Sea E un espacio vectorial euclídeo, $\mathbf{x} \in E$, $U \subset E$ subespacio de E . Se dice que \mathbf{x} es ortogonal a U y escribimos $\mathbf{x} \perp U$ si \mathbf{x} es ortogonal a todos los vectores de U , esto es

$$\mathbf{x} \perp U \iff \forall \mathbf{y} \in U : \mathbf{x} \perp \mathbf{y}$$

Teorema 87. Sea E un espacio vectorial euclídeo, $U = L(\mathbf{y}_1, \dots, \mathbf{y}_k)$ subespacio de E . Entonces, si $\mathbf{x} \in E$, se verifica

$$\mathbf{x} \perp U \iff \begin{cases} \mathbf{x} \perp \mathbf{y}_1 \\ \vdots \\ \mathbf{x} \perp \mathbf{y}_k \end{cases}$$

Demostración. \implies obvia.

\Leftarrow Sea $\mathbf{y} \in U$; $\mathbf{y} = \alpha_1 \mathbf{y}_1 + \dots + \alpha_k \mathbf{y}_k$;

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \alpha_1 \mathbf{y}_1 + \dots + \alpha_k \mathbf{y}_k \rangle = \alpha_1 \langle \mathbf{x}, \mathbf{y}_1 \rangle + \dots + \alpha_k \langle \mathbf{x}, \mathbf{y}_k \rangle = 0 .$$

Proposición 40. Sea E espacio vectorial euclídeo, $\dim E = n$, y sea $U \subset E$ un subespacio. Entonces el conjunto

$$U^\perp = \{ \mathbf{x} \in E \mid \mathbf{x} \perp U \}$$

es un subespacio vectorial de E , y

$$E = U \oplus U^\perp .$$

Demostración. Sean $\mathbf{x}_1, \mathbf{x}_2 \in U^\perp$, $\lambda_1, \lambda_2 \in \mathbb{R}$, $\mathbf{y} \in U$. Entonces

$$\langle \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2, \mathbf{y} \rangle = \lambda_1 \langle \mathbf{x}_1, \mathbf{y} \rangle + \lambda_2 \langle \mathbf{x}_2, \mathbf{y} \rangle = 0 \implies \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \in U^\perp .$$

Sea ahora $\{ \mathbf{u}_1, \dots, \mathbf{u}_k \}$ una base ortogonal de U , y amplíemos dicha base a una base ortogonal: $\mathcal{B} = \{ \mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n \}$. Demostremos que $\{ \mathbf{u}_{k+1}, \dots, \mathbf{u}_n \}$ es una base de U^\perp . Evidentemente, $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n \in U^\perp$ por construcción, y son linealmente independientes. Demostremos que generan U^\perp . Sea $\mathbf{z} \in U^\perp$. Entonces, siendo \mathbf{z} un vector de E ,

$$\mathbf{z} = z_1 \mathbf{u}_1 + \dots + z_n \mathbf{u}_n, \quad z_i = \frac{\langle \mathbf{z}, \mathbf{u}_i \rangle}{\| \mathbf{u}_i \|^2} .$$

Puesto que $\mathbf{z} \in U^\perp$, $z_1 = \dots = z_k = 0$, o sea

$$\mathbf{z} = z_{k+1} \mathbf{u}_{k+1} + \dots + z_n \mathbf{u}_n \implies U^\perp = L(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n).$$

□

Definición 122. El subespacio $U^\perp = \{ \mathbf{x} \in E \mid \mathbf{x} \perp U \}$ se llamará *complemento ortogonal de U* .

Determinemos U^\perp . Si $\dim U = k$, y $\{ \mathbf{u}_1, \dots, \mathbf{u}_k \}$ es una base de U , entonces

$$\mathbf{x} \in U^\perp \iff \begin{cases} \langle \mathbf{x}, \mathbf{u}_1 \rangle = 0 \\ \vdots \\ \langle \mathbf{x}, \mathbf{u}_k \rangle = 0 \end{cases}$$

Ejemplo 5.17. Sea \mathbb{R}^4 dotado del producto escalar canónico $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \dots + x_4 y_4$. Sea $U = \text{lin}\{(0, 0, 1, 1)\}$. Entonces

$$\mathbf{x} = (x_1, x_2, x_3, x_4) \in U^\perp \iff \langle (x_1, x_2, x_3, x_4), (0, 0, 1, 1) \rangle = 0 \iff x_3 + x_4 = 0$$

Luego U^\perp es descrito por la ecuación cartesiana $x_3 + x_4 = 0$.

Ejemplo 5.18. Sea \mathbb{R}^4 dotado de un producto escalar cuya matriz de Gram es

$$A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & 1 & 3 & 2 \\ 1 & 0 & 2 & 4 \end{pmatrix}$$

Sea $U = \text{lin}\{(0, 0, 1, 0), (0, 0, 0, 1)\}$. Determinemos U^\perp .

$$(x_1, x_2, x_3, x_4) \in U^\perp \iff \begin{cases} \langle (x_1, x_2, x_3, x_4), (0, 0, 1, 0) \rangle = 0 \\ \langle (x_1, x_2, x_3, x_4), (0, 0, 0, 1) \rangle = 0 \end{cases}$$

$$(x_1, x_2, x_3, x_4) \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & 1 & 3 & 2 \\ 1 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = 0 \iff$$

$$(x_1, x_2, x_3, x_4) \begin{pmatrix} 1 \\ 1 \\ 3 \\ 2 \end{pmatrix} = 0 \iff x_1 + x_2 + 3x_3 + 2x_4 = 0$$

$$(x_1, x_2, x_3, x_4) \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & 1 & 3 & 2 \\ 1 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 0 \iff$$

$$(x_1, x_2, x_3, x_4) \begin{pmatrix} 1 \\ 0 \\ 2 \\ 4 \end{pmatrix} = 0 \iff x_1 + 2x_3 + 4x_4 = 0$$

Hemos por tanto obtenido el sistema

$$\begin{cases} x_1 + x_2 + 3x_3 + 2x_4 = 0 \\ x_1 + 2x_3 + 4x_4 = 0 \end{cases} \iff \begin{cases} x_2 = -x_3 + 2x_4 \\ x_1 = -2x_3 - 4x_4 \end{cases}$$

La solución general del sistema es

$$\{(-2\alpha - 4\beta, -\alpha + 2\beta, \alpha, \beta) \mid \alpha, \beta \in \mathbb{R}\}.$$

En definitiva,

$$U^\perp = \text{lin}\{(-2, -1, 1, 0), (-4, 2, 0, 1)\}.$$

5.11 Proyección ortogonal

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $U \subset E$ un subespacio. Como consecuencia de la Proposición 40, todo vector $\mathbf{x} \in E$ podrá escribirse de forma única como

$$\mathbf{x} = \mathbf{u} + \mathbf{v}, \quad \mathbf{u} \in U, \quad \mathbf{v} \in U^\perp .$$

Diremos entonces que \mathbf{u} es la **proyección ortogonal** de \mathbf{x} sobre U y lo denotaremos con

$$\mathbf{u} = \pi_U(\mathbf{x}) .$$

Podemos interpretar \mathbf{u} como el único vector tal que $\mathbf{u} \in U$ y $\mathbf{x} - \mathbf{u} \in U^\perp$. Entonces, podemos introducir la aplicación

$$\pi_U : E \longrightarrow U, \quad \pi_U(\mathbf{x}) := \mathbf{u}$$

es decir, la aplicación que asocia a \mathbf{x} su proyección ortogonal sobre U . La aplicación π_U es denominada **proyector ortogonal sobre U**. Observamos que, como consecuencia inmediata de la definición de proyector ortogonal, valen las propiedades siguientes:

- (a) $\text{im}\pi_U = U$;
- (b) $\ker\pi_U = U^\perp$
- (c) $\pi_U^2 = \pi_U$.

La proposición siguiente nos permite escribir explícitamente un proyector ortogonal sobre un subespacio a partir de una base del subespacio mismo.

Proposición 41. *Si $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ es una base ortogonal de U , entonces $\forall \mathbf{x} \in E$ se verifica*

$$\pi_U(\mathbf{x}) = \frac{\langle \mathbf{u}_1, \mathbf{x} \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 + \dots + \frac{\langle \mathbf{u}_k, \mathbf{x} \rangle}{\|\mathbf{u}_k\|^2} \mathbf{u}_k . \quad (5.26)$$

Teorema 88. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, y $U \subset E$ un subespacio. Si $\mathbf{x} \in E$, entonces $\mathbf{u} := \pi_U(\mathbf{x})$ es el único vector en U que hace mínima la distancia*

$$d(\mathbf{x}, \mathbf{w}) = \|\mathbf{x} - \mathbf{w}\|, \quad \mathbf{w} \in U ,$$

es decir

$$\forall \mathbf{w} \in U, \quad \mathbf{u} \neq \mathbf{w}, \quad \|\mathbf{x} - \mathbf{u}\| \leq \|\mathbf{x} - \mathbf{w}\| .$$

Demostración. Sea $\mathbf{u} = \pi_U(\mathbf{x})$. Entonces, si $\mathbf{w} \in U$, tenemos que $\mathbf{u} - \mathbf{w} \in U$, y al mismo tiempo, $\mathbf{x} - \mathbf{u} \in U^\perp$. Por tanto,

$$(\mathbf{x} - \mathbf{u}) \perp (\mathbf{u} - \mathbf{w}) .$$

Ahora bien, utilizando el teorema de Pitágoras, tenemos

$$\|\mathbf{x} - \mathbf{w}\|^2 = \|\mathbf{x} - \mathbf{u} + \mathbf{u} - \mathbf{w}\|^2 = \|\mathbf{x} - \mathbf{u}\|^2 + \|\mathbf{u} - \mathbf{w}\|^2 \geq \|\mathbf{x} - \mathbf{u}\|^2.$$

La igualdad se realiza si y solo si

$$\|\mathbf{u} - \mathbf{w}\| = 0 \implies \mathbf{u} = \mathbf{w}.$$

□

Ejercicio 5.19. Sea \mathbb{R}^3 dotado del producto escalar canónico. Calcular la proyección del vector $\mathbf{x} = (1, 0, 0)$ sobre el subespacio

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = 0\}.$$

Resolución. Una base de U es dada por $B = \{(1, -1, 0), (0, 0, 1)\}$. Sea $(x, y, z) \in U^\perp$. Entonces

$$\begin{cases} \langle (x, y, z), (1, -1, 0) \rangle = 0 \\ \langle (x, y, z), (0, 0, 1) \rangle = 0 \end{cases} \iff \begin{cases} x - y = 0 \\ z = 0 \end{cases}$$

Por tanto, $U^\perp = \text{lin}\{(1, 1, 0)\}$. Escribimos el vector $\mathbf{x} = (1, 0, 0)$ en la forma $\mathbf{x} = \mathbf{y} + \mathbf{z}$, con $\mathbf{y} \in U$, $\mathbf{z} \in U^\perp$. Tenemos

$$(1, 0, 0) = \alpha_1(1, -1, 0) + \alpha_2(0, 0, 1) + \alpha_3(1, 1, 0)$$

es decir

$$\begin{cases} \alpha_1 + \alpha_3 = 1, \\ -\alpha_1 + \alpha_3 = 0 \\ \alpha_2 = 0 \end{cases} \iff \begin{cases} \alpha_1 = \frac{1}{2}, \\ \alpha_2 = 0 \\ \alpha_3 = \frac{1}{2} \end{cases}.$$

En definitiva,

$$\mathbf{y} = \frac{1}{2}(1, -1, 0) = \left(\frac{1}{2}, -\frac{1}{2}, 0\right), \quad \mathbf{z} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$$

5.12 Producto vectorial en \mathbb{R}^3

Definición 123. Sea \mathbb{R}^3 dotado del producto escalar canónico, $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$. Se define **producto vectorial** de \mathbf{x} e \mathbf{y} a la aplicación lineal

$$\mathbf{x} \times \mathbf{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \quad (5.27)$$

$$= \left(\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}, \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \right). \quad (5.28)$$

Vale el siguiente

Teorema 89. Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. El vector $\mathbf{x} \times \mathbf{y}$ es el único vector que verifica las propiedades siguientes:

- (1) $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \alpha$
- (2) $(\mathbf{x} \times \mathbf{y}) \perp \mathbf{x}, \quad (\mathbf{x} \times \mathbf{y}) \perp \mathbf{y}$
- (3) Si \mathbf{x} no es paralelo a \mathbf{y} , la base $C = \{\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}\}$ tiene orientación positiva respecto de la base canónica, es decir la matriz de cambio de la base canónica a la base C tiene determinante positivo.

Otras propiedades. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$, tenemos

- (a) $\mathbf{x} \times \mathbf{y} = \mathbf{0} \iff \mathbf{x}$ e \mathbf{y} son linealmente independientes
- (b) $(\lambda \mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (\lambda \mathbf{y}) = \lambda(\mathbf{x} \times \mathbf{y}) \quad \forall \lambda \in \mathbb{R}$
- (c) $(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}$
- (d) $(\mathbf{x} \times \mathbf{y}) = -(\mathbf{y} \times \mathbf{x})$

5.13 Formas sesquilineales y cuadráticas en espacios vectoriales complejos

5.13.1 Definición

Sea V un espacio vectorial sobre \mathbb{C} . Una forma sesquilineal es una aplicación:

$$\omega : V \times V \implies \mathbb{C}$$

tal que $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \forall \lambda \in \mathbb{C}$,

$$\omega(\mathbf{x}, \mathbf{y} + \mathbf{z}) = \omega(\mathbf{x}, \mathbf{y}) + \omega(\mathbf{x}, \mathbf{z}) \quad (5.29)$$

$$\omega(\mathbf{x}, \lambda \mathbf{y}) = \lambda \omega(\mathbf{x}, \mathbf{y}) \quad (5.30)$$

$$\omega(\mathbf{x} + \mathbf{y}, \mathbf{z}) = \omega(\mathbf{x}, \mathbf{z}) + \omega(\mathbf{y}, \mathbf{z}) \quad (5.31)$$

$$\omega(\lambda \mathbf{x}, \mathbf{y}) = \bar{\lambda} \omega(\mathbf{x}, \mathbf{y}) \quad (5.32)$$

Definición 124. Una forma sesquilineal se dirá *Hermítica* si

$$\omega(\mathbf{x}, \mathbf{y}) = \overline{\omega(\mathbf{y}, \mathbf{x})} \quad (5.33)$$

Una forma sesquilineal se dirá *antihermítica* si

$$\omega(\mathbf{x}, \mathbf{y}) = -\overline{\omega(\mathbf{y}, \mathbf{x})} \quad (5.34)$$

Denotaremos con $S_q(V, \mathbb{C})$ el conjunto de todas las formas sesquilineales sobre \mathbb{C} .

Las propiedades siguientes son una consecuencia directa de las definiciones anteriores.

Proposición 42. $(S_q(V, \mathbb{C}), +, \cdot)$ es un espacio vectorial sobre \mathbb{C} .

Proposición 43. $\omega \in S_q(V, \mathbb{C})$ es hermitica $\iff i\omega$ es antihermitica.

Definición 125. Llamaremos forma cuadrática hermitica asociada a la forma sesquilineal hermitica φ a la aplicación

$$\begin{aligned} \Phi : V &\longrightarrow \mathbb{R} \\ \forall \mathbf{x} \in V, \quad \Phi(\mathbf{x}) &= \varphi(\mathbf{x}, \mathbf{x}) . \end{aligned}$$

Proposición 44. Sea V un espacio vectorial sobre \mathbb{C} . $\forall \mathbf{x} \in V, \forall \lambda \in \mathbb{C}$:

$$\Phi(\lambda\mathbf{x}) = |\lambda|^2 \Phi(\mathbf{x}) .$$

Demostración.

$$\Phi(\lambda\mathbf{x}) = \varphi(\lambda\mathbf{x}, \lambda\mathbf{x}) = \lambda\bar{\lambda}\varphi(\mathbf{x}, \mathbf{x}) = |\lambda|^2 \Phi(\mathbf{x}) .$$

□

Definición 126. Una matriz $A \in M_n(\mathbb{C})$ se dice hermitica si

$$A^t = \bar{A} .$$

Proposición 45. La matriz de una forma sesquilineal hermitica $\varphi \in S_q(V, \mathbb{C})$ es hermitica.

Demostración. Si $\mathbf{e}_1, \dots, \mathbf{e}_n$ es una base de V , tenemos:

$$a_{ji} = \varphi(\mathbf{e}_j, \mathbf{e}_i) = \overline{\varphi(\mathbf{e}_i, \mathbf{e}_j)} = \bar{a}_{ij}, \quad i, j = 1, \dots, n.$$

Entonces $A^t = \bar{A}$.

