

Métodos Matemáticos de la Física

Carlos B. Brizozzo

20 de septiembre de 2012

Índice general

1. Algebra Lineal	7
1.1. Espacios lineales	7
1.1.1. Propiedades fundamentales	7
1.1.2. Independencia lineal	8
1.1.3. Dimensión de un EV	9
1.1.4. Base de un EV	9
1.1.5. Componentes de un vector	10
1.1.6. Subespacios lineales	11
1.2. Operadores lineales	12
1.2.1. Operadores lineales	12
1.2.2. Componentes de un operador lineal	14
1.2.3. Operaciones elementales	15
1.2.4. Conmutatividad	16
1.2.5. Inversa	16
1.3. Matrices	17
1.3.1. Propiedades fundamentales	18
1.3.2. Definiciones	19
1.3.3. Matrices notables	20
1.3.4. Operación por bloques	21
1.4. Transformaciones de coordenadas	22
1.4.1. Matriz de transformación	22
1.4.2. Covarianza y contravarianza	23
1.4.3. Componentes de un operador	23
1.4.4. Transformaciones de semejanza	24
1.5. Formas y espacio dual	24
1.5.1. Formas	24
1.5.2. Espacio dual	25
1.5.3. Componentes de una forma	25
1.5.4. Transformaciones de coordenadas en V^*	26
1.6. Producto interno, métrica y norma	27
1.6.1. Producto interno	27
1.6.2. Métrica	27
1.6.3. Norma	28
1.6.4. De V a V^* y vuelta con el producto interno	28

1.6.5. Correspondencia geométrica	29
1.7. Autovalores y autovectores	30
1.7.1. Autovectores a derecha	31
1.7.2. Autovectores a izquierda	33
1.7.3. Diagonalización de un operador	34
1.7.4. Operadores Hermitianos	35
1.7.5. Operadores Hermitianos con autovalores degenerados	36
1.7.6. Diagonalización simultánea de operadores Hermitianos	39
1.7.7. Operadores normales	41
1.8. Forma de Jordan	41
1.8.1. Ejemplos ilustrativos	41
1.8.2. Descomposición primaria	43
1.8.3. Forma normal de Jordan	45
1.8.4. Reducción a la forma normal de Jordan	46
1.9. Tensores	63
1.9.1. Definición de tensor	63
1.9.2. Espacio tensorial	65
1.9.3. Producto tensorial	66
1.9.4. Base y componentes	66
1.9.5. Cambio de base	68
1.9.6. Contracción de índices	68
1.9.7. Simetría	69
1.9.8. Producto exterior	71
1.9.9. Densidades tensoriales	71
1.9.10. Tensor adjunto	72
1.9.11. Ley del cociente	73
1.10. Coordenadas curvilíneas	73
1.10.1. Cambios de coordenadas locales	74
1.10.2. Vectores contravariantes	75
1.10.3. Vectores covariantes	75
1.10.4. Tensores	76
1.10.5. Densidades tensoriales	76
1.10.6. Integrales multidimensionales	77
1.10.7. Tensor métrico	80
1.10.8. Ascenso y descenso de índices	80
1.10.9. Producto escalar y norma	80
1.10.10. Símbolos de Christoffel	81
1.10.11. Derivación covariante	81
1.10.12. Operadores diferenciales	83
1.10.13. Componentes físicas de un vector	85
2. Ecuaciones Diferenciales Ordinarias	87
2.1. EDO de primer orden	87
2.2. EDO lineales de segundo orden	87
2.2.1. Coeficientes analíticos	87
2.2.2. La Ecuación de Legendre	90

ÍNDICE GENERAL

5

2.2.3. Puntos singulares regulares	93
2.2.4. La ecuación de Bessel	100
2.2.5. Problemas de contorno	103

Capítulo 1

Algebra Lineal

1.1. Espacios lineales

En Física usamos habitualmente vectores para representar magnitudes tales como posición, velocidad, fuerza, campos eléctrico y magnético, etc. Por ello es de interés formalizar y organizar algunos conceptos que quizá ya estemos habituados a usar.

Comenzaremos repasando conceptos básicos de espacios vectoriales.

1.1.1. Propiedades fundamentales

Llamaremos *espacio lineal* o *espacio vectorial* (y abreviaremos EV) a un conjunto V de objetos (llamados *vectores*), $V = \{\vec{x}, \vec{y}, \vec{z}, \dots\}$ (más un *cuerpo escalar*,¹ que denotaremos² por \mathbb{C}), que es *cerrado* bajo las siguientes operaciones:³

1. **Suma:** $\vec{a}, \vec{b} \in V \implies \vec{a} + \vec{b} = \vec{c} \in V$, que es:

- a) asociativa: $(\vec{a} + \vec{b}) + \vec{c} \equiv \vec{a} + (\vec{b} + \vec{c})$;
- b) conmutativa: $\vec{a} + \vec{b} \equiv \vec{b} + \vec{a}$.

2. **Producto por un escalar:** $\vec{a} \in V, \lambda \in \mathbb{C} \implies \lambda\vec{a} \in V$, que es:

a) distributiva:

- respecto a los vectores: $\lambda(\vec{a} + \vec{b}) \equiv \lambda\vec{a} + \lambda\vec{b}$;
- respecto a los escalares: $(\lambda + \mu)\vec{a} \equiv \lambda\vec{a} + \mu\vec{a}$

¹En algunos casos basta que los escalares tengan estructura de anillo (sin inversa multiplicativa), como ser en el reticulado $V = \mathbb{Z}^n$ con “cuerpo” (anillo) \mathbb{Z} .

²Para simplificar, usaremos \mathbb{C} tanto como notación genérica para un cuerpo escalar, como para denotar el cuerpo complejo, ya que en casos particulares el contexto dejará claro si el cuerpo son los reales, los complejos, etc.

³El símbolo \equiv denotará que ambos lados de la ecuación son iguales para todo valor posible de las variables que aparezcan, es decir $x + y \equiv y + x$ será lo mismo que $x + y = y + x \forall x, y$.

b) asociativa: $\lambda(\mu\vec{a}) \equiv (\lambda\mu)\vec{a}$

Suponemos además que:

- Existe un (único) *elemento nulo* en V : $\exists! \vec{0} \in V$ t.q. $\vec{a} + \vec{0} \equiv \vec{a}$.
- Existe un (único) *elemento unidad* en \mathbb{C} : $\exists! 1 \in \mathbb{C}$ t.q. $1\vec{a} \equiv \vec{a}$.
- Para cada vector, existe un (único) *elemento inverso* (aditivo) en V : $\forall \vec{a} \in V \exists! -\vec{a} \in V$ t.q. $\vec{a} + (-\vec{a}) = \vec{0}$.

Cualquier conjunto V que posea todas estas propiedades, será para nosotros un EV.

Ejemplo 1.1.1. Los espacios tridimensionales de vectores reales, \mathbb{R}^3 , y complejos, \mathbb{C}^3 , y el reticulado Cartesiano tridimensional (red cúbica) \mathbb{Z}^3 , son EV, con cuerpos \mathbb{R} , \mathbb{C} y \mathbb{Z} , respectivamente (este último “cuerpo” es sólo un anillo: carece de inversa multiplicativa).

Ejemplo 1.1.2. Los polinomios en una variable real x de grado a lo sumo N forman un EV, $V = \{p_N(x)\}$.

Ejemplo 1.1.3. Las funciones a valores reales definidas sobre el intervalo real $[a, b]$ forman un EV, $V = \{f : [a, b] \rightarrow \mathbb{R}\}$. Sin embargo, si restringimos a las funciones a valores positivos no obtenemos un EV.

Ejercicio 1.1.1. Muestre que $\{f : [a, b] \rightarrow \mathbb{R}_+\}$ no es un EV.

1.1.2. Independencia lineal

Un conjunto cualquiera de n vectores no nulos, $\{\vec{a}_i\}_{i=1}^n$, $\vec{a}_i \in V$, $\vec{a}_i \neq \vec{0}$, $i = 1, \dots, n$ se llama *linealmente independiente* (abreviamos LI) si la única combinación lineal de ellos que se anula es la trivial, es decir si⁴

$$\lambda_i \vec{a}_i = \vec{0} \implies \lambda_i \equiv 0, \quad \lambda_i \in \mathbb{C}, \quad i = 1, \dots, n. \quad (1.1)$$

Si un conjunto de vectores no es LI, lo llamaremos *linealmente dependiente* (y abreviaremos LD).

Ejemplo 1.1.4. En \mathbb{R}^3 , si \hat{x} , \hat{y} y \hat{z} son los versores Cartesianos usuales, $\{\hat{x}, \hat{y}\}$, $\{\hat{x}, \hat{x} + \hat{y}\}$ y $\{\hat{x}, \hat{x} + \hat{y}, \hat{x} + \hat{y} + \hat{z}\}$ son LI, pero $\{\hat{x}, \hat{y}, \hat{x} + \hat{y}\}$ es LD.

Ejemplo 1.1.5. En $\{p_5(x)\}$, $\{1, x, x^2\}$ es LI, pero $\{x, x^2, 2x + 3x^2\}$ es LD.

Ejemplo 1.1.6. En $\{f : [0, 2\pi] \rightarrow \mathbb{R}\}$, $\{\sin x, \cos x\}$ es LI.

Ejercicio 1.1.2. Demuestre las afirmaciones precedentes.

⁴De ahora en mas, salvo nota explícita en contrario, asumiremos el *convenio de suma* sobre índices repetidos: $a_i b_i := \sum_i a_i b_i$.

1.1.3. Dimensión de un EV

La *dimensión* de un EV se define como el máximo número de vectores LI que podemos hallar en ese EV. Diremos entonces que la dimensión de V es n si podemos hallar un conjunto $\{\vec{x}_i\}_{i=1}^n$ que sea LI, pero no podemos construir ningún conjunto $\{\vec{x}_i\}_{i=1}^{n+1}$ que sea LI.

La dimensión de un EV V se denota por $\dim(V)$. Una definición compacta es⁵

$$\dim(V) := \max_{\mathbb{N}_0} \{n \in \mathbb{N}_0 / \{\vec{x}_i\}_{i=1}^n \subset V \text{ LI}\}. \quad (1.2)$$

Nótese que la dimensión de un EV no necesita ser finita.

Ejemplo 1.1.7. $\dim(\mathbb{R}^3) = \dim(\mathbb{C}^3) = \dim(\mathbb{Z}^3) = 3$

Ejemplo 1.1.8. $\dim(\{p_N(x)\}) = N + 1$.

Ejemplo 1.1.9. $\dim(\{\sin(nx), \cos(nx), x \in [0, 2\pi], n \in \mathbb{N}\}) = \aleph_0$; el símbolo \aleph_0 , que se lee “aleph cero”, denota la cardinalidad de los naturales, también llamada “infinito numerable”.

Ejemplo 1.1.10. $\dim(\{f : [0, 1] \rightarrow \mathbb{R}\}) = \aleph_1$; el símbolo \aleph_1 , que se lee “aleph uno”, denota la cardinalidad de los reales, también llamada “infinito no numerable”.

Ejercicio 1.1.3. Demuestre las afirmaciones precedentes. Una forma de mostrar la última es considerar el conjunto de las funciones

$$f_a(x) = \begin{cases} 1 & \text{si } x < a, \\ 0 & \text{si } x \geq a, \end{cases}$$

con un parámetro real $a \in [0, 1]$, que es un subconjunto de $\{f : [0, 1] \rightarrow \mathbb{R}\}$, y probar que $f_a(x)$ y $f_{a'}(x)$ son LI si $a' \neq a$.

1.1.4. Base de un EV

Si $\dim(V) = n < \infty$, todo $\{\vec{e}_i\}_{i=1}^n \subset V$ LI se llama *base* de V . Los \vec{e}_i se llaman *vectores base*. Notemos que ninguno de ellos puede ser nulo.

Demostraremos ahora que cualquier vector puede ser escrito como una *única* combinación lineal de los vectores de una base dada.

Teorema 1.1.1. Sea $\dim(V) = n < \infty$, $\{\vec{e}_i\}_{i=1}^n$ base de V . Entonces $\forall \vec{x} \neq \vec{0} \in V, \exists! \{x^i\}_{i=1}^n \subset \mathbb{C}$ t.q. $\vec{x} = x^i \vec{e}_i$.

Demostración.

Existencia: Notemos $\vec{x} \neq \vec{0} \implies \{\vec{e}_i\}_{i=1}^n \cup \{\vec{x}\}$ LD. Luego existe solución no trivial para $\lambda_i \vec{e}_i + \mu \vec{x} = \vec{0}$ con $\mu \neq 0$ (porque si $\mu = 0$ la única solución sería la trivial). Sin pérdida de generalidad podemos asumir $\mu = 1$ (dividiendo todo por μ). Luego $\exists \lambda_i \neq 0$ t.q. $\vec{x} = -\lambda_i \vec{e}_i$ y basta tomar $x^i \equiv -\lambda_i$.

⁵La notación $:=$ indica que lo que está a la izquierda se *define* como lo que está a la derecha; la situación inversa se denotará por $=:$

Unicidad: Supongamos $\vec{x} = x^i \vec{e}_i = y^i \vec{e}_i$. Luego $\vec{0} = \vec{x} - \vec{x} = x^i \vec{e}_i - y^i \vec{e}_i = (x^i - y^i) \vec{e}_i$. Pero $\{\vec{e}_i\}_{i=1}^n$ base de V , luego $x^i - y^i \equiv 0$, es decir $x^i \equiv y^i$.

□

Este resultado no se extiende directamente a EVs de dimensión infinita, sino que se reduce a la desigualdad de Cauchy-Schwartz, que será tratada en los Problemas.

Ejemplo 1.1.11. En \mathbb{R}^3 , si \hat{x} , \hat{y} y \hat{z} son los versores Cartesianos usuales, $\{\hat{x}, \hat{y}, \hat{z}\}$ es base, y $\{\hat{x}, \hat{x} + \hat{y}, \hat{x} + \hat{y} + \hat{z}\}$ también.

Ejemplo 1.1.12. En $\{p_N(x)\}$, $\{1, x, x^2, \dots, x^N\}$ es base.

Ejercicio 1.1.4. Demuestre las afirmaciones precedentes.

Ejemplo 1.1.13. En $\{f : [0, 2\pi] \rightarrow \mathbb{R}, f$ periódica y de variación acotada $\}$, $\{1\} \cup \{\sin(nx), \cos(nx), x \in [0, 2\pi], n \in \mathbb{N}\}$ es base, resultado que conocemos del desarrollo en serie de Fourier (Teorema de Dirichlet).

1.1.5. Componentes de un vector

Dado un EV V de dimensión n (finita), una base $\{\vec{e}_i\}_{i=1}^n$ y un vector $\vec{x} \neq \vec{0} \in V$, los coeficientes x^i de la (única) combinación lineal que expresa

$$\vec{x} = x^i \vec{e}_i \quad (1.3)$$

se llaman *componentes* de \vec{x} en la base $\{\vec{e}_i\}_{i=1}^n$. El Teorema 1.1.1 muestra que los \vec{e}_i , $i = 1, \dots, n$ forman un *conjunto completo*, ya que los x^i existen para *cualquier* vector $\vec{x} \neq \vec{0}$ de V . También muestra que, dada una base, las componentes de un vector en esa base son únicas, de donde es directo deducir los siguientes corolarios:

Corolario. *Dos vectores son iguales si y sólo si sus componentes en una dada base son iguales.*

Corolario. *En cualquier base, las componentes del vector nulo $\vec{x} = \vec{0}$ son $x^i \equiv 0$.*

Ejercicio 1.1.5. Demuestre los dos corolarios precedentes del Teorema 1.1.1.

Ejemplo 1.1.14. Las componentes de $a_0 + a_1 x + a_2 x^2 + \dots + a_N x^N \in \{p_N(x)\}$, en la base $\{1, x, x^2, \dots, x^N\}$, son simplemente sus coeficientes $a_0, a_1, a_2, \dots, a_N$.

Ejemplo 1.1.15. En $\{f : [0, 2\pi] \rightarrow \mathbb{R}, f$ periódica y de variación acotada $\}$, la función definida por $f(x) = x$, $-\pi < x < \pi$, $f(x + 2\pi) \equiv f(x)$ puede escribirse en la base $\{1\} \cup \{\sin(nx), \cos(nx), x \in [0, 2\pi], n \in \mathbb{N}\}$ como $f(x) = \sum_{n=1}^{\infty} (2(-1)^{n+1}/n) \sin(nx)$.

Ejercicio 1.1.6. Demuestre las afirmaciones precedentes.

1.1.6. Subespacios lineales

El tema de subespacios lineales es muy extenso, y la mayor parte de él cae fuera de los objetivos de este curso. Aquí daremos sólo unas pocas definiciones que necesitaremos mas adelante.

Dado un EV V , ciertos subconjuntos (propios o no) de sus elementos pueden ser, por derecho propio, EVs. Diremos que W es un *subespacio lineal* o *subespacio vectorial* de V si:

- $\forall \vec{x} \in W, \vec{x} \in V;$
- W es un EV.

Ejemplo 1.1.16. \forall EV V , $\{\vec{0}\}$ y V son subespacios de V .

Ejemplo 1.1.17. El plano (x, y) es un subespacio de \mathbb{R}^3 , y el eje z también.

Ejemplo 1.1.18. $\{p_3(x)\}$ es un subespacio de $\{p_5(x)\}$.

Ejemplo 1.1.19. $\mathcal{C}^N[\mathbb{R}]$, el conjunto de las funciones de \mathbb{R} en \mathbb{R} continuamente diferenciables N veces, es un subespacio de $\mathcal{C}^M[\mathbb{R}]$ si $M \geq N$.

Ejercicio 1.1.7. Demuestre las afirmaciones precedentes.

Dado un conjunto de vectores $\{\vec{x}_i\}_{i=1}^k$ de V , no necesariamente LI, el conjunto de sus combinaciones lineales

$$\text{span}(\{\vec{x}_i\}_{i=1}^k) := \{\lambda_i \vec{x}_i, \lambda_i \in \mathbb{C}\} \quad (1.4)$$

es un subespacio de V , y

$$\dim(\text{span}(\{\vec{x}_i\}_{i=1}^k)) \leq k \quad (= k \text{ si } \{\vec{x}_i\} \text{ LI}). \quad (1.5)$$

Ejemplo 1.1.20. En \mathbb{R}^3 con los versores cartesianos usuales, $\text{span}(\hat{x})$ es el eje x , y $\text{span}(\{\hat{x}, \hat{y}, \hat{x} + \hat{y}\})$ es el plano (x, y) .

Ejemplo 1.1.21. En $\{p_5(x)\}$, $\text{span}(1, x, x^2) = \{p_2(x)\}$

Ejercicio 1.1.8. Demuestre la propiedad (1.5).

Dados dos subespacios U y W de un EV V , su *intersección*

$$U \cap W := \{\vec{x} \in V / \vec{x} \in U \wedge \vec{x} \in W\} \quad (1.6)$$

y su *suma vectorial*

$$U + W := \{\vec{x} + \vec{y} \in V / \vec{x} \in U \wedge \vec{y} \in W\} \quad (1.7)$$

son también subespacios de V . Dos subespacios U y W se llaman *disjuntos* si

$$U \cap W = \{\vec{0}\}, \quad (1.8)$$

es decir si no contienen vectores comunes a ambos, salvo el vector nulo.

Ejercicio 1.1.9. Muestre que dos vectores pertenecientes a sendos subespacios disjuntos son linealmente independientes.

Una noción importante es la de suma directa de subespacios. Consideremos una colección *finita* de subespacios V_i de V , $i = 1, \dots, r$, tales que

- $V_i \cap V_j = \{\vec{0}\}$ si $i \neq j$, es decir son todos disjuntos dos a dos;
- $\forall \vec{x} \in V \exists! \vec{x}_i \in V_i / \vec{x} = \sum_{i=1}^r \vec{x}_i$, es decir que todo vector de V puede descomponerse de forma *única* como suma de “contribuciones” provenientes de cada subespacio.

Entonces decimos que V es *suma directa*⁶ de sus subespacios V_i , y lo denotamos

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_r = \bigoplus_{i=1}^r V_i. \quad (1.9)$$

Es inmediato que en ese caso

$$\dim(V) = \sum_{i=1}^r \dim(V_i). \quad (1.10)$$

Ejemplo 1.1.22. \mathbb{R}^3 puede descomponerse como suma directa del plano (x, y) y el eje z , o también

$$\mathbb{R}^3 = \text{span}(\hat{x}) \oplus \text{span}(\hat{x} + \hat{y}) \oplus \text{span}(\hat{x} + \hat{y} + \hat{z}),$$

de donde podemos ver que la descomposición de un EV en suma directa de subespacios disjuntos dista de ser única.

Ejercicio 1.1.10. Demuestre la propiedad (1.10).

1.2. Operadores lineales

Los operadores son aplicaciones que toman elementos de un espacio V y devuelven elementos correspondientes pertenecientes a otro espacio W (en general distinto). Estudiaremos algunos casos y propiedades importantes de operadores de un EV a otro.

1.2.1. Operadores lineales

Llamaremos *función vectorial* u *operador* de un EV V en otro EV W a una aplicación $\mathcal{A} : V \rightarrow W$, que a cada elemento \vec{x} de V le hace corresponder un elemento $\mathcal{A}(\vec{x})$ de W , es decir

$$V \ni \vec{x} \xrightarrow{\mathcal{A}} \mathcal{A}(\vec{x}) \in W. \quad (1.11)$$

⁶La generalización de esta idea a espacios de dimensión infinita se llama *producto directo*.

Llamaremos *operador lineal* a un \mathcal{A} tal que

$$\mathcal{A}(\lambda \vec{x} + \mu \vec{y}) \equiv \lambda \mathcal{A}(\vec{x}) + \mu \mathcal{A}(\vec{y}), \quad (1.12)$$

es decir, tal que su acción sea distributiva respecto a la suma en V , y “transparente” a la multiplicación por un escalar.⁷

Ejemplo 1.2.1. El operador que rota (un ángulo dado) alrededor del eje z : $\mathcal{R}_z : \mathbb{R}^3 \rightarrow \mathbb{R}$; el operador que proyecta sobre el plano (x, y) : $\mathcal{P}_z : \mathbb{R}^3 \rightarrow \mathbb{R}^2$.