□

Teorema 90. Sea V un espacio vectorial sobre \mathbb{C} , $\dim V = n$, sea $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de V y $\varphi \in S_q(V, \mathbb{C})$. Entonces $\forall \mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$, tenemos

$$\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \bar{x}_i y_j , \quad (5.35)$$

donde $a_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j)$. Equivalentemente:

$$\varphi(\mathbf{x}, \mathbf{y}) = \bar{X}^t A Y = Y^t A^t \bar{X} \quad (5.36)$$

Demostración. Sean $\mathbf{x}, \mathbf{y} \in V$. Tenemos

$$\varphi(\mathbf{x}, \mathbf{y}) = \varphi\left(\sum_{i=1}^n x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j\right) = \sum_{i=1}^n \sum_{j=1}^n \bar{x}_i y_j \varphi(\mathbf{e}_i, \mathbf{e}_j) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \bar{x}_i y_j$$

Pasando a las matrices asociadas, y teniendo en cuenta que $\varphi(\mathbf{x}, \mathbf{y})$ es un escalar, y coincide con su traspuesta, obtenemos la fórmula (5.36). □

Está claro que, si Φ es la forma cuadrática asociada a $\varphi \in S_q(V, \mathbb{C})$, tenemos inmediatamente

$$\Phi(\mathbf{x}) = \bar{X}^t A X = X^t A^t \bar{X} . \quad (5.37)$$

5.13.2 Teorema de los autovalores reales para matrices hermíticas

Teorema 91. *Una matriz hermítica de orden n posee n autovalores reales.*

Demostración. Sea V un espacio vectorial sobre \mathbb{C} , y A la matriz de una forma sesquilineal Hermítica $\varphi \in S_q(V, \mathbb{C})$. Tenemos $\varphi(\mathbf{x}, \mathbf{y}) = \bar{X}^t A Y$. Si $\Phi(\mathbf{x})$ es la forma cuadrática asociada,

$$\Phi(\mathbf{x}) = \bar{X}^t A X.$$

La ecuación característica $\det(A - \lambda \mathbb{1}_n) = 0$, siendo una ecuación algebraica de grado n con coeficientes en \mathbb{C} , posee n raíces complejas. Sea entonces $\lambda \in \mathbb{C}$ un autovalor de A y X el vector columna de las componentes de un autovector correspondiente. Tenemos

$$A X = \lambda X$$

y

$$\bar{X}^t A X = \bar{X}^t \lambda X = \lambda \bar{X}^t X. \quad (5.38)$$

Por otro lado, tomando las transpuestas de estas matrices y teniendo en cuenta que $A^t = \bar{A}$, tenemos

$$X^t A^t \bar{X} = X^t \bar{A} \bar{X} = X^t \overline{(\lambda X)} = \bar{\lambda} X^t \bar{X},$$

es decir

$$X^t \bar{A} \bar{X} = \bar{\lambda} X^t \bar{X}.$$

Tomando la transpuesta de primer y segundo miembro, tenemos

$$\bar{X}^t \bar{A}^t X = \bar{\lambda} \bar{X}^t X$$

que nos da ($\bar{A}^t = A$)

$$\bar{X}^t A X = \bar{\lambda} \bar{X}^t X$$

Por tanto, comparando esta ecuación con la ec. (5.38), es decir $\bar{X}^t A X = \lambda \bar{X}^t X$, llegamos a la relación

$$\lambda(\bar{X}^t X) = \bar{\lambda}(\bar{X}^t X)$$

Como $X \neq O$, se tiene $\bar{X}^t X \neq O$, y en definitiva

$$\lambda = \bar{\lambda}.$$

□

Una matriz hermítica real es obviamente una matriz simétrica: la condición $A^t = \bar{A}$ se convierte en $A^t = A$. Por tanto, como consecuencia del Teorema 91 tenemos el

Corolario 15. *Una matriz simétrica real posee n autovalores reales.*

Teorema 92. *Sea V un espacio vectorial sobre \mathbb{C} , $\varphi \in S_q(V, \mathbb{C})$, y A la matriz de φ respecto de una base \mathcal{B} de V . Sea \mathcal{B}' otra base de V y P la matriz de cambio de la base \mathcal{B} a la base \mathcal{B}' . Entonces la matriz de φ respecto de la base \mathcal{B}' es*

$$A' = \bar{P}^t A P. \quad (5.39)$$

5.14 Productos escalares hermíticos

Definición 127. Un espacio vectorial E sobre \mathbb{C} se dirá espacio hermítico o unitario si es dotado de una forma sesquilineal hermítica definida positiva

$$\langle \cdot, \cdot \rangle : E \times E \longrightarrow \mathbb{C},$$

que también se denomina **producto escalar hermítico**, y satisface las propiedades siguientes:

$\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in E, \forall \lambda \in \mathbb{C}$:

$$\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle \quad (5.40)$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \quad (5.41)$$

$$\langle \mathbf{x}, \lambda \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle \quad (5.42)$$

$$\langle \lambda \mathbf{x}, \mathbf{y} \rangle = \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle \quad (5.43)$$

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0, \quad \text{y} \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}. \quad (5.44)$$

La aplicación

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

se dirá la norma asociada al producto escalar hermítico $\langle \cdot, \cdot \rangle$. Como consecuencia de la Proposición (44) tenemos

$$\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\| \quad \forall \mathbf{x} \in E, \quad \lambda \in \mathbb{C}$$

Ejemplo 5.20. Consideremos el espacio hermítico \mathbb{C}^n dotado de la aplicación

$$\langle \mathbf{x}, \mathbf{y} \rangle = \bar{x}_1 y_1 + \dots + \bar{x}_n y_n \quad (5.45)$$

Se comprueba fácilmente que se trata de un producto escalar hermítico, denominado producto escalar canónico (o estándar) en \mathbb{C}^n . La norma inducida por este producto escalar es

$$\|\mathbf{x}\| = \sqrt{|x_1|^2 + \dots + |x_n|^2} \quad (5.46)$$

5.14.1 Identidad de polarización

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio con producto escalar. Si E es euclídeo,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 \right) \quad (5.47)$$

Si E es hermítico, tenemos

$$(a) \quad \operatorname{Re}(\langle \mathbf{x}, \mathbf{y} \rangle) = \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 \right) \quad (5.48)$$

$$(b) \quad \operatorname{Im}(\langle \mathbf{x}, \mathbf{y} \rangle) = \frac{1}{4} \left(\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2 \right) \quad (5.49)$$

Demostración. La ec. (5.47) ya está demostrada.

$$\begin{aligned}
 (a) \quad \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \\
 &= \|\mathbf{x}\|^2 + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \|\mathbf{y}\|^2 \\
 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \\
 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\operatorname{Re}(\langle \mathbf{x}, \mathbf{y} \rangle) \\
 \|\mathbf{x} - \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\operatorname{Re}(\langle \mathbf{x}, \mathbf{y} \rangle) \implies \\
 4\operatorname{Re}(\langle \mathbf{x}, \mathbf{y} \rangle) &= \|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2
 \end{aligned}$$

$$\begin{aligned}
 (b) \quad \|\mathbf{x} + i\mathbf{y}\|^2 &= \langle \mathbf{x} + i\mathbf{y}, \mathbf{x} + i\mathbf{y} \rangle = \\
 &= \|\mathbf{x}\|^2 + |i|^2 \|\mathbf{y}\|^2 + i\langle \mathbf{x}, \mathbf{y} \rangle - i\langle \mathbf{y}, \mathbf{x} \rangle \\
 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + i(\langle \mathbf{x}, \mathbf{y} \rangle - \overline{\langle \mathbf{x}, \mathbf{y} \rangle}) \\
 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\operatorname{Im}(\langle \mathbf{x}, \mathbf{y} \rangle) \\
 \|\mathbf{x} - i\mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\operatorname{Im}(\langle \mathbf{x}, \mathbf{y} \rangle) \implies \\
 4\operatorname{Im}(\langle \mathbf{x}, \mathbf{y} \rangle) &= \|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2
 \end{aligned}$$

□

5.14.2 Desigualdad de Cauchy-Schwarz y triangular en un espacio hermítico

Sea E un espacio Hermítico. Tenemos la desigualdad

$$\forall \mathbf{x}, \mathbf{y} \in E : \quad |\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

conocida como desigualdad de Cauchy-Schwarz. La desigualdad triangular de la norma se escribe formalmente de la misma manera:

$$\forall \mathbf{x}, \mathbf{y} \in E : \quad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| .$$

Mencionamos también el

Teorema 93 (Riesz). *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial unitario, $\omega \in E^* = L(E, \mathbb{C})$. Entonces existe un único vector $\mathbf{v} \in E$ tal que*

$$\forall \mathbf{x} \in E : \quad \omega(\mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle .$$

5.14.3 Método de Gram-Schmidt en espacios hermíticos

El método de ortonormalización de Gram-Schmidt también es válido en el caso de un espacio hermítico. Las relaciones (5.26) siguen siendo válidas cuando $\mathbb{K} = \mathbb{C}$, teniendo en cuenta que $\langle \cdot, \cdot \rangle$ es en este caso un producto escalar hermítico.

5.15 Ejercicios

Ejercicio 5.21. En \mathbb{R}^4 , dotado del producto escalar canónico, sean el vector $\mathbf{x} = (5, 1, 5, 1)$ y los subespacios

$$V_1 = \{(x, y, z, w) | x = z, y = w\}, \quad V_2 = \text{lin}\{(1, 0, 0, 0), (0, 1, 0, 0)\}.$$

¿Qué ángulo forma $\mathbf{x}_1 = \pi_{V_1}(\mathbf{x})$ con $\mathbf{x}_2 = \pi_{V_2}(\mathbf{x})$?

Resolución. Una base de V_1 es dada por

$$V_1 = \text{lin}\{(1, 0, 1, 0), (0, 1, 0, 1)\}$$

Tenemos

$$\begin{aligned} \pi_{V_1}(\mathbf{x}) &= \frac{\langle (1, 0, 1, 0), (5, 1, 5, 1) \rangle}{2} (1, 0, 1, 0) + \frac{\langle (0, 1, 0, 1), (5, 1, 5, 1) \rangle}{2} (0, 1, 0, 1) \\ &= (5, 0, 5, 0) + (0, 1, 0, 1) = (5, 1, 5, 1) = \mathbf{x}_1 \end{aligned}$$

$$\begin{aligned} \pi_{V_2}(\mathbf{x}) &= \langle (1, 0, 0, 0), (5, 1, 5, 1) \rangle (1, 0, 0, 0) + \langle (0, 1, 0, 0), (5, 1, 5, 1) \rangle (0, 1, 0, 0) \\ &= (5, 0, 0, 0) + (0, 1, 0, 0) = (5, 1, 0, 0) = \mathbf{x}_2 \end{aligned}$$

Entonces el ángulo entre \mathbf{x}_1 y \mathbf{x}_2 es

$$\alpha = \arccos \frac{\langle (5, 1, 5, 1), (5, 1, 0, 0) \rangle}{\sqrt{52}\sqrt{26}} = \arccos \frac{1}{\sqrt{2}} = \frac{\pi}{4}.$$

Ejercicio 5.22.

(a) Comprobar que la aplicación $f : \mathbb{C}^3 \times \mathbb{C}^3 \rightarrow \mathbb{C}$ dada por

$$f[(x_1, x_2, x_3), (y_1, y_2, y_3)] = \bar{x}_1 y_1 - i \bar{x}_1 y_2 + \bar{x}_3 y_2$$

es una forma sesquilineal sobre \mathbb{C}^3 .

(b) Determinar la matriz de f en la base canónica de \mathbb{C}^3 .

(c) Hallar la expresión de f en la base \mathcal{B} de \mathbb{C}^3 formada por los vectores

$$a_1 = (i, 0, 0), \quad a_2 = (0, 1, i), \quad a_3 = (0, 1, -i).$$

Resolución.

(a) Comprobemos la sesquilinealidad de f :

$$\begin{aligned} f(\mathbf{x} + \mathbf{y}, \mathbf{z}) &= f((x_1, x_2, x_3) + (y_1, y_2, y_3), (z_1, z_2, z_3)) = (\bar{x}_1 + \bar{y}_1)z_1 - i(\bar{x}_1 + \bar{y}_1)z_2 + (\bar{x}_3 + \bar{y}_3)z_2 \\ &= \bar{x}_1 z_1 - i \bar{x}_1 z_2 + \bar{x}_3 z_2 + \bar{y}_1 z_1 - i \bar{y}_1 z_2 + \bar{y}_3 z_2 \\ &= f((x_1, x_2, x_3), (z_1, z_2, z_3)) + f((y_1, y_2, y_3), (z_1, z_2, z_3)) = f(\mathbf{x}, \mathbf{z}) + f(\mathbf{y}, \mathbf{z}). \end{aligned}$$

$$\begin{aligned}
f(\mathbf{x}, \mathbf{y} + \mathbf{z}) &= f((x_1, x_2, x_3), (y_1, y_2, y_3) + (z_1, z_2, z_3)) = \bar{x}_1(y_1 + z_1) - i\bar{x}_1(y_2 + z_2) + \bar{x}_3(y_2 + z_2) \\
&= \bar{x}_1y_1 - i\bar{x}_1y_2 + \bar{x}_3y_2 + \bar{x}_1z_1 - i\bar{x}_1z_2 + \bar{x}_3z_2 \\
&= f((x_1, x_2, x_3), (y_1, y_2, y_3)) + f((x_1, x_2, x_3), (z_1, z_2, z_3)) = f(\mathbf{x}, \mathbf{y}) + f(\mathbf{x}, \mathbf{z}) .
\end{aligned}$$

$$f(\lambda\mathbf{x}, \mathbf{y}) = \bar{\lambda}\bar{x}_1y_1 - i\bar{\lambda}\bar{x}_1y_2 + \bar{\lambda}\bar{x}_3y_2 = \bar{\lambda}f(\mathbf{x}, \mathbf{y})$$

$$f(\mathbf{x}, \lambda\mathbf{y}) = \lambda\bar{x}_1y_1 - i\lambda\bar{x}_1y_2 + \lambda\bar{x}_3y_2 = \lambda f(\mathbf{x}, \mathbf{y})$$

(b) Los elementos de la matriz A asociada a f respecto de la base canónica de \mathbb{C}^3 son $a_{ij} = f(\mathbf{e}_i, \mathbf{e}_j)$. Tenemos

$$A = \begin{pmatrix} 1 & -i & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

(c) La matriz de cambio de la base canónica a la base \mathcal{B} es

$$P = \begin{pmatrix} i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & i & -i \end{pmatrix}$$

La matriz de f en \mathcal{B} es

$$A' = \bar{P}^t A P ,$$

con

$$A' = \begin{pmatrix} 1 & -1 & -1 \\ 0 & -i & -i \\ 0 & i & i \end{pmatrix}$$

Chapter 6

Aplicaciones lineales entre espacios con producto escalar

6.1 Aplicación adjunta

Introducimos la noción de aplicación adjunta de una aplicación lineal, que es fundamental en muchas ramas de la matemática y de la física teórica: por ejemplo en mecánica cuántica. Ante todo, enunciamos el siguiente

Teorema 94. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$. Sea $f \in L(E)$. Existe una única aplicación lineal*

$$f^\dagger : E \rightarrow E$$

que verifica la propiedad siguiente:

$$\forall \mathbf{x}, \mathbf{y} \in E : \quad \langle \mathbf{x}, f(\mathbf{y}) \rangle = \langle f^\dagger(\mathbf{x}), \mathbf{y} \rangle \quad (6.1)$$

Llamaremos **aplicación adjunta** de f a la aplicación lineal $f^\dagger : E \rightarrow E$ definida por la relación (6.1). La aplicación adjunta posee algunas propiedades importantes.

Teorema 95. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, y $f \in L(E)$. Valen las propiedades siguientes:*

$$(1) \quad (f^\dagger)^\dagger = f \quad (6.2)$$

$$(2) \quad (f + g)^\dagger = f^\dagger + g^\dagger \quad (6.3)$$

$$(3) \quad (\lambda f)^\dagger = \bar{\lambda} f^\dagger \quad (6.4)$$

$$(4) \quad (f \circ g)^\dagger = g^\dagger \circ f^\dagger \quad (6.5)$$

Demostración. $\forall \mathbf{x}, \mathbf{y} \in E$:

(1)

$$\langle \mathbf{x}, f(\mathbf{y}) \rangle = \langle f^\dagger(\mathbf{x}), \mathbf{y} \rangle = \overline{\langle \mathbf{y}, f^\dagger(\mathbf{x}) \rangle} = \overline{\langle (f^\dagger)^\dagger(\mathbf{y}), \mathbf{x} \rangle} = \langle \mathbf{x}, (f^\dagger)^\dagger(\mathbf{y}) \rangle$$

(2)

$$\begin{aligned} \langle (f+g)^\dagger(\mathbf{x}), \mathbf{y} \rangle &= \langle \mathbf{x}, (f+g)(\mathbf{y}) \rangle = \langle \mathbf{x}, f(\mathbf{y}) + g(\mathbf{y}) \rangle \\ &= \langle \mathbf{x}, f(\mathbf{y}) \rangle + \langle \mathbf{x}, g(\mathbf{y}) \rangle = \langle f^\dagger(\mathbf{x}), \mathbf{y} \rangle + \langle g^\dagger(\mathbf{x}), \mathbf{y} \rangle \end{aligned}$$

(3)

$$\begin{aligned} \langle (\lambda f)^\dagger(\mathbf{x}), \mathbf{y} \rangle &= \langle \mathbf{x}, \lambda f(\mathbf{y}) \rangle = \lambda \langle \mathbf{x}, f(\mathbf{y}) \rangle \\ &= \lambda \langle f^\dagger(\mathbf{x}), \mathbf{y} \rangle = \langle \bar{\lambda} f^\dagger(\mathbf{x}), \mathbf{y} \rangle \end{aligned}$$

(4)

$$\begin{aligned} \langle (f \circ g)^\dagger(\mathbf{x}), \mathbf{y} \rangle &= \langle \mathbf{x}, (f \circ g)(\mathbf{y}) \rangle = \langle \mathbf{x}, f(g(\mathbf{y})) \rangle = \langle \mathbf{x}, f(\mathbf{z}) \rangle, \quad \mathbf{z} := g(\mathbf{y}) \\ &= \langle f^\dagger(\mathbf{x}), \mathbf{z} \rangle = \langle f^\dagger(\mathbf{x}), g(\mathbf{y}) \rangle = \langle g^\dagger(f^\dagger(\mathbf{x})), \mathbf{y} \rangle = \langle (g^\dagger \circ f^\dagger)(\mathbf{x}), \mathbf{y} \rangle \end{aligned}$$

□

La proposición siguiente nos permite calcular la inversa de la aplicación adjunta, si existe, de una forma directa.