Ejemplo 1.2.2. El operador que deriva respecto a x : $\partial_x : \{p_N(x)\} \rightarrow \{p_{N-1}(x)\}$; el operador que integra respecto a x : $\int dx : \{p_{N-1}(x)\} \rightarrow \{p_N(x)\}$.

Dado un operador $\mathcal{A} : V \rightarrow W$ y un vector $\vec{x} \in V$, llamaremos a $\vec{y} = \mathcal{A}\vec{x}$ la *imagen* de \vec{x} bajo \mathcal{A} , y a \vec{x} una *preimagen* de \vec{y} bajo \mathcal{A} . Para \mathcal{A} lineal, la imagen es única, pero la preimagen no necesariamente lo es.

Ejemplo 1.2.3. En \mathbb{R}^3 , la imagen de $2\hat{x} + \hat{y} - 3\hat{z}$ bajo \mathcal{P}_z es $2\hat{x} + \hat{y}$, pero cualquier vector de la forma $2\hat{x} + \hat{y} + \lambda\hat{z}$ con $\lambda \in \mathbb{C}$ es preimagen de $2\hat{x} + \hat{y}$ bajo \mathcal{P}_z .

Ejemplo 1.2.4. nx^{n-1} es la imagen de x^n bajo ∂_x , pero cualquier polinomio de la forma $a + x^n$ con $a \in \mathbb{R}$ es preimagen de nx^{n-1} bajo ∂_x .

Dado un operador lineal $\mathcal{A} : V \rightarrow W$, el conjunto de vectores de V que son llevados al vector nulo en W se conoce como el *núcleo (kernel)* de \mathcal{A} , y se denota por

$$\ker(\mathcal{A}) := \{\vec{x} \in V / \mathcal{A}\vec{x} = \vec{0} \in W\}. \quad (1.13)$$

Es fácil ver que

$$\dim(\ker(\mathcal{A})) \geq \dim(V) - \dim(W). \quad (1.14)$$

Ejemplo 1.2.5. En \mathbb{R}^3 , $\ker(\mathcal{P}_z) = \text{span}(\hat{z})$.

Ejemplo 1.2.6. En $\{p_N(x)\}$, $\ker(\partial_x) = \{p_0(x)\}$.

Ejercicio 1.2.1. Demuestre que $\ker(\mathcal{A})$ es un subespacio de V .

Ejercicio 1.2.2. Demuestre la desigualdad (1.14).

Dado un operador lineal $\mathcal{A} : V \rightarrow V$, un subespacio W de V se llama *invariante bajo \mathcal{A}* si $\mathcal{A}\vec{x} \in W \forall \vec{x} \in W$. En particular, $\ker(\mathcal{A})$ es invariante bajo \mathcal{A} .

Ejemplo 1.2.7. En \mathbb{R}^3 , $\text{span}(\{\hat{x}, \hat{y}\})$ y $\text{span}(\hat{z})$ son invariantes bajo \mathcal{R}_z .

Ejemplo 1.2.8. En $\mathcal{C}^1[\mathbb{R}]$, $\{e^{ax}, a \in \mathbb{R}\}$ es invariante bajo ∂_x .

Ejercicio 1.2.3. Demuestre las afirmaciones precedentes.

⁷En adelante, cuando un operador sea lineal omitiremos los paréntesis del argumento siempre que ello no cree ambigüedad, poniendo simplemente $\mathcal{A}\vec{x}$ en lugar de $\mathcal{A}(\vec{x})$; ya que esta notación recuerda a la de producto usual, que comparte las propiedades de linealidad.

Una clase de operadores lineales íntimamente relacionados a los subespacios de un EV V , y que usaremos mas adelante, son los llamados proyectores. Dado un subespacio W de V , el *proyector sobre W* se define por

$$\mathcal{P}_W \vec{x} = \begin{cases} \vec{x} & \text{si } \vec{x} \in W, \\ \vec{0} & \text{si } \vec{x} \notin W, \end{cases} \quad (1.15)$$

y es el operador que “extrae la parte de \vec{x} que está en W ”. Es inmediato que

$$\dim(\ker(\mathcal{P}_W)) = \dim(V) - \dim(W). \quad (1.16)$$

Ejercicio 1.2.4. Muestre que \mathcal{P}_W es *idempotente*: $\mathcal{P}_W^2 = \mathcal{P}_W$.

Si un dado subespacio es invariante bajo un operador lineal, entonces la acción del operador sobre los vectores de ese subespacio puede estudiarse ignorando el resto del espacio vectorial, lo que en general simplifica el problema. La formalización de esta idea lleva a definir la restricción de un operador a un subespacio: dado un operador lineal $\mathcal{A} : V \rightarrow V$ y un subespacio W invariante bajo \mathcal{A} , la *restricción de \mathcal{A} a W* se define como el operador $\mathcal{A}/W : W \rightarrow W$ que actúa sobre los elementos de W exactamente como lo haría \mathcal{A} . Debe notarse que \mathcal{A}/W actúa sobre W y *no* sobre V . En las aplicaciones que desarrollaremos más adelante, sin embargo, será mas conveniente pensar a \mathcal{A}/W como un operador sobre V , para lo cual lo redefiniremos

$$\mathcal{A}/W := \mathcal{A}\mathcal{P}_W. \quad (1.17)$$

1.2.2. Componentes de un operador lineal

Veremos ahora que un operador lineal puede descomponerse, en forma análoga a un vector, en sus componentes respecto a bases dadas.

Sea $\{\vec{e}_i\}_{i=1}^n$ base de un EV V [$\dim(V) = n$], y sea $\vec{x} = x^i \vec{e}_i \in V$.

Sea $\{\vec{f}_j\}_{j=1}^m$ base de un EV W [$\dim(W) = m$], y sea $\vec{y} = y^j \vec{f}_j \in W$.

Sea $\mathcal{A} : V \rightarrow W$ un operador lineal tal que $\vec{y} = \mathcal{A}\vec{x}$.

Por linealidad, $\mathcal{A}(\vec{x}) = \mathcal{A}(x^i \vec{e}_i) = x^i \mathcal{A}\vec{e}_i$.

Pero para cada \vec{e}_i , $\mathcal{A}\vec{e}_i$ es un vector en W ; luego $\exists!$ coeficientes A_i^j t.q. $\mathcal{A}\vec{e}_i = A_i^j \vec{f}_j$; entonces $y^j \vec{f}_j = A_i^j x^i \vec{f}_j$.

Como los \vec{f}_j forman base de W , podemos igualar las componentes: $y^j = A_i^j x^i$.

Vemos entonces que dado un operador lineal $\mathcal{A} : V \rightarrow W$, su acción sobre un vector cualquiera $\vec{x} \in V$ queda completamente determinada dando los coeficientes A_i^j , que se denominan *componentes* de \mathcal{A} en el par de bases $\{\vec{e}_i\}$, $\{\vec{f}_j\}$. Dicho de otra forma, un operador lineal queda completamente caracterizado por su acción sobre cada uno de los vectores de una base.

Debemos notar que, así como los vectores tienen en general un significado físico independiente del sistema de coordenadas (es decir, de la base en que expresamos sus componentes), los operadores también tienen un significado físico independiente del sistema de coordenadas. Volveremos sobre esta idea en la Sección 1.4.

Ejemplo 1.2.9. Consideremos una rotación de ángulo θ (en sentido positivo) alrededor del eje z en \mathbb{R}^3 , y llamemos \mathcal{R} al correspondiente operador lineal. En este caso $V = W = \mathbb{R}^3$, y podemos tomar $\vec{f}_i \equiv \vec{e}_i$, con \vec{e}_1 los versores Cartesianos usuales. Sea $\vec{x} = x^i \vec{e}_i$ un vector cualquiera en \mathbb{R}^3 ; el vector rotado será $\vec{y} = \mathcal{R}\vec{x} = x^i \mathcal{R}\vec{e}_i$, con

$$\mathcal{R}\vec{e}_i = \begin{cases} \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2 & \text{si } i = 1, \\ -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2 & \text{si } i = 2, \\ \vec{e}_3 & \text{si } i = 3. \end{cases}$$

Las componentes de \mathcal{R} en el par de bases $\{\vec{e}_i\}$, $\{\vec{e}_i\}$ serán entonces los elementos R_i^j de la matriz

$$[R_i^j] = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

y las componentes del vector rotado serán $y^j = R_i^j x^i$.

Ejemplo 1.2.10. Consideremos el operador $\partial_x : \{p_3(x)\} \rightarrow \{p_2(x)\}$; sus componentes en el par de bases $\{1, x, x^2, x^3\}$ de $\{p_3(x)\}$ y $\{1, x, x^2\}$ de $\{p_2(x)\}$ serán los elementos de la matriz

$$[\partial_i^j] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix},$$

y el resultado de aplicarlo al elemento $ax - bx^3 \in \{p_3(x)\}$ será

$$\partial_x(ax - bx^3) = a - 3bx^2.$$

1.2.3. Operaciones elementales

Las operaciones elementales con operadores lineales se definen a partir de sus acciones sobre los vectores de un EV:

Igualdad: Dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow W$ se dicen *iguales* si $\mathcal{A}\vec{x} \equiv \mathcal{B}\vec{x}$, y lo denotamos $\mathcal{A} = \mathcal{B}$.

Suma: Dados dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow W$, su *suma* $\mathcal{C} := \mathcal{A} + \mathcal{B}$ se define como el operador lineal que realiza la asignación $\mathcal{C}\vec{x} \equiv \mathcal{A}\vec{x} + \mathcal{B}\vec{x}$.

Producto: Dados dos operadores lineales $\mathcal{A} : V \rightarrow W$, $\mathcal{B} : W \rightarrow U$, su *composición o producto* $\mathcal{C} := \mathcal{B} \circ \mathcal{A} = \mathcal{B}\mathcal{A}$ se define como el operador lineal $\mathcal{C} : V \rightarrow U$ que realiza la asignación $\mathcal{C}\vec{x} \equiv \mathcal{B}(\mathcal{A}\vec{x})$.

Producto por un escalar: Dado un operador lineal $\mathcal{A} : V \rightarrow W$, su *producto por un escalar* $\lambda \in \mathbb{C}$ se define como el operador lineal $\lambda\mathcal{A} : V \rightarrow W$ que realiza la asignación $(\lambda\mathcal{A})\vec{x} \equiv \lambda(\mathcal{A}\vec{x})$.

Identidad: El *operador identidad* $\mathcal{I} : V \rightarrow V$ se define como el que realiza la asignación $\mathcal{I}\vec{x} \equiv \vec{x}$.

Las consecuencias de estas definiciones sobre las componentes de los operadores lineales, serán tratadas en el contexto de matrices en la Sección 1.3.

1.2.4. Conmutatividad

Diremos que dos operadores lineales \mathcal{A} y \mathcal{B} *comutan* si $\mathcal{AB} = \mathcal{BA}$. Para ello es necesario que ambas composiciones tengan sentido, así que es condición necesaria que $V \xrightarrow{\mathcal{A}} V \xrightarrow{\mathcal{B}} V$, pero no es condición suficiente. En general, $\mathcal{AB} \neq \mathcal{BA}$.

Ejercicio 1.2.5. Usando las definiciones de las operaciones elementales entre operadores lineales, muestre que $\mathcal{AB} = \mathcal{BA}$ requiere $V \xrightarrow{\mathcal{A}} V \xrightarrow{\mathcal{B}} V$.

Ejemplo 1.2.11. Consideremos nuevamente rotaciones en \mathbb{R}^3 . En particular, sea \mathcal{R}_x una rotación de ángulo $\pi/2$ en sentido positivo alrededor del eje x , y sea \mathcal{R}_y una rotación de ángulo $\pi/2$ en sentido positivo alrededor del eje y . Es fácil ver que $\mathcal{R}_x\hat{y} = \hat{z}$ y $\mathcal{R}_y\hat{z} = \hat{x}$, así que $\mathcal{R}_y\mathcal{R}_x\hat{y} = \hat{x}$. Por otro lado, $\mathcal{R}_y\hat{y} = \hat{y}$ y $\mathcal{R}_x\hat{y} = \hat{z}$, de modo que $\mathcal{R}_x\mathcal{R}_y\hat{y} = \hat{z}$. Como hay al menos un vector (\hat{y}) cuyas imágenes bajo $\mathcal{R}_y\mathcal{R}_x$ y bajo $\mathcal{R}_x\mathcal{R}_y$ son distintas, la definición de igualdad entre operadores nos dice que $\mathcal{R}_y\mathcal{R}_x \neq \mathcal{R}_x\mathcal{R}_y$, es decir, *no comutan*.

Dados dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow VB$ definimos su *comutador* como

$$[\mathcal{A}, \mathcal{B}] = \mathcal{AB} - \mathcal{BA}. \quad (1.18)$$

Es evidente que $[\mathcal{B}, \mathcal{A}] = -[\mathcal{A}, \mathcal{B}]$, $[\mathcal{A}, \mathcal{A}] = 0$ y $[\mathcal{A}, \mathcal{I}] = 0$.

1.2.5. Inversa

Un operador lineal lleva elementos de un EV a otro, o posiblemente al mismo, así que cabe preguntarse si existirá otro operador que “deshaga” lo que este hizo. Esto nos lleva a la definición de inversa de un operador. Dado $\mathcal{A} : V \rightarrow W$ y $\mathcal{B} : W \rightarrow V$ diremos que \mathcal{B} es la *inversa* de \mathcal{A} si

$$\mathcal{BA} = \mathcal{I}, \quad (1.19)$$

y usualmente denotaremos $\mathcal{B} = \mathcal{A}^{-1}$. Nótese que la inversa de $\mathcal{A} : V \rightarrow W$ existe sólo si $\dim(W) \leq \dim(V)$, pero esta condición no es suficiente.

Ejemplo 1.2.12. Consideremos $\mathcal{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dado por

$$\mathcal{A}\vec{e}_i = \begin{cases} \vec{e}_1 + \vec{e}_3 & \text{si } i = 1, \\ \vec{e}_2 + \vec{e}_3 & \text{si } i = 2, \end{cases}$$

y sea $\mathcal{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definido como

$$\mathcal{P}\vec{e}_i = \begin{cases} \vec{e}_1 & \text{si } i = 1, \\ \vec{e}_2 & \text{si } i = 2, \\ \vec{0} & \text{si } i = 3. \end{cases}$$

Entonces es evidente que $\mathcal{P}\mathcal{A} = \mathcal{I} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, ya que

$$\mathcal{P}\mathcal{A}\vec{e}_i = \begin{cases} \mathcal{P}(\vec{e}_1 + \vec{e}_3) = \vec{e}_1 & \text{si } i = 1, \\ \mathcal{P}(\vec{e}_2 + \vec{e}_3) = \vec{e}_2 & \text{si } i = 2, \end{cases}$$

y podemos poner $\mathcal{P} = \mathcal{A}^{-1}$. Sin embargo $\mathcal{A}\mathcal{P} \neq \mathcal{I} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, ya que por ejemplo $\mathcal{A}\mathcal{P}(\vec{e}_1 + \lambda\vec{e}_3) = \mathcal{A}\vec{e}_1 = \vec{e}_1 + \vec{e}_3 \forall \lambda \in \mathbb{C}$. Es más, podemos ver que $\#\mathcal{B} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tal que $\mathcal{B}\mathcal{P} = \mathcal{I}$, justamente debido a que $\mathcal{P}(\lambda\vec{e}_3) = \vec{0} \forall \lambda$: la información sobre la componente \vec{e}_3 de un vector dado, es destruida irrecuperablemente por la acción de \mathcal{P} , y no habrá operador que la recupere.

Ejemplo 1.2.13. Consideremos el espacio $V = \mathcal{C}^1[\mathbb{R}]$ y los operadores derivada ∂_x e integral $\int dx$. Es fácil ver que $\partial_x \int dx = \mathcal{I}$, sin embargo $\int dx \partial_x \neq \mathcal{I}$, como se comprueba aplicándolo a la función constante.

De los ejemplos anteriores se desprende que

$$\mathcal{B}\mathcal{A} = \mathcal{I} \quad \Leftrightarrow \quad \mathcal{B}\mathcal{A} = \mathcal{I}. \quad (1.20)$$

También vemos que el problema para invertir un operador surge cuando el kernel del operador es no trivial, lo que nos lleva a enunciar el siguiente

Teorema 1.2.1. *Un operador lineal \mathcal{A} es invertible si $\dim(\ker(\mathcal{A})) = 0$.*

La demostración puede consultarse en cualquier buen texto de Álgebra Lineal.

1.3. Matrices

Sea $\mathcal{A} : V \rightarrow W$ un operador lineal, y A_i^j sus componentes en el par de bases $\{\vec{e}_i\}_{i=1}^n$ de V y $\{\vec{f}_j\}_{j=1}^m$ de W . Podemos escribir estas componentes como una matriz $m \times n$ con elementos A_i^j , en la forma

$$\mathbf{A} = [A_i^j] := \begin{pmatrix} A_1^1 & A_2^1 & \dots & A_n^1 \\ A_1^2 & A_2^2 & \dots & A_n^2 \\ \vdots & \vdots & & \vdots \\ A_1^m & A_2^m & \dots & A_n^m \end{pmatrix} \quad (\text{m filas por n columnas}). \quad (1.21)$$

De la misma forma en que usualmente representamos a un vector por la n -upla formada por sus componentes en una base dada, representaremos en general a un operador lineal por la matriz $m \times n$ formada por sus componentes en un par de bases dadas, y nos referiremos a esa matriz como la *representación* del operador en esas bases.

1.3.1. Propiedades fundamentales

La definición y propiedades fundamentales de las matrices y las operaciones entre ellas, pueden deducirse de aquellas de los operadores lineales que representan:

- Por la definición de **igualdad** entre operadores lineales, dos matrices serán *iguales*, $A = B$, si sus componentes o *elementos* son iguales: $A_i^j \equiv B_i^j$.
- Por la definición de **suma** entr operadores lineales, la *suma* de dos matrices A y B satisface $[A + B]_i^j \equiv A_i^j + B_i^j$; es decir, los elementos de la suma son la suma de los elementos.
- Por la definición de **producto** entr operadores lineales, el *producto* de dos matrices A y B satisface $[AB]_i^j \equiv A_i^k B_k^j$; es decir, el elemento $_{ij}^j$ de AB es la suma de los productos de los elementos de la fila j de A por los correspondientes elementos de la columna i de B . Evidentemente, el número de *columnas* de A debe ser igual al número de *filas* de B para que el producto AB tenga sentido. Recuperamos así la conocida regla de multiplicación entre matrices.
- Por la definición de **producto por un escalar** de un operador lineal, el *producto por un escalar* λ de una matriz A satisface $[\lambda A]_i^j = \lambda A_i^j$; es decir, multiplicar a una matriz por un escalar multiplica cada uno de sus elementos por ese mismo escalar.
- Por la definición del **operador identidad** \mathcal{I} , sus componentes en cualquier base resultan $I_i^j \equiv \delta_i^j$, donde δ_i^j es la *delta de Kronecker*; por lo tanto la matriz que representa a \mathcal{I} es la *matriz identidad*

$$\mathbf{I} := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

La matriz identidad *siempre* es cuadrada; cuando sea necesario indicar su dimensión, lo haremos por medio de un subíndice: \mathbf{I}_n indica la matriz identidad $n \times n$.

Ejercicio 1.3.1. Demuestre las propiedades precedentes.

Ejemplo 1.3.1. Como ayuda para el ejercicio anterior, para la suma: Sean $\mathcal{A}, \mathcal{B} : V \rightarrow W$ operadores lineales, $\{\vec{e}_i\}_{i=1}^n$ base de V , $\{\vec{f}_j\}_{j=1}^m$ base de W , $\vec{x} = x^i \vec{e}_i \in V$. Sabemos que $(\mathcal{A} + \mathcal{B})\vec{x} = \mathcal{A}\vec{x} + \mathcal{B}\vec{x} \forall \vec{x} \in V$. Luego $[\mathcal{A} + \mathcal{B}]_i^j x^i \vec{f}_j = A_i^j x^i \vec{f}_j + B_i^j x^i \vec{f}_j = (A_i^j + B_i^j) x^i \vec{f}_j$. Como $\{\vec{f}_j\}$ es base, podemos igualar las componentes: $[\mathcal{A} + \mathcal{B}]_i^j x^i = (A_i^j + B_i^j) x^i \forall j$. Pero esta igualdad debe valer $\forall \vec{x}$, es decir $\forall x^i$, luego $[\mathcal{A} + \mathcal{B}]_i^j \equiv (A_i^j + B_i^j)_i$ Q.E.D.

Debemos destacar que, de igual modo que la composición de operadores lineales (de la que deriva sus propiedades), el producto de matrices en general *no es conmutativo*, es decir $\mathbf{A}\mathbf{B} \neq \mathbf{B}\mathbf{A}$. Dos matrices *comutan*, esto es $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$, si los correspondientes operadores lo hacen, es decir si $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$ (y por ello mismo, si dos matrices comutan, lo hacen en cualquier base).

Notemos que si $\vec{x} = x^i \vec{e}_i \in V$ [$\dim(V) = n$], podemos ordenar las componentes x^i de \vec{x} como una *matriz columna* $n \times 1$

$$\mathbf{x} = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

Si además $\vec{y} = \mathcal{A}\vec{x} = y^j \vec{f}_j \in W$ [$\dim(W) = m$], entonces $y^j = A^j_i x^i$, y vemos que la notación de matriz columna es consistente con la regla de producto entre matrices:

$$\mathbf{y} = \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} A^1_1 & A^1_2 & \dots & A^1_n \\ A^2_1 & A^2_2 & \dots & A^2_n \\ \vdots & \vdots & & \vdots \\ A^m_1 & A^m_2 & \dots & A^m_n \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \mathbf{Ax}.$$

1.3.2. Definiciones

Dada una matriz \mathbf{A} , existen diversas matrices relacionadas a ella;⁸ algunas de las mas comunes son:

Matriz	Notación	Componentes	Ejemplo
\mathbf{A}	A^j_i		$\begin{pmatrix} 1 & i \\ 1+i & 2 \end{pmatrix}$
Conjugada	\mathbf{A}^*	$[\mathbf{A}^*]^j_i = (A^j_i)^*$	$\begin{pmatrix} 1 & -i \\ 1-i & 2 \end{pmatrix}$
Transpuesta	\mathbf{A}^T	$[\mathbf{A}^T]^j_i = A^i_j$	$\begin{pmatrix} 1 & 1+i \\ i & 2 \end{pmatrix}$
Adjunta	\mathbf{A}^\dagger	$[\mathbf{A}^\dagger]^j_i = (A^i_j)^*$	$\begin{pmatrix} 1 & 1-i \\ -i & 2 \end{pmatrix}$
Inversa	\mathbf{A}^{-1}	$[\mathbf{A}^{-1}]^j_i = \frac{\text{cof } A^i_j}{\det \mathbf{A}}$	$\frac{1}{3-i} \begin{pmatrix} 2 & -i \\ -1-i & 1 \end{pmatrix}$

En particular, si \mathbf{A} es la matriz que representa a un operador lineal \mathcal{A} , entonces la *matriz inversa* \mathbf{A}^{-1} es la matriz que representa al operador inverso \mathcal{A}^{-1} ; esto puede verificarse fácilmente escribiendo ambos operadores en componentes. De la misma forma que el producto de un operador por su inversa resulta en el operador identidad, $\mathcal{A}^{-1}\mathcal{A} = \mathcal{I}$, el producto de una matriz por su inversa resulta en la matriz identidad, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Como se recordará, la matriz inversa de \mathbf{A} existe si \mathbf{A} es *no singular*, es decir si $\det(\mathbf{A}) \neq 0$, de donde podemos escribir

$$\det(\mathbf{A}) \neq 0 \Leftrightarrow \dim(\ker(\mathcal{A})) = 0. \quad (1.22)$$

⁸Restringimos lo que sigue al caso $\dim(W) = \dim(V)$, es decir a matrices *cuadradas*.

La fórmula para los elementos de A^{-1} se justifica como sigue: Recordemos que el *cofactor* del elemento A_{ij}^j se define como

$$\text{cof } A_{ij}^j := (-1)^{i+j} \times \det[\text{A sin si fila } j \text{ y columna } i].$$

Desarrollando $\det(A)$ por la fila j vemos que⁹

$$\det(A) = A_{ii}^j \text{ cof } A_{ii}^j,$$

luego

$$[AA^{-1}]_i^j = A_{ik}^j [A^{-1}]_k^k = A_{ik}^j \frac{\text{cof } A_{ik}^i}{\det A} = \delta_{ik}^j \equiv [I]_i^j,$$

ya que si $i \neq j$, $A_{ik}^j \text{ cof } A_{ik}^i = \det[\text{matriz con filas } j \text{ e } i \text{ iguales}] = 0$, mientras que si $i = j$, $A_{ii}^j \text{ cof } A_{ii}^i = \det(A)$ (desarrollado por la fila j).

1.3.3. Matrices notables

Dependiendo de sus propiedades, diversas matrices cuadradas reciben nombres especiales. Algunos de los mas usuales son:

Propiedad	Definición	Ejemplo	Notas
Real	$A^* = A$	$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$	
Simétrica	$A^T = A$	$\begin{pmatrix} i & 1 \\ 1 & 2 \end{pmatrix}$	
Antisimétrica	$A^T = -A$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	singular, $A_{\frac{i}{2}}^{\frac{i}{2}} \equiv 0$
Autoadjunta	$A^\dagger = A$	$\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$	$\det(A)$ real
Ortogonal	$A^{-1} = A^T$	$\begin{pmatrix} \cos \theta & \operatorname{sen} \theta \\ -\operatorname{sen} \theta & \cos \theta \end{pmatrix}$	$\det(A) = \pm 1$
Unitaria	$A^{-1} = A^\dagger$	$\frac{1}{2} \begin{pmatrix} 1 & i \\ -i & -1 \end{pmatrix}$	$ \det(A) = 1$
Diagonal	$A_{ii}^j = 0 \forall i \neq j$	$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$	
Idempotente	$A^2 = A$	$\frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$	$\det(A) = 1$
Nilpotente	$A^2 = 0$	$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$	singular

Ejercicio 1.3.2. Demuestre que:

1. $(AB^{-1}) = B^{-1}A^{-1}$.
2. $(AB^T) = B^T A^T$.
3. $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$.
4. $\det(AB) = \det(BA) = \det(A)\det(B)$.

⁹Mientras se utiliza el convenio de suma sobre índices repetidos, la aparición de índices *subrayados* significa que *no sumamos* sobre ellos; en la fórmula del texto, ponemos \underline{j} para indicar que no sumamos sobre j .

1.3.4. Operación por bloques

En los cálculos con matrices, muchas veces podemos simplificar las operaciones subdividiendo las matrices involucradas en *bloques* o submatrices. Por ejemplo, sean

$$\mathbf{A} := \begin{pmatrix} A_1^1 & A_2^1 & \vdots & A_3^1 & A_4^1 & A_5^1 \\ A_1^2 & A_2^2 & \vdots & A_3^2 & A_4^2 & A_5^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ A_1^3 & A_2^3 & \vdots & A_3^3 & A_4^3 & A_5^3 \\ A_1^4 & A_2^4 & \vdots & A_3^4 & A_4^4 & A_5^4 \\ A_1^5 & A_2^5 & \vdots & A_3^5 & A_4^5 & A_5^5 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1^1 & \mathbf{A}_2^1 \\ \mathbf{A}_1^2 & \mathbf{A}_2^2 \end{pmatrix},$$

$$\mathbf{B} := \begin{pmatrix} B_1^1 & B_2^1 & \vdots & B_3^1 & B_4^1 & B_5^1 \\ B_1^2 & B_2^2 & \vdots & B_3^2 & B_4^2 & B_5^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ B_1^3 & B_2^3 & \vdots & B_3^3 & B_4^3 & B_5^3 \\ B_1^4 & B_2^4 & \vdots & B_3^4 & B_4^4 & B_5^4 \\ B_1^5 & B_2^5 & \vdots & B_3^5 & B_4^5 & B_5^5 \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1^1 & \mathbf{B}_2^1 \\ \mathbf{B}_1^2 & \mathbf{B}_2^2 \end{pmatrix},$$

con

$$\mathbf{A}_1^1 = \begin{pmatrix} A_1^1 & A_2^1 \\ A_1^2 & A_2^2 \end{pmatrix}, \quad \mathbf{A}_2^1 = \begin{pmatrix} A_3^1 & A_4^1 & A_5^1 \\ A_3^2 & A_4^2 & A_5^2 \end{pmatrix}, \quad \text{etc...},$$

y similarmente para \mathbf{B}_i^j . Entonces podemos calcular el producto \mathbf{AB} como

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_1^1 & \mathbf{A}_2^1 \\ \mathbf{A}_1^2 & \mathbf{A}_2^2 \end{pmatrix} \begin{pmatrix} \mathbf{B}_1^1 & \mathbf{B}_2^1 \\ \mathbf{B}_1^2 & \mathbf{B}_2^2 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1^1 \mathbf{B}_1^1 + \mathbf{A}_2^1 \mathbf{B}_1^2 & \mathbf{A}_1^1 \mathbf{B}_2^1 + \mathbf{A}_2^1 \mathbf{B}_2^2 \\ \mathbf{A}_1^2 \mathbf{B}_1^1 + \mathbf{A}_2^2 \mathbf{B}_1^2 & \mathbf{A}_1^2 \mathbf{B}_2^1 + \mathbf{A}_2^2 \mathbf{B}_2^2 \end{pmatrix}.$$

Notemos que es necesario, para proceder de esta manera, que cada uno de los productos y sumas matriciales que aparezcan tenga sentido, es decir, los bloques en que subdividimos las matrices deben tener números de filas y columnas consistentes. En particular, los bloques que yacen sobre la diagonal principal deben ser siempre cuadrados.

Así como definimos una matriz *triangular* cuando todos los elementos por encima (o por debajo) de la diagonal principal eran nulos, diremos que una matriz es *triangular por bloques* si por ejemplo

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1^1 & 0 & \cdots & 0 \\ \mathbf{A}_1^2 & \mathbf{A}_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_1^n & \mathbf{A}_2^n & \cdots & \mathbf{A}_n^n \end{pmatrix}.$$

El cálculo del determinante se simplifica notablemente en este caso, ya que resulta

$$\det(\mathbf{A}) = \prod_{j=1}^n \det(\mathbf{A}_{\underline{j}}^{\underline{j}}).$$

1.4. Transformaciones de coordenadas

En Física suponemos que la descripción que hagamos de un problema, las leyes físicas que formulemos, etc., deben ser independientes de en qué sistema de coordenadas estamos trabajando. Por ejemplo, cuando decimos que en una onda electromagnética los vectores campo eléctrico y campo magnético son ortogonales, o cuando calculamos los niveles de energía del electrón de un átomo de hidrógeno, no esperamos que estos resultados cambien si cambiamos de sistema de coordenadas. Sin embargo, cuando tratamos con entes como vectores (p.ej. campos eléctrico y magnético en Electromagnetismo) u operadores (p.ej. el operador energía en Mecánica Cuántica) necesitamos, por una cuestión práctica, describirlos por sus componentes en algún sistema de coordenadas dado, es decir, trabajamos con su *representación* en algún sistema de coordenadas. Y cuando cambiamos a otro sistema de coordenadas, estas componentes cambian. Es importante, entonces, saber *cómo* cambian, de modo de poder distinguir si una propiedad dada, es propiedad de los vectores u operadores, o sólo de su representación en un dado sistema de coordenadas.

Comenzaremos estudiando qué ocurre con las componentes de un vector ante un cambio de base, y veremos luego qué pasa con los operadores.

1.4.1. Matriz de transformación

Sea V un EV, $\{\vec{e}_i\}_{i=1}^n$ una base de V , y $\vec{x} = x^i \vec{e}_i \in V$. Sea $\{\vec{e}'_j\}_{j=1}^n$ otra base de V . Entonces, por ser $\{\vec{e}_i\}$ base de V , para cada $\vec{e}'_j \exists! \gamma_j^i \neq 0$ tales que

$$\vec{e}'_j = \vec{e}_i \gamma_j^i. \quad (1.23)$$

Las componentes γ_j^i de los vectores \vec{e}'_j en la base $\{\vec{e}_i\}$ son los elementos de la llamada *matriz de transformación* $\boldsymbol{\gamma}$ ($n \times n$), que representa el *cambio de base* (o transformación de coordenadas)

$$\{\vec{e}_i\}_{i=1}^n \xrightarrow{\boldsymbol{\gamma}} \{\vec{e}'_j\}_{j=1}^n.$$

Si pensamos a los \vec{e}_i y los \vec{e}'_j como columnas de las matrices ($n \times n$) \mathbf{E} y \mathbf{E}' respectivamente, es decir poniendo

$$\mathbf{E} := \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}, \quad \mathbf{E}' := \begin{pmatrix} \vec{e}'_1 & \vec{e}'_2 & \dots & \vec{e}'_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix},$$

podemos poner (1.24) en notación matricial como

$$\mathbf{E}' = \mathbf{E}\boldsymbol{\gamma}. \quad (1.24)$$

Notemos que, por ser base de V , tanto $\{\vec{e}_i\}$ como $\{\vec{e}'_j\}$ son LI. Luego la matriz $\boldsymbol{\gamma}$ es *no singular*, y por lo tanto tiene inversa $\boldsymbol{\gamma}^{-1}$, la cual evidentemente representa el cambio de base inverso

$$\mathbf{E} = \mathbf{E}'\boldsymbol{\gamma}^{-1}.$$

1.4.2. Covarianza y contravarianza

Siguiendo la idea de que un dado vector \vec{x} tiene un significado físico propio, independiente de la base en que lo representemos, pediremos que su desarrollo en las bases $\{\vec{e}_i\}$ y $\{\vec{e}'_j\}$ representen el mismo vector, es decir

$$\vec{x} = x^i \vec{e}_i = x'^j \vec{e}'_j.$$

Substituyendo $\vec{e}'_j = \vec{e}_i \gamma^i_j$ e igualando componentes en la base $\{\vec{e}_i\}$, obtenemos $x^i = x'^j \gamma^i_j$, y aplicando la transformación inversa resulta $x'^j = [\boldsymbol{\gamma}^{-1}]^j_i x^i$. Ordenando las componentes de \vec{x} y \vec{x}' como matrices columna, podemos poner entonces en notación matricial

$$\mathbf{x}' = \boldsymbol{\gamma}^{-1} \mathbf{x}. \quad (1.25)$$

Comparando las ecuaciones (1.24) y (1.25) vemos que los vectores de una base y las componentes de un vector en esa base, transforman de maneras exactamente opuestas. Por convención, se llama *covariante* a todo objeto que, ante cambios de base, transforme como los vectores de la base. Todo objeto que transforme como las componentes de un vector se llama *contravariante*.

La notación que hemos adoptado para los índices es consistente con esta distinción: los *subíndices* corresponden a objetos covariantes, y los *superíndices* a objetos contravariantes.

1.4.3. Componentes de un operador

Veremos ahora que las componentes de un operador lineal en un par de bases dadas también poseen propiedades de covarianza y contravarianza bien definidas.

Sea V un EV, $\{\vec{e}_i\}_{i=1}^n \xrightarrow{\boldsymbol{\gamma}} \{\vec{e}'_j\}_{j=1}^n$ bases de V , $\boldsymbol{\gamma}$ la matriz de transformación entre ellas, y $\vec{x} = x^i \vec{e}_i = x'^j \vec{e}'_j \in V$.

Sea W otro EV, $\{\vec{f}_l\}_{l=1}^m \xrightarrow{\boldsymbol{\delta}} \{\vec{f}'_k\}_{k=1}^m$ bases de W , $\boldsymbol{\delta}$ la matriz de transformación entre ellas, y $\vec{y} = y^l \vec{f}_l = y'^k \vec{f}'_k \in W$.

Sea $\mathcal{A} : V \rightarrow W$ un operador lineal tal que $\vec{y} = \mathcal{A}\vec{x}$, de modo que $y^l = A^l_i x^i$ y $y'^k = A'^k_j x'^j$.

Las propiedades de transformación de las componentes de \vec{x} e \vec{y} nos permiten deducir las de las componentes de \mathcal{A} , dado que

$$y'^k = [\boldsymbol{\delta}^{-1}]^k_l y^l = [\boldsymbol{\delta}^{-1}]^k_l A^l_i x^i = [\boldsymbol{\delta}^{-1}]^k_l A^l_i \gamma^i_j x'^j \equiv A'^k_j x'^j \quad \forall \vec{x}.$$

Luego, igualando componentes

$$A'^k_j = [\boldsymbol{\delta}^{-1}]^k_l A^l_i \gamma^i_j,$$

o en notación matricial,

$$\mathbf{A}' = \boldsymbol{\delta}^{-1} \mathbf{A} \boldsymbol{\gamma}. \quad (1.26)$$

Vemos entonces que A^l_i es *contravariante* en el superíndice l y *covariante* en el subíndice i , consistentemente con la notación adoptada.

1.4.4. Transformaciones de semejanza

Un caso particular muy importante de transformación de coordenadas, es aquel donde $W = V$. En ese caso $\vec{f}_i \equiv \vec{e}_i$ y $\vec{f}'_k \equiv \vec{e}'_k$, $\delta = \gamma$; luego A es $n \times n$ y

$$A' = \gamma^{-1} A \gamma, \quad A = \gamma A' \gamma^{-1}. \quad (1.27)$$

Llamaremos *transformación de semejanza* a toda transformación lineal tal que $A' = S^{-1}AS$ para todo operador lineal representado por las matrices A o A' . Evidentemente, la matriz de transformación S es no singular.

La importancia de las transformaciones de semejanza reside en que muchas propiedades de las matrices y de las operaciones entre ellas son *invariantes* bajo estas transformaciones.

Ejemplo 1.4.1. $(\lambda A)' = S^{-1}(\lambda A)S = \lambda S^{-1}AS = \lambda A'$.

Ejemplo 1.4.2. $(AB)' = S^{-1}(AB)S = S^{-1}ASS^{-1}BS = A'B'$.

Ejercicio 1.4.1. Considere la transformación de semejanza $A' = S^{-1}AS$. Muestre que

1. $\det(A') = \det(A)$.
2. $\text{Tr}(A') = \text{Tr}(A)$.
3. $f(A') = S^{-1}f(A)S \quad \forall f$ desarrollable en serie de potencias.
4. A Hermitiana $\implies A'$ Hermitiana.
5. A unitaria $\implies A'$ unitaria.
6. A ortogonal $\implies A'$ ortogonal.
7. $AB = BA \implies A'B' = B'A'$.
8. $B = A^{-1} \implies B' = (A')^{-1}$.

1.5. Formas y espacio dual

Vamos a introducir ahora la importante noción de espacio dual de un EV.

1.5.1. Formas

Sea V un EV con cuerpo escalar \mathbb{C} .

Una *forma* (o 1-forma) es una aplicación $\phi : V \rightarrow \mathbb{C}$, es decir una aplicación que lleva elementos del espacio vectorial al cuerpo escalar,

$$V \ni \vec{a} \xrightarrow{\phi} \phi(\vec{a}) \in \mathbb{C}. \quad (1.28)$$

Análogamente, una *n-forma* es una aplicación $\phi : \underbrace{V \oplus V \oplus \cdots \oplus V}_{n \text{ veces}} \rightarrow \mathbb{C}$, es decir una aplicación que lleva *n*-tuplas de elementos del espacio vectorial al cuerpo escalar,

$$\underbrace{V \oplus V \oplus \cdots \oplus V}_{n \text{ veces}} \ni (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n) \xrightarrow{\phi} \phi(\vec{x}) \in \mathbb{C}. \quad (1.29)$$

Llamaremos *forma lineal* a cualquier forma que verifique

$$\phi(\lambda \vec{a} + \mu \vec{b}) \equiv \lambda \phi(\vec{a}) + \mu \phi(\vec{b}). \quad (1.30)$$

En general y salvo nota explícita en contrario, asumiremos que trabajamos con formas lineales.

1.5.2. Espacio dual

Dado un EV V con cuerpo escalar \mathbb{C} , definimos el *espacio dual* V^* de V como el espacio de todas las formas lineales definidas sobre V , es decir

$$V^* := \{\phi : V \rightarrow \mathbb{C} / \phi \text{ lineal}\}. \quad (1.31)$$

El dual de un EV es también un EV, y tiene su misma dimensión. Por ello, en adelante adoptaremos la notación $\tilde{\phi}$ para sus elementos.

1.5.3. Componentes de una forma

Veremos que las formas lineales, igual que los vectores, pueden representarse adecuadamente dando sus componentes en una base.

Sea $\tilde{\phi} \in V^*$, $\{\vec{e}_i\}_{i=1}^n$ base de V , $\vec{x} = x^i \vec{e}_i \in V$. Por linealidad, $\tilde{\phi}(\vec{x}) = \tilde{\phi}(\vec{e}_i)x^i$. De acuerdo a la idea general de que la acción de una aplicación lineal sobre un EV queda completamente especificada si damos su acción sobre cada vector de una base, definiremos las *componentes* de la forma $\tilde{\phi}$ como

$$\phi_i := \tilde{\phi}(\vec{e}_i), \quad (1.32)$$

y ahora podemos poner

$$\tilde{\phi}(\vec{x}) = \phi_i x^i. \quad (1.33)$$

Sin embargo, nos gustaría completar la analogía con lo realizado para vectores de un EV respecto a su desarrollo en una base dada, definiendo un desarrollo análogo para una forma. Para ello necesitamos definir una base de V^* , lo que podemos conseguir definiendo las formas \tilde{e}^i tales que

$$\tilde{e}^i(\vec{e}_j) = \delta_j^i. \quad (1.34)$$

Ahora podemos poner

$$\tilde{\phi} = \phi_i \tilde{e}^i \quad (1.35)$$

y consistentemente tendremos

$$\tilde{\phi}(\vec{x}) = \phi_i \tilde{e}^i(x^j \vec{e}_j) = \phi_i \tilde{e}^i(\vec{e}_j) x^j = \phi_i \delta_j^i x^j = \phi_i x^i.$$

La base $\{\tilde{e}^i\}_{i=1}^n$ de V^* así construida se conoce como la *base dual* de $\{\vec{e}_i\}_{i=1}^n$.

Empleando la notación matricial, vemos que podemos ordenar las componentes ϕ_i de $\tilde{\phi}$ como una *matriz fila* $1 \times n$

$$\boldsymbol{\phi} = (\phi_1, \dots, \phi_n),$$

de modo que un uso consistente de la notación nos permite expresar $\tilde{\phi}(\vec{x})$ como el producto matricial

$$\tilde{\phi}(\vec{x}) = \boldsymbol{\phi} \mathbf{x} = (\phi_1, \dots, \phi_n) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

1.5.4. Transformaciones de coordenadas en V^*

Las formas, al igual que los vectores y operadores, tienen “existencia” independiente de la base en la cual las representemos, pero sus componentes cambiarán si cambiamos de una base a otra. Veremos ahora cómo una transformación de coordenadas en V induce una transformación relacionada en su dual V^* .

Sea un vector $\vec{x} = x^i \vec{e}_i = x'^j \vec{e}'_j \in V$, con $\vec{e}'_j = \vec{e}_i \gamma^i_j$, de modo que $x'^j = [\boldsymbol{\gamma}^{-1}]^j_i x^i$; y sea $\tilde{\phi}(\vec{x}) = \phi_i x^i = \phi'_j x'^j$, con $\tilde{\phi} \in V^*$. Por linealidad, $\tilde{\phi}(\vec{x}) = \tilde{\phi}(x'^j \vec{e}'_j) = \tilde{\phi}(\vec{e}'_j) x'^j = \phi'_j x'^j = \phi'_j [\boldsymbol{\gamma}^{-1}]^j_i x^i = \phi_i x^i$; como esta relación debe valer $\forall \vec{x}$, podemos deducir que $\phi'_j [\boldsymbol{\gamma}^{-1}]^j_i = \phi_i$, es decir

$$\phi'_j = \phi_i \gamma^i_j, \quad (1.36)$$

o en notación matricial

$$\boldsymbol{\phi}' = \boldsymbol{\phi} \boldsymbol{\gamma}. \quad (1.37)$$

La ley de transformación para la base $\{\tilde{e}^i\}_{i=1}^n$ de V^* puede deducirse ahora fácilmente del hecho de que

$$\tilde{\phi} = \phi_i \tilde{e}^i = \phi'_j \tilde{e}'^j,$$

y resulta ser

$$\tilde{e}'^j = [\boldsymbol{\gamma}^{-1}]^j_i \tilde{e}^i. \quad (1.38)$$

Vemos entonces que las componentes de una forma son *covariantes* y la base de V^* es *contravariante*, exactamente al revés que en V .