Proposición 46. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico. Si f es biyectiva, también lo es f^\dagger , y*

$$(f^\dagger)^{-1} = (f^{-1})^\dagger .$$

Demostración. Observamos que, si f es biyectiva, $f \circ f^{-1} = I$, siendo $I \in L(E)$ la aplicación identidad: $\forall \mathbf{x} \in E: I(\mathbf{x}) = \mathbf{x}$. También se comprueba inmediatamente que $I^\dagger = I$. Entonces

$$(f \circ f^{-1})^\dagger = I = (f^{-1})^\dagger \circ f^\dagger \implies (f^\dagger)^{-1} = (f^{-1})^\dagger .$$

□

6.2 Núcleo de f e imagen de la aplicación adjunta

Lema 9. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, y $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base ortogonal de E . Si $\mathbf{x} \in E$ es ortogonal a los vectores de \mathcal{B} , entonces $\mathbf{x} = \mathbf{0}$.*

Demostración. Tenemos la descomposición en coeficientes de Fourier

$$\mathbf{x} = \frac{\langle \mathbf{e}_1, \mathbf{x} \rangle}{\|\mathbf{e}_1\|^2} \mathbf{e}_1 + \dots + \frac{\langle \mathbf{e}_n, \mathbf{x} \rangle}{\|\mathbf{e}_n\|^2} \mathbf{e}_n = \mathbf{0} .$$

□

El núcleo de una aplicación f y la imagen de la aplicación adjunta están directamente relacionados. Tenemos que

$$\begin{aligned} \ker f &= \{\mathbf{v} \in E \mid f(\mathbf{v}) = \mathbf{0}\} = \{\mathbf{v} \in E \mid \forall \mathbf{w} \in E : \langle \mathbf{w}, f(\mathbf{v}) \rangle = 0\} \\ &= \{\mathbf{v} \in E \mid \forall \mathbf{w} \in E : \langle f^\dagger(\mathbf{w}), \mathbf{v} \rangle = 0\} = (\operatorname{im} f^\dagger)^\perp. \end{aligned}$$

Entonces, dado que, como consecuencia de la Proposición 40 tenemos que

$$E = (\operatorname{im} f^\dagger) \oplus (\operatorname{im} f^\dagger)^\perp,$$

entonces el espacio E se descompone en la suma directa del núcleo de f y de la imagen de su aplicación adjunta:

$$E = \ker f \oplus \operatorname{im} f^\dagger.$$

□

Ejercicio 6.1. Sea el operador lineal $f \in \mathbb{C}^3$, representado en la base canónica de \mathbb{C}^3 por la matriz

$$\mathcal{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & i & i \end{pmatrix}$$

Siendo la base canónica ortonormal, la matriz asociada al operador adjunto f^\dagger es

$$\mathcal{A}^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -i \\ 0 & 1 & -i \end{pmatrix}$$

La imagen de f es

$$\operatorname{im} f = \{(y_1, y_2, y_3) \in \mathbb{C}^3 \mid iy_2 = y_3\} = \operatorname{lin}\{(1, 0, 0), (0, 1, i)\}$$

y su núcleo es

$$\ker f = \{(x_1, x_2, x_3) \in \mathbb{C}^3 \mid x_1 + x_2 = 0, x_2 + x_3 = 0\} = \operatorname{lin}\{(1, -1, 1)\}$$

Hallemos ahora el núcleo y la imagen de f^\dagger :

$$\begin{aligned} \ker f^\dagger &= \{(x_1, x_2, x_3) \in \mathbb{C}^3 \mid x_1 = 0, x_2 - ix_3 = 0\} = \operatorname{lin}\{(0, i, 1)\} = (\operatorname{im} f)^\perp \\ \operatorname{im} f^\dagger &= \{(y_1, y_2, y_3) \in \mathbb{C}^3 \mid y_1 - y_2 + y_3 = 0\} = \operatorname{lin}\{(1, 1, 0), (0, 1, 1)\} = (\ker f)^\perp \end{aligned}$$

6.3 Representación matricial de la aplicación adjunta

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$. Sea $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base ortonormal de E , y $A = (a_{ij}) \in M_n(\mathbb{C})$ la matriz asociada a $f \in L(E)$ respecto de la base \mathcal{B} . Tenemos

$$f(\mathbf{e}_i) = \sum_{k=1}^n a_{ki} \mathbf{e}_k.$$

Sea $A' = (a'_{ij})$ la matriz asociada a f^\dagger respecto de la base \mathcal{B} :

$$f^\dagger(\mathbf{e}_i) = \sum_{k=1}^n a'_{ki} \mathbf{e}_k$$

Por definición, $\forall \mathbf{x}, \mathbf{y} \in E$:

$$\langle \mathbf{x}, f(\mathbf{y}) \rangle = \langle f^\dagger(\mathbf{x}), \mathbf{y} \rangle .$$

En particular,

$$\begin{aligned} \langle \mathbf{e}_i, f(\mathbf{e}_j) \rangle = \langle f^\dagger(\mathbf{e}_i), \mathbf{e}_j \rangle &\iff \left\langle \mathbf{e}_i, \sum_{k=1}^n a_{kj} \mathbf{e}_k \right\rangle = \left\langle \sum_{k=1}^n a'_{ki} \mathbf{e}_k, \mathbf{e}_j \right\rangle &\iff \\ \sum_{k=1}^n a_{kj} \langle \mathbf{e}_i, \mathbf{e}_k \rangle = \sum_{k=1}^n \overline{a'_{ki}} \langle \mathbf{e}_k, \mathbf{e}_j \rangle &\iff \sum_{k=1}^n a_{kj} \delta_{ik} = \sum_{k=1}^n \overline{a'_{ki}} \delta_{kj} &\iff \\ a_{ij} = \overline{a'_{ji}} &\iff \overline{a_{ij}} = a'_{ji} \end{aligned}$$

y por tanto, tomando la transpuesta de ambos miembros,

$$a'_{ij} = \overline{a_{ji}}$$

Ahora bien si denotamos la matriz \mathcal{A}' con el símbolo \mathcal{A}^\dagger , hemos llegado a la importante relación

$$\mathcal{A}^\dagger = (\overline{\mathcal{A}})^t . \quad (6.6)$$

Definición 128. Sea $\mathcal{A} \in M_n(\mathbb{C})$. Llamaremos *matriz adjunta* de \mathcal{A} a la matriz $\mathcal{A}^\dagger = (\overline{\mathcal{A}})^t$.

Observación 6.1. Nótese que, dada una aplicación $f \in L(E)$, la matriz asociada a la aplicación adjunta f^\dagger es $\mathcal{A}^\dagger = (\overline{\mathcal{A}})^t$ sólo en el caso en que la **base** considerada es **ortonormal**.

6.4 Operadores normales, autoadjuntos y unitarios

En esta sección vamos a introducir algunas nociones muy básicas y, al mismo tiempo, fundamentales de la teoría de los operadores lineales sobre un espacio vectorial.

6.4.1 Definiciones

Definición 129. Sea V un espacio vectorial sobre \mathbb{K} . Llamaremos **operador lineal** a cada endomorfismo de V .

Para denotar la acción de un operador lineal, a menudo utilizaremos la notación siguiente: si

$$A : E \rightarrow E$$

es un operador lineal, escribiremos $A\mathbf{x}$ en lugar de $A(\mathbf{x})$. También se utiliza comúnmente en estos contextos la notación $f(\mathbf{x})$.

Definición 130. Sean $f, g \in L(E)$ dos operadores de un espacio vectorial E . Llamaremos conmutador de f y g al operador

$$[f, g] := f \circ g - g \circ f$$

La definición anterior se extiende al caso de las matrices asociadas.

Es común utilizar la notación ff^\dagger en lugar de $f \circ f^\dagger$.

Definición 131. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, y sea f un operador en E . Diremos que f es un operador **normal** si

$$ff^\dagger = f^\dagger f. \quad (6.7)$$

Diremos que f es un operador **autoadjunto** si coincide con su adjunto:

$$f = f^\dagger \quad (6.8)$$

Diremos que A es un operador **unitario** si

$$ff^\dagger = f^\dagger f = \mathbf{1}. \quad (6.9)$$

Como consecuencia inmediata de la definición anterior, tenemos que si f es autoadjunto, entonces

$$\langle \mathbf{x}, f(\mathbf{y}) \rangle = \langle f(\mathbf{x}), \mathbf{y} \rangle \quad (6.10)$$

mientras que si es unitario,

$$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \langle f(f^\dagger(\mathbf{x})), \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \quad (6.11)$$

Las nociones anteriores se extienden de forma obvia al caso de las matrices asociadas a los operadores normales. Con más precisión, se usa la siguiente

Definición 132. Una matriz $M \in M_n(\mathbb{C})$ se dirá **normal**, **autoadjunta** o **unitaria** si M satisface las relaciones (6.7), (6.8), (6.9) respectivamente.

Ejemplos 6.2.

(a) La matriz

$$\mathcal{N} = \begin{pmatrix} 1 & 0 & 1+i \\ 0 & i & 0 \\ -1-i & 0 & 1 \end{pmatrix}$$

es **normal**. La matriz adjunta es

$$\mathcal{N}^\dagger = \begin{pmatrix} 1 & 0 & -1+i \\ 0 & -i & 0 \\ 1-i & 0 & 1 \end{pmatrix},$$

y se comprueba que $\mathcal{N}\mathcal{N}^\dagger = \mathcal{N}^\dagger\mathcal{N} = \begin{pmatrix} 3 & 0 & 2i \\ 0 & 1 & 0 \\ -2i & 0 & 3 \end{pmatrix}$,

por tanto \mathcal{N} es normal.

(b) La matriz

$$\mathcal{A} = \begin{pmatrix} 1 & i & 1 \\ -i & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

es autoadjunta: $\mathcal{A} = \mathcal{A}^\dagger$.

(c) La matriz

$$\mathcal{U} = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 & -\sqrt{2} & 0 \\ -i & -\sqrt{2}i & \sqrt{3}i \\ 1 & \sqrt{2} & \sqrt{3} \end{pmatrix}$$

es unitaria: tenemos que

$$\mathcal{U}^\dagger = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 & i & 1 \\ -\sqrt{2} & \sqrt{2}i & \sqrt{2} \\ 0 & -\sqrt{3}i & \sqrt{3} \end{pmatrix}$$

y $\mathcal{U}\mathcal{U}^\dagger = \mathcal{U}^\dagger\mathcal{U} = \mathbf{1}$.

Proposición 47. Sea E un espacio vectorial hermítico. Entonces, respecto de una **base ortonormal**:

- (a) La matriz asociada a un operador normal es normal
- (b) La matriz asociada a un operador autoadjunto es autoadjunta
- (c) La matriz asociada a un operador unitario es unitaria.

Observación 6.2. Los operadores autoadjuntos y unitarios son normales, pero existen operadores normales que no son ni autoadjuntos ni unitarios.

Estudiemos ante todo el caso más general de los operadores normales, y en particular su teoría espectral. Para ello, consideremos algunos Lemas preliminares.

6.4.2 Propiedades generales

Lema 10. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio hermítico, $\dim E = n$, y sean H, G dos operadores de V tales que

$$HG = GH.$$

Entonces existe un vector no nulo $\mathbf{y} \in V$ que es simultáneamente autovector de H y G .

Demostración. Como E es un espacio vectorial sobre \mathbb{C} de dimensión finita, entonces el polinomio característico de H posee al menos una raíz en \mathbb{C} . Entonces, $\exists \mathbf{x} \in E$ tal que

$$H\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{x} \in E, \quad \mathbf{x} \neq \mathbf{0}, \quad \lambda \in \mathbb{C}$$

Como H y G conmutan, entonces también $HG^k = G^kH$. Por ejemplo, si $k = 2$, tenemos:

$$HG^2 = (HG)G = (GH)G = G(GH) = G^2H$$

y análogamente para $k > 2$. Por tanto los vectores $G\mathbf{x}$, $G^2\mathbf{x}$, $G^3\mathbf{x}$ son también autovectores de H respecto del mismo autovalor λ :

$$HG^k\mathbf{x} = G^kH\mathbf{x} = \lambda G^k\mathbf{x}, \quad k = 0, 1, 2, \dots$$

Sin embargo, los autovectores \mathbf{x} , $G\mathbf{x}$, \dots , $G^j\mathbf{x}$, \dots no pueden ser todos independientes, dado que E es un espacio de dimensión finita: entonces existe $r \in \mathbb{N}$ tal que $G^{r+1}\mathbf{x}$ depende linealmente de los vectores \mathbf{x} , $G\mathbf{x}$, $G^2\mathbf{x}$, \dots , $G^r\mathbf{x}$, es decir $G^{r+1}\mathbf{x} \in \text{lin}\{\mathbf{x}, G\mathbf{x}, G^2\mathbf{x}, \dots, G^r\mathbf{x}\}$. Ahora bien, el subespacio

$$S = \text{lin}\{\mathbf{x}, G\mathbf{x}, G^2\mathbf{x}, \dots, G^r\mathbf{x}\}$$

es invariante por el operador G (o G -invariante):

$$GS \subseteq S$$

El subespacio S está integrado por autovectores de H relativos al autovalor λ . Dado que S tiene dimensión finita, entonces el polinomio característico de la restricción del operador $G : S \rightarrow S$ posee al menos una raíz en \mathbb{C} . Por tanto, existe en S un autovector de G que es también autovector de H . \square

Lema 11. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$. Sea f un operador en E . Si W es un subespacio invariante por f (es decir, $f(W) \subseteq W$), entonces el subespacio W^\perp es invariante por f^\dagger .*

Demostración. Sea $\mathbf{x} \in W^\perp$. Entonces $\forall \mathbf{v} \in E$:

$$\langle \mathbf{v}, f^\dagger(\mathbf{x}) \rangle = \langle f(\mathbf{v}), \mathbf{x} \rangle .$$

Siendo W un subespacio f -invariante, si en particular $\mathbf{v} \in W$, $f(\mathbf{v}) \in W$, luego

$$\langle \mathbf{v}, f^\dagger(\mathbf{x}) \rangle = \langle f(\mathbf{v}), \mathbf{x} \rangle = 0 \implies f^\dagger(\mathbf{x}) \in W^\perp .$$

\square

Proposición 48. [3] *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico y f un operador **normal** en E . Entonces*

$$\forall \mathbf{x} \in E, \quad \|f(\mathbf{x})\| = \|f^\dagger(\mathbf{x})\|$$

Demostración. Tenemos

$$\|f(\mathbf{x})\|^2 = \langle f(\mathbf{x}), f(\mathbf{x}) \rangle = \langle f^\dagger(f(\mathbf{x})), \mathbf{x} \rangle = \langle f(f^\dagger(\mathbf{x})), \mathbf{x} \rangle = \langle f^\dagger(\mathbf{x}), f^\dagger(\mathbf{x}) \rangle = \|f^\dagger(\mathbf{x})\|^2$$

\square

Proposición 49. [3] $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico y f un operador **normal** en E , $\mathbf{x} \in E$, $\lambda \in \mathbb{C}$. Entonces

$$f(\mathbf{x}) = \lambda \mathbf{x} \quad \Longleftrightarrow \quad f^\dagger(\mathbf{x}) = \bar{\lambda} \mathbf{x}$$

Demostración. Desde el Teorema (95) deducimos que el endomorfismo $f - \lambda \mathcal{I}$ tiene por adjunto $f^\dagger - \bar{\lambda} \mathcal{I}$. Además,

$$\begin{aligned} (f - \lambda \mathcal{I}) \circ (f^\dagger - \bar{\lambda} \mathcal{I}) &= f \circ f^\dagger - \bar{\lambda} f - \lambda f^\dagger + \lambda \bar{\lambda} \mathcal{I} = \\ &= f^\dagger \circ f - \lambda f^\dagger - \bar{\lambda} f + \lambda \bar{\lambda} \mathcal{I} = \\ &= (f^\dagger - \bar{\lambda} \mathcal{I}) \circ (f - \lambda \mathcal{I}) \end{aligned}$$

Por tanto, $f - \lambda \mathcal{I}$ es normal. Como consecuencia de la Proposición 48, resulta que

$$\|(f - \lambda \mathcal{I})(\mathbf{x})\| = \|(f^\dagger - \bar{\lambda} \mathcal{I})(\mathbf{x})\|$$

o sea,

$$\|f(\mathbf{x}) - \lambda \mathbf{x}\| = \|f^\dagger(\mathbf{x}) - \bar{\lambda} \mathbf{x}\| .$$

Por tanto,

$$f(\mathbf{x}) - \lambda \mathbf{x} = \mathbf{0} \quad \Longleftrightarrow \quad f^\dagger(\mathbf{x}) - \bar{\lambda} \mathbf{x} = \mathbf{0} .$$

□

Corolario 16. Si f es un operador **normal** en un espacio hermítico $(E, \langle \cdot, \cdot \rangle)$, entonces autovectores correspondientes a autovalores distintos son ortogonales.

Demostración. Sean $\mathbf{v}_1, \mathbf{v}_2$ autovectores correspondientes a λ_1, λ_2 , con $\lambda_1 \neq \lambda_2$:

$$\lambda_1 \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = \langle \mathbf{v}_2, \lambda_1 \mathbf{v}_1 \rangle = \langle \mathbf{v}_2, f(\mathbf{v}_1) \rangle = \langle f^\dagger(\mathbf{v}_2), \mathbf{v}_1 \rangle = \langle \bar{\lambda}_2 \mathbf{v}_2, \mathbf{v}_1 \rangle = \lambda_2 \langle \mathbf{v}_2, \mathbf{v}_1 \rangle$$

Por tanto

$$\lambda_1 \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = \lambda_2 \langle \mathbf{v}_2, \mathbf{v}_1 \rangle \quad \text{y} \quad \lambda_1 \neq \lambda_2 \implies \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = 0 .$$

□

6.5 Teoría espectral para operadores normales

Teorema 96. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, y f un operador normal de E . Entonces, existe una base ortonormal de E formada por autovectores de f .