1.6. Producto interno, métrica y norma

Probablemente nos hayamos habituado a la definición geométrica de la norma de un vector como su “largo”, o del producto interno entre dos vectores como “el producto de las normas por el coseno del ángulo comprendido”, y en realidad así fue como estos conceptos surgieron históricamente. Sin embargo, estas definiciones son difícilmente extensibles a espacios menos “geométricos” como $\{p_N(x)\}$ o $C^\infty[\mathbb{R}]$. En esta Sección veremos cómo introducir una noción mucho más general (aunque más abstracta) de producto interno y norma, que sin embargo se reducirá a la ya conocida para espacios “geométricos”. Como bonificación, obtendremos una forma de definir un mapa 1 a 1 y sobre entre V y V^* , es decir que podremos poner en correspondencia cada vector con una forma y viceversa.

1.6.1. Producto interno

El *producto interno* o *producto escalar* entre dos vectores \vec{a} y \vec{b} de un mismo espacio vectorial V se define como una 2-forma $\Phi : V \oplus V \rightarrow \mathbb{C}$, es decir

$$V \oplus V \ni (\vec{a}, \vec{b}) \xrightarrow{\Phi} \Phi(\vec{a}, \vec{b}) \in \mathbb{C} \quad (1.39)$$

que satisface

1. $\Phi(\vec{a}, \vec{b}) = \Phi^*(\vec{b}, \vec{a})$;
2. $\Phi(\vec{a}, \lambda\vec{b}) = \lambda\Phi(\vec{a}, \vec{b})$; $\Phi(\lambda\vec{a}, \vec{b}) = \lambda^*\Phi(\vec{a}, \vec{b})$;
3. $\Phi(\vec{a}, \vec{a}) \geq 0$ (y obviamente debe ser real); $\Phi(\vec{a}, \vec{a}) = 0 \implies \vec{a} = \vec{0}$.

Cualquier 2-forma que cumpla estas condiciones será un “buen” producto interno sobre V .

1.6.2. Métrica

Consideremos ahora una base $\{\vec{e}_i\}_{i=1}^n$ de V y dos vectores $\vec{x} = x^i \vec{e}_i$ e $\vec{y} = y^i \vec{e}_i$. Por las propiedades de linealidad del producto interno, tendremos $\Phi(\vec{x}, \vec{y}) = x^{*i}\Phi(\vec{e}_i, \vec{e}_j)y^j$. De acuerdo a la idea de que el efecto de una aplicación queda completamente especificado dando su efecto sobre los vectores de una base, definimos

$$g_{ij} := \Phi(\vec{e}_i, \vec{e}_j). \quad (1.40)$$

Los elementos g_{ij} se conocen como *elementos de la métrica*, y podemos ordenarlos en una matriz

$$\mathbf{g} := [g_{ij}] \quad (1.41)$$

que se conoce como la *métrica*. Las propiedades del producto interno imponen que esta matriz $n \times n$ sea Hermitiana (simétrica si es real) y definida positiva.¹⁰ Dada una métrica podemos calcular el producto interno de dos vectores

¹⁰En las teorías relativistas se utilizan métricas que no son definidas positivas, por ejemplo la métrica de Minkowski de la Relatividad Especial, que es $g_{ij} = \text{diag}(1, 1, 1, -1)$.

cualesquiera a partir simplemente de sus componentes, de acuerdo a

$$\vec{x} \cdot \vec{y} := \Phi(\vec{x}, \vec{y}) = x^{*i} g_{ij} y^j, \quad (1.42)$$

o en notación matricial

$$\vec{x} \cdot \vec{y} := \mathbf{x}^\dagger \mathbf{g} \mathbf{y}. \quad (1.43)$$

1.6.3. Norma

Una vez definido un producto interno (es decir, dada una métrica) podemos usarlo para definir la *norma* de un vector como

$$|\vec{x}| := \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x^{*i} g_{ij} x^j}. \quad (1.44)$$

Las propiedades del producto interno garantizan que esta norma estará bien definida, aún para vectores complejos, y que sólo será nula para el vector nulo.

Es más, la norma así definida es una función *monótonamente creciente* de las normas (complejas) de cada componente de un vector. Esto permite asociarla a nuestra noción “intuitiva” de la norma de un vector como una medida de su “distancia al origen”, y permite también definir la *distancia* entre dos vectores como

$$\text{dist}(\vec{x}, \vec{y}) := |\vec{x} - \vec{y}|. \quad (1.45)$$

1.6.4. De V a V^* y vuelta con el producto interno

Consideremos un espacio vectorial V y su dual V^* , respectivamente con bases $\{\vec{e}_i\}_{i=1}^n$ y $\{\vec{e}^i\}_{i=1}^n$, y un dado producto interno Φ que podemos especificar dando la métrica \mathbf{g} . Recordemos que Φ es una 2-forma, es decir una aplicación que toma dos vectores de V y devuelve un escalar. Si tomamos un vector dado $\vec{x} \in V$, podemos pensar en la aplicación $\Phi(\vec{x}, \cdot)$ que obtenemos de Φ “llenando” su “primera casilla” con \vec{x} , y dejando la segunda “libre” para recibir otro vector cualquiera y devolver un escalar. Entonces vemos que $\Phi(\vec{x}, \cdot)$ es una aplicación (lineal) de V a los escalares, y por lo tanto es una forma, es decir un elemento de V , que denotaremos por

$$\tilde{\phi}_{\vec{x}} := \Phi(\vec{x}, \cdot). \quad (1.46)$$

Vemos así que Φ mapea cada elemento \vec{x} de V en un elemento $\tilde{\phi}_{\vec{x}}$ de V^* .

Esto nos permite replantear el concepto de producto interno como un “producto” entre un elemento de V y uno de V^* : para tomar el producto interno $\vec{x} \cdot \vec{y}$, primero obtenemos la forma $\tilde{\phi}_{\vec{x}}$ y luego se la aplicamos a \vec{y} .

Las componentes de $\tilde{\phi}_{\vec{x}}$ pueden deducirse del hecho que $\vec{x} \cdot \vec{y} = x^{*i} g_{ij} y^j$ y que para cualquier forma ϕ , $\tilde{\phi}(\vec{y}) = \phi_j y^j$, de lo que obtenemos

$$\phi_{\vec{x}j} := x^{*i} g_{ij}. \quad (1.47)$$

La métrica, entonces, lleva las componentes de un vector a las de la correspondiente forma.

Siguiendo con la misma idea, podemos ahora deducir cuál es la base $\{\tilde{e}^j\}_{j=1}^n$ de V^* en la cual $\tilde{\phi}_{\vec{x}}$ tiene las componentes dadas, ya que por un lado debemos tener

$$\tilde{\phi}_{\vec{x}} = \phi_{\vec{x}j}\tilde{e}^j, \quad (1.48)$$

de modo que

$$\tilde{\phi}_{\vec{x}}(\vec{y}) = \phi_{\vec{x}k}\tilde{e}^k(\vec{e}_j)y^j = x^{*i}g_{ik}\tilde{e}^k(\vec{e}_j)y^j,$$

y por otro lado

$$\vec{x} \cdot \vec{y} = x^{*i}g_{ij}y^j.$$

Comparando estas expresiones resulta

$$x^{*i}g_{ik}\tilde{e}^k(\vec{e}_j) = x^{*i}g_{ij},$$

de donde concluimos que

$$\tilde{e}^k(\vec{e}_j) = \delta_j^k,$$

es decir, $\{\tilde{e}^i\}_{i=1}^n$ es la *base dual* (o *adjunta*) de $\{\vec{e}_i\}_{i=1}^n$.

La expresión $\vec{x} \cdot \vec{y} := \mathbf{x}^\dagger \mathbf{g} \mathbf{y}$ para el producto interno de dos vectores nos provee una manera sencilla de completar estas correspondencias, interpretando el producto $\mathbf{x}^\dagger \mathbf{g} \mathbf{y}$ de dos formas alternativas:

- Teníamos las componentes de dos vectores como matrices columna \mathbf{x} e \mathbf{y} ; hallamos las componentes de la forma correspondiente a \mathbf{x} como la matriz fila $\mathbf{x}^\dagger \mathbf{g}$, y multiplicamos por \mathbf{y} para obtener el escalar $\mathbf{x}^\dagger \mathbf{g} \mathbf{y}$.
- Teníamos las componentes de dos formas como matrices fila \mathbf{x}^\dagger e \mathbf{y}^\dagger ; hallamos las componentes del vector correspondiente a \mathbf{y}^\dagger como la matriz columna $\mathbf{g} \mathbf{y}$, y multiplicamos por \mathbf{x}^\dagger para obtener el escalar $\mathbf{x}^\dagger \mathbf{g} \mathbf{y}$.

Notemos que con la última interpretación, ¡acabamos de definir el producto interno entre dos formas! El hecho de que la métrica, siendo definida positiva, sea no singular, y por lo tanto invertible, es lo que hace que la correspondencia entre vectores y formas así definida sea 1 a 1 y sobre, de modo que podemos poner

$$V \ni \vec{x} \xleftrightarrow{\Phi} \tilde{\phi}_{\vec{x}} \in V^*. \quad (1.49)$$

1.6.5. Correspondencia geométrica

Inicialmente, los vectores fueron introducidos como artificios para describir la posición de un punto en el espacio tridimensional. Tomando un punto arbitrario como origen, hacemos pasar por él tres rectas mutuamente ortogonales (en sentido geométrico), los *ejes Cartesianos*. El vector que da la posición de un punto será el segmento orientado del origen a ese punto, y sus *componentes Cartesianas* serán sus proyecciones geométricas sobre cada eje,

$$x^i := |\vec{x}| \cos \theta_i,$$

donde θ_i son los ángulos entre el vector y cada eje, y la norma $|\vec{x}|$ es su largo.

Si colocamos sobre cada eje vectores de largo 1, los *versores Cartesianos* \vec{e}_i , y definimos el producto escalar entre dos vectores \vec{x} e \vec{y} por

$$\vec{x} \cdot \vec{y} := |\vec{x}| |\vec{y}| \cos \theta,$$

donde θ es el ángulo comprendido entre ellos, las componentes Cartesianas de \vec{x} pueden escribirse $x^i = \vec{e}_i \cdot \vec{x}$, y la trigonometría nos dice que $|\vec{x}| = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2} = \sqrt{\vec{x} \cdot \vec{x}}$. En particular, el producto escalar entre dos vectores geométricamente ortogonales es nulo, de modo que

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij}.$$

Comparando estas expresiones para la norma y el producto escalar (y en particular el producto escalar entre los versores) con las obtenidas previamente, vemos que resultan consistentes si definimos la *métrica Cartesiana*

$$g_{ij} := \delta_{ij}$$

Por definición (herencia de su origen geométrico), un sistema de coordenadas con esta métrica se llama *ortonormal* (los versores son *ortogonales* y de *norma unidad*). Es de destacar que, siendo la métrica la matriz identidad $g = I$, en coordenadas ortonormales la distinción entre objetos covariantes y contravariantes desaparece, y el mapa de V a V^* es trivial.

1.7. Autovalores y autovectores

Consideremos el siguiente problema: dado un operador lineal $\mathcal{A} : V \rightarrow V$, ¿cuándo existen vectores en V que sean *invariantes* bajo \mathcal{A} ? Es decir, ¿bajo qué condiciones

$$\exists \vec{x} \in V / \mathcal{A}\vec{x} = \lambda \vec{x} \quad (1.50)$$

para algún $\lambda \in \mathbb{C}$?

La importancia de este problema reside en que cada vector \vec{x} que satisface (1.50) genera un subespacio de V *invariante* bajo \mathcal{A} . Si podemos dividir el espacio vectorial en una suma directa de subespacios invariantes, el estudio de las propiedades del operador se simplificará notablemente, ya que podemos limitarnos a estudiar su acción sobre los vectores de cada subespacio por separado.

Si $\mathcal{A}\vec{x} = \lambda \vec{x}$ con $\vec{x} \neq \vec{0}$, λ se llama un *autovalor* del operador \mathcal{A} , y \vec{x} se llama un *autovector* asociado (o perteneciente) a ese autovalor.

Ejemplo 1.7.1. Sea $\mathcal{R} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ el operador que rota en un ángulo θ alrededor del eje z ; entonces cualquier vector $\vec{a} = a\hat{z}$ es un autovector de \mathcal{R} con autovalor 1: $\mathcal{R}\vec{a} = \vec{a}$.

Ejemplo 1.7.2. Sea $\partial_x : \mathcal{C}^\infty[\mathbb{R}] \rightarrow \mathcal{C}^\infty[\mathbb{R}]$; entonces e^{ax} son autovectores (llamados en este caso *autofunciones*) con autovalor a : $\partial_x e^{ax} = ae^{ax}$.

Notemos que si \vec{x} es un autovector del operador \mathcal{A} con autovalor λ , cualquier múltiplo escalar de \vec{x} también lo es. Es más, si $\vec{x}_1, \dots, \vec{x}_r$ son todos autovectores de \mathcal{A} con autovalor λ , cualquier combinación lineal de ellos también lo es.

1.7.1. Autovectores a derecha

Sea $\mathcal{A} : V \rightarrow V$, y sea A su representación matricial en una base dada $\{\vec{e}_i\}$. En notación matricial, el problema (1.50) puede escribirse

$$Av = \lambda v \implies (A - \lambda I)v = 0. \quad (1.51)$$

Esta última ecuación matricial es un sistema homogéneo de ecuaciones lineales para las componentes del vector \vec{v} , y tendrá solución no trivial $\vec{v} \neq \vec{0}$ si y solo si la matriz $(A - \lambda I)$ es no invertible, o sea singular. Debemos exigir entonces que

$$\det(A - \lambda I) = 0. \quad (1.52)$$

Definimos el *polinomio característico* del operador \mathcal{A} como

$$\varphi_A(\lambda) := \det(\lambda I - A) = \sum_{j=0}^n \alpha_j \lambda^{n-j}, \quad (1.53)$$

donde $n = \dim(V)$. Este es evidentemente un polinomio en λ de grado n igual a la dimensión del espacio, que es la dimensión de la matriz A . Sus coeficientes α_j se llaman *invariantes algebraicos* (o *invariantes escalares*) del operador \mathcal{A} , por motivos que pronto se harán evidentes.

De (1.52) y (1.53) vemos que si λ es un autovalor, el polinomio característico satisface la llamada *ecuación secular*

$$\varphi_A(\lambda) = 0.$$

El Teorema Fundamental del Álgebra nos garantiza que *siempre* existen n raíces complejas $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ de esta ecuación, de modo que el polinomio característico puede factorizarse, como cualquier otro polinomio, en la forma

$$\varphi_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n), \quad \lambda_1, \dots, \lambda_n \in \mathbb{C}.$$

Estas raíces no son necesariamente todas distintas (pese a nuestra notación), y puede que ni siquiera estén en el cuerpo escalar asociado a nuestro espacio vectorial.

Ejemplo 1.7.3. Consideremos el operador $R_z : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ que efectúa una rotación de ángulo θ en sentido positivo alrededor del eje z . En la base Cartesiana usual, su representación matricial es

$$R_z = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Por lo tanto

$$\varphi_{R_z}(\lambda) = (\lambda^2 - 2 \cos(\theta)\lambda + 1)(\lambda - 1)$$

y los autovalores son $e^{\pm i\theta}$ y 1.

Para λ igual a cada una de las raíces λ_i de la ecuación secular, la ecuación (1.51) tendrá (al menos) una solución no trivial $v_i \neq 0$, es decir

$$\mathbf{A}v_i = \lambda_i v_i. \quad (1.54)$$

Por lo tanto, las raíces $\lambda_1, \dots, \lambda_n$ de la ecuación secular serán los autovalores del operador \mathbf{A} , y las correspondientes soluciones no triviales v_i de (1.54) serán las representaciones matriciales (como matrices columna), en la base $\{\vec{e}_i\}$, de los autovectores \vec{v}_i de \mathbf{A} asociados al correspondiente autovalor λ_i . El conjunto de autovalores de un operador se llama su *espectro*.

Vamos ahora a enunciar un primer resultado general, para el caso relativamente frecuente en que los autovalores son todos distintos.

Teorema 1.7.1. *Si $\wp_{\mathbf{A}}(\lambda)$ tiene n raíces λ_i todas distintas, con $n = \dim(V)$, entonces los autovectores asociados \vec{v}_i forman una base¹¹ del espacio vectorial V .*

Demostración. Demostraremos la independencia lineal de $\{\vec{v}_i\}_{i=1}^n$ por inducción. Por simplicidad suspenderemos momentáneamente el convenio de suma sobre índices repetidos.

- $\{\vec{v}_1\}$ LI es trivial: $a^1 \vec{v}_1 = \vec{0} \implies a^1 = 0$, ya que suponemos $\vec{v}_1 \neq \vec{0}$.
- Supongamos $\{\vec{v}_i\}_{i=1}^{k-1}$ LI (para un dado $k \leq n$), y sean $a^i \in \mathbb{C}$ tales que $\sum_{i=1}^k a^i \vec{v}_i = \vec{0}$. Luego

$$\mathbf{A} \left(\sum_{i=1}^k a^i \vec{v}_i \right) = \sum_{i=1}^k a^i \mathbf{A}(\vec{v}_i) = \sum_{i=1}^k a^i \lambda_i \vec{v}_i = \vec{0}.$$

Por otro lado

$$\sum_{i=1}^k a^i \vec{v}_i = \vec{0} \implies \lambda_k \sum_{i=1}^k a^i \vec{v}_i = \vec{0}.$$

Restando miembro a miembro

$$\sum_{i=1}^k a^i \lambda_i \vec{v}_i - \lambda_k \sum_{i=1}^k a^i \vec{v}_i = \sum_{i=1}^{k-1} a^i (\lambda_i - \lambda_k) \vec{v}_i = \vec{0}.$$

Pero $\{\vec{v}_i\}_{i=1}^{k-1}$ LI $\implies a^i (\lambda_i - \lambda_k) = 0$, $i = 1, \dots, k-1$. Luego

$$\sum_{i=1}^k a^i \vec{v}_i = a^k \vec{v}_k = \vec{0} \implies a^k = 0, \text{ ya que suponemos } \vec{v}_k \neq \vec{0}.$$

Por lo tanto hemos probado que $a^i = 0$, $i = 1, \dots, k$, es decir

$$\{\vec{v}_i\}_{i=1}^{k-1} \text{ LI} \implies \{\vec{v}_i\}_{i=1}^k \text{ LI.}$$

¹¹No necesariamente ortonormal.

□

Notemos que la validez de este Teorema *no* está restringida a espacios de dimensión finita.¹²

1.7.2. Autovectores a izquierda

Consideremos un operador $\mathcal{A} : V \rightarrow V$ tal que $\mathcal{A}\vec{v}_i = \lambda_i \vec{v}_i$, con n autovalores λ_i [$n = \dim(V)$]. Podemos expresar esta relación matricialmente como

$$\mathbf{Av}_i = \lambda_i \mathbf{v}_i, \quad i = 1, \dots, n. \quad (1.55)$$

Busquemos ahora autovalores y autovectores del operador adjunto \mathcal{A}^\dagger , cuya representación matricial es la matriz adjunta \mathbf{A}^\dagger ; es decir busquemos $\mu \in \mathbb{C}$ y $\vec{u} \in V$ tales que $\mathcal{A}^\dagger \vec{u} = \mu \vec{u}$, o matricialmente

$$\mathbf{A}^\dagger \mathbf{u} = \mu \mathbf{u}, \quad \mathbf{u} \neq 0. \quad (1.56)$$

Como antes, esta ecuación puede ponerse como $(\mathbf{A}^\dagger - \mu \mathbf{I})\mathbf{u} = 0$, y la existencia de solución no trivial requiere $\det(\mu \mathbf{I} - \mathbf{A}^\dagger) = 0$. Pero

$$\det(\mu \mathbf{I} - \mathbf{A}^\dagger) = \det[(\mu^* \mathbf{I} - \mathbf{A})^\dagger] = [\det(\mu^* \mathbf{I} - \mathbf{A})]^* = \left(\sum_{j=0}^n \alpha_j \mu^{*n-j} \right)^* = (\varphi_{\mathbf{A}}(\mu^*))^*,$$

de modo que

$$\det(\mu \mathbf{I} - \mathbf{A}^\dagger) = 0 \implies \varphi_{\mathbf{A}}(\mu^*) = 0.$$

Luego μ^* debe ser igual a alguna de las raíces λ_i de $\varphi_{\mathbf{A}}(\lambda)$, es decir $\mu = \lambda_i^*$. De lo anterior podemos concluir que los autovalores μ_i de \mathbf{A}^\dagger son los conjugados de los autovalores de \mathbf{A} ,

$$\mu_i \equiv \lambda_i^*.$$

Las matrices columna \mathbf{u}_i que representan a los correspondientes autovectores \vec{u}_i son soluciones de

$$\mathbf{A}^\dagger \mathbf{u}_i = \lambda_i^* \mathbf{u}_i \quad (1.57)$$

Si los autovalores λ_i son todos distintos, los μ_i también resultan todos diferentes entre sí, y puede demostrarse que los autovectores a izquierda \vec{u}_i forman base de V , de la misma forma que se hizo para los autovectores a derecha \vec{v}_i .

El nombre de *autovectores a izquierda* se debe a que las componentes de \vec{u}_i , en notación matricial, satisfacen la ecuación

$$\mathbf{u}_i^\dagger \mathbf{A} = \lambda_i \mathbf{u}_i^\dagger, \quad (1.58)$$

obtenida de la anterior tomando el adjunto Hermitiano de ambos miembros, y poniendo

$$\mathbf{u}_i^\dagger := (\mathbf{u}_i)^\dagger.$$

¹²Para espacios de dimensión infinita nuestra demostración sólo prueba la independencia lineal de los autovectores; el concepto de base requiere restricciones adicionales sobre los vectores “admisibles” como parte del espacio.

Notemos que esto nos permite interpretar al operador \mathcal{A} , representado por la matriz A , como actuando “hacia adelante” sobre elementos de V representados como matrices columna, o actuando “hacia atrás” sobre elementos de V^* representados como matrices fila.

1.7.3. Diagonalización de un operador

Sea un operador $\mathcal{A} : V \rightarrow V$, y sea A su representación matricial en una dada base $\{\vec{e}_i\}_{i=1}^n$ de V ; sean $Av_i = \lambda_i v_i$, $A^\dagger u_i = \lambda_i^* u_i$, con n autovalores λ_i todos distintos [$n = \dim(V)$]. Entonces

$$\begin{aligned} Av_i &= \lambda_i v_i & \Rightarrow & & u^{j\dagger} Av_i &= \lambda_i u^{j\dagger} v_i \\ u^{j\dagger} A &= \lambda_j u^{j\dagger} & \Rightarrow & & u^{j\dagger} Av_i &= \lambda_j u^{j\dagger} v_i \end{aligned}$$

Como $\lambda_i \neq \lambda_j$ si $i \neq j$, entonces comparando las dos expresiones de la derecha vemos que $u^{j\dagger} v_i = 0$ si $i \neq j$. Por otro lado, podemos multiplicar tanto a $u^{j\dagger}$ como a v_i por escalares arbitrarios sin alterar nada de lo que antecede, es decir, siempre podemos *normalizarlos* de modo que

$$\lambda_i \neq \lambda_j \text{ si } i \neq j \implies u^{j\dagger} v_i = \delta_i^j. \quad (1.59)$$

Ahora bien, la matriz fila $u^{j\dagger}$ puede pensarse como la representación matricial de un elemento \tilde{u}^j de V^* , igual que la matriz columna v_i es la representación matricial del elemento \vec{v}_i de V . En estos términos, la relación anterior se escribe

$$\lambda_i \neq \lambda_j \text{ si } i \neq j \implies \tilde{u}^j(\vec{v}_i) = \delta_i^j. \quad (1.60)$$

Esto quiere decir que las bases $\{\vec{v}_i\}$ de V y $\{\tilde{u}^j\}$ de V^* son *bases adjuntas*. Las correspondientes matrices columna v_i y fila $u^{j\dagger}$ dan sus componentes en la base $\{\vec{e}_i\}$ de V .

Si ahora queremos interpretar a $u^{j\dagger} v_i$ como el producto escalar de los autovectores \vec{u}_j y \vec{v}_i , debemos pensar que primero obtenemos de \vec{u}_j la matriz columna u_j de sus componentes en la base $\{\vec{e}_i\}$, luego tomamos su adjunta Hermitiana $u^{j\dagger} = (u_j)^\dagger$, y finalmente decimos que ésta es la matriz fila que representa al correspondiente elemento \tilde{u}^j de V^* . La consistencia con lo expuesto en la Sección 1.6 nos muestra entonces que debemos considerar a $\{\vec{e}_i\}$ una base ortonormal de V ; es decir, debemos asumir que la métrica, en la base $\{\vec{e}_i\}$, es la identidad.¹³ Con esta consideración, podremos entonces poner

$$\vec{u}_j \cdot \vec{v}_i = \delta_{ji} \quad (1.61)$$

y decir que los autovectores a izquierda y a derecha son dos a dos ortogonales. Nótese que no pedimos que $\vec{v}_j \cdot \vec{v}_i = \delta_{ji}$ ni que $\vec{u}_j \cdot \vec{u}_i = \delta_{ji}$, y en general *no ocurre así*.

De lo anterior se desprende que $u^{j\dagger} Av_i = u^{j\dagger} v_i \lambda_i = \delta_i^j \lambda_i$ son los elementos de la matriz diagonal $\text{diag}(\lambda_1, \dots, \lambda_n)$. Sea U la matriz $n \times n$ que tiene a las

¹³Por cierto, la métrica en la base de los $\{\vec{e}_i\}$ en general *no* será la identidad.

matrices fila $u^{j\dagger}$ por filas, y sea V la matriz $n \times n$ que tiene a las matrices columna v_i por columnas, es decir

$$U := \begin{pmatrix} u^{1\dagger} & \longrightarrow \\ \vdots & \longrightarrow \\ u^{n\dagger} & \longrightarrow \end{pmatrix}, \quad V := \begin{pmatrix} v_1 & \cdots & v_n \\ \downarrow & & \downarrow \end{pmatrix}.$$

Luego

$$UV = \begin{pmatrix} u^{1\dagger}v_1 & \cdots & u^{1\dagger}v_n \\ \vdots & & \vdots \\ u^{n\dagger}v_1 & \cdots & u^{n\dagger}v_n \end{pmatrix} = I,$$

de modo que $= V^{-1}$, y

$$UAV = V^{-1}AV = \begin{pmatrix} u^{1\dagger}Av_1 & \cdots & u^{1\dagger}Av_n \\ \vdots & & \vdots \\ u^{n\dagger}Av_1 & \cdots & u^{n\dagger}Av_n \end{pmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Podemos resumir los resultados obtenidos hasta ahora en el siguiente

Teorema 1.7.2. *Sea $\mathcal{A} : V \rightarrow V$ un operador lineal sobre un espacio vectorial V de dimensión n , que tiene n autovalores diferentes $\lambda_1, \dots, \lambda_n$; sean \vec{v}_i y \vec{u}_j los autovectores a derecha y a izquierda (respectivamente) asociados a ellos; y sean \bar{u}^j los elementos de V^* asignados a \vec{u}_j a través de la métrica identidad. Entonces:*

1. $\{\vec{v}_i\}_{i=1}^n$ es base de V , $\{\bar{u}^j\}_{j=1}^n$ es base de V^* , y son bases adjuntas; los \vec{v}_i y los \vec{u}_j son ortonormales dos a dos.
2. Las representaciones matriciales v_i de \vec{v}_i son las columnas de una matriz V , y las representaciones matriciales $u^{j\dagger}$ de \bar{u}^j son las columnas de su inversa V^{-1} , tales que la transformación de semejanza representada por V diagonaliza A : $V^{-1}AV = \text{diag}(\lambda_1, \dots, \lambda_n)$.

1.7.4. Operadores Hermitianos

Una clase de operadores muy importantes en Física son aquellos llamados Hermitianos, es decir aquellos cuya representación matricial es una matriz Hermitiana $A^\dagger = A$. En particular, en Mecánica Cuántica los operadores Hermitianos representan *observables* físicos, sus autovalores son los valores “permitidos” de dichos observables, y sus autovectores (o autofunciones) representan los estados cuánticos correspondientes.

Supongamos que $\mathcal{A} : V \rightarrow V$ sea Hermitiano, A su representación matricial en una dada base $\{\vec{e}_i\}_{i=1}^n$ de V , y $Av_i = \lambda_i v_i$, $A^\dagger u_i = \lambda_i^* u_i$. Entonces

1. $A^\dagger = A \implies \det(\lambda I - A) = \det(\lambda I - A^\dagger) \forall \lambda$; pero, factorizando el polinomio característico, vemos que

$$\begin{aligned}\det(\lambda I - A) &= (\lambda - \lambda_1) \dots (\lambda - \lambda_n), \\ \det(\lambda I - A^\dagger) &= (\lambda - \lambda_1^*) \dots (\lambda - \lambda_n^*).\end{aligned}$$

Por lo tanto, podemos concluir que

$$\lambda_i \equiv \lambda_i^*,$$

es decir, los autovalores de un operador Hermitiano *son todos reales*.

2. Por el resultado precedente,

$$A^\dagger = A \implies \left\{ \begin{array}{l} Av_i = \lambda_i v_i \\ A^\dagger u_i = \lambda_i^* u_i \end{array} \right\} \implies u_i \equiv v_i,$$

es decir, los autovectores a derecha e izquierda serán idénticos, $\vec{u}_i \equiv \vec{v}_i$, y $u_j^\dagger \equiv (v_j)^\dagger$.

3. Asumiendo la métrica identidad en la base $\{\vec{e}_i\}_{i=1}^n$ de V , y normalizando los autovectores adecuadamente, si los autovalores λ_i son todos diferentes, entonces

$$\vec{v}_i \cdot \vec{v}_j = \delta_{ij}.$$

Con estos resultados, y los del Teorema 1.7.2, hemos demostrado el siguiente

Teorema 1.7.3. *Sea $A : V \rightarrow V$ un operador lineal Hermitiano sobre un espacio vectorial V de dimensión n , con autovalores $\lambda_1, \dots, \lambda_n$; sean \vec{v}_i y \vec{u}_i los autovectores a derecha y a izquierda (respectivamente) asociados a ellos; y sean \vec{u}^j los elementos de V^* asignados a \vec{u}_j a través de la métrica identidad. Entonces:*

1. $\lambda_i^* \equiv \lambda_i$, es decir $\lambda_i \in \mathbb{R}$, $i = 1, \dots, n$, y $\vec{u}_i \equiv \vec{v}_i$.
2. Si λ_i son todos distintos, $\{\vec{v}_i\}_{i=1}^n$ es base de V , $\{\vec{v}^i\}_{i=1}^n$ es base de V^* , y son bases adjuntas y ortonormales.
3. Si λ_i son todos distintos, las representaciones matriciales v_i de \vec{v}_i son las columnas de una matriz unitaria V , y la transformación de semejanza representada por V diagonaliza A : $V^{-1}AV = V^\dagger AV = \text{diag}(\lambda_1, \dots, \lambda_n)$.

1.7.5. Operadores Hermitianos con autovalores degenerados

Si los autovalores de un operador Hermitiano no son todos diferentes, nuestra demostración de que los autovectores forman base se desmorona. Veremos ahora que, aún en ese caso, podemos construir una base con los autovectores de un operador Hermitiano.

Sea $\mathcal{A} : V \rightarrow V$ Hermitiano, $\dim(V) = n$. El polinomio característico de \mathcal{A} siempre puede escribirse como

$$\varphi_{\mathcal{A}}(\lambda) = (\lambda - \lambda_1)^{q_1}(\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r}, \quad \sum_{i=1}^r q_i = n.$$

Notemos que ahora denotamos por $\lambda_1, \lambda_2, \dots, \lambda_r$ las raíces *distintas* de $\varphi_{\mathcal{A}}$. Para cada raíz λ_i , el número de veces q_i que aparece repetida se llama su *multiplicidad*. Si $q_i > 1$, el autovalor λ_i se llama *degenerado*.

Es fácil ver que $\mathcal{A}^\dagger = \mathcal{A} \implies \lambda_i^* \equiv \lambda_i \implies \lambda_i \in \mathbb{R}$, $i = 1, \dots, n$, como antes.

Para cada autovalor distinto λ_i , sabemos que $\det(\lambda_i I - \mathcal{A}) = 0$. Por lo tanto, existe *al menos una* solución \vec{v}_i no trivial de $\mathcal{A}\vec{v}_i = \lambda_i \vec{v}_i$ para cada $i = 1, \dots, r$, y el correspondiente autovector a izquierda satisface $\vec{u}_i = \vec{v}_i$ como antes.

Construiremos ahora un procedimiento iterativo para obtener una base de autovectores.

1. Consideraremos el autovalor λ_1 , y supongamos que hemos encontrado un autovector \vec{v}_1 asociado a él. Procedamos a normalizarlo, es decir, hagamos $v^{1\dagger} v_1 = 1$.
2. Construyamos una base ortonormal $\{\vec{e}_i\}_{i=1}^n$ con $e_1 := v_1$, $e^{i\dagger} e_j = \delta_j^i$. Para ello podemos tomar cualquier conjunto LI de n vectores que contenga a \vec{v}_1 y ortonormalizarlo por el procedimiento de Gram-Schmidt.
3. Construyamos las matrices $n \times n$

$$V := \begin{pmatrix} v_1 & e_2 & \dots & e_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}, \quad V^\dagger := \begin{pmatrix} v^{1\dagger} & \longrightarrow \\ e^{2\dagger} & \longrightarrow \\ \vdots & \\ e^{n\dagger} & \longrightarrow \end{pmatrix},$$

que resultan no singulares por ser $\{\vec{e}_i\}_{i=1}^n$ base. Tendremos $V^\dagger V = I$, luego V es unitaria, $V^{-1} = V^\dagger$.

4. Construyamos

$$V^\dagger A V = \begin{pmatrix} v^{1\dagger} A v_1 & v^{1\dagger} A e_2 & \dots & v^{1\dagger} A e_n \\ e^{2\dagger} A v_1 & e^{2\dagger} A e_2 & \dots & e^{2\dagger} A e_n \\ \vdots & \vdots & & \vdots \\ e^{n\dagger} A v_1 & e^{n\dagger} A e_2 & \dots & e^{n\dagger} A e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 v^{1\dagger} v_1 & \lambda_1 v^{1\dagger} e_2 & \dots & \lambda_1 v^{1\dagger} e_n \\ \lambda_1 e^{2\dagger} v_1 & A'^{22} & \dots & A'^{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 e^{n\dagger} v_1 & A'^{n2} & \dots & A'^{nn} \end{pmatrix},$$

donde hemos definido $A'^i_j := e^{i\dagger} A e_j$. Luego

$$V^\dagger A V = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix},$$

con A' la matriz $(n-1) \times (n-1)$ con elementos A'^i_j .

5. Notemos que

$$\det(\lambda I - A) = \det[V^{-1}(\lambda I - A)V] = \det(\lambda I - V^{-1}AV)$$

$$= \begin{vmatrix} \lambda - \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \lambda I' - A' & \\ 0 & & & \end{vmatrix} = (\lambda - \lambda_1) \det(\lambda I' - A'),$$

con I' la matriz identidad $(n-1) \times (n-1)$. Pero

$$\begin{aligned} \det(\lambda I - A) &= (\lambda - \lambda_1)^{q_1} (\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r}, \\ \therefore \det(\lambda I' - A') &= (\lambda - \lambda_1)^{q_1-1} (\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r}. \end{aligned}$$

6. Si $q_1 > 1$, $\det(\lambda I' - A') = 0$; por lo tanto existe al menos una solución no trivial v'_2 de $A'v'_2 = \lambda_1 v'_2$. Podemos repetir entonces todo el proceso anterior, construyendo una segunda matriz unitaria

$$V' = \left(\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & v'_2 & e'_3 & \cdots & e'_n \\ \vdots & \downarrow & \downarrow & & \downarrow \end{array} \right)$$

tal que

$$V'^{-1}V^{-1}AVV' = \left(\begin{array}{cc|c} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & \\ \hline 0 & & A'' \end{array} \right),$$

con A'' una matriz $(n-2) \times (n-2)$.

7. Repitiendo el proceso q_1 veces, terminamos por construir una matriz unitaria $S_1 := VV'V''\dots V^{(q_1-1)}$ tal que

$$S_1 AS_1 = \left(\begin{array}{cc|c} \lambda_1 & 0 & \\ \ddots & & 0 \\ 0 & \lambda_1 & \\ \hline 0 & & A^{(q_1)} \end{array} \right).$$

El mismo procedimiento puede ahora aplicarse, por turno, a cada uno de los restantes autovalores distintos. Por lo tanto, hemos demostrado el siguiente

Teorema 1.7.4. *Sea $A : V \rightarrow V$ un operador lineal Hermitiano sobre un espacio vectorial V de dimensión n , con*

$$\varphi_A(\lambda) = (\lambda - \lambda_1)^{q_1} (\lambda - \lambda_2)^{q_2} \dots (\lambda - \lambda_r)^{q_r}, \quad \sum_{i=1}^r q_i = n.$$

Entonces:

1. $\lambda_i \in \mathbb{R}$, $i = 1, \dots, r$, y cada autovalor diferente λ_i hay asociados q_i autovectores LI.
2. Los autovectores de \mathcal{A} forman una base ortonormal de V , por lo tanto \mathcal{A} es diagonalizable. En particular, $\mathbf{V} := (\begin{smallmatrix} v_1 & \cdots & v_n \\ \downarrow & \cdots & \downarrow \end{smallmatrix})$ es unitaria, y

$$\mathbf{V}^\dagger \mathbf{A} \mathbf{V} = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{q_1 \text{ veces}}, \underbrace{\lambda_2, \dots, \lambda_2}_{q_2 \text{ veces}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{q_r \text{ veces}}).$$

Notemos que a cada autovalor diferente λ_i hay asociado un *subespaci invariant* de \mathcal{A} que, *en este caso*, resulta ser $\ker(\lambda_i \mathbf{l} - \mathbf{A})$.

Si bien la demostración de este Teorema ha sido constructiva, el procedimiento seguido para construir la base donde la representación de \mathcal{A} es diagonal, es engorroso y nada conveniente en las aplicaciones. En la práctica, una vez que sabemos que a cada autovalor λ_i con multiplicidad q_i le corresponden q_i autovectores LI, sabemos que la solución general de

$$\mathbf{A} \mathbf{v}_i = \lambda_i \mathbf{v}_i$$

contendrá q_i parámetros libres. Haciendo sucesivamente cada uno de estos parámetros igual a 1 y los demás a cero, obtenemos q_i autovectores LI, que luego podemos ortonormalizar por el procedimiento de Gram-Schmidt. (Los autovectores correspondientes a autovalores diferentes serán automáticamente ortogonales entre sí).

1.7.6. Diagonalización simultánea de operadores Hermitianos

Ahora que ya sabemos que un operador Hermitiano siempre es diagonalizable, consideremos el siguiente problema: dados dos operadores lineales $\mathcal{A}, \mathcal{B} : V \rightarrow V$, ambos Hermitianos y por consiguiente diagonalizables *por separado*, ¿existirá alguna matriz \mathbf{V} no singular tal que

$$\mathcal{A}' = \mathbf{V}^{-1} \mathcal{A} \mathbf{V} \quad \text{y} \quad \mathcal{B}' = \mathbf{V}^{-1} \mathcal{B} \mathbf{V}$$

sean *ambos* diagonales?

Supongamos que tal matriz exista. Entonces, por ser diagonales,

$$\mathcal{A}' \mathcal{B}' = \mathcal{B}' \mathcal{A}',$$

y por consiguiente

$$\mathcal{A} \mathcal{B} = \mathcal{B} \mathcal{A},$$

ya que la comutatividad es propiedad de los *operadores* (véase el Ejercicio 1.4.1). Es decir, si \mathcal{A} y \mathcal{B} son diagonalizables *simultáneamente*, entonces comutan.

Supongamos ahora que \mathcal{A} y \mathcal{B} comutan.

1. Como A es Hermitiana, existe alguna matriz V no singular tal que $A' = V^{-1}AV$ es diagonal, digamos $A' = \text{diag}(a_1, \dots, a_n)$; luego $A'^i{}_j = a_i \delta^i_j$.
2. Sea $B' = V^{-1}BV$; entonces como $AB = BA \implies A'B' = B'A'$, tendremos $[A'B']^i{}_j \equiv [B'A']^i{}_j$, de donde

$$\left. \begin{aligned} [A'B']^i{}_j &= A'^i{}_k B'^k{}_j = a_i \delta^i_k B'^k{}_j = a_i B'^i{}_j \\ [B'A']^i{}_j &= B'^i{}_l A'^l{}_j = B'^i{}_l a_l \delta^l_j = a_i B'^i{}_j \end{aligned} \right\} \implies a_i B'^i{}_j = a_i B'^i{}_j.$$

3. Para todo i y j tales que $a_i \neq a_j$, la ecuación anterior implica $B'^i{}_j = 0$. Ordenemos los elementos de A' de modo que $a_1 = a_2 = \dots = a_{q_1} = \lambda_1$, el primer autovalor de A ($q_1 \leq n$), y $a_{q_1+1}, \dots, a_n \neq \lambda_1$. Entonces

$$A' = \begin{pmatrix} \lambda_1 I_{q_1} & 0 \\ 0 & A'_2 \end{pmatrix}, \quad B' = \begin{pmatrix} B'_1 & 0 \\ 0 & B'_2 \end{pmatrix},$$

con I_{q_1} la identidad $q_1 \times q_1$, $A'_2 = \text{diag}(a_{q_1+1}, \dots, a_n)$, B'_1 una matriz $q_1 \times q_1$, y B'_2 una matriz $(n - q_1) \times (n - q_1)$.

4. B Hermitiana $\implies B'$ Hermitiana $\implies \exists W_1$ unitaria, $q_1 \times q_1$, tal que $B''_1 := W_1^{-1} B'_1 W_1$ es *diagonal*. Definamos

$$W = \begin{pmatrix} W_1 & 0 \\ 0 & I_{n-q_1} \end{pmatrix} \implies W_1^{-1} = \begin{pmatrix} W_1^{-1} & 0 \\ 0 & I_{n-q_1} \end{pmatrix},$$

con I_{n-q_1} la identidad $(n - q_1) \times (n - q_1)$. Entonces (operando por bloques)

$$\begin{aligned} A'' &:= W^{-1} A' W = \begin{pmatrix} W_1^{-1} & 0 \\ 0 & I_{n-q_1} \end{pmatrix} \begin{pmatrix} \lambda_1 I_{q_1} & 0 \\ 0 & A'_2 \end{pmatrix} \begin{pmatrix} W_1 & 0 \\ 0 & I_{n-q_1} \end{pmatrix} \\ &= \begin{pmatrix} W_1^{-1} \lambda_1 I_{q_1} W_1 & 0 \\ 0 & I_{n-q_1} A'_2 I_{n-q_1} \end{pmatrix} = \begin{pmatrix} \lambda_1 I_{q_1} & 0 \\ 0 & A'_2 \end{pmatrix}, \end{aligned}$$

que es una matriz diagonal, y

$$\begin{aligned} B'' &:= W^{-1} B' W = \begin{pmatrix} W_1^{-1} & 0 \\ 0 & I_{n-q_1} \end{pmatrix} \begin{pmatrix} B'_1 & 0 \\ 0 & B'_2 \end{pmatrix} \begin{pmatrix} W_1 & 0 \\ 0 & I_{n-q_1} \end{pmatrix} \\ &= \begin{pmatrix} W_1^{-1} B'_1 W_1 & 0 \\ 0 & I_{n-q_1} B'_2 I_{n-q_1} \end{pmatrix} = \begin{pmatrix} B''_1 & 0 \\ 0 & B''_2 \end{pmatrix}, \end{aligned}$$

con B''_1 una matriz *diagonal* $q_1 \times q_1$, y $B''_2 = B'_2$.