Demostración. Como f es normal, $ff^\dagger = f^\dagger f$ (es decir, $[f, f^\dagger] = 0$), entonces por el Lema 10 existe un autovector no nulo común a f y f^\dagger :

$$\exists \mathbf{x} \in V \quad \text{t.q.} \quad f(\mathbf{x}) = \lambda_1 \mathbf{x}, \quad f^\dagger(\mathbf{x}) = \mu \mathbf{x}$$

Desde la Proposición 49 sabemos que $\bar{\lambda}_1 = \mu$. Sea entonces

$$\mathbf{u}_1 = \frac{\mathbf{x}}{\|\mathbf{x}\|}.$$

Claramente, \mathbf{u}_1 es un autovector de norma unitaria de f y f^\dagger . Sea S_1 el subespacio generado por \mathbf{u}_1 : $S_1 = \text{lin}\{\mathbf{u}_1\}$. Por construcción, es invariante por f y f^\dagger . Utilizando el Lema 11, deducimos que S_1^\perp es invariante, al mismo tiempo, por f^\dagger y $(f^\dagger)^\dagger = f$.

Entonces, si restringimos f al subespacio S_1^\perp , como f y f^\dagger conmutan, en virtud del Lema 10 obtenemos que S_1^\perp contiene un autovector \mathbf{v}_2 común a f y f^\dagger ,

$$f(\mathbf{v}_2) = \lambda_2 \mathbf{v}_2, \quad f^\dagger(\mathbf{v}_2) = \bar{\lambda}_2 \mathbf{v}_2.$$

Sea ahora

$$\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}.$$

Además, tenemos

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0.$$

como consecuencia del Corolario 16.

Podemos introducir el subespacio $S_2 = \text{lin}\{\mathbf{u}_1, \mathbf{u}_2\}$. Iterando el procedimiento, obtenemos al final el subespacio $S_n \equiv E$. En definitiva, hemos construido una base de f formada por autovectores de f ortonormales. \square

Observación 6.3. Nótese que la base ortonormal de autovectores de f también es por construcción una base ortonormal de autovectores de f^\dagger

Vale también el vice versa del teorema anterior.

Teorema 97. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, f un operador de E y $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ una base ortonormal de E formada por autovectores de f . Entonces f es un operador normal.

Demostración. Tenemos

$$f(\mathbf{u}_i) = \lambda_i \mathbf{u}_i.$$

Descomponiendo el vector $f^\dagger(\mathbf{u}_i)$ en la base ortonormal \mathcal{B} , tenemos

$$\begin{aligned} f^\dagger(\mathbf{u}_i) &= \sum_{j=1}^n \langle \mathbf{u}_j, f^\dagger(\mathbf{u}_i) \rangle \mathbf{u}_j = \sum_{j=1}^n \langle f(\mathbf{u}_j), \mathbf{u}_i \rangle \mathbf{u}_j = \sum_{j=1}^n \langle \lambda_j \mathbf{u}_j, \mathbf{u}_i \rangle \mathbf{u}_j = \\ &= \sum_{j=1}^n \bar{\lambda}_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle \mathbf{u}_j = \sum_{j=1}^n \bar{\lambda}_j \delta_{ji} \mathbf{u}_j = \bar{\lambda}_i \mathbf{u}_i \end{aligned}$$

por tanto, \mathbf{u}_i es autovector de f^\dagger con autovalor $\bar{\lambda}_i$. Entonces

$$f(f^\dagger(\mathbf{u}_i)) = \bar{\lambda}_i f(\mathbf{u}_i) = \bar{\lambda}_i \lambda_i \mathbf{u}_i = \lambda_i f^\dagger(\mathbf{u}_i) = f^\dagger(f(\mathbf{u}_i)), \quad i = 1, \dots, n$$

luego

$$f f^\dagger = f^\dagger f,$$

dado que estos operadores coinciden sobre una base de E . □

Proposición 50. *Sea $M \in M_n(\mathbb{C})$ una matriz normal. Entonces, existe una matriz unitaria U tal que*

$$U^\dagger M U = D$$

con D matriz diagonal.

6.6 Teoría espectral de los operadores autoadjuntos

En el caso de operadores autoadjuntos, la teoría espectral se puede desarrollar de una forma análoga.

Lema 12. *Los autovalores de un operador autoadjunto son reales.*

Demostración. Sea λ un autovalor de f , \mathbf{x} un autovector relativo a λ . Entonces siendo $f = f^\dagger$,

$$\bar{\lambda} \langle \mathbf{x}, \mathbf{x} \rangle = \langle f(\mathbf{x}), \mathbf{x} \rangle = \langle \mathbf{x}, f(\mathbf{x}) \rangle = \lambda \langle \mathbf{x}, \mathbf{x} \rangle$$

es decir

$$\lambda = \bar{\lambda}$$

□

Teorema 98 (Teorema espectral para operadores autoadjuntos). *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, y f un operador de E . El operador f es autoadjunto si y solo si sus autovalores son reales y existe una base ortonormal formada por autovectores de f .*

Proposición 51. *Si M es una matriz autoadjunta, M es diagonalizable mediante una matriz unitaria U :*

$$U^\dagger M U = D$$

con D matriz diagonal real.

6.7 Teoría espectral para operadores unitarios

Lema 13. *Los autovalores de un operador unitario tienen módulo igual a 1.*

Demostración. Sea f unitario, y sea \mathbf{x} un autovector de f con autovalor λ . Entonces:

$$f(\mathbf{x}) = \lambda \mathbf{x} \implies f^\dagger(f(\mathbf{x})) = \lambda f^\dagger(\mathbf{x}) = \lambda \bar{\lambda} \mathbf{x} = |\lambda|^2 \mathbf{x}$$

Como $f^\dagger f = \mathcal{I}$, tenemos

$$\mathbf{x} = |\lambda|^2 \mathbf{x} \implies |\lambda| = 1.$$

□

Teorema 99 (Teorema espectral para operadores unitarios). Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, y f un operador de E . El operador f es unitario si y solo si sus autovalores son de módulo unitario y existe una base ortonormal formada por autovectores de f .

Proposición 52. Si M es una matriz unitaria, M es diagonalizable mediante una matriz unitaria U :

$$U^\dagger M U = D$$

con D matriz diagonal con elementos de módulo 1 en la diagonal principal.

6.8 Otras propiedades de los operadores unitarios

El siguiente teorema ilustra las principales propiedades de los operadores unitarios, que también justifican su considerable relevancia en contextos físicos.

Teorema 100. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, y u un operador de E . Las propiedades siguientes son equivalentes;

- (1) u es unitario
- (2) u es normal y todos sus autovalores tienen módulo unitario
- (3) $\forall \mathbf{x}, \mathbf{y} \in E: \langle u(\mathbf{x}), u(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$
- (4) $\forall \mathbf{x} \in E, \|u(\mathbf{x})\| = \|\mathbf{x}\|$
- (5) u transforma bases ortonormales en bases ortonormales .

En una base ortonormal, la matriz U asociada a un operador unitario $u \in L(E)$ es unitaria ($U U^\dagger = U^\dagger U = \mathbf{1}$). Como consecuencia obvia de la definición de matriz unitaria, tenemos que las matrices unitarias son invertibles, y vale la importante relación

$$U^{-1} = U^\dagger . \quad (6.12)$$

Las matrices unitarias cumplen

$$|\det U|^2 = 1$$

El conjunto de todas las matrices unitarias $U \in \mathcal{M}_n(\mathbb{C})$ forma un grupo respecto del producto de matrices, llamado **grupo unitario** y es denotado con $U(n)$:

$$U(n) = \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = U U^\dagger = \mathbf{1}\} \quad (6.13)$$

El **grupo especial unitario** es formado por las matrices de $U(n)$ que tienen determinante igual a 1:

$$SU(n) = \{U \in U(n) \mid \det U = 1\} \quad (6.14)$$

Ambos grupos son fundamentales en teoría de las partículas elementales.

6.9 Diagonalización simultanea de operadores normales que conmutan

Proposición 53. *Sea E un espacio vectorial hermítico, $\dim E = n$ y f, g dos operadores normales que conmutan: $[f, g] = 0$. Supongamos que f tenga n autovalores distintos. Entonces g es diagonalizable y todos los autovectores de f son también autovectores de g .*

Demostración. Sea $\mathbf{v} \neq \mathbf{0}$ autovector de f relativo al autovalor λ :

$$f(\mathbf{v}) = \lambda \mathbf{v} \implies g(f(\mathbf{v})) = g(\lambda \mathbf{v}) = \lambda g(\mathbf{v})$$

Como f y g conmutan, tenemos $f(g(\mathbf{v})) = g(f(\mathbf{v})) = \lambda g(\mathbf{v})$, así que $g(\mathbf{v})$ es autovector de f con autovalor λ . Ahora, tenemos dos posibilidades.

- (1) $g(\mathbf{v}) = \mathbf{0}$, y por tanto \mathbf{v} es también autovector de g con autovalor $\lambda = 0$.
- (2) $g(\mathbf{v}) \neq \mathbf{0}$. En este caso, siendo $g(\mathbf{v})$ autovector de f , $g(\mathbf{v}) \in V_\lambda$. Como $\dim V_\lambda = 1$, entonces $g(\mathbf{v})$ tiene que ser proporcional a \mathbf{v} : $g(\mathbf{v}) = \mu \mathbf{v}$, y por tanto $g(\mathbf{v})$ es autovector de g con autovalor μ . \square

Observación 6.4. La hipótesis que los autovalores de f son todos distintos es crucial. Si no se cumple, en general el resultado no es cierto.

Proposición 54. *Si $f, g \in L(E)$ son operadores normales, entonces existe una base ortonormal de E formada por vectores propios comunes de f y g .*

6.10 Proyectores ortogonales en espacios hermíticos

6.10.1 Descomposición espectral

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, y f un operador normal de E , con espectro

$$\text{spec}(f) = \{\lambda_1, \dots, \lambda_p\}, \quad \lambda_i \neq \lambda_j, \quad i \neq j, \quad i, j = 1, \dots, p \leq n.$$

Sea $V_i := \ker(f - \lambda_i \mathcal{I})$, $i = 1, \dots, p$, el autoespacio correspondiente al autovalor λ_i . Como $\lambda_i \neq \lambda_j$, entonces $V_i \perp V_j$, $i \neq j$. Además, sabemos que

$$E = V_1 \oplus \dots \oplus V_p$$

Sea

$$\pi_i : E \longrightarrow V_i, \quad \mathbf{x} \longrightarrow \mathbf{x}_i.$$

Tenemos la descomposición

$$\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_p,$$

donde $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0$ si $i \neq j$. El operador π_i se dirá el proyector ortogonal sobre el autoespacio V_i .

Teorema 101 (descomposición espectral). Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, y f un operador normal de E . Los proyectores ortogonales asociados a f satisfacen las propiedades siguientes:

- (1) $\pi_i^\dagger = \pi_i$
- (2) $\pi_i \pi_j = \delta_{ij} \pi_i$
- (3) $\pi_1 + \dots + \pi_p = \mathcal{I}$
- (4) $\lambda_1 \pi_1 + \dots + \lambda_p \pi_p = f$

Demostración.

(1)

$$\langle \mathbf{x}, \pi_i \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y}_i \rangle = \langle \mathbf{x}_i, \mathbf{y}_i \rangle = \langle \pi_i \mathbf{x}, \mathbf{y}_i \rangle = \langle \pi_i \mathbf{x}, \mathbf{y} \rangle$$

es decir $\pi_i^\dagger = \pi_i$.

(2)

$$\begin{aligned} \pi_i^2 \mathbf{x} &= \pi_i(\mathbf{x}_i) = \mathbf{x}_i, & \text{es decir} & \quad \pi_i^2 = \pi_i \quad \text{idempotencia} \\ \pi_i \pi_j(\mathbf{x}) &= \pi_i(\mathbf{x}_j) = 0 \quad \text{si } i \neq j. \end{aligned}$$

Entonces

$$\pi_i \pi_j(\mathbf{x}) = \delta_{ij} \pi_i(\mathbf{x}).$$

(3)

$$(\pi_1 + \dots + \pi_p)\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_p = \mathbf{x}$$

Por tanto

$$\pi_1 + \dots + \pi_p = \mathcal{I}.$$

(4)

$$\begin{aligned} (\lambda_1 \pi_1 + \dots + \lambda_p \pi_p)(\mathbf{x}) &= \lambda_1 \mathbf{x}_1 + \dots + \lambda_p \mathbf{x}_p = f(\mathbf{x}_1) + \dots + f(\mathbf{x}_p) \\ &= f(\mathbf{x}_1 + \dots + \mathbf{x}_p) = f(\mathbf{x}) \end{aligned}$$

es decir

$$f = \lambda_1 \pi_1 + \dots + \lambda_p \pi_p. \quad (6.15)$$

□

A la fórmula (6.15) se da el nombre de **descomposición espectral** del operador f .

El siguiente teorema relaciona las propiedades de los proyectores con las de los operadores normales.

Teorema 102. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermítico, $\dim E = n$, y f un operador de E . Sea $\{\pi_1, \dots, \pi_p\}$ una familia de proyectores ortogonales idempotentes y autoadjuntos, y $\{\lambda_1, \dots, \lambda_p\}$ un conjunto de escalares tales que

- (1) $\pi_i \pi_j = 0, \quad i \neq j$
- (2) $\pi_1 + \dots + \pi_p = \mathcal{I}$
- (3) $\lambda_1 \pi_1 + \dots + \lambda_p \pi_p = f$

Entonces, f es un operador normal.

Corolario 17. Si los escalares $\{\lambda_1, \dots, \lambda_p\}$ son reales, f es autoadjunto.

Corolario 18. Si los escalares $\{\lambda_1, \dots, \lambda_p\}$ tienen módulo igual a 1, f es unitario.

6.10.2 Cálculo de proyectores ortogonales

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial hermitico, $\dim E = n$, f un operador normal de E y $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ una base ortonormal de E formada por autovectores de f . Sean V_1, \dots, V_p los autoespacios de f , $\dim V_i = r_i$, $i = 1, \dots, p$. Supongamos que los vectores de la base sean ordenados de forma tal que

$$\begin{aligned} V_1 &= \text{lin}\{\mathbf{u}_1, \dots, \mathbf{u}_{r_1}\} \\ V_2 &= \text{lin}\{\mathbf{u}_{r_1+1}, \dots, \mathbf{u}_{r_1+r_2}\} \\ &\vdots \\ V_k &= \text{lin}\{\mathbf{u}_{r_1+\dots+r_{k-1}+1}, \dots, \mathbf{u}_{r_1+\dots+r_k}\} \\ &\vdots \\ V_p &= \text{lin}\{\mathbf{u}_{r_1+\dots+r_{p-1}+1}, \dots, \mathbf{u}_n\} . \end{aligned}$$

Si $\mathbf{x} \in E$, entonces la acción del proyector k -ésimo es dada por

$$\pi_k(\mathbf{x}) = \sum_{j=r_1+\dots+r_{k-1}+1}^{r_1+\dots+r_k} \langle \mathbf{u}_j, \mathbf{x} \rangle \mathbf{u}_j \quad (6.16)$$

Si U_j es el vector columna de las coordenadas de \mathbf{u}_j en una base asignada, entonces la representación matricial del proyector sobre el autoespacio V_k es dada por

$$\pi_k = \sum_{j=r_1+\dots+r_{k-1}+1}^{r_1+\dots+r_k} U_j U_j^\dagger \quad (6.17)$$

En particular,

$$\mathbb{1} = \sum_{i=1}^n U_i U_i^\dagger \quad (6.18)$$

mientras que la descomposición espectral de f se escribe como

$$f = \sum_{i=1}^n \lambda_i U_i U_i^\dagger . \quad (6.19)$$

6.11 Ejemplos

Ejercicio 6.3. Sea

$$F = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

la matriz asociada a un operador lineal $f \in L(\mathbb{C}^2)$, dotado del producto escalar estándar, respecto de la base canónica de \mathbb{C}^2 . Comprobar que F es una matriz normal y determinar su descomposición espectral.

Resolución. Observamos que, respecto de la base canónica (que es una base ortonormal)

$$F^\dagger = \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$$

La matriz F es normal:

$$FF^\dagger = F^\dagger F = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Determinemos sus autovalores.

$$\begin{vmatrix} 1 - \lambda & i \\ i & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 + 1 = 0 \implies \lambda_{1/2} = 1 \pm i$$

Determinemos el autoespacio asociado a $\lambda_1 = 1 + i$. Tenemos

$$\begin{pmatrix} -i & i \\ i & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{cases} -ix_1 + ix_2 = 0 \\ ix_1 - ix_2 = 0 \end{cases} \implies x_1 = x_2$$

Entonces V_{λ_1} está generado por $\mathbf{e}_1 = (1, 1)$. Normalizándolo, y escribiéndolo en forma de vector columna, tenemos

$$U_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

El proyector asociado es

$$\pi_1 = U_1 U_1^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Determinemos el autoespacio asociado a $\lambda_2 = 1 - i$. Tenemos

$$\begin{pmatrix} i & i \\ i & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies ix_1 + ix_2 = 0 \implies x_1 = -x_2$$

$$U_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Tenemos el proyector

$$\pi_2 = U_2 U_2^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

En definitiva,

$$F = \lambda_1 \pi_1 + \lambda_2 \pi_2 = \frac{1+i}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1-i}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

Ejercicio 6.4. Sea \mathbb{R}^3 dotado del producto escalar canónico y sea el operador cuya matriz asociada respecto de la base canónica de \mathbb{R}^3 es

$$H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Demostrar que es autoadjunto y determinar su descomposición espectral.