5. Repitiendo este proceso para cada uno de los autovalores distintos de A , terminamos construyendo una transformación de semejanza que diagonaliza simultáneamente a A y B .

Hemos demostrado entonces el siguiente

Teorema 1.7.5. *Sean $A, B : V \rightarrow V$ Hermitianos. Entonces $\exists S$ t.q. $S^{-1}AS$ y $S^{-1}BS$ son ambas diagonales, si y sólo si A y B comutan.*

1.7.7. Operadores normales

Una clase más general de operadores, que incluye a los operadores Hermitianos como caso particular, son los operadores *normales*, que son aquellos que commutan con su adjunto Hermitiano: $\mathcal{A}\mathcal{A}^\dagger = \mathcal{A}^\dagger\mathcal{A}$.

Notemos que para toda matriz A podemos escribir

$$A = X + iY, \text{ con } X := \frac{1}{2}(A + A^\dagger), Y := \frac{1}{2i}(A - A^\dagger),$$

descomposición que guarda cierta semejanza con la de un número complejo en sus partes real e imaginaria. Es inmediato de la definición que tanto X como Y son matrices Hermitianas, por lo que podremos diagonalizarlas simultáneamente si comutan. Pero

$$XY - YX = \frac{1}{4i}[(A + A^\dagger)(A - A^\dagger) - (A - A^\dagger)(A + A^\dagger)] = \frac{1}{2i}[A^\dagger A - AA^\dagger],$$

de donde deducimos el siguiente

Teorema 1.7.6. *Todo operador lineal normal es diagonalizable.*

Debemos destacar que la conversa no es cierta: bién puede existir una transformación de semejanza que diagonalice A pero no diagonalice ni a X ni a Y *por separado*, como muestra el siguiente

Ejemplo 1.7.4. Consideremos la matriz $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Es evidente que podemos diagonalizarla, ya que sus dos autovalores (1 y 2) son diferentes. Sin embargo, no es normal: $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$, pero $\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}$.

1.8. Forma de Jordan

De lo expuesto en la Sección precedente queda claro que un operador será diagonalizable si todos sus autovalores son diferentes, o si es normal (este último caso incluye a los operadores Hermitianos). La diagonalizabilidad resulta una característica altamente deseable, ya que en la base en que su representación es una matriz diagonal, la acción de un operador resulta evidente.

Resulta natural, entonces, preguntarnos cuál es la clase más general de operadores de los que podremos decir, *a priori* que serán diagonalizables, y cuál será la forma más simple a la que podremos reducir la representación de un operador que no lo sea. Como veremos en lo que sigue, las respuestas a ambas preguntas están íntimamente relacionadas.

1.8.1. Ejemplos ilustrativos

En un sentido trivial, la respuesta a la pregunta de cuándo un operador podrá diagonalizarse es “cuando exista una base del espacio vectorial en la que su representación sea una matriz diagonal”. Sin embargo, si una matriz es diagonal, sus autovectores son simplemente los vectores de la base donde adopta

esa forma (y los autovalores son obviamente los elementos de la diagonal); y si los autovectores forman base, es evidente que la representación del operador en esa base será una matriz diagonal. Pero el que los autovectores formen base es independiente de la base en que expresemos el operador, así que por ahora podemos expresar el siguiente

Lema 1.8.1. *Un operador $\mathcal{A} : V \rightarrow V$ es diagonalizable sii sus autovectores forman base de V .*

Esto nos indica que, si queremos hallar operadores no diagonalizables, debemos buscar operadores cuyos autovectores *no formen base*. Demos algunos ejemplos:

Ejemplo 1.8.1. Consideremos un operador $\mathcal{A} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que en una dada base se representa por la matriz $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Su polinomio característico es $\varphi_A(\lambda) = (\lambda - 1)^2$, de modo que tiene un autovalor $\lambda_1 = 1$ con multiplicidad 2. Pero la ecuación de autovectores $(A - \lambda_1 I)v = 0$ tiene en este caso una sola solución LI, ya que $(A - \lambda_1 I) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; poniendo $v = \begin{pmatrix} x \\ y \end{pmatrix}$ tenemos que $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies y = 0$, de modo que $\ker(A - \lambda_1 I) = \text{span}\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$.

Dicho de otro modo, el subespacio de \mathbb{R}^2 desarrollado por los autovectores (uno, en este caso) es unidimensional. Sin embargo, cabe preguntarse qué ocurre con los vectores que *no* pertenecen a este subespacio. En particular, vemos que $(A - \lambda_1 I)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, es decir, los vectores del subespacio complementario de $\ker(A - \lambda_1 I)$ son llevados a $\ker(A - \lambda_1 I)$ por acción de $(A - \lambda_1 I)$. Una vez allí, una segunda aplicación de $(A - \lambda_1 I)$ los enviará al vector nulo, ya que en efecto $(A - \lambda_1 I)^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, de modo que $\ker[(A - \lambda_1 I)^2] = \mathbb{R}^2$, el espacio completo.

Notemos para terminar que la ecuación $(A - \lambda_1 I)^2 v = 0$ *sí* tiene dos soluciones LI, por ejemplo $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Ejemplo 1.8.2. Consideremos ahora un operador $\mathcal{A} : \mathbb{R}^4 \rightarrow \mathbb{R}^4$, que en alguna base se representa

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Su polinomio característico es $\varphi_A(\lambda) = (\lambda - 1)^2(\lambda - 2)^2$, de modo que tiene dos autovalores diferentes $\lambda_1 = 1$ y $\lambda_2 = 2$, ambos con multiplicidad 2. Tenemos

$$A - \lambda_1 I = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A - \lambda_2 I = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

luego

$$\ker(A - \lambda_1 I) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}, \quad \ker(A - \lambda_2 I) = \text{span} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Nuevamente vemos que los autovectores no forman base, pero

$$(A - \lambda_1 I) \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (A - \lambda_2 I) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Notando que

$$(A - \lambda_1 I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (A - \lambda_2 I)^2 = \begin{pmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

tenemos

$$\ker[(A - \lambda_1 I)^2] = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}, \quad \ker[(A - \lambda_2 I)^2] = \text{span} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

y vemos que son *subespacios invariantes* bajo \mathcal{A} .

1.8.2. Descomposición primaria

De los ejemplos precedentes podemos ver aparecer un patrón, que ahora vamos a formalizar.

Supongamos que V es suma directa de cierta colección de subespacios V_i , $i = 1, \dots, r$, y denotemos por \mathcal{P}_i el proyector sobre V_i . Es inmediato que

- $\mathcal{P}_i^2 \equiv \mathcal{P}_i$;
- $\mathcal{P}_i \mathcal{P}_j = 0$ si $i \neq j$;
- $\sum_{i=1}^r \mathcal{P}_i = \mathcal{I}$.

También es inmediato que $\forall \vec{x} \in V$,

- $\vec{x}_i := \mathcal{P}_i \vec{x} \in V_i$;
- $\vec{x} = \sum_{i=1}^r \vec{x}_i$.

Las restricciones de un operador $\mathcal{A} : V \rightarrow V$ a cada subespacio V_i las denotaremos por $\mathcal{A}_i := \mathcal{A}/V_i$, y las identificaremos con los correspondientes operadores $\mathcal{A}\mathcal{P}_i : V \rightarrow V$ (ver el comentario final de la Sección 1.2.1). De lo anterior se desprende que

$$\mathcal{A}\vec{x} \equiv \sum_{i=1}^r \mathcal{A}_i \vec{x}_i. \tag{1.62}$$

Si cada subespacio V_i es además invariante bajo \mathcal{A} , es decir si $\vec{x}_i \in V_i \implies \mathcal{A}\vec{x}_i \equiv \mathcal{A}_i \vec{x}_i \in V_i$, $i = 1, \dots, r$, diremos que \mathcal{A} es *suma directa* de sus restricciones \mathcal{A}_i a cada subespacio, y lo denominaremos

$$\mathcal{A} = \bigoplus_{i=1}^r \mathcal{A}_i. \tag{1.63}$$

Una consecuencia inmediata de lo anterior es que, si tenemos bases B_i de cada V_i y formamos una base $B = \{B_1, \dots, B_r\}$ de V colecciónándolas, entonces en la

base B la matriz que representa al operador \mathcal{A} toma la forma *diagonal por bloques*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_r \end{pmatrix}, \quad (1.64)$$

donde cada \mathbf{A}_i es una matriz $\dim(V_i) \times \dim(V_i)$, que podemos pensar como la representación matricial de \mathcal{A}_i en la base B_i .¹⁴

Supongamos ahora que $\mathcal{A} : V \rightarrow V$ es un operador lineal, con $\dim(V) = n$, y que $\varphi_{\mathcal{A}}(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{q_i}$, $\sum_{i=1}^r q_i = n$. Definimos el *subespacio característico* V_i asociado al autovalor λ_i como

$$V_i := \ker[(\mathbf{A} - \lambda_i I)^{q_i}] \quad (1.65)$$

(nótese que $\dim(V_i) = q_i$). Enunciaremos, sin demostración, el siguiente

Teorema 1.8.2 (Descomposición Primaria). *Sea $\mathcal{A} : V \rightarrow V$ un operador lineal, entonces:*

1. *V es suma directa de los subespacios característicos V_i de \mathcal{A} : $V = \bigoplus_{i=1}^r V_i$.*
2. *Cada V_i es invariante bajo \mathcal{A} , y por consiguiente $\mathcal{A} = \bigoplus_{i=1}^r \mathcal{A}/V_i$.*
3. *El polinomio característico de \mathcal{A}/V_i es $\varphi_{\mathcal{A}/V_i}(\lambda) = (\lambda - \lambda_i)^{q_i}$.*
4. *Cada \mathcal{A}_i puede descomponerse como $\mathcal{A}_i = \mathcal{D}_i + \mathcal{N}_i$, donde $\mathcal{D}_i := \lambda_i \mathcal{P}_i$, que es diagonalizable, y $\mathcal{N}_i := (\mathcal{A}_i - \lambda_i I)\mathcal{P}_i$, que es nilpotente de índice $\leq q_i$.*¹⁵

Es sencillo demostrar que todos los operadores \mathcal{D}_i y \mathcal{N}_i comutan entre sí.

Las primeras tres afirmaciones del Teorema anterior implican que, para todo operador lineal $\mathcal{A} : V \rightarrow V$, existirá una base de V donde su representación matricial adopte una forma diagonal por bloques como la mostrada en (1.64), y que cada bloque representará la restricción de \mathcal{A} al correspondiente subespacio invariante. La cuarta afirmación nos dice que, además, existirá una base donde

¹⁴Con la definición “habitual” de \mathcal{A}/V_i , tenemos $\mathcal{A}_i : V_i \rightarrow V_i$, y las representaciones matriciales de cada \mathcal{A}_i en la correspondiente base B_i de V_i son las matrices \mathbf{A}_i ; sin embargo, si identificamos (como lo hacemos aquí) a \mathcal{A}/V_i con $\mathcal{A}\mathcal{P}_i$, estos son operadores de V en V , y sus representaciones matriciales en la base B de V son

$$\tilde{\mathbf{A}}_1 = \begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \tilde{\mathbf{A}}_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \dots, \quad \tilde{\mathbf{A}}_r = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_r \end{pmatrix}.$$

La ventaja de este enfoque es que podemos escribir $\mathbf{A} = \sum_{i=1}^r \tilde{\mathbf{A}}_i$.

¹⁵Un operador \mathcal{N} se dice nilpotente de índice q si $\mathcal{N}^q = 0$ pero $\mathcal{N}^{q-1} \neq 0$.

cada bloque A_i adopte la forma triangular¹⁶

$$A_i = \begin{pmatrix} \lambda_1 & x & x & \cdots & x & x \\ 0 & \lambda_1 & x & \cdots & x & x \\ 0 & 0 & \lambda_1 & \cdots & x & x \\ \vdots & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_1 & x \\ 0 & 0 & 0 & \cdots & 0 & \lambda_1 \end{pmatrix}, \quad (1.66)$$

donde las “x” representan elementos posiblemente no nulos, que dan la representación matricial del nilpotente N_i .

1.8.3. Forma normal de Jordan

Vamos a analizar ahora en mas detalle la estructura de cada uno de los bloques de la descomposición primaria de un operador lineal. Para ello deberemos introducir primero algo de nomenclatura.

Un operador lineal $\mathcal{A} : V \rightarrow V$ se llama *nilcíclico* de orden n [= dim(V)] si existe una base $\{\vec{e}_1, \dots, \vec{e}_n\}$ de V tal que

$$\begin{aligned} \mathcal{A}\vec{e}_i &= \vec{e}_{i-1}, \quad i = n, n-1, \dots, 2, \\ \mathcal{A}\vec{e}_1 &= \vec{0}. \end{aligned}$$

Una tal base se llama *cíclica*, y en ella la representación matricial de \mathcal{A} adopta la forma

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Notemos que para un operador nilcíclico, $\varphi_A(\lambda) = \lambda^n$, y que su único autovector (en la base cíclica) es \vec{e}_1 , con autovalor cero. También se verifica que todo operador nilcíclico de orden n es un operador nilpotente de índice n .

Ejercicio 1.8.1. Demuestre las afirmaciones precedentes.

Enunciaremos ahora sin demostración el siguiente

Teorema 1.8.3. *Sea $\mathcal{N} : V \rightarrow V$ un operador nilpotente sobre un espacio V de dimensión finita n . Entonces existen subespacios V_i , $i = 1, \dots, r$, tales que*

1. *V es su suma directa: $V = \bigoplus_{i=1}^r V_i$;*

¹⁶Por convención se asume *triangular superior*.

2. las restricciones \mathcal{N}/V_i de \mathcal{N} a cada uno de ellos son operadores nilcíclicos de orden $\dim(V_i)$;

3. \mathcal{N} es suma directa de sus restricciones: $\mathcal{N} = \bigoplus_{i=1}^r \mathcal{N}/V_i$.

Esto significa que, dado un operador nilpotente \mathcal{N} sobre un espacio V , existirá una base de V donde su representación matricial adopte la forma diagonal por bloques

$$\mathbf{N} = \begin{pmatrix} \mathbf{N}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{N}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{N}_r \end{pmatrix}, \quad (1.67)$$

donde cada matriz \mathbf{N}_i representa un operador nilcíclico de orden $\dim(V_i)$ igual a la dimensión del subespacio V_i sobre el que actúa.

Aplicando el teorema precedente a la descomposición primaria de un operador $\mathcal{A} : V \rightarrow V$, vemos que siempre existirá alguna base de V donde cada bloque \mathbf{A}_i adopte la forma particularmente simple

$$\mathbf{A}_i = \begin{pmatrix} \mathbf{J}_{(\lambda_i)t_{i1}} & 0 & \cdots & 0 \\ 0 & \mathbf{J}_{(\lambda_i)t_{i2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}_{(\lambda_i)t_{ir_i}} \end{pmatrix}, \quad (1.68)$$

donde cada *bloque de Jordan* $\mathbf{J}_{(\lambda_i)t_{ij}}$ será una matriz $t_{ij} \times t_{ij}$ de la forma

$$\mathbf{J}_{(\lambda_i)t_{ij}} = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}, \quad (1.69)$$

con el autovalor λ_i en la diagonal principal, unos en la diagonal inmediata superior, y ceros en todas las demás posiciones. También se desprende de lo anterior que a *cada* bloque de Jordan $\mathbf{J}_{(\lambda_i)t_{ij}}$ habrá asociado un *único* autovector de \mathcal{A} de autovalor λ_i .

1.8.4. Reducción a la forma normal de Jordan

Enfrentaremos ahora el problema siguiente: dado un operador lineal $\mathcal{A} : V \rightarrow V$, y conocida su representación matricial \mathbf{A} en una dada base $\{\vec{e}_i\}_{i=1}^n$ de V , encontrar: (i) su forma normal de Jordan; (ii) la transformación de semejanza

que lleva \mathbf{A} a esa forma. Daremos un método constructivo para resolver este problema, que puede entenderse como un algoritmo.¹⁷

Descomposición primaria

Lo primero que debemos hacer es encontrar una base de V que implemente la descomposición primaria de \mathbf{A} . Para ello deberemos encontrar los autovalores de \mathbf{A} y los subespacios característicos asociados a cada uno.

1. Construimos el polinomio característico

$$\varphi_{\mathbf{A}}(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}),$$

resolvemos la ecuación secular

$$\varphi_{\mathbf{A}}(\lambda) = 0$$

para encontrar los autovalores λ_i con sus respectivas multiplicidades q_i , y factorizamos el polinomio característico en la forma

$$\varphi_{\mathbf{A}}(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{q_i}.$$

2. Para cada autovalor distinto λ_i , determinamos la matriz columna \mathbf{v}_i , solución general de la ecuación

$$(\mathbf{A} - \lambda_i \mathbf{I})^{q_i} \mathbf{v}_i = 0.$$

Esta solución contendrá q_i parámetros libres, así que de ella podremos obtener q_i columnas $\mathbf{v}_{i1}, \dots, \mathbf{v}_{iq_i}$ LI,¹⁸ que desarrollan el subespacio característico V_i de \mathbf{A} asociado al autovalor λ_i .

3. De manera automática (por construcción), todas las columnas obtenidas serán LI entre sí. Construimos entonces con las columnas \mathbf{v}_{il} la matriz

$$\mathbf{T} := \begin{pmatrix} \mathbf{v}_{11} & \cdots & \mathbf{v}_{rq_r} \\ \downarrow & & \downarrow \end{pmatrix},$$

y calculamos su inversa \mathbf{T}^{-1} . La transformación de semejanza definida por estas matrices llevará \mathbf{A} a la forma diagonal por bloques

$$\mathbf{A}' = \mathbf{T}^{-1} \mathbf{A} \mathbf{T} = \begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_r \end{pmatrix}.$$

¹⁷Este procedimiento *no es exactamente* el descripto en la clase teórica correspondiente, sino una versión algo más breve y, a mi entender, más clara. Toda crítica al respecto será particularmente bienvenida.

¹⁸Por ejemplo, poniendo sucesivamente cada parámetro libre igual a 1 y los demás a 0.

Hasta aquí, habremos realizado la *descomposición primaria* de \mathbf{A} .¹⁹ Debe notarse la fuerte similitud entre el procedimiento llevado a cabo aquí y el seguido usualmente para diagonalizar una matriz diagonalizable, siendo la principal diferencia que en lugar de trabajar con *autovectores*, soluciones de $(\mathbf{A} - \lambda_i I)\mathbf{v}_i = 0$, aquí trabajamos con las soluciones de $(\mathbf{A} - \lambda_i I)^{q_i}\mathbf{v}_i = 0$.

Reducción de cada bloque a la forma normal

Pasamos ahora a trabajar en la reducción de cada bloque \mathbf{A}_i a su forma normal de Jordan. Notemos primero que, al trabajar con cada bloque \mathbf{A}_i , ya no necesitamos trabajar en el espacio completo V , sino sólo en el correspondiente subespacio característico V_i . Esto quiere decir que si estamos trabajando con el bloque \mathbf{A}_i , nuestras filas y columnas sólo contendrán elementos distintos de cero en las posiciones correspondientes a ese bloque, es decir serán de la forma

$$\mathbf{v} = (\underbrace{0, \dots, 0}_{q_1 + \dots + q_{i-1} \text{ veces}}, \underbrace{v, \dots, v}_{q_i \text{ elementos no nulos}}, \underbrace{0, \dots, 0}_{q_{i+1} + \dots + q_r \text{ veces}})^T,$$

$$\mathbf{u}^\dagger = (\underbrace{0, \dots, 0}_{q_1 + \dots + q_{i-1} \text{ veces}}, \underbrace{u, \dots, u}_{q_i \text{ elementos no nulos}}, \underbrace{0, \dots, 0}_{q_{i+1} + \dots + q_r \text{ veces}}).$$

En la práctica, entonces, para cada bloque \mathbf{A}_i trabajaremos con columnas y filas de tamaño q_i (los únicos elementos posiblemente no nulos), y recién al final ordenaremos estos elementos en filas y columnas de tamaño n .

Como se verá a continuación, el procedimiento a seguir con cada bloque \mathbf{A}_i consiste en ir construyendo, por un método iterativo, las bases cíclicas de cada uno de los nilcíclicos contenidos en \mathbf{A}_i .

¹⁹Alternativamente, en el paso (2) podemos determinar también cada matriz fila $\mathbf{u}^{i\dagger}$, solución general de la ecuación

$$\mathbf{u}^{i\dagger}(\mathbf{A} - \lambda_i I)^{q_i} = 0,$$

que contendrá q_i parámetros libres, y obtener de ella q_i filas $\mathbf{u}^{i1\dagger}, \dots, \mathbf{u}^{iq_i\dagger}$ LI, que desarrollan el subespacio dual de V_i . Por construcción, las filas de un subespacio V_i y las columnas de cualquier otro subespacio *distinto* V_j , $i \neq j$, serán mutuamente ortogonales:

$$\mathbf{u}^{il\dagger} \mathbf{v}_{jk} = 0 \quad \forall l = 1, \dots, q_i, k = 1, \dots, q_j \quad \text{si } i \neq j.$$

Dentro de cada subespacio V_i , podemos aplicar a las correspondientes filas y columnas el procedimiento de Gram-Schmidt para convertirlas en mutuamente ortonormales (si ya no lo fueran). Obtenemos así un conjunto de n columnas \mathbf{v}_{il} y otro de n filas $\mathbf{u}^{il\dagger}$, con $i = 1, \dots, r$, $l = 1, \dots, q_i$, que son mutuamente *ortonormales*:

$$\mathbf{u}^{il\dagger} \mathbf{v}_{jk} = \delta_j^i \delta_k^l.$$

La matriz \mathbf{T} se construye entonces como en el paso (3), pero su inversa se obtiene simplemente poniendo

$$\mathbf{T}^{-1} := \begin{pmatrix} \mathbf{u}^{11\dagger} & \longrightarrow \\ \vdots & \longrightarrow \\ \mathbf{u}^{rq_r\dagger} & \longrightarrow \end{pmatrix}.$$

Cuál de las dos alternativas resulta más conveniente en la práctica, dependerá de la matriz \mathbf{A} .