Resolución. Observamos que $H^t = H$, es decir H es simétrico. La ecuación característica asociada es

$$\lambda^3 - 3\lambda - 2 = 0 \implies (\lambda + 1)^2(\lambda - 2) = 0, \quad \lambda_1 = -1, \quad \lambda_2 = 2$$

Halleemos el autoespacio V_{λ_1} . Tenemos

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies x_1 + x_2 + x_3 = 0$$

Por tanto $V_1 = \{(-\alpha - \beta, \alpha, \beta), \alpha, \beta \in \mathbb{R}\}$, así que $V_1 = \text{lin}\{(1, 0, -1), (1, -1, 0)\}$. Determinemos una base ortonormal de V_1 mediante un procedimiento directo (es decir, sin utilizar el de Gram-Schmidt). El primer vector de la base es

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, 0, -1).$$

Observamos ahora que los vectores de V_1 son de la forma $(-\alpha - \beta, \alpha, \beta)$. Impongamos que $\langle \mathbf{u}_1, (-\alpha - \beta, \alpha, \beta) \rangle = 0 \implies -\alpha - \beta - \beta = 0 \implies \beta = -\alpha/2$. Está claro que el vector $\mathbf{v}_2 = (1, -2, 1)$ es ortogonal a \mathbf{u}_1 . Entonces, si $\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{1}{\sqrt{6}}(1, -2, 1)$, tenemos que $\{\mathbf{u}_1, \mathbf{u}_2\}$ forman una base ortonormal de V_1 .

El proyector sobre el primer autoespacio $\pi_1 := \pi_{V_1}$ es dado por

$$\begin{aligned} \pi_1 &= U_1 U_1^\dagger + U_2 U_2^\dagger = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} (1 \ 0 \ -1) + \frac{1}{6} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} (1 \ -2 \ 1) = \\ &= \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \end{aligned}$$

Halleemos el autoespacio V_{λ_2} . Tenemos

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{cases} -2x_1 + x_2 + x_3 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ x_1 + x_2 - 2x_3 = 0 \end{cases}$$

Entonces $V_{\lambda_2} = \{\alpha, \alpha, \alpha\}$, y una base de V_{λ_2} es dada por el vector normalizado $\mathbf{u}_3 = \frac{1}{\sqrt{3}}(1, 1, 1)$. El proyector sobre V_{λ_2} es dado por

$$\pi_2 = U_3 U_3^\dagger = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} (1 \ 1 \ 1) = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} .$$

En definitiva

$$H = \lambda_1 \pi_1 + \lambda_2 \pi_2 = -\frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} + \frac{2}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} .$$

Ejercicio 6.5. Sea $P_2(\mathbb{R})$ el espacio vectorial de los polinomios con coeficientes reales de grado menor o igual a 2. Sea

$$\varphi : P_2 \times P_2 \longrightarrow \mathbb{R}$$

la aplicación

$$\varphi(p(x), q(x)) = p(0)q(0) + p'(0)q'(0) + p''(0)q''(0)$$

- (a) Demostrar que φ es producto escalar en P_2 .
- (b) Si $A = \text{lin}\{(x, x^2)\}$, determinar A^\perp .
- (c) Calcular el ángulo formado por los polinomios

$$p(x) = x^2 \quad \text{y} \quad q(x) = x .$$

Resolución.

(a) Observamos que φ es una forma bilineal simétrica definida positiva. Por ejemplo,

$$\begin{aligned} \varphi(p_1(x) + p_2(x), q(x)) &= (p_1(0) + p_2(0))q(0) + (p_1'(0) + p_2'(0))q'(0) + (p_1''(0) + p_2''(0))q''(0) = \\ &= p_1(0)q(0) + p_1'(0)q'(0) + p_1''(0)q''(0) \\ &+ p_2(0)q(0) + p_2'(0)q'(0) + p_2''(0)q''(0) = \\ &= \varphi(p_1(x), q(x)) + \varphi(p_2(x), q(x)) \end{aligned}$$

etc. La simetría es obvia:

$$\varphi(p(x), q(x)) = \varphi(q(x), p(x)),$$

etc. También observamos que

$$\varphi(p(x), p(x)) = [p(0)]^2 + [p'(0)]^2 + [p''(0)]^2 \geq 0$$

y

$$\varphi(p(x), p(x)) = 0 \iff p(0) = p'(0) = p''(0) = 0$$

lo que implica $p(x) = 0$. Por tanto φ es un producto escalar en P_2 .

(b) Sea $A = \text{lin}\{(x, x^2)\} = \{ax + bx^2 \mid a, b \in \mathbb{R}\}$. Tenemos que

$$A^\perp = \{p(x) \in P_2 \mid \varphi(p(x), x) = 0, \varphi(p(x), x^2) = 0\}$$

Entonces:

$$\begin{cases} \varphi(p(x), x) = 0 & \iff p'(0) = 0 \\ \varphi(p(x), x^2) = 0 & \iff p''(0) = 0 \end{cases} \implies p(x) = \text{const}$$

Concluimos que

$$A^\perp = \text{lin}\{1\}.$$

(c) El ángulo entre dos polinomios de P_2 es dado por

$$\cos \alpha = \frac{\langle p(x), q(x) \rangle}{\|p(x)\| \|q(x)\|}$$

Entonces

$$p(x) = x^2, \quad q(x) = x, \quad \langle x^2, x \rangle = \varphi(x^2, x) = 0$$

Por tanto $\alpha = \frac{\pi}{2}$.

6.12 Operadores positivos

Definición 133. *Un operador autoadjunto se dirá positivo (no negativo) si todos sus autovalores son positivos (no negativos)*

Proposición 55. *Los operadores positivos satisfacen las propiedades siguientes:*

- h positivo $\iff \forall \mathbf{x} \neq \mathbf{0}, \langle \mathbf{x}, h(\mathbf{x}) \rangle > 0$.
- h positivo $\iff h^{-1}$ positivo
- h positivo $\iff \exists s \in L(E)$ positivo tal que $s^2 = h$

6.13 Operadores simétricos en espacios vectoriales reales

Sea E un espacio vectorial euclídeo. Un operador f tal que

$$f = f^\dagger$$

se dirá simétrico (en lugar de autoadjunto).

Proposición 56. *Sea E un espacio vectorial euclídeo, y f un operador normal en E . Si $\mathbf{x} \in E$ es un autovector de f con autovalor λ , entonces \mathbf{x} también es autovector de f^\dagger con el mismo autovalor.*

Teorema 103. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo de dimensión finita. Un operador f es simétrico en E si y sólo si existe una base ortonormal de E formada por autovectores de f .

La demostración del teorema anterior utiliza la complejización de V . Un resultado interesante es el

Teorema 104. Toda matriz simétrica real es diagonalizable mediante una transformación ortogonal:

$$P^t AP = D .$$

6.13.1 Descomposición espectral de un operador simétrico

Los operadores simétricos admiten una descomposición espectral en proyectores

$$\pi_i : E \longrightarrow V_i$$

con

$$\pi_i \pi_j = \delta_{ij} \pi_i \quad (6.20)$$

$$\pi_1 + \dots + \pi_p = \mathbb{1} \quad (6.21)$$

$$\lambda_1 \pi_1 + \dots + \lambda_p \pi_p = f, \quad \lambda_1, \dots, \lambda_p \in \mathbb{R} . \quad (6.22)$$

6.14 Complejización de un espacio vectorial real

6.14.1 Espacio complejificado

Sea V un espacio vectorial real. La complejización $V_{\mathbb{C}}$ de V es el espacio

$$V_{\mathbb{C}} = V \times V$$

dotado de las leyes de composición suma

$$+ : V_{\mathbb{C}} \times V_{\mathbb{C}} \longrightarrow V_{\mathbb{C}},$$

$$(\mathbf{x}_1, \mathbf{y}_1) + (\mathbf{x}_2, \mathbf{y}_2) = (\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}_1 + \mathbf{y}_2)$$

y producto por escalares

$$\cdot : V_{\mathbb{C}} \times V_{\mathbb{C}} \longrightarrow V_{\mathbb{C}}$$

$$(a + bi)(\mathbf{x}, \mathbf{y}) = (a\mathbf{x} - b\mathbf{y}, b\mathbf{x} + a\mathbf{y}), \quad a, b, \in \mathbb{R}$$

Evidentemente $i(\mathbf{x}, \mathbf{0}) = (\mathbf{0}, \mathbf{x})$. Introducimos la notación

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x} + i\mathbf{y} .$$

Si $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ es una base de V , entonces $\mathcal{B}_c = \{(\mathbf{v}_1, \mathbf{0}), \dots, (\mathbf{v}_n, \mathbf{0})\}$ es una base de $V_{\mathbb{C}}$ que, en la notación anterior, se suele seguir escribiendo $\mathcal{B}_c = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

6.14.2 Complejización de un operador lineal

Sea $f \in L(V)$. Se define el operador $f_{\mathbb{C}} \in L(V_{\mathbb{C}})$ mediante la relación

$$f_{\mathbb{C}}(\mathbf{x} + i\mathbf{y}) = f(\mathbf{x}) + if(\mathbf{y})$$

Si F es la matriz de f en la base \mathcal{B} , también será la matriz de $f_{\mathbb{C}}$ en la base $\mathcal{B}_{\mathbb{C}}$. Como consecuencia, el polinomio característico de $f_{\mathbb{C}}$ es igual a él de f , y por tanto tiene coeficientes reales. Si λ es autovalor real de $f_{\mathbb{C}}$, también es autovalor de f :

$$f_{\mathbb{C}}(\mathbf{x} + i\mathbf{y}) = \lambda(\mathbf{x} + i\mathbf{y}) \iff f(\mathbf{x}) + if(\mathbf{y}) = \lambda\mathbf{x} + i\lambda\mathbf{y} \iff \begin{cases} f(\mathbf{x}) = \lambda\mathbf{x} \\ f(\mathbf{y}) = \lambda\mathbf{y} \end{cases}$$

6.14.3 Complejización de un producto escalar real

Sea $\langle \cdot, \cdot \rangle$ un producto escalar real. Entonces, podemos introducir su complejificación

$$\langle \cdot, \cdot \rangle : V_{\mathbb{C}} \times V_{\mathbb{C}} \longrightarrow \mathbb{C}$$

definida así:

$$\langle \mathbf{x}_1 + i\mathbf{y}_1, \mathbf{x}_2 + i\mathbf{y}_2 \rangle_{\mathbb{C}} = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle + \langle \mathbf{y}_1, \mathbf{y}_2 \rangle + i(\langle \mathbf{x}_1, \mathbf{y}_2 \rangle - \langle \mathbf{y}_1, \mathbf{x}_2 \rangle)$$

6.15 Operadores ortogonales en espacios euclídeos

Sea E un espacio euclídeo. Un operador $f \in L(E)$ tal que $ff^{\dagger} = f^{\dagger}f = \mathcal{I}$ se dirá operador ortogonal. En otras palabras, la unitariedad en espacios hermíticos se convierte en la ortogonalidad en espacios euclídeos. Un operador ortogonal posee un cierto número de propiedades interesantes.

(1) f es ortogonal \iff conserva el producto escalar \iff conserva la norma asociada al p.e. \iff transforma bases ortonormales en bases ortonormales (igual que en el caso hermítico).

(2) En general un operador ortogonal no es diagonalizable, porque ahora no se puede garantizar que su polinomio característico tenga tantas raíces reales como orden tiene este polinomio.

(3) En una base ortonormal $F^t = F^{-1}$.

Las matrices ortogonales cumplen la simple propiedad $(\det O)^2 = 1$.

Definición 134. Se llama grupo ortogonal al conjunto de matrices:

$$O(n) = \{O \in M_n(\mathbb{R}) \mid O^t O = \mathbf{1}\}$$

Se llama grupo especial ortogonal al conjunto de matrices

$$SO(n) = \{O \in O(n) \mid \det O = 1\}$$

Observación 6.5. Las matrices de $SO(n)$ representan rotaciones. Las de determinante -1 representan reflexiones, o reflexiones compuestas con rotaciones, y no forman subgrupo.

6.16 Diagonalización por bloques de un operador ortogonal

6.16.1 Consideraciones generales

Si f es un operador ortogonal en un espacio euclídeo, y tiene autovalores reales, entonces $\lambda_i = \pm 1$. En general, si la dimensión del espacio vectorial E es impar, entonces existe al menos un autovalor real.

Al complejificar, el operador $f_{\mathbb{C}}$ unitario tendrá además autovalores complejos con $|\lambda| = 1 \implies$

$$\lambda = \cos \theta - i \sin \theta, \quad \bar{\lambda} = \cos \theta + i \sin \theta.$$

Si $\mathbf{x} + i\mathbf{y}$ es autovector con autovalor λ , entonces $\mathbf{x} - i\mathbf{y}$ es autovector con autovalor $\bar{\lambda}$:

$$f_{\mathbb{C}}(\mathbf{x} + i\mathbf{y}) = \lambda(\mathbf{x} + i\mathbf{y}) \implies f_{\mathbb{C}}(\mathbf{x} - i\mathbf{y}) = \bar{\lambda}(\mathbf{x} - i\mathbf{y})$$

Si $\theta \neq 0, \pi$, $\lambda \neq \bar{\lambda}$ y autovectores correspondientes a autovalores distintos son ortogonales:

$$\begin{aligned} \langle \mathbf{x} + i\mathbf{y}, \mathbf{x} - i\mathbf{y} \rangle_c &= 0 = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{y}, \mathbf{y} \rangle - 2i\langle \mathbf{x}, \mathbf{y} \rangle \implies \\ \langle \mathbf{x}, \mathbf{y} \rangle &= 0 \quad \text{y} \quad \langle \mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{y} \rangle \end{aligned}$$

Entonces los vectores \mathbf{x} , \mathbf{y} (normalizados) forman una base ortonormal de un subespacio invariante (no propio) de dimensión dos, ya que

$$f(\mathbf{x}) = \cos \vartheta \mathbf{x} + \sin \vartheta \mathbf{y}, \quad f(\mathbf{y}) = -\sin \vartheta \mathbf{x} + \cos \vartheta \mathbf{y}$$

Supongamos que el operador f ortogonal admite un conjunto ortonormal de autovectores $\{\mathbf{u}_1, \dots, \mathbf{u}_l\}$ de autovalor 1, un conjunto ortonormal de autovectores $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ de autovalor -1 , y $f_{\mathbb{C}}$ tiene además p autovalores no reales distintos, $\lambda_i = \cos \theta_i - i \sin \theta_i$ ($i = 1, \dots, p$), y sus complejos conjugados, con $(\mathbf{x}_i + i\mathbf{y}_i)$ autovector (de norma 1) con autovalor λ_i , tal que $l + k + 2p = n = \dim E$. Entonces en la base de E dada por $\{\mathbf{u}_1, \dots, \mathbf{u}_l, \mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{x}_1, \mathbf{y}_1, \dots, \mathbf{x}_p, \mathbf{y}_p\}$ la

matriz que representa a f es **diagonal por bloques** y dada por

$$F = \begin{pmatrix} 1 & 0 & \dots & & & & & & 0 \\ 0 & \ddots & 0 & \dots & & & & & \vdots \\ & & 1 & & & & & & \\ \vdots & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & \cos \vartheta_1 & -\sin \vartheta_1 & \\ & & & & & & \sin \vartheta_1 & \cos \vartheta_1 & \\ \vdots & & & & & & & & \ddots \\ 0 & \dots & & & & & \dots & \sin \vartheta_p & \cos \vartheta_p \end{pmatrix}$$

6.16.2 Caso $n = 2$

Sea $F \in O(2)$:

$$F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad FF^t = F^tF = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

La condición $F^tF = 1$ nos proporciona las condiciones

$$a^2 + c^2 = 1 = b^2 + d^2 \quad (6.23)$$

$$ab + cd = 0 \quad (6.24)$$

Claramente, siendo $a^2 + c^2 = 1$, existe un número real θ en el intervalo $[0, 2\pi)$ tal que $a = \cos \theta$, $c = \sin \theta$. Análogamente, $b^2 + d^2 = 1 \implies b = \sin \theta'$, $d = \cos \theta'$, con $\theta' \in [0, 2\pi)$.

Ahora bien, la relación $ab + cd = 0 \implies \cos \theta \sin \theta' + \sin \theta \cos \theta' = 0 \iff$

$$\sin(\theta + \theta') = 0 \iff \begin{cases} \theta + \theta' = 0 \\ \theta + \theta' = \pi \end{cases}.$$

Por tanto, tenemos las dos soluciones

$$\theta' = -\theta, \quad \theta' = \pi - \theta.$$

Entonces, distinguimos dos casos:

(1)

$$F = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad (6.25)$$

(2)

$$F = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}. \quad (6.26)$$

Analicemos el caso (1). Observamos que la forma (6.25) para $\theta \neq 0$ no admite vectores invariantes (salvo el vector nulo), es decir vectores \mathbf{v} tales que $f(\mathbf{v}) = \mathbf{v}$,

siendo f el operador ortogonal asociado a F en la base canónica. Esto se ve fácilmente porque

$$\begin{vmatrix} \cos \theta - 1 & -\sin \theta \\ \sin \theta & \cos \theta - 1 \end{vmatrix} = \cos^2 \theta - 2 \cos \theta + 1 + \sin^2 \theta = 2(1 - \cos \theta)$$

que en el intervalo considerado se anula sólo para $\theta = 0$. Para este valor, $F = \mathbf{1}$, que trivialmente deja cualquier vector invariante. La matriz F en la forma (6.25) representa entonces una rotación de un ángulo θ en sentido antihorario (positivo). Tenemos

$$\det F = 1, \quad \text{Tr } F = 2 \cos \theta. \quad (6.27)$$

Toda matriz ortogonal 2×2 de determinante igual a 1 tiene esta forma (en una base ortonormal).