4. Para cada bloque A_i debemos comenzar encontrando todos los autovectores LI correspondientes al autovalor λ_i . Para ello, determinamos la matriz columna v ($|tamaño q_i|$), solución general de la ecuación

$$(A_i - \lambda_i I)v = 0.$$

Esta solución contendrá $m_i \leq q_i$ parámetros libres, así que de ella podremos obtener m_i columnas v_1, \dots, v_{m_i} LI,²⁰ que desarrollan el subespacio $\ker(A_i - \lambda_i I)$.²¹ En este paso, es altamente conveniente ortonormalizar estas columnas (p.ej. por el procedimiento de Gram-Schmidt), y en lo que sigue *asumiremos que así se hizo*, tras lo cual construimos la columna

$$v = a_1 v_1 + \dots + a_{m_i} v_{m_i},$$

con a_1, \dots, a_{m_i} parámetros libres y v_1, \dots, v_{m_i} mutuamente ortonormales.

5. Buscamos la columna w_1 , solución general de la ecuación

$$(A_i - \lambda_i I)w_1 = v.$$

Supondremos primero que este sistema de ecuaciones *no impone ninguna condición* sobre los parámetros a_s para que exista solución. Entonces w_1 contendrá $2m_i$ parámetros libres,²² ya que como toda solución general de un sistema *inhomogéneo* lleva sumada la solución general del sistema *homogéneo*; de estos parámetros, m_i serán los a_s y los otros m_i serán nuevos, llamémoslos b_s .²³

Para seguir, buscamos una columna w_2 , solución general de

$$(A_i - \lambda_i I)w_2 = w_1,$$

e iteramos el procedimiento anterior. Otra vez aparecerán m_i nuevos parámetros, llamémoslos c_i . Es posible que a medida que avancemos algunos de los nuevos parámetros vayan quedando determinados en términos de otros o de los originales a_s , pero mientras no aparezca una condición sobre los parámetros *originales* a_s , seguimos iterando este paso.²⁴

²⁰Por ejemplo, poniendo sucesivamente cada parámetro libre igual a 1 y los demás a 0.

²¹Recordemos que $\dim(\ker(A_i - \lambda_i I)) = m_i$, el número de autovectores LI asociados a λ_i .

²²Obviamente este caso presupone $2m_i \leq q_i$.

²³Estos nuevos parámetros b_s simplemente aportan la solución general del sistema homogéneo, que en principio podríamos pensar que no queremos y que deberíamos descartar poniendo $b_s \equiv 0$; porque estaríamos buscando una preimagen de $\ker(A_i - \lambda_i I)$ bajo $(A_i - \lambda_i I)$ que “no contega nada” que ya estuviera en $\ker(A_i - \lambda_i I)$ (lo que llamaríamos una preimagen “propia”). Sin embargo, pese a que estamos construyendo una base del subespacio invariante correspondiente a A_i , generalmente ella no será ortogonal, de modo que “no estar” en $\ker(A_i - \lambda_i I)$ *no implica* no tener proyección sobre (o componentes en) él.

²⁴Así, vamos generando en cada paso subespacios que son llevados cada uno al anterior [y finalmente a $\ker(A_i - \lambda_i I)$] por sucesivas aplicaciones de $(A_i - \lambda_i I)$.

6. Supongamos en cambio que en una dada iteración del paso anterior determinamos (por primera vez) que el sistema

$$(\mathbf{A}_i - \lambda_i \mathbf{I})\mathbf{w}_t = \mathbf{w}_{t-1}$$

no tiene solución a menos que se cumpla cierta relación entre los parámetros *originales* a_s presentes en \mathbf{w}_{t-1} , la que adoptará la forma genérica

$$k_1 a_1 + \cdots + k_{m_i} a_{m_i} = 0,$$

con k_1, \dots, k_{m_i} dados.

Por una parte, esto quiere decir que existe un autovector particular en

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_{m_i} \mathbf{v}_{m_i}$$

que debemos eliminar para poder seguir adelante. Este autovector tendrá justamente la forma

$$\mathbf{v}_{10} = k_1 \mathbf{v}_1 + \cdots + k_{m_i} \mathbf{v}_{m_i},$$

y sus sucesivas preimágenes $\mathbf{v}_{11}, \mathbf{v}_{12}, \dots, \mathbf{v}_{1(t-1)}$ se obtienen análogamente, substituyendo cada a_s por el correspondiente k_s en $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{t-1}$, respectivamente. A este autovector particular (y sus preimágenes) tendremos asociado un bloque de Jordan $J_{(\lambda_i)t}$ de tamaño $t \times t$, que actuará sobre el subespacio $V_{i1} := \text{span}\{\mathbf{v}_{10}, \mathbf{v}_{11}, \dots, \mathbf{v}_{1(t-1)}\}$ de V_i .²⁵

Por otra parte, deberemos imponer la condición

$$k_1 a_1 + \cdots + k_{m_i} a_{m_i} = 0$$

a

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_{m_i} \mathbf{v}_{m_i},$$

eliminando así \mathbf{v}_{10} y obteniendo un subespacio de $\ker(\mathbf{A}_i - \lambda_i \mathbf{I})$ disjunto de $\text{span}\{\mathbf{v}_{10}\}$. Llamaremos $\mathbf{v}', \mathbf{w}'_1, \mathbf{w}'_2, \dots, \mathbf{w}'_{t-1}$ al resultado de imponer la condición en \mathbf{v} y sus sucesivas preimágenes, todos los cuales tendrán ahora $m_i - 1$ parámetros a_s libres. Veremos que ahora la ecuación

$$(\mathbf{A}_i - \lambda_i \mathbf{I})\mathbf{w}'_t = \mathbf{w}'_{t-1}$$

sí tiene solución. Continuaremos entonces con el proceso iterativo, pero con un parámetro original a_s menos.²⁶

²⁵Nótese que $\{\mathbf{v}_{10}, \mathbf{v}_{11}, \dots, \mathbf{v}_{1(t-1)}\}$ forman, por construcción, una base cíclica de V_i .

²⁶Si hubiéramos determinado que el sistema

$$(\mathbf{A}_i - \lambda_i \mathbf{I})\mathbf{w}_t = \mathbf{w}_{t-1}$$

no tiene solución a menos que se cumplan *dos* relaciones entre los a_s presentes en \mathbf{w}_{t-1} , de la forma

$$k_1 a_1 + \cdots + k_{m_i} a_{m_i} = 0,$$

$$h_1 a_1 + \cdots + h_{m_i} a_{m_i} = 0,$$

A medida que procedamos, los pasos (5) y (6) se irán alternando, e irán quedando cada vez menos parámetros *originales* a_s libres, es decir sin haber sido eliminados en términos de los demás. El proceso iterativo de cada bloque A_i terminará cuando, quedando $p \geq 1$ parámetros originales libres (digamos a_1, \dots, a_p), obtengamos p condiciones

$$\begin{aligned} k_1 a_1 + \dots + k_p a_p &= 0, \\ h_1 a_1 + \dots + h_p a_p &= 0, \\ &\vdots \\ f_1 a_1 + \dots + f_p a_p &= 0 \end{aligned}$$

que impongan $a_1 = \dots = a_p = 0$. Determinaremos entonces los últimos p autovectores particulares y sus respectivas preimágenes, y los respectivos bloques de Jordan.

Una vez determinados todos los autovectores particulares y sus preimágenes, pueden quedar sin determinar parámetros b_s, c_s , etc. de las contribuciones de la solución homogénea incorporadas en cada iteración del paso 5. Procedemos entonces a “hacer limpieza” simplemente anulándolos.²⁷

Finalmente, una vez completados los pasos 4 a 6 para cada bloque A_i de la descomposición primaria, procedemos como sigue:

Construcción de la transformación de semejanza

7. En cada subespacio V_i construimos la matriz $q_i \times q_i$

$$S_i := \begin{pmatrix} v_{10} & v_{11} & \cdots & v_{1(t_1-1)} & v_{20} & v_{21} & \cdots & v_{2(t_2-1)} & \cdots \text{etc.} \cdots \\ \downarrow & \downarrow & \cdots & \downarrow & \downarrow & \downarrow & \cdots & \downarrow & \cdots \text{etc.} \cdots \end{pmatrix},$$

(k_1, \dots, k_{m_i} y h_1, \dots, h_{m_i} dados), procedemos de igual forma, pero determinando ahora *dos* autovectores particulares

$$\begin{aligned} v_{10} &= k_1 v_1 + \dots + k_{m_i} v_{m_i}, \\ v_{20} &= h_1 v_1 + \dots + h_{m_i} v_{m_i}, \end{aligned}$$

cada uno con su respectivo bloque de Jordan $J_{(\lambda_i)t}$ asociado, de tamaño $t \times t$, actuando sobre sendos subespacios generados por v_{10} y v_{20} y sus respectivas preimágenes. Imponiendo las dos condiciones a v y sus preimágenes (dejando $m_i - 2$ parámetros a_s libres en ellos), la ecuación

$$(A_i - \lambda_i I)w'_t = w'_{t-1}$$

tendrá solución, y continuaremos entonces con el proceso iterativo, con dos parámetros originales a_s menos. La generalización al caso de tres o más condiciones es directa.

²⁷Ya que, por ejemplo, los b_s aún libres tras finalizar la iteración de los pasos 5 y 6 sólo aportan a w_1 vectores que ya están en v , a w_2 vectores que ya están en w_1 , etc.; los c_s libres sólo aportan a w_2 vectores que ya están en v , a w_3 vectores que ya están en w_1 , etc.; etc. Como se verá en los ejemplos que siguen, muchas veces podemos descartar estos parámetros tan pronto estemos seguros que ya no serán necesarios en ulteriores iteraciones.

y su inversa S_i^{-1} ; la transformación de semejanza generada por ellas dejará el bloque A_i en la forma

$$S_i^{-1}A_iS_i = \begin{pmatrix} J_{(\lambda_i)t_1} & & & \\ & J_{(\lambda_i)t_2} & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix}.$$

8. Construimos ahora las matrices $n \times n$

$$S := \begin{pmatrix} S_1 & 0 & \cdots & 0 \\ 0 & S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_r \end{pmatrix}, \quad S^{-1} := \begin{pmatrix} S_1^{-1} & 0 & \cdots & 0 \\ 0 & S_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_r^{-1} \end{pmatrix}.$$

Aplicando la transformación de semejanza generada por ellas a la descomposición primaria A' de A , tendremos

$$A'' = S^{-1}T^{-1}ATS$$

en su forma normal de Jordan.

Ejemplo 1.8.3. Presentaremos primero un ejemplo relativamente simple para fijar algunas ideas. Utilizaremos la misma numeración del texto precedente, para facilitar el seguimiento de cada paso.

Sea la matriz

$$A = \begin{pmatrix} 2 & 1 & -1 & 0 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Descomposición primaria

Un rápido examen de ésta matriz muestra que sólo tenemos un único autovector $\lambda = 2$ con multiplicidad 5, por lo que ya estamos en una base que implementa la descomposición primaria. Los pasos 1 a 3 pueden darse entonces por completados.

Reducción a forma normal

Necesitaremos

$$(A - \lambda I)x = \begin{pmatrix} 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_2 - x_3 + x_5 \\ x_4 + x_5 \\ x_5 \\ x_5 \\ 0 \end{pmatrix}.$$

4. Los autovectores vienen dados por

$$\begin{pmatrix} x_2 - x_3 + x_5 \\ x_4 + x_5 \\ x_5 \\ x_5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_2 \\ 0 \\ 0 \end{pmatrix}.$$

Hay dos autovectores LI; ortonormalizándolos ponemos

$$\mathbf{v} = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a_2 \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

5. Buscamos una primera preimagen \mathbf{w}_1 de \mathbf{v} , solución general de

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{w}_1 = \mathbf{v} \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2/\sqrt{2} \\ a_2/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}.$$

La tercera y cuarta ecuaciones muestran que no habrá solución a menos que $a_2 = 0$.

6. Tenemos la condición

$$k_1 a_1 + k_2 a_2 = 0 \text{ con } k_1 = 0, k_2 = 1.$$

Por un lado, entonces, substituimos a_1 por k_1 y a_2 por k_2 en \mathbf{v} y determinamos

$$\mathbf{v}_{10} = k_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + k_2 \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

que tiene asociado un bloque de Jordan $J_{(2)1}$.

Por otro lado, imponemos la condición $a_2 = 0$ en \mathbf{v} y obtenemos

$$\mathbf{v}' = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0 \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Volvemos entonces al paso anterior.

5. Buscamos una primera preimagen \mathbf{w}'_1 de \mathbf{v}' , solución general de

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{w}'_1 = \mathbf{v}' \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \mathbf{w}'_1 = \begin{pmatrix} b_1 \\ b_2 \\ b_2 - a_1 \\ 0 \\ 0 \end{pmatrix},$$

que no impone condición alguna sobre a_1 .

Buscamos una segunda preimagen w'_2 de w'_1 , solución general de

$$(A - \lambda I)w'_2 = w'_1 \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_2 - a_1 \\ 0 \\ 0 \end{pmatrix}.$$

Vemos que la tercera y cuarta ecuaciones imponen $b_2 = a_1$, pero no hay condición impuesta sobre a_1 . Tampoco se impone ninguna condición sobre b_1 , así que asumiremos el riesgo de descartarlo ahora, poniendo $b_1 = 0$. Recordemos que debemos substituir estos valores de b_1 y b_2 también en w'_1 , así que escribimos

$$w'_1 = \begin{pmatrix} 0 \\ a_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w'_2 = \begin{pmatrix} c_1 \\ c_2 \\ c_2 \\ a_1 \\ 0 \end{pmatrix}.$$

Buscamos una tercera preimagen w'_3 de w'_2 , solución general de

$$(A - \lambda I)w'_3 = w'_2 \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_2 \\ a_1 \\ 0 \end{pmatrix}.$$

Vemos que la tercera y cuarta ecuaciones imponen $c_2 = a_1$, pero no hay condición impuesta sobre a_1 . Tampoco se impone ninguna condición sobre c_1 , así que asumiremos el riesgo de descartarlo ahora, poniendo $c_1 = 0$. Substituyendo estos valores de c_1 y c_2 también en w'_2 , escribimos

$$w'_2 = \begin{pmatrix} 0 \\ a_1 \\ a_1 \\ a_1 \\ 0 \end{pmatrix}, \quad w'_3 = \begin{pmatrix} d_1 \\ d_2 \\ d_2 + a_1 \\ 0 \\ a_1 \end{pmatrix}.$$

Buscamos una cuarta preimagen w'_4 de w'_3 , solución general de

$$(A - \lambda I)w'_4 = w'_3 \implies \begin{pmatrix} w_2 - w_3 + w_5 \\ w_4 + w_5 \\ w_5 \\ w_5 \\ 0 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_2 + a_1 \\ 0 \\ a_1 \end{pmatrix}.$$

Vemos que la tercera y cuarta ecuaciones imponen $d_2 = -a_1$, y que no hay condición sobre d_1 , así que tomamos $d_1 = 0$ y escribimos

$$\mathbf{w}'_3 = \begin{pmatrix} 0 \\ -a_1 \\ 0 \\ 0 \\ a_1 \end{pmatrix}.$$

Vemos también que la quinta ecuación impone la condición $a_1 = 0$.²⁸

6. Tenemos la condición

$$k_1 a_1 = 0 \text{ con } k_1 = 1.$$

Por un lado, substituimos a_1 por k_1 en \mathbf{v}' , \mathbf{w}'_1 , \mathbf{w}'_2 y \mathbf{w}'_3 , y determinamos

$$\mathbf{v}_{20} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{w}'_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{w}'_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{w}'_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

que tienen asociado un bloque de Jordan $J_{(2)4}$.

Por otro lado, imponemos la condición a_2 en \mathbf{v} y obtenemos

$$\mathbf{v}' = 0,$$

el subespacio nulo. Esto indica que hemos terminado con este bloque de la descomposición primaria.

Construcción de la transformación de semejanza

7. Procedemos ahora a construir las matrices

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/\sqrt{2} & 0 & 1 & 1 & -1 \\ 1/\sqrt{2} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{S}^{-1} = \begin{pmatrix} 0 & 0 & \sqrt{2} & -\sqrt{2} & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

y obtenemos

$$\mathbf{S}^{-1} \mathbf{A} \mathbf{S} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix},$$

²⁸El lector debe convencerse de que esta condición no es “salvable”, es decir no hay valores posibles de d_1 y d_2 que permitan no anular a_1 ; es más, debe convencerse también de que, aún si no hubiéramos descartado b_1 y c_1 , la condición $a_1 = 0$ persistiría.

que ya está en forma normal de Jordan.

Ejercicio 1.8.2. Muestre que si en el ejemplo precedente, elegimos otros valores distintos para b_1 , c_1 y d_1 (que en el ejemplo fueron igualados a cero), la forma normal de Jordan resulta la misma, y por consiguiente sólo hemos obtenido *otra base cíclica diferente* para el subespacio donde actúa el bloque de Jordan $J_{(2)4}$.

Ejemplo 1.8.4. Vamos a presentar ahora un ejemplo relativamente completo del uso del método para reducir una matriz a su forma normal de Jordan. De nuevo utilizaremos la misma numeración, para facilitar el seguimiento de cada paso.

Sea la matriz

$$A = \begin{pmatrix} -10 & -6 & 7 & -12 & 2 & 8 & 8 & -5 \\ -9 & -4 & 5 & -11 & 3 & 9 & 7 & -6 \\ 0 & 1 & 3 & 3 & -3 & -3 & -1 & 3 \\ 19 & 12 & -10 & 25 & -7 & -18 & -15 & 13 \\ 12 & 7 & -6 & 15 & -3 & -11 & -9 & 8 \\ -8 & -4 & 5 & -8 & 1 & 6 & 5 & -3 \\ 0 & 0 & -1 & -2 & 2 & 2 & 2 & -1 \\ -14 & -9 & 7 & -18 & 6 & 13 & 11 & -8 \end{pmatrix}.$$

Descomposición primaria

1. El polinomio característico es

$$\varphi_A(\lambda) = \lambda^8 - 11\lambda^7 + 52\lambda^6 - 138\lambda^5 + 225\lambda^4 - 231\lambda^3 + 146\lambda^2 - 52\lambda + 8,$$

que factorizado queda

$$\varphi_A(\lambda) = (\lambda - 2)^3(\lambda - 1)^5.$$

Tenemos entonces

$$\begin{aligned} \lambda_1 &= 2, & q_1 &= 3, \\ \lambda_2 &= 1, & q_2 &= 5. \end{aligned}$$

2. La solución general de $(A - \lambda_1 I)^3 v = 0$ resulta ser

$$\ker[(A - \lambda_1 I)^3] = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \\ -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

de donde podemos leer las columnas v_{11}, v_{12} y v_{13} , que son LI. Análogamente, la solución general de $(A - \lambda_2 I)^5 v = 0$ resulta ser

$$\ker[(A - \lambda_2 I)^5] = \text{span} \left\{ \begin{pmatrix} -4 \\ -1 \\ -4 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 2 \\ 0 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 1 \\ -4 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 2 \end{pmatrix} \right\},$$

de donde podemos leer las columnas $v_{21}, v_{22}, v_{23}, v_{24}$ y v_{25} , que son LI entre sí y de las anteriores.

3. Formamos la matriz

$$T = \begin{pmatrix} 1 & 1 & 0 & -4 & 0 & 2 & 2 & -4 \\ 2 & 1 & -1 & -1 & 3 & 0 & 1 & 1 \\ 1 & 0 & 1 & -4 & 2 & 2 & 2 & -4 \\ -1 & -1 & 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

con las columnas encontradas en el paso 2, determinamos su inversa

$$T^{-1} = \begin{pmatrix} -7 & -4 & 4 & -8 & 2 & 6 & 5 & -4 \\ -2 & -2 & 0 & -5 & 3 & 4 & 3 & -3 \\ -6 & -4 & 3 & -8 & 3 & 6 & 5 & -4 \\ -\frac{3}{2} & -1 & \frac{1}{2} & -2 & 1 & 2 & \frac{3}{2} & -\frac{3}{2} \\ \frac{3}{2} & 1 & -1 & \frac{3}{2} & 0 & -1 & -1 & \frac{1}{2} \\ \frac{7}{2} & 4 & -4 & 8 & -2 & -5 & -5 & 4 \\ 1 & 1 & 0 & \frac{5}{2} & -\frac{3}{2} & -2 & -1 & \frac{3}{2} \\ 3 & 2 & -\frac{3}{2} & 4 & -\frac{3}{2} & -3 & -\frac{5}{2} & \frac{5}{2} \end{pmatrix},$$

y aplicamos a A la transformación de semejanza así definida, obteniendo

$$A' = T^{-1}AT = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}.$$

Reducción a forma normal del bloque A_1 . Trabajamos primero con

$$A_1 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad (A_1 - \lambda_1 I) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

4. La solución general de $(A_1 - \lambda_1 I)v = 0$ resulta ser

$$\ker(A_1 - \lambda_1 I) = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Estas columnas ya son ortonormales; ponemos entonces

$$v = a_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

5. Buscamos una primera preimagen w_1 , solución general de

$$(A_1 - \lambda_1 I)w_1 = v.$$

La tercera ecuación de este sistema muestra que no habrá solución a menos que $k_1 a_1 + k_2 a_2 = 0$, con $k_1 = 0$ y $k_2 = 1$, es decir, a menos que $a_2 = 0$.

6. Substituyendo a_1 por k_1 y a_2 por k_2 en v , determinamos entonces un primer autovector

$$v_{10} = k_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

asociado a un bloque de Jordan $J_{(2)1}$, y que no tiene ninguna preimagen.

Para seguir debemos eliminar v_{10} de v . Para ello substituimos $a_2 = 0$ en v , obteniendo

$$v' = a_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

5. Buscamos su primera preimagen w'_1 , solución general de

$$(A_1 - \lambda_1 I)w'_1 = v'.$$

Obtenemos $w'_1 = (a_1, b_1, b_2)^T$; un vistazo a la iteración siguiente muestra que no se impone condición sobre los nuevos parámetros b_1 y b_2 , así que los descartamos y escribimos

$$w'_1 = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix}.$$

Buscamos ahora una segunda preimagen w'_2 , solución general de

$$(A_1 - \lambda_1 I)w'_2 = w'_1.$$

La segunda ecuación de este sistema muestra que no habrá solución a menos que $k_1 a_1 = 0$, con $k_1 = 1$, es decir a menos que $a_1 = 0$.