Vamos ahora a analizar el caso (2). Ante todo

$$\det F = -1, \quad \text{Tr } F = 0. \quad (6.28)$$

La ecuación característica es

$$\begin{vmatrix} \cos \theta - \lambda & \sin \theta \\ \sin \theta & -\cos \theta - \lambda \end{vmatrix} = \lambda^2 - 1 = 0$$

Por tanto, $\lambda = \pm 1$. Los autovectores (normalizados) correspondientes son

$$\lambda = +1, \quad \mathbf{u}_1 = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}, \quad \lambda = -1, \quad \mathbf{u}_2 = \begin{pmatrix} \sin \frac{\theta}{2} \\ -\cos \frac{\theta}{2} \end{pmatrix} \quad (6.29)$$

Entonces \mathbf{u}_1 es fijo. En la base ortonormal $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$, la matriz de F es

$$F = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Este operador representa una reflexión: hay una recta que queda invariante, y que corresponde al vector \mathbf{u}_1 . El resto de vectores sufre una reflexión con respecto a esta recta (en particular el \mathbf{u}_2 que simplemente cambia de signo).

Toda matriz ortogonal (2×2) de determinante igual a -1 es el producto de una matriz que representa una rotación (con determinante positivo) por la matriz de la segunda forma canónica (6.26).

Ejemplos 6.6.

(I) Sea

$$F = \frac{1}{2} \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}, \quad FF^t = \frac{1}{4} \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Tenemos $\det F = 1$, $\text{Tr } F = \sqrt{3}$. El operador F representa entonces una rotación en el plano \mathbb{R}^2 , de un ángulo $\theta = \arccos \frac{\sqrt{3}}{2} = \frac{\pi}{6}$.

(II) Sea

$$F = \frac{1}{3} \begin{pmatrix} 1 & 2\sqrt{2} \\ 2\sqrt{2} & -1 \end{pmatrix}, \quad FF^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

en este caso, $\det F = -1$, $\text{Tr } F = 0$. Vamos a hallar los autovectores de F utilizando las fórmulas (6.29). Tenemos

$$\frac{1}{3} = \cos \theta = 2 \cos^2 \frac{\theta}{2} - 1 = 1 - 2 \sin^2 \frac{\theta}{2}$$

Por tanto,

$$\cos^2 \frac{\theta}{2} = \frac{2}{3}, \quad \sin^2 \frac{\theta}{2} = \frac{1}{3}, \quad \sin \frac{\theta}{2} = \pm \frac{1}{\sqrt{3}}, \quad \cos \frac{\theta}{2} = \pm \frac{\sqrt{6}}{3}$$

Entonces

$$\mathbf{u}_1 = \frac{1}{3} \begin{pmatrix} \sqrt{6} \\ \sqrt{3} \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{3} \begin{pmatrix} \sqrt{3} \\ -\sqrt{6} \end{pmatrix}$$

Tenemos $F\mathbf{u}_1 = \mathbf{u}_1$, $F\mathbf{u}_2 = -\mathbf{u}_2$. El operador F representa una reflexión con respecto al eje \mathbf{u}_1 .

6.16.3 Caso $n = 3$

Una matriz ortogonal de orden impar admite como autovalor $\lambda = 1$ o bien $\lambda = -1$. Sea $F \in O(3)$. Consideremos el caso en que $\lambda = 1$ sea autovalor, y sea V_1 el autoespacio asociado.

(ia) Sea $\dim V_1 = 1$. Sea \mathbf{u} el autovector (normalizado) asociado, $V_1 = \text{lin}\{\mathbf{u}\}$. El operador ortogonal, después de complejificar, admite autovectores $\mathbf{x} \pm i\mathbf{y}$ que corresponden a autovalores complejos conjugados. En la base ortonormal $\{\mathbf{u}, \mathbf{x}, \mathbf{y}\}$ adaptada a la suma directa $\mathbb{R}^3 = V_1 \oplus V_1^\perp$, tenemos

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad (6.30)$$

Entonces

$$\det F = 1, \quad \text{Tr } F = 1 + 2 \cos \theta$$

Tenemos una rotación, de un ángulo θ , con eje de giro dado por la recta generada por el vector \mathbf{u} .

(iia) Si $\dim V_1 = 2$, la matriz tomará la forma

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (6.31)$$

que representa una simetría respecto del plano correspondiente a V_1 , con $\det F = -1$.

(iiia) Si $\dim V_1 = 3$, tenemos la identidad: $F = \mathbf{1}$.

Consideremos ahora el caso en que $\lambda = -1$ sea autovalor, y sea V_{-1} el autoespacio asociado.

(ib) Sea $\dim V_{-1} = 1$. Sea \mathbf{u} un autovector asociado normalizado, $V_1 = \text{lin}\{\mathbf{u}\}$. Respecto de una base ortonormal cuyo primer vector es \mathbf{u} , tenemos

$$F = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad (6.32)$$

que puede verse como el producto

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (6.33)$$

Por tanto, se trata de la composición de una rotación con eje de giro la recta generada por \mathbf{u} , y ángulo θ , con una simetría respecto del plano perpendicular a dicha recta.

(iib) Si $\dim V_{-1} = 2$, se puede mostrar que regresaríamos a uno de los casos discutidos en el apartado relativo a $\lambda = +1$.

(iiib) Si $\dim V_{-1} = 3$, tenemos $F = -\mathbf{1}$.

6.17 Fórmula de Rodrigues

La matriz de una rotación en \mathbb{R}^3 de eje de giro \mathbf{u} , **normalizado**, y ángulo de rotación θ , se puede escribir mediante la fórmula de O. Rodrigues:

$$R_{ij} = \cos \theta \delta_{ij} + (1 - \cos \theta) u_i u_j - \sin \theta \epsilon_{ijk} u_k \quad (6.34)$$

donde $\mathbf{u} = (u_1, u_2, u_3)$, y

$$\epsilon_{ijk} = \begin{cases} +1 & \text{si } (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2) \\ -1 & \text{si } (i, j, k) = (1, 3, 2), (2, 1, 3), (3, 2, 1) \\ 0 & \text{si } i = j, j = k \text{ o } k = i \end{cases} \quad (6.35)$$

es el **símbolo de Levi-Civita**. En forma matricial, si $U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$, tenemos

$$R = \cos \theta \mathbf{1} + (1 - \cos \theta) U U^t + \sin \theta \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}. \quad (6.36)$$

Dado un vector $\mathbf{x} \in \mathbb{R}^3$, podemos determinar su imagen \mathbf{x}' bajo una rotación de un ángulo θ asignado:

$$\mathbf{x}' = R(\mathbf{x}) = \cos \theta \mathbf{x} + (1 - \cos \theta) \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{u} + \sin \theta \mathbf{u} \times \mathbf{x} . \quad (6.37)$$

Por otro lado, podemos definir el ángulo ϕ entre \mathbf{x} y su transformado \mathbf{x}' bajo la rotación R de ángulo θ y eje \mathbf{u} . Para determinar el ángulo ϕ es suficiente observar que

$$\langle \mathbf{x}, R(\mathbf{x}) \rangle = \|\mathbf{x}\|^2 \cos \phi$$

donde hemos tenido en cuenta que $\|R(\mathbf{x})\| = \|\mathbf{x}\|$ (nótese que este ángulo es nulo para los vectores del eje \mathbf{u}).

Chapter 7

Diagonalización de formas cuadráticas reales

7.1 Reducción a forma diagonal

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, y $\Phi : E \rightarrow \mathbb{R}$ la forma cuadrática asociada al producto escalar:

$$\Phi(\mathbf{x}) = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2.$$

La forma cuadrática es definida positiva, dado que

$$\Phi(\mathbf{x}) \geq 0 \quad \text{y} \quad \Phi(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}$$

En general, como ya hemos visto, una forma cuadrática puede ser **definida**, **semidefinida** o bien **indefinida**.

Ejemplos 7.1.

(1)

$$\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \Phi(\mathbf{x}) = x^2 + y^2$$

Claramente, Φ es *definida positiva*.

(2)

$$\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad \Phi(\mathbf{x}) = -x^2 - y^2 - z^2$$

Φ es *definida negativa*.

(3)

$$\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \Phi(\mathbf{x}) = x^2 + y^2 + 2xy$$

Φ es *semidefinida positiva*, ya que para cada vector no nulo $\mathbf{x} = (x, y) \in \mathbb{R}^2$,

$$\Phi(\mathbf{x}) = (x + y)^2 \geq 0.$$

En particular, existen vectores no nulos que anulan Φ : por ejemplo,

$$\mathbf{x} = (2, -2), \quad \Phi(2, -2) = 0.$$

$$(4) \quad \Phi : \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad \Phi(\mathbf{x}) = -x^2 - y^2 + z^2$$

Φ es *indefinida*, dado que

$$\Phi[(0, 0, 1)] = 1 > 0 \quad \Phi[(0, 1, 0)] = -1 < 0 \quad \text{y} \quad \Phi[(1, 0, 1)] = 0 .$$

En general el problema de clasificar las formas cuadráticas en un espacio euclídeo de dimensión arbitraria no es obvio. La clasificación se obtiene realizando un cambio de base adecuado, mediante el cual la matriz de Φ se diagonaliza (siendo simétrica):

$$A = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & d_n \end{pmatrix}$$

Por tanto, la forma cuadrática toma **forma diagonal**

$$\Phi(\mathbf{x}) = d_1 x_1^2 + \dots + d_n x_n^2 .$$

Es importante subrayar que la expresión de una forma cuadrática como suma de cuadrados no es única.

En general, si f es un endomorfismo del espacio E , con matriz asociada simétrica: $A = A^t$, entonces existe en E una base ortonormal formada por autovectores de A . También sabemos que una matriz simétrica real se puede diagonalizar mediante una matriz de cambio ortogonal: $\exists P$ tal que $P^t A P = D$, con $P^t = P^{-1}$.

Proposición 57. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\Phi \in Q(E, \mathbb{R})$ y A la matriz asociada a Φ respecto de una base de E . Si $\lambda_1, \dots, \lambda_n$ son los autovalores de A , Φ admite la forma diagonal*

$$\Phi(\mathbf{x}) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2 . \quad (7.1)$$

Demostración. Como A es una matriz real simétrica, sabemos que $\exists P$ ortogonal tal que

$$P^t A P = D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_n \end{pmatrix}$$

Las columnas de P son las componentes de los vectores de una base ortonormal $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ de autovectores (cuya existencia tenemos garantizada), respecto de la base original. Además, las componentes X de un vector de E en la base original y las componentes Y en la de autovectores están relacionadas por

$$X = P Y$$

Entonces

$$\begin{aligned}\Phi(\mathbf{x}) &= X^t A X = (PY)^t A (PY) = Y^t P^t A P Y = Y^t D Y = \\ &= \lambda_1 y_1^2 + \dots + \lambda_n y_n^2.\end{aligned}$$

Por tanto, los autovalores de A son los coeficientes de la forma diagonal de la forma cuadrática Φ . \square

Sin embargo, existen otras bases que no son bases ortonormales de autovectores, en las que la matriz de la forma canónica también es diagonal.

7.2 Método de Gauss-Lagrange

Se puede reducir una forma cuadrática Φ a forma diagonal mediante un sendo cambio lineal de variables, que se determina **completando cuadrados**. El método, conceptualmente muy sencillo, se puede ilustrar fácilmente con un ejemplo: sea

$$\Phi[(x_1, x_2, x_3)] = 2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3.$$

Mediante operaciones algebraicas elementales, tenemos:

$$\begin{aligned}\Phi[(x_1, x_2, x_3)] &= 2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3 = 2(x_1^2 - 2x_1x_2) + x_2^2 - 4x_2x_3 = \\ &= 2(x_1^2 - 2x_1x_2 + x_2^2) - 2x_2^2 + x_2^2 - 4x_2x_3 = \\ &= 2(x_1 - x_2)^2 - x_2^2 - 4x_2x_3 = \\ &= 2(x_1 - x_2)^2 - (x_2^2 + 4x_2x_3 + 4x_3^2) + 4x_3^2 = \\ &= 2(x_1 - x_2)^2 - (x_2 + 2x_3)^2 + 4x_3^2\end{aligned}$$

Sea entonces

$$\begin{cases} y_1 = x_1 - x_2 \\ y_2 = x_2 + 2x_3 \\ y_3 = x_3 \end{cases} \iff \begin{cases} x_1 = y_1 + y_2 - 2y_3 \\ x_2 = y_2 - 2y_3 \\ x_3 = y_3 \end{cases}$$

relaciones que proporcionan el cambio de variables lineal deseado. En las nuevas variables,

$$\Phi[(y_1, y_2, y_3)] = 2y_1^2 - y_2^2 + 4y_3^2.$$

En general, el método se puede plantear de la forma siguiente. Sea una forma cuadrática

$$\Phi(\mathbf{x}) = \sum_{i=1}^n a_{ii} x_i^2 + \sum_{i=1}^n \sum_{j \neq i}^n a_{ij} x_i x_j$$

Elegimos un término diagonal, por ejemplo el i -ésimo, y escribimos $\Phi(\mathbf{x})$ en la forma

$$\Phi(\mathbf{x}) = a_{ii} \left(x_i + \sum_{j \neq i}^n \frac{a_{ij}}{2a_{ii}} x_j \right)^2 + \Phi_2(\mathbf{x})$$

donde $\Phi_2(\mathbf{x})$ es una forma cuadrática que no depende de x_i . Introducimos las coordenadas

$$\tilde{x}_i = x_i + \sum_{j \neq i}^n \frac{a_{ij}}{2a_{ii}} x_j, \quad \tilde{x}_k = x_k, \quad k \neq i$$

En estas nuevas coordenadas,

$$\Phi(\mathbf{x}) = a_{ii}\tilde{x}_i^2 + \Phi_2(\mathbf{x}),$$

donde $\Phi_2(\mathbf{x})$ depende de $n-1$ coordenadas; iterando el procedimiento, aplicándolo a Φ_2, Φ_3, \dots , llegamos a una suma de cuadrados.

7.3 Ley de inercia de Sylvester

Demostraremos que el número de coeficientes positivos y de coeficientes negativos de una forma diagonal de una forma cuadrática Φ es invariante, es decir es independiente del proceso de diagonalización.

Lema 14. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo y $\Phi \in Q(E, \mathbb{R})$. El número de coeficientes no nulos de una forma diagonal de Φ es igual a su rango.*

Demostración. Como consecuencia directa de la Proposición 57, si $\lambda_1, \dots, \lambda_r$ son los autovalores no nulos de A , entonces

$$\Phi(\mathbf{x}) = \lambda_1 x_1^2 + \dots + \lambda_r x_r^2. \quad (7.2)$$

Por tanto, el rango de $D = P^t A P$ es igual a r que también es el rango de A y entonces de Φ . Cualquier otra forma diagonal de A tendrá el mismo rango, y por tanto Φ tendrá en cada forma diagonal el mismo número de coeficientes no nulos. \square

Teorema 105 (Ley de inercia). *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\dim E = n$ y $\Phi \in Q(E, \mathbb{R})$ una forma cuadrática de rango $r \leq n$. Si D_1 y D_2 son matrices diagonales asociadas a Φ respecto de dos bases distintas, el número p de coeficientes positivos y el número q de coeficientes negativos que aparecen en las formas diagonales de Φ es el mismo.*

Demostración. Utilizando el Lema anterior, el rango de Φ es $r = p + q$. Es suficiente demostrar que p es el mismo sea cual sea la base en que diagonalizamos D . Sean $\mathcal{B}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_{p_1}, \mathbf{v}_{p_1+1}, \dots, \mathbf{v}_n\}$ y $\mathcal{B}_2 = \{\mathbf{w}_1, \dots, \mathbf{w}_{p_2}, \mathbf{w}_{p_2+1}, \dots, \mathbf{w}_n\}$ bases de E para las cuales Φ es diagonal, y sean p_1 y p_2 el número de elementos positivos en las respectivas formas diagonales de Φ . Las matrices asociadas a Φ respecto de \mathcal{B}_1 y \mathcal{B}_2 son

$$\begin{pmatrix} \Phi(\mathbf{v}_1) & 0 & \dots & 0 \\ 0 & \Phi(\mathbf{v}_2) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Phi(\mathbf{v}_n) \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \Phi(\mathbf{w}_1) & 0 & \dots & 0 \\ 0 & \Phi(\mathbf{w}_2) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Phi(\mathbf{w}_n) \end{pmatrix}$$

Por hipótesis,

$$\Phi(\mathbf{v}_1) > 0, \dots, \Phi(\mathbf{v}_{p_1}) > 0, \quad \Phi(\mathbf{v}_{p_1+1}) \leq 0, \dots, \Phi(\mathbf{v}_n) \leq 0$$

y

$$\Phi(\mathbf{w}_1) > 0, \dots, \Phi(\mathbf{w}_{p_2}) > 0, \quad \Phi(\mathbf{w}_{p_2+1}) \leq 0, \dots, \Phi(\mathbf{w}_n) \leq 0.$$

Sean

$$V = \text{lin}\{\mathbf{v}_1, \dots, \mathbf{v}_{p_1}\}, \quad W = \text{lin}\{\mathbf{w}_{p_2+1}, \dots, \mathbf{w}_n\}$$

Entonces

$$\forall \mathbf{x} \in V, \mathbf{x} \neq \mathbf{0}, \Phi(\mathbf{x}) > 0, \quad \forall \mathbf{y} \in W, \mathbf{y} \neq \mathbf{0}, \Phi(\mathbf{y}) \leq 0.$$

Por tanto $V \cap W = \{\mathbf{0}\}$, luego

$$\dim(V + W) = \dim V + \dim W = p_1 + n - p_2.$$

Por otro lado

$$V \cup W \subseteq E \implies p_1 + n - p_2 \leq n \implies p_1 \leq p_2.$$

Ahora bien, por simetría, intercambiando los papeles de p_1 y p_2 , tenemos que $p_2 \leq p_1$. En definitiva, $p_1 = p_2$. \square

Definición 135. Llamaremos *signatura* de Φ al par

$$sg(\Phi) = (p, q)$$

Definición 136. Llamaremos *índices de inercia* de Φ al triple (p, q, c) , siendo c el número de coeficientes nulos en una forma diagonal de Φ .