6. Substituyendo a_1 por k_1 en v' y su preimagen w'_1 , determinamos un segundo autovector v_{20} y su preimagen v_{21} ,

$$v_{20} = k_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_{21} = \begin{pmatrix} k_1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

asociados a un bloque de Jordan $J_{(2)2}$, y que no tienen otra preimagen.

Para seguir debemos eliminar v_{20} de v' . Para ello substituimos $a_1 = 0$ en v' , obteniendo el vector nulo, lo que indica que hemos terminado con el bloque asociado a λ_1 .

Reducción a forma normal del bloque A_2 . Pasamos ahora a trabajar con

$$A_2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ -\frac{1}{2} & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}, \quad (A_2 - \lambda_2 I) = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

4. La solución general de $(A_2 - \lambda_2 I)v = 0$ resulta ser

$$\ker(A_2 - \lambda_2 I) = \text{span} \left\{ \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix}, \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix} \right\}.$$

Estas columnas ya son ortonormales; ponemos entonces

$$v = a_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix} + a_2 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix}.$$

5. Buscamos su primera preimagen w_1 , solución general de

$$(A_1 - \lambda_1 I)w_1 = v.$$

Obtenemos $w_1 = (b_1 + b_2 + \frac{1}{2}\sqrt{2}(a_1 + a_2), -b_1 + b_2 + \frac{1}{2}\sqrt{2}(a_1 + a_2), 2b_1 - 2b_2 + \frac{1}{3}\sqrt{6}(-a_1 + a_2), b_1 + b_2, b_1 - b_2)^T$; un vistazo a la iteración siguiente

muestra que no se impondrá condición sobre los nuevos parámetros b_1 y b_2 , así que los descartamos y escribimos

$$\mathbf{w}_1 = \begin{pmatrix} \frac{1}{2}\sqrt{2}(a_1 + a_2) \\ \frac{1}{2}\sqrt{2}(a_1 + a_2) \\ \frac{1}{3}\sqrt{6}(-a_1 + a_2) \\ 0 \\ 0 \end{pmatrix}.$$

Buscamos ahora una segunda preimagen \mathbf{w}_2 , solución general de

$$(\mathbf{A}_1 - \lambda_1 \mathbf{I})\mathbf{w}_2 = \mathbf{w}_1.$$

La primera y cuarta ecuaciones de este sistema muestran que no habrá solución a menos que $k_1 a_1 + k_2 a_2 = 0$, con $k_1 = 1$ y $k_2 = 1$, es decir, a menos que $a_2 = -a_1$ (la segunda y quinta ecuaciones muestran lo mismo).

6. Substituyendo a_1 por k_1 y a_2 por k_2 en \mathbf{v} y su preimagen \mathbf{w}_1 , determinamos entonces un primer autovector \mathbf{v}_{10} y su preimagen \mathbf{v}_{11} ,

$$\mathbf{v}_{10} = k_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix} + k_2 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \\ \sqrt{2} \\ 0 \end{pmatrix}, \quad \mathbf{v}_{11} = \begin{pmatrix} \frac{1}{2}\sqrt{2}(k_1 + k_2) \\ \frac{1}{2}\sqrt{2}(k_1 + k_2) \\ \frac{1}{3}\sqrt{6}(-k_1 + k_2) \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

asociados a un bloque de Jordan $J_{(1)2}$, y que no tienen otra preimagen.

Para seguir debemos eliminar \mathbf{v}_{10} de \mathbf{v} . Para ello substituimos $a_2 = -a_1$ en \mathbf{v} , obteniendo

$$\mathbf{v}' = a_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{pmatrix} - a_1 \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \\ -\frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{1}{3}\sqrt{6}a_1 \\ \frac{2}{3}\sqrt{6}a_1 \\ 0 \\ \frac{1}{3}\sqrt{6}a_1 \end{pmatrix}.$$

Haciendo la misma sustitución en \mathbf{w}_1 , obtenemos la primera preimagen de \mathbf{v}' ,

$$\mathbf{w}'_1 = \begin{pmatrix} 0 \\ 0 \\ -\frac{2}{3}\sqrt{6}a_1 \\ 0 \\ 0 \end{pmatrix}.$$

5. Buscamos ahora una preimagen \mathbf{w}'_2 , solución general de

$$(\mathbf{A}_1 - \lambda_1 \mathbf{I})\mathbf{w}'_2 = \mathbf{w}'_1.$$

Obtenemos $w'_2 = (b_1 + b_2 + \frac{2}{3}\sqrt{6}a_1, -b_1 + b_2, 2b_1 - 2b_2 + \frac{2}{3}\sqrt{6}a_1, b_1 + b_2, b_1 - b_2)^T$; un vistazo a la iteración siguiente muestra que no se impondrá condición sobre los nuevos parámetros b_1 y b_2 , así que los descartamos y escribimos

$$w'_2 = \begin{pmatrix} \frac{2}{3}\sqrt{6}a_1 \\ 0 \\ \frac{2}{3}\sqrt{6}a_1 \\ 0 \\ 0 \end{pmatrix}.$$

Buscamos la siguiente preimagen, solución general de

$$(A_1 - \lambda_1 I)w'_3 = w'_2.$$

La primera y cuarta ecuaciones de este sistema muestran que no habrá solución a menos que $k_1 a_1 = 0$, con $k_1 = 1$, es decir, a menos que $a_1 = 0$.

6. Substituyendo a_1 por $k_1 = 1$ en v' y sus preimágenes w'_1 y w'_2 , respectivamente, determinamos un segundo autovector v_{20} y sus preimágenes v_{21} y v_{22} ,

$$v_{20} = \begin{pmatrix} 0 \\ -\frac{1}{3}\sqrt{6} \\ \frac{2}{3}\sqrt{6} \\ 0 \\ \frac{1}{3}\sqrt{6} \end{pmatrix}, \quad v_{21} = \begin{pmatrix} 0 \\ 0 \\ -\frac{2}{3}\sqrt{6} \\ 0 \\ 0 \end{pmatrix}, \quad v_{22} = \begin{pmatrix} \frac{2}{3}\sqrt{6} \\ 0 \\ \frac{2}{3}\sqrt{6} \\ 0 \\ 0 \end{pmatrix},$$

asociados a un bloque de Jordan $J_{(1)3}$, y que ya no tienen otra preimagen.

Para seguir debemos eliminar v_{20} de v' . Para ello substituimos $a_1 = 0$ en v' , obteniendo el vector nulo, lo que indica que hemos terminado con el bloque asociado a λ_2 .

Construcción de la transformación de semejanza

7. Recolectamos los autovectores y sus preimágenes asociados a λ_1 en la matriz S_1 y calculamos su inversa, obteniendo

$$S_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad S_1^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Aplicando la transformación de semejanza así definida a A_1 , queda

$$S_1^{-1}A_1S_1 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

También recolectamos los autovectores y sus preimágenes asociados a λ_2

en la matriz S_2 y calculamos su inversa, obteniendo

$$S_2 = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 0 & 0 & \frac{2}{3}\sqrt{6} \\ 0 & \sqrt{2} & -\frac{1}{3}\sqrt{6} & 0 & 0 \\ 0 & 0 & \frac{2}{3}\sqrt{6} & -\frac{2}{3}\sqrt{6} & \frac{2}{3}\sqrt{6} \\ \sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3}\sqrt{6} & 0 & 0 \end{pmatrix},$$

$$S_2^{-1} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2}\sqrt{2} & 0 \\ 0 & \frac{1}{2}\sqrt{2} & 0 & 0 & \frac{1}{2}\sqrt{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2}\sqrt{6} \\ \frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & \frac{1}{4}\sqrt{6} \\ \frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} & 0 & -\frac{1}{4}\sqrt{6} & -\frac{1}{4}\sqrt{6} \end{pmatrix}.$$

Aplicando la transformación de semejanza así definida a A_2 , queda

$$S_2^{-1}A_2S_2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

8. Finalmente, recolectamos las matrices S_1 y S_2 y sus inversas en las matrices

$$S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} S_1^{-1} & 0 \\ 0 & S_2^{-1} \end{pmatrix},$$

y aplicamos la transformación de semejanza así definida a A' , quedando

$$A'' = S^{-1}A'S = S^{-1}T^{-1}ATS = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

que está en su forma normal de Jordan. Esta última transformación es sólo formal, ya que en realidad nos limitamos a colocar los bloques $S_1^{-1}A_1S_1$ y $S_2^{-1}A_2S_2$ en la diagonal de A'' , y completar el resto con ceros.

Para finalizar, debemos destacar algunos hechos importantes que probablemente no hayan quedado explicitados.

1. Cuando diagonalizamos una matriz diagonalizable, estamos habituados a que los autovectores pueden renormalizarse arbitrariamente. Aquí eso sigue siendo cierto respecto de los autovectores, pero si multiplicamos

un dado autovector por una constante, deberemos multiplicar todas sus preimágenes por *la misma* constante, ya que estas últimas son soluciones de sucesivos sistemas inhomogéneos, el primero de los cuales tiene al autovector como inhomogeneidad.

2. También cuando diagonalizamos una matriz diagonalizable, estamos habituados a que los autovectores correspondientes a un mismo autovalor pueden recombinarse (linealmente) en forma arbitraria. Para una matriz no diagonalizable esto en general *es falso*, ya que cada autovector tiene asociado un bloque de Jordan en general distinto. Sólo podremos recombinar libremente autovectores al mismo autovalor y que tengan asociados bloques de Jordan *de igual tamaño*.
3. Finalmente, también de nuestra experiencia con matrices diagonalizables puede habernos quedado la idea de que dos matrices serán semejantes si tienen el mismo espectro (o, lo que es igual, los mismos invariantes algebraicos). Para matrices no diagonalizables ello *no es cierto*: debe especificarse, además, la estructura de bloques de Jordan.

1.9. Tensores

Los tensores aparecen frecuentemente en Física, representando por ejemplo entes como momento cuadrupolar o gradiente del campo eléctrico, momentos de inercia, propiedades elásticas de sólidos, propiedades de transporte de fluidos, etc. En esta Sección unificaremos algunos conceptos ya introducidos y daremos una formulación general de los mismos. En el camino, veremos que ya nos hemos encontrado, sin saberlo, con diversas instancias de tensores, que ahora reconoceremos como tales.

1.9.1. Definición de tensor

Ya nos hemos encontrado con el concepto de una 1-forma (lineal) $\tilde{\omega}$, que es una aplicación de un espacio vectorial V al cuerpo escalar \mathbb{C} ,

$$\tilde{\omega} : V \rightarrow \mathbb{C}, \quad V \ni \vec{u} \xrightarrow{\tilde{\omega}} \tilde{\omega}(\vec{u}) \in \mathbb{C}.$$

Menos familiar es el hecho de que podemos pensar a un vector \vec{u} como una aplicación lineal del espacio dual V^* al cuerpo escalar: si definimos

$$\vec{u}(\tilde{\omega}) := \tilde{\omega}(\vec{u}),$$

vemos que efectivamente

$$\vec{u} : V^* \rightarrow \mathbb{C}, \quad V^* \ni \vec{u} \xrightarrow{\tilde{\omega}} \vec{u}(\tilde{\omega}) \in \mathbb{C}.$$

Otro ente en cierto modo semejante que ya hemos encontrado es el producto interno, que definimos como una 2-forma (lineal) que lleva dos vectores de V al

cuerpo escalar,

$$\Phi : V \times V \rightarrow \mathbb{C}, \quad V \times V \ni (\vec{u}, \vec{v}) \xrightarrow{\Phi} \Phi(\vec{u}, \vec{v}) \in \mathbb{C}.$$

Aquí $V \times V$ representa el *producto Cartesiano* del espacio V consigo mismo, es decir, el EV formado por todos los pares ordenados de vectores (\vec{u}, \vec{v}) , con $\vec{u}, \vec{v} \in V$.²⁹

Aún otro ente con características similares es un operador lineal \mathcal{A} de V en V . A primera vista esto parece contradictorio, ya que $\mathcal{A} : V \rightarrow V$, es decir no va al cuerpo escalar. Sin embargo, dado $\vec{u} \in V$, también $\mathcal{A}\vec{u} \in V$, de modo que podemos aplicarlo a una 1-forma $\tilde{\omega} \in V^*$ para obtener un escalar:

$$(\mathcal{A}\vec{u})(\tilde{\omega}) \equiv \tilde{\omega}(\mathcal{A}\vec{u}) \in \mathbb{C}.$$

Podemos entonces pensar en el operador \mathcal{A} como una aplicación lineal que toma un vector \vec{u} de V y una 1-forma $\tilde{\omega}$ de V^* y devuelve un escalar. Haciendo un ligero abuso de notación (aprovechándonos de la linealidad) desecharémos los paréntesis, y escribiremos

$$\mathcal{A} : V^* \times V \rightarrow \mathbb{C}, \quad V^* \times V \ni (\tilde{\omega}, \vec{u}) \xrightarrow{\mathcal{A}} \tilde{\omega}\mathcal{A}\vec{u} \in \mathbb{C}.$$

La notación $\tilde{\omega}\mathcal{A}\vec{u}$ es altamente sugestiva, y nos permite dar *tres* interpretaciones diferentes de la operación que indica: (i) la aplicación \mathcal{A} toma un vector \vec{u} y una forma $\tilde{\omega}$, y devuelve un escalar ($\mathcal{A} : V^* \times V \rightarrow \mathbb{C}$); (ii) el operador \mathcal{A} actúa “hacia adelante” sobre el vector \vec{u} , devolviendo otro vector sobre el que actúa $\tilde{\omega}$ para dar un escalar ($\mathcal{A} : V \rightarrow V$); y (iii) el operador \mathcal{A} actúa “hacia atrás” sobre la forma $\tilde{\omega}$, devolviendo otra forma sobre la que actúa \vec{u} para dar un escalar ($\mathcal{A} : V^* \rightarrow V^*$).

La generalización más o menos obvia de estos casos particulares es una aplicación lineal que tome un cierto número de formas de V^* y un cierto número de vectores de V y devuelva un escalar. Para formalizarla, definiremos primero

$$\Pi_r^s := \underbrace{V^* \times V^* \times \cdots \times V^*}_{r \text{ veces}} \times \underbrace{V \times V \times \cdots \times V}_{s \text{ veces}}. \quad (1.70)$$

Definiremos ahora un *tensor* T de tipo (r, s) como una aplicación

$$T : \Pi_r^s \rightarrow \mathbb{C} \quad (1.71)$$

lineal en todos sus argumentos, y escribiremos

$$T(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) = t \in \mathbb{C}, \quad \tilde{\omega}^1, \dots, \tilde{\omega}^r \in V^*, \quad \vec{u}_1, \dots, \vec{u}_s \in V.$$

El *rango* de un tensor se define como el número total de argumentos que toma la aplicación, es decir $r + s$.

Revisando los casos particulares del principio, vemos que:

²⁹En espacios de dimensión *finita*, el producto Cartesiano $V \times W$ y la suma directa $V \oplus W$ de dos espacios vectoriales V y W son equivalentes, y podemos usar indistintamente uno o la otra. Sin embargo, esto no es válido en espacios de dimensión infinita; la noción que se generaliza correctamente a este caso es la de producto Cartesiano, que es la que usaremos aquí.

- las 1-formas son tensores de tipo $(0, 1)$, rango 1;
- los vectores son tensores de tipo $(1, 0)$, rango 1;
- el producto interno es un tensor de tipo $(0, 2)$, rango 2;
- los operadores lineales son tensores de tipo $(1, 1)$, rango 2.

1.9.2. Espacio tensorial

Volviendo al concepto de 1-forma como aplicación lineal de V en \mathbb{C} , recordemos que lo usamos para definir el espacio dual V^* como el EV de todas las 1-formas lineales sobre V , es decir

$$V^* = \{\tilde{\omega} : V \rightarrow \mathbb{C}, \tilde{\omega} \text{ lineal}\}.$$

Para completar esta definición, debemos agregar al conjunto de las 1-formas las operaciones de *suma* de 1-formas y de *producto por un escalar*, con las propiedades usuales.

De modo análogo, y pensando ahora en los vectores como aplicaciones lineales de V^* en \mathbb{C} , podemos definir V como

$$V = \{\vec{u} : V^* \rightarrow \mathbb{C}, \vec{u} \text{ lineal}\},$$

agregando las operaciones de suma de vectores y de producto de un vector por un escalar, también con las propiedades usuales.

Asimismo, podemos definir el espacio consistente en todos los operadores lineales \mathcal{A} de V en V , como el conjunto

$$\{\mathcal{A} : V^* \times V \rightarrow \mathbb{C}, \mathcal{A} \text{ lineal}\},$$

agregando las operaciones de suma entre operadores lineales y de producto de un operador lineal por un escalar, y este será un espacio vectorial.

Para generalizar esta idea, debemos definir las siguientes operaciones:

- **Suma de tensores:** Dados dos tensores T y S , ambos de tipo (r, s) , su *suma* es el tensor

$$(T + S)(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) \\ := T(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) + S(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s),$$

también de tipo (r, s) .

- **Producto por un escalar:** Dado un tensor T de tipo (r, s) , su *producto por un escalar* $\lambda \in \mathbb{C}$ es el tensor

$$(\lambda T)(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) := \lambda T(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s),$$

también de tipo (r, s) .

Estas operaciones tendrán las mismas propiedades (distributividad, etc.) que las definidas para vectores. Dadas estas operaciones, definimos el *espacio tensorial* V_r^s como

$$V_r^s := \{T : \Pi_r^s \rightarrow \mathbb{C}, T \text{ lineal}\}, \quad (1.72)$$

y es un espacio vectorial.

1.9.3. Producto tensorial

Un concepto útil que nos permite definir tensores sobre espacios “más grandes” a partir de otros tensores sobre espacios “más chicos” es el de producto tensorial. Dados dos tensores $T_1 \in V_{r_1}^{s_1}$ y $T_2 \in V_{r_2}^{s_2}$ en espacios tensoriales $V_{r_1}^{s_1}$ y $V_{r_2}^{s_2}$ (posiblemente distintos), su *producto tensorial* T es el nuevo tensor

$$T := T_1 \otimes T_2 \in V_{r_1+r_2}^{s_1+s_2} \quad (1.73)$$

definido por

$$\begin{aligned} T(\tilde{\omega}^1, \dots, \tilde{\omega}^{r_1}, \tilde{\tau}^1, \dots, \tilde{\tau}^{r_2}; \vec{u}_1, \dots, \vec{u}_{s_1}, \vec{v}_1, \dots, \vec{v}_{s_2}) \\ = T_1(\tilde{\omega}^1, \dots, \tilde{\omega}^{r_1}; \vec{u}_1, \dots, \vec{u}_{s_1}) T_2(\tilde{\tau}^1, \dots, \tilde{\tau}^{r_2}; \vec{v}_1, \dots, \vec{v}_{s_2}), \end{aligned} \quad (1.74)$$

donde $\tilde{\omega}^1, \dots, \tilde{\omega}^{r_1}, \tilde{\tau}^1, \dots, \tilde{\tau}^{r_2} \in V^*$ y $\vec{u}_1, \dots, \vec{u}_{s_1}, \vec{v}_1, \dots, \vec{v}_{s_2} \in V$.

Ejemplo 1.9.1. Dados dos vectores $\vec{u}, \vec{v} \in V$, definamos $T := \vec{u} \otimes \vec{v}$; entonces

$$T(\tilde{\omega}, \tilde{\tau}) = \vec{u}(\tilde{\omega}) \vec{v}(\tilde{\tau}) \quad \forall \tilde{\omega}, \tilde{\tau} \in V^*,$$

y vemos que T es un tensor de tipo $(2, 0)$.

Ejemplo 1.9.2. Dados un vector $\vec{u} \in V$ y una forma $\tilde{\omega} \in V^*$, definamos $T := \vec{u} \otimes \tilde{\omega}$; entonces

$$T(\tilde{\tau}, \vec{v}) = \vec{u}(\tilde{\tau}) \tilde{\omega}(\vec{v}) \quad \forall \tilde{\tau} \in V^*, \vec{v} \in V,$$

y vemos que T es un tensor de tipo $(1, 1)$.

1.9.4. Base y componentes

Consideremos un EV V con una base $\{\vec{e}_i\}_{i=1}^n$, y su dual V^* con la base dual $\{\tilde{e}^j\}_{j=1}^n$. Recordemos cómo definimos las componentes ω_i de una forma $\tilde{\omega} \in V^*$ por $\omega_i = \tilde{\omega}(\vec{e}_i)$. Análogamente podríamos definir las componentes x^j de un vector $\vec{x} \in V$ por $x^j = \vec{x}(\tilde{e}^j)$. También podríamos definir las componentes de un operador lineal \mathcal{A} como $A^j{}_i = \tilde{e}^j \mathcal{A} \vec{e}_i$, o las componentes Φ_{ij} del producto interno Φ como $\Phi_{ij} = \Phi(\vec{e}_i, \vec{e}_j)$ (estas últimas serían los elementos de la métrica g_{ij}). Veríamos que todas estas definiciones son consistentes con las propiedades que hemos ido hallando para las componentes de los diversos entes.

Vamos a generalizar entonces esa idea, definiendo las *componentes* $S_{j_1 \dots j_s}^{i_1 \dots i_r}$ de un tensor $S \in V_r^s$ en las bases $\{\vec{e}_i\}_{i=1}^n$ y $\{\tilde{e}^j\}_{j=1}^n$, como

$$S_{j_1 \dots j_s}^{i_1 \dots i_r} := S(\tilde{e}^{i_1}, \dots, \tilde{e}^{i_r}; \vec{e}_{j_1}, \dots, \vec{e}_{j_s}). \quad (1.75)$$

De manera análoga a lo que ocurría para las componentes de vectores, operadores, etc., las propiedades de linealidad de los tensores harán que las componentes de una suma sean la suma de las componentes, y las componentes del producto por un escalar sean el producto del escalar por las componentes.

Consideremos ahora formas $\tilde{\omega}^1 = \omega_{i_1} \tilde{e}^{i_1}, \dots, \tilde{\omega}^r = \omega_{i_r} \tilde{e}^{i_r}$ en V^* , y vectores $\vec{u}_1 = u^{j_1} \vec