La noción de signatura de una forma cuadrática es útil a la hora de proceder con su clasificación,

Teorema 106. Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo, $\dim E = n$ y $\Phi \in Q(E, \mathbb{R})$ una forma cuadrática de rango $r \leq n$. Entonces

- (1) Φ es definida positiva $\iff sg(\Phi) = (n, 0)$.
- (2) Φ es definida negativa $\iff sg(\Phi) = (0, n)$.
- (3) Φ es semidefinida positiva $\iff sg(\Phi) = (r, 0), \quad r < n$.
- (4) Φ es semidefinida negativa $\iff sg(\Phi) = (0, r), \quad r < n$.

Demostración.

- (1) Sea Φ definida positiva, $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base de E y

$$A = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & d_n \end{pmatrix}$$

la matriz asociada a Φ respecto de \mathcal{B} . Por construcción, $d_i = \Phi(\mathbf{e}_i)$. Entonces, siendo Φ definida positiva, $d_1 > 0, \dots, d_n > 0$ y por tanto $sg(\Phi) = (n, 0)$. Vice versa, supongamos que $sg(\Phi) = (n, 0)$. Entonces, se tiene $d_1 > 0, \dots, d_n > 0$. Por tanto, si $\mathbf{x} \in E$, $\mathbf{x} = (x_1, \dots, x_n)$, se verifica

$$\mathbf{x} \neq \mathbf{0}, \quad \Phi(\mathbf{x}) = d_1 x_1^2 + \dots + d_n x_n^2 > 0, \quad \Phi(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}$$

es decir, $\Phi(\mathbf{x})$ es definida positiva. De forma completamente análoga se demuestran los demás casos. \square

Ejercicio 7.2. Diagonalizar la forma cuadrática

$$\Phi[(x_1, x_2, x_3)] = 2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3$$

mediante una matriz de cambio de base ortogonal.

Resolución.

La matriz asociada a Φ es

$$A = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}$$

La ecuación característica es

$$\det(A - \lambda \mathbf{1}) = 0 \iff \lambda^3 - 3\lambda^2 - 6\lambda + 8 = 0 \iff (\lambda - 1)(\lambda + 2)(\lambda - 4) = 0$$

Tenemos

$$\lambda_1 = 1, \quad \begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & -2 \\ 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} x_1 - 2x_2 = 0 \\ -2x_1 - 2x_3 = 0 \\ -2x_2 - x_3 = 0 \end{cases} \iff \begin{cases} x_1 = -x_3 \\ x_2 = -\frac{1}{2}x_3 \end{cases}$$

Obtenemos

$$V_{\lambda_1} = \text{lin}\{(2, 1, -2)\}.$$

Un vector normalizado de este subespacio es $\mathbf{u}_1 = \frac{1}{3}(2, 1, -2)$. Análogamente,

$$\lambda_2 = -2, \quad \mathbf{u}_2 = \frac{1}{3}(1, 2, 2)$$

$$\lambda_3 = 4, \quad \mathbf{u}_3 = \frac{1}{3}(2, -2, 1)$$

Los vectores $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ forman una base ortonormal de \mathbb{R}^3 . La matriz ortogonal de cambio de base

$$P = \frac{1}{3} \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ -2 & 2 & 1 \end{pmatrix}$$

nos proporciona el cambio de variable

$$X = PY \iff \begin{cases} x_1 = \frac{1}{3}(2y_1 + y_2 + 2y_3) \\ x_2 = \frac{1}{3}(y_1 + 2y_2 - 2y_3) \\ x_3 = \frac{1}{3}(-2y_1 + 2y_2 + y_3) \end{cases}$$

mediante el cual la forma Φ asume la forma diagonal:

$$\Phi[(y_1, y_2, y_3)] = y_1^2 - 2y_2^2 + 4y_3^2.$$

7.4 Diagonalización por congruencia de una forma cuadrática

Sean A y B dos matrices congruentes: \exists una matriz regular P tal que

$$B = P^t A P$$

Como P^t es invertible, se puede obtener como el producto de un cierto número de matrices elementales:

$$P^t = E_k \cdot E_{k-1} \cdot \dots \cdot E_1 \quad (7.3)$$

Entonces

$$B = P^t A P = (E_k \cdot \dots \cdot E_1) A (E_k \cdot \dots \cdot E_1)^t = E_k \cdot \dots \cdot E_1 \cdot A \cdot E_1^t \cdot \dots \cdot E_k^t \quad (7.4)$$

Entonces tenemos los resultados siguientes.

Proposición 58. *Dos matrices cuadradas $A, B \in M_n$ son congruentes si y sólo si una puede obtenerse a partir de la otra efectuando transformaciones elementales, las mismas por filas que por columnas.*

Proposición 59. *Toda matriz simétrica real es congruente a una matriz diagonal en cuya diagonal sólo aparecen los valores 0, 1 y -1 .*

Corolario 19. *Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio euclídeo y $\Phi \in Q(E, \mathbb{R})$ una forma cuadrática, con matriz asociada A . Entonces:*

$$\begin{aligned} \Phi \text{ es definida positiva} &\iff A \text{ es congruente a la matriz identidad } \mathbb{1} \\ \Phi \text{ es definida negativa} &\iff A \text{ es congruente a la matriz } -\mathbb{1} \end{aligned}$$

La discusión anterior nos permite introducir la noción de forma canónica de una forma cuadrática.

Definición 137. *Llamaremos **forma canónica** de la forma cuadrática Φ a su expresión como suma de cuadrados, en que en los coeficientes de la suma sólo aparecen los valores 0, 1 y -1 .*

Ejercicio 7.3. Diagonalizar por congruencia la forma cuadrática

$$\Phi[(x_1, x_2, x_3)] = 2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3$$

Resolución. Tenemos

$$\begin{aligned} A &= \begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix} \xrightarrow{2^\circ f \rightarrow 1^\circ f + 2^\circ f} \begin{pmatrix} 2 & -2 & 0 \\ 0 & -1 & -2 \\ 0 & -2 & 0 \end{pmatrix} \xrightarrow{2^\circ c \rightarrow 1^\circ c + 2^\circ c} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & -2 & 0 \end{pmatrix} \\ &\xrightarrow{3^\circ f \rightarrow (-2) \cdot 2^\circ f + 3^\circ f} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & 4 \end{pmatrix} \xrightarrow{3^\circ c \rightarrow (-2) \cdot 2^\circ c + 3^\circ c} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \end{aligned}$$

Repetimos las mismas operaciones por columnas, sobre la matriz identidad. Tenemos

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{2^\circ c \rightarrow 1^\circ c + 2^\circ c} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{3^\circ c \rightarrow (-2) \cdot 2^\circ c + 3^\circ c} \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

Hemos obtenido que

$$P = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \quad P^t A P = D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

La base (no ortonormal) en la que la forma cuadrática se diagonaliza es

$$\mathcal{B} = \{(1, 0, 0), (1, 1, 0), (-2, -2, 1)\}.$$

La relación entre las nuevas coordenadas y las antiguas es $X = PY$:

$$\begin{cases} x_1 = y_1 + y_2 - 2y_3 \\ x_2 = y_2 - 2y_3 \\ x_3 = y_3 \end{cases}$$

y en las nuevas coordenadas la forma cuadrática se diagonaliza:

$$\Phi[(y_1, y_2, y_3)] = 2y_1^2 - y_2^2 + 4y_3^2$$

Si queremos escribir la forma cuadrática en forma canónica, podemos seguir adelante con transformaciones elementales:

$$\begin{aligned} &\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \xrightarrow{1^\circ f \rightarrow \frac{1}{\sqrt{2}} 1^\circ f} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \xrightarrow{1^\circ c \rightarrow \frac{1}{\sqrt{2}} 1^\circ c} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \\ &\xrightarrow{3^\circ f \rightarrow \frac{1}{2} 3^\circ f} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \xrightarrow{3^\circ c \rightarrow \frac{1}{2} 3^\circ c} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Repetiendo las operaciones por columna sobre la matriz P , obtenemos

$$\begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{1^\circ c \rightarrow \frac{1}{\sqrt{2}} 1^\circ c} \begin{pmatrix} \frac{1}{\sqrt{2}} & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{3^\circ c \rightarrow \frac{1}{2} 3^\circ c} \begin{pmatrix} \frac{1}{\sqrt{2}} & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

En definitiva, obtenemos

$$\tilde{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad \tilde{P}^t A \tilde{P} = \tilde{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

En las coordenadas $X = \tilde{P}Z$,

$$\begin{cases} x_1 = \frac{z_1}{\sqrt{2}} + z_2 - z_3 \\ x_2 = z_2 - z_3 \\ x_3 = \frac{z_3}{2} \end{cases}$$

la forma cuadrática Φ adquiere su forma canónica:

$$\Phi[(z_1, z_2, z_3)] = z_1^2 - z_2^2 + z_3^2.$$

Ejercicio 7.4. Sea

$$\Phi(\mathbf{x}) = x_1^2 + 3x_2^2 + 7x_3^2 + 2x_1x_2 + 4x_1x_3 + 4x_2x_3.$$

Diagonalizar Φ por congruencia.

Resolución. Tenemos

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 7 \end{pmatrix} \xrightarrow{2^\circ f \rightarrow 2^\circ f - 1^\circ f} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 2 & 2 & 7 \end{pmatrix} \xrightarrow{3^\circ f \rightarrow 3^\circ f - 2 \cdot 1^\circ f} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Efectuamos las mismas operaciones con las columnas:

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \xrightarrow{2^\circ c \rightarrow 2^\circ c - 1^\circ c} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \xrightarrow{3^\circ c \rightarrow 3^\circ c - 2 \cdot 1^\circ c} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Ahora bien, repetimos las operaciones por columna a la matriz identidad:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{2^\circ c \rightarrow 2^\circ c - 1^\circ c} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{3^\circ c \rightarrow 3^\circ c - 2 \cdot 1^\circ c} \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Entonces hemos construido las matrices

$$P = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P^t = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$

En definitiva,

$$D = P^t A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

La base (no ortonormal) en la que la forma cuadrática se diagonaliza es

$$\mathcal{B} = \{(1, 0, 0), (-1, 1, 0), (-2, 0, 1)\}.$$

En las coordenadas

$$\begin{cases} x_1 = y_1 - y_2 - 2y_3 \\ x_2 = y_2 \\ x_3 = y_3 \end{cases}$$

la forma cuadrática se diagonaliza:

$$\Phi[(y_1, y_2, y_3)] = y_1^2 + 2y_2^2 + 3y_3^2.$$

También podemos llegar a la forma canónica de Φ siguiendo el procedimiento ilustrado en el ejercicio anterior.

7.4.1 Criterio de Sylvester

Teorema 107. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica. Llamaremos submatrices principales de A a las submatrices de A situadas en sus esquinas superiores izquierdas. Con más precisión, si A :*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix}$$

entonces las submatrices principales son

$$A_k = (a_{ij})_{i,j=1,\dots,k}$$

es decir

$$A_1 = a_{11} \quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \dots \quad A_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

Teorema 108. Sea Φ una forma cuadrática real y A la matriz asociada a Φ respecto de una base dada. Entonces:

$$\begin{aligned} \Phi \text{ es definida positiva} &\iff \det(A_k) > 0, \forall k = 1, \dots, n \\ \Phi \text{ es definida negativa} &\iff (-1)^k \det(A_k) > 0, \forall k = 1, \dots, n \end{aligned}$$

Ejercicio 7.5.

Sea la forma cuadrática en \mathbb{R}^4

$$\Phi(\mathbf{x}) = ax_1^2 + bx_2^2 + ax_3^2 + bx_4^2 + 2x_1x_3 + 2x_2x_4, \quad a, b \in \mathbb{R} .$$

Determinar para qué valores reales de a y b la forma Φ es definida positiva.

Resolución.

La matriz asociada a Φ es

$$A = \begin{pmatrix} a & 0 & 1 & 0 \\ 0 & b & 0 & 1 \\ 1 & 0 & a & 0 \\ 0 & 1 & 0 & b \end{pmatrix}$$

La forma es definida positiva si y sólo si las submatrices principales de A son positivas:

- (i) $a > 0$
- (ii) $ab > 0$
- (iii)

$$\begin{vmatrix} a & 0 & 1 \\ 0 & b & 0 \\ 1 & 0 & a \end{vmatrix} = a^2b - b = b(a^2 - 1) > 0$$

- (iv)

$$a \begin{vmatrix} b & 0 & 1 \\ 0 & a & 0 \\ 1 & 0 & b \end{vmatrix} + \begin{vmatrix} 0 & b & 1 \\ 1 & 0 & 0 \\ 0 & 1 & b \end{vmatrix} > 0$$

Teniendo en cuenta las cuatro condiciones anteriores, obtenemos

$$a > 0, \quad b > 0, \quad b(a^2 - 1) > 0 \implies a > 1, \quad a(ab^2 - a) - b^2 + 1 = (a^2 - 1)(b^2 - 1) > 0 \implies b > 1$$

En definitiva, Φ es definida positiva $\iff a > 1, b > 1$.

7.5 Cónicas

7.5.1 Secciones cónicas

Intersecando un cono con un plano obtenemos secciones cónicas. Con más precisión, sea un cono de ecuación

$$x^2 + y^2 - z^2 = 0$$

y consideremos distintas intersecciones:

(1) Con el plano $\pi : z = 1$. Llegamos a la ecuación

$$x^2 + y^2 = 1$$

ecuación de una circunferencia (caso particular de una elipse).

(2) Con el plano $\pi : y = 1$. Se obtiene

$$x^2 - z^2 = -1$$

ecuación de una hipérbola.

(3) Con el plano $\pi : x - z = 1$. La ecuación correspondiente es

$$y^2 = -2x + 1$$

ecuación de una parábola.

Observación 7.1. Con el procedimiento anterior hemos obtenido cónicas no degeneradas. Al intersecar un cono con un plano también podemos obtener cónicas degeneradas.

(4) Sea $\pi : z = 0$. Obtenemos

$$x^2 + y^2 = 0$$

ecuación que define el punto origen.

(5) Sea $\pi : y = 0$. Obtenemos

$$x^2 - z^2 = 0$$

es decir,

$$(x - z)(x + z) = 0$$

ecuación verificada por los puntos de dos rectas que se cortan en el origen.

(6) Sea $\pi : x - z = 0$. La ecuación correspondiente es

$$y^2 = 0$$

ecuación que representa una recta en que cada punto es una solución doble (recta doble).

7.5.2 Ecuación general de una cónica

Desde un punto de vista algebraico, una cónica se puede interpretar como el lugar geométrico de los puntos del plano que verifican una ecuación de segundo grado en dos variables del tipo siguiente:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0 \quad (7.5)$$

Introducimos las matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad B = (2b_1 \quad 2b_2), \quad X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (7.6)$$

Entonces la ecuación (7.7) se escribe

$$X^tAX + BX + c = 0 \quad (7.7)$$

En la notación

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{pmatrix}, \quad \tilde{X} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (7.8)$$

tenemos la forma

$$\tilde{X}^t\tilde{A}\tilde{X} = 0 \quad (7.9)$$

Ejemplos 7.6.

(1) **Elipse**

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

(2) **Hipérbola**

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

(3) **Parábola**

$$y^2 = 2px$$

7.6 Ecuación reducida de una cónica

Como A es una matriz real simétrica, entonces es diagonalizable por semejanza ortogonal (lo mismo que congruencia): existen D matriz diagonal y P matriz ortogonal (que puede elegirse con $\det P = 1$), tales que

$$P^tAP = D.$$

Entonces, el cambio de variable $X = PX'$ representa una rotación en la que el origen de los ejes O queda fijo. En las nuevas coordenadas $(')$, la ecuación (7.7) se escribe en la forma

$$(X')^tDX' + BPX' + c = 0 \quad (7.10)$$

Otro paso consiste en eliminar los términos lineales en x e y , si $\lambda = 0$ no es valor propio de A , o al menos uno de ellos si $\lambda = 0$ es valor propio de A . Esto se realiza “completando cuadrados” ($\lambda \neq 0$):

$$\begin{aligned} x^2 + 2b_1x &= (x + b_1)^2 - b_1^2 \\ y^2 + 2b_2y &= (y + b_2)^2 - b_2^2 \end{aligned}$$

Por tanto, mediante la sustitución (traslación)

$$\begin{cases} \tilde{x} = x + b_1 \\ \tilde{y} = y + b_2 \end{cases} \quad (7.11)$$

quedan eliminados los elementos lineales de x e y . Si $\lambda = 0$ es valor propio, sólo podemos completar cuadrados para una variable; la otra se puede modificar para obtener un término independiente nulo.

En definitiva, podemos pasar de la ecuación general de una cónica a una ecuación reducida, mediante dos movimientos rígidos: una rotación y una traslación. Como resultado de la discusión anterior, podemos enunciar el siguiente

Teorema 109. *La ecuación general de una cónica*

$$X^t A X + B X + c$$

se puede transformar en una ecuación reducida mediante un cambio de sistema de referencia dado por

$$X'' = P^t X + X_0 \quad (7.12)$$

con P matriz ortogonal, $\det P = 1$ y X_0 vector constante.

Ecuaciones reducidas

$\alpha^2 x^2 + \beta^2 y^2 = c^2$	elipse real
$\alpha^2 x^2 + \beta^2 y^2 = -c^2$	elipse imaginaria
$\alpha^2 x^2 - \beta^2 y^2 = \pm c^2$	hipérbola
$y^2 = 2px$	parábola
$\alpha^2 x^2 + \beta^2 y^2 = 0$	un punto
$\alpha^2 x^2 - \beta^2 y^2 = 0$	par de rectas que se cortan
$y^2 = 0$	recta doble
$y^2 = c^2$	par de rectas reales paralelas
$y^2 = -c^2$	par de rectas imaginarias paralelas

7.7 Invariantes métricos de las cónicas

Regresemos a la ecuación general de una cónica:

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0$$

Introducimos las cantidades

$$I_1 = \text{tr} A = a_{11} + a_{22} \quad (7.13)$$

$$I_2 = \det A = a_{11}a_{22} - a_{12}^2 \quad (7.14)$$

$$I_3 = \det \tilde{A} \quad (7.15)$$

Vale el siguiente

Teorema 110. *Las cantidades I_1 , I_2 , I_3 quedan invariantes al cambiar de un sistema de coordenadas a otro mediante traslaciones o rotaciones.*

Estas cantidades son denominadas **invariantes métricos** de una cónica y nos permiten identificar el tipo de cónica.

7.8 Clasificación de las cónicas por invariantes

Como mostraremos más adelante, vale la clasificación siguiente.

- (I) Sea $I_3 \neq 0$. Tenemos cónicas no degeneradas.
 - (a) $I_2 > 0$ Elipse. Si $I_1 I_3 < 0$ la elipse es real. Si $I_1 I_3 > 0$ es imaginaria.
 - (b) $I_2 < 0$ Hipérbola
 - (c) $I_2 = 0$. Parábola
- (II) Sea $I_3 = 0$. Tenemos cónicas degeneradas.
 - (a) $I_2 > 0$ Un punto
 - (b) $I_2 < 0$ dos rectas secantes
 - (c) $I_2 = 0$ Dos rectas paralelas (o una doble), o ningún punto real

7.8.1 Centro de simetría de cónicas

Dada una cónica de ecuación

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0 \quad (7.16)$$

supongamos que $I_3 \neq 0$, es decir la cónica es no degenerada. Además, supondremos $I_2 \neq 0$, es decir, la cónica es una elipse o bien una hipérbola. Determinemos entonces su centro de simetría.

Mediante consideraciones geométricas elementales se puede demostrar que las ecuaciones del centro de simetría son dadas por

$$\begin{cases} a_{11}x + a_{12}y + b_1 = 0 \\ a_{12}x + a_{22}y + b_2 = 0 \end{cases} \quad (7.17)$$

o sea, en forma matricial, por la relación $AX_c = B$, siendo X_c el vector columna de las coordenadas del centro de simetría.

Las ecuaciones del centro de simetría también se escriben usualmente en la forma equivalente

$$\frac{\partial f(x, y)}{\partial x} = 0, \quad \frac{\partial f(x, y)}{\partial y} = 0 \quad (7.18)$$

Observamos que este sistema admite solución única $P = (x_c, y_c)$, puesto que por hipótesis, $I_2 \neq 0$. El punto P es el centro de simetría de la cónica considerada.

7.8.2 Reducción a forma canónica. Caso $I_2 \neq 0$

Los autovalores de A , que denotamos con λ_1 y λ_2 , son reales y distintos de cero. Los autovectores correspondientes serán los vectores directores de los ejes de la cónica.

Sea P la matriz cuyas columnas representan las componentes de los autovectores ortonormales correspondientes a λ_1 y λ_2 , elegidos con el convenio

$$P = [\mathbf{u}_1 | \mathbf{u}_2], \quad \det P = +1 .$$

Tenemos

$$A = PDP^{-1}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Sea el cambio de coordenadas

$$X = X_c + PX' \tag{7.19}$$

donde X_c denota las coordenadas del centro de la cónica, $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$. Este cambio de coordenadas lleva la ecuación a su forma reducida

$$\lambda_1(x')^2 + \lambda_2(y')^2 + \gamma = 0 \tag{7.20}$$

donde γ es una cierta constante. Observamos que la matriz completa asociada a la cónica

$$\tilde{A} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \gamma \end{pmatrix}$$

tiene ahora determinante $\det(\tilde{A}) = \lambda_1 \lambda_2 \gamma$. Como este determinante es el invariante I_3 , y siendo $I_1 = \lambda_1 + \lambda_2$, $I_2 = \lambda_1 \lambda_2$, deducimos que $\gamma = \frac{I_3}{I_2}$. Por tanto obtenemos la forma

$$\lambda_1(x')^2 + \lambda_2(y')^2 + \frac{I_3}{I_2} = 0, \quad I_2 \neq 0. \tag{7.21}$$

Distinguimos dos casos.

(a) Sea $I_2 > 0$. Entonces λ_1 y λ_2 tienen el mismo signo, que es el signo de I_1 . Si $I_1 > 0$, $I_3 < 0$ (o bien $I_1 < 0$ e $I_3 > 0$), entonces la cónica es una elipse. Si $I_3 = 0$, tenemos un punto. Si $I_1 > 0$, $I_3 > 0$ (o bien $I_1 < 0$, $I_3 < 0$), tenemos el conjunto vacío. En resumen, obtenemos una elipse si $I_1 I_3 < 0$, mientras que si $I_1 I_3 > 0$ tenemos el conjunto vacío (elipse imaginaria). Si $I_3 = 0$ tenemos un punto (véase la clasificación de la sección 7.8).

(b) Sea $I_2 < 0$. Ahora λ_1 y λ_2 tienen signos diferentes. Tenemos una hipérbola si $I_3 \neq 0$, mientras que si $I_3 = 0$ tenemos dos rectas secantes.

Ejemplo 7.7. Sea la ecuación

$$5x^2 + 6xy + 5y^2 - 32x - 32y + 56 = 0$$

Tenemos

$$A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

Los invariantes son

$$\begin{aligned} I_1 &= \operatorname{tr} A = 10 \\ I_2 &= \det A = 16 \neq 0 \\ I_3 &= \det \tilde{A} = \begin{vmatrix} 5 & 3 & -16 \\ 3 & 5 & -16 \\ -16 & -16 & 56 \end{vmatrix} = -128 < 0 \end{aligned}$$

Por tanto, $I_3 \neq 0$, $I_2 > 0$, $I_1 I_3 < 0$: se trata de una elipse real. Determinemos su centro de simetría:

$$\begin{cases} 10x + 6y - 32 = 0 \\ 6x + 10y - 32 = 0 \end{cases} \implies (x_c, y_c) = (2, 2)$$

Determinemos autovalores y autovectores de A .

$$\begin{vmatrix} 5 - \lambda & 3 \\ 3 & 5 - \lambda \end{vmatrix} = 0 \iff (5 - \lambda)^2 - 9 = 0 \iff \lambda^2 - 10\lambda + 16 = 0$$

Por tanto, $\lambda_1 = 2$, $\lambda_2 = 8$. Una base ortonormal de autovectores es dada por $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, -1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, 1)$. La matriz ortogonal

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \det P = +1$$

nos permite escribir el cambio de variable

$$\begin{aligned} X = PX' + X_c &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix} \iff \\ \begin{cases} x = \frac{1}{\sqrt{2}}(x' + y') + 2 \\ y = \frac{1}{\sqrt{2}}(-x' + y') + 2 \end{cases} \end{aligned}$$

Llegamos a la forma canónica

$$2(x')^2 + 8(y')^2 + \frac{-128}{16} = 0$$

es decir

$$\frac{(x')^2}{4} + (y')^2 = 1$$

Ejemplo 7.8. Sea la ecuación

$$x^2 + 4xy + y^2 + 6x + 4y + \frac{2}{3} = 0$$

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & \frac{2}{3} \end{pmatrix}$$

Los invariantes son

$$\begin{aligned} I_1 &= \operatorname{tr} A = 2 \\ I_2 &= \det A = -3 \\ I_3 &= \det \tilde{A} = 9 \end{aligned}$$

Entonces, $I_3 \neq 0$, $I_2 < 0$: se trata de una hipérbola.

Ecuaciones del centro de simetría:

$$\begin{cases} 2x + 4y + 6 = 0 \\ 4x + 2y + 4 = 0 \end{cases} \implies (x_c, y_c) = \left(-\frac{1}{3}, -\frac{4}{3}\right)$$

Los autovalores de A son $\lambda_1 = 3$, $\lambda_2 = -1$, y los autovectores normalizados correspondientes son $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, 1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(-1, 1)$. Tenemos:

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \det P = +1 .$$

Cambio de variable:

$$\begin{aligned} X = PX' + X_c &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} -\frac{1}{3} \\ -\frac{4}{3} \end{pmatrix} \iff \\ &\begin{cases} x = \frac{1}{\sqrt{2}}(x' - y') - \frac{1}{3} \\ y = \frac{1}{\sqrt{2}}(x' + y') - \frac{4}{3} \end{cases} \end{aligned}$$

Obtenemos la forma reducida $3(x')^2 - (y')^2 - 3 = 0$, es decir

$$(x')^2 - \frac{(y')^2}{3} = 1 .$$

7.8.3 Reducción a forma canónica. Caso $I_2 = 0$

Supongamos ahora que uno de los valores propios de A sea nulo, y que el otro sea distinto de cero (si ambos fueran nulos, no tendríamos una curva de segundo grado).

Con más precisión, sean $\lambda_1 = 0$, $\lambda_2 = \operatorname{tr}(A) \neq 0$. Entonces $I_2 = 0$.

Como antes, sea P la matriz de los autovectores ortonormales correspondientes a $\lambda_1 = 0$ y $\lambda_2 = \lambda$, elegidos con el convenio

$$P = [\mathbf{u}_1 | \mathbf{u}_2], \quad \det P = +1 .$$

Sea $X = PX'$. Esta rotación diagonaliza la parte cuadrática de la ecuación de la cónica: tenemos

$$(X')^t DX' + BPX' + c = 0 \tag{7.22}$$

Luego, completando cuadrados, determinamos la traslación que lleva la ecuación a su forma canónica. Con más precisión, en las coordenadas $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$, la ecuación de la cónica asume la forma

$$\lambda y'^2 + \beta_1 x' + \beta_2 y' + \delta = 0, \quad \lambda \neq 0. \quad (7.23)$$

Distinguimos dos casos.

(a) $\beta_1 \neq 0$

Completando cuadrados, obtenemos

$$\lambda \left(y' + \frac{\beta_2}{2\lambda} \right)^2 + \beta_1 \left(x' + \frac{\delta}{\beta_1} - \frac{\beta_2^2}{4\lambda\beta_1} \right) = 0$$

Introducimos la traslación

$$\tilde{x} = x' + \frac{\delta}{\beta_1} - \frac{\beta_2^2}{4\lambda\beta_1}, \quad \tilde{y} = y' + \frac{\beta_2}{2\lambda}$$

En las nuevas variables, la ecuación de la cónica es

$$\tilde{y}^2 + \frac{\beta_1}{\lambda} \tilde{x} = 0 \quad (7.24)$$

que representa una parábola. En términos de invariantes, tenemos $I_3 \neq 0$ (como se deduce de la forma de la ecuación reducida) e $I_2 = 0$.

(b) $\beta_1 = 0$

La ecuación (7.23) toma ahora la forma

$$\lambda y'^2 + \beta_2 y' + \delta = 0$$

Completando cuadrados, obtenemos

$$\lambda \left(y' + \frac{\beta_2}{2\lambda} \right)^2 + \delta - \frac{\beta_2^2}{4\lambda} = 0$$

Mediante la traslación

$$\tilde{x} = x', \quad \tilde{y} = y' + \frac{\beta_2}{2\lambda}$$

Llegamos a la forma reducida

$$\tilde{y}^2 = -\delta', \quad \delta' = \frac{1}{\lambda} \left(\delta - \frac{\beta_2^2}{4\lambda} \right) \quad (7.25)$$

En definitiva, (i) si $\delta' < 0$, obtenemos dos rectas paralelas $\tilde{y} = \pm\sqrt{|\delta'|}$

(ii) si $\delta' = 0$, obtenemos la recta $y = 0$ (iii) Si $\delta' > 0$, no tenemos puntos reales.

Ejercicio 7.9. Sea la ecuación

$$x^2 - 2xy + y^2 + 4x - 6y + 1 = 0.$$

Tenemos

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 1 \end{pmatrix}$$

Claramente $\text{tr}(A) = 2$. Los autovalores de A son $\lambda_1 = 0$ y $\lambda_2 = 2$. Entonces

$$I_1 = 2, \quad I_2 = 0, \quad I_3 = \det(\tilde{A}) = -1 < 0.$$

En definitiva, siendo $I_3 \neq 0$, la cónica es no degenerada, y como $I_2 = 0$ se trata de una parábola. Una base ortonormal de autovectores es dada por

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, 1), \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}}(-1, 1).$$

Tenemos:

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \det P = +1$$

Sea el cambio de variable

$$X = PX' \iff \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \iff \begin{cases} x = \frac{1}{\sqrt{2}}(x' - y') \\ y = \frac{1}{\sqrt{2}}(x' + y') \end{cases}$$

La cónica asume la forma

$$2(y')^2 - \sqrt{2}(x' + 5y') + 1 = 0$$

Vamos a completar cuadrados:

$$2 \left(y' - \frac{5}{2\sqrt{2}} \right)^2 - \frac{25}{4} - \sqrt{2}x' + 1 = 0 \iff 2 \left(y' - \frac{5}{2\sqrt{2}} \right)^2 - \sqrt{2} \left(x' + \frac{21}{4\sqrt{2}} \right) = 0$$

Entonces mediante la traslación

$$\tilde{x} = x' + \frac{21}{4\sqrt{2}}, \quad \tilde{y} = y' - \frac{5}{2\sqrt{2}}$$

llegamos a la forma canónica

$$2\tilde{y}^2 - \sqrt{2}\tilde{x} = 0$$

Ejercicio 7.10. Sea la ecuación

$$x^2 + 4xy + 4y^2 - 3x + 2y + 1 = 0$$

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1 & 2 & -\frac{3}{2} \\ 2 & 4 & 1 \\ -\frac{3}{2} & 1 & 1 \end{pmatrix}$$

Los invariantes son

$$\begin{aligned} I_1 &= \operatorname{tr} A = 5 \\ I_2 &= \det A = 0 \\ I_3 &= \det \tilde{A} = -16 \end{aligned}$$

Entonces, $I_3 \neq 0$, $I_2 = 0$: se trata de una parábola.

Los autovalores de A son $\lambda_1 = 5$, $\lambda_2 = 0$, y los autovectores normalizados correspondientes son $\mathbf{u}_1 = \frac{1}{\sqrt{5}}(2, -1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{5}}(1, 2)$. Tenemos:

$$P = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}, \quad \det P = +1$$

Cambio de variable (rotación):

$$X = PX' \iff \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \iff \begin{cases} x = \frac{1}{\sqrt{2}}(2x' + y') \\ y = \frac{1}{\sqrt{2}}(-x' + 2y') \end{cases}$$

Consideremos la transformación

$$X'DX' + BPX' + c = 0$$

siendo $c = a_{33} = 1$. Tenemos:

$$\begin{aligned} 5(y')^2 + (-3 \quad 2) \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + 1 &= 0 \iff \\ 5(y')^2 + \frac{1}{\sqrt{5}}(-6x' - 3y' - 2x' + 4y') + 1 &= 0 \iff \\ 5(y')^2 + \frac{y'}{\sqrt{5}} - \frac{8}{\sqrt{5}}x' + 1 &= 0 \iff 5 \left((y')^2 + \frac{1}{5\sqrt{5}} \right) - \frac{8}{\sqrt{5}}x' + 1 = 0 \iff \\ 5 \left(y' + \frac{1}{10\sqrt{5}} \right)^2 - \frac{1}{100} - \frac{8}{\sqrt{5}}x' + 1 &= 0 \iff 5 \left(y' + \frac{1}{10\sqrt{5}} \right)^2 - \frac{8}{\sqrt{5}}x' + \frac{99}{100} = 0 \iff \\ 5 \left(y' + \frac{1}{10\sqrt{5}} \right)^2 - \frac{8}{\sqrt{5}} \left(x' - \frac{\sqrt{5}}{8} \frac{99}{100} \right) &= 0 \end{aligned}$$

Por tanto, introduciendo el cambio de variables

$$\begin{cases} y'' = y' + \frac{1}{10\sqrt{5}} \\ x'' = x' - \frac{99\sqrt{5}}{800} \end{cases}$$

obtenemos la forma reducida $5(y'')^2 - \frac{8}{\sqrt{5}}x'' = 0$, es decir

$$(y'')^2 = \frac{8}{5\sqrt{5}}x''.$$

Bibliography

- [1] S. Lang, *Algebra*.
- [2] L. Merino, E. Santos, *Álgebra lineal con métodos elementales*, Paraninfo, 2006.
- [3] J. Rojo, *Álgebra lineal*, McGraw-Hill, 2001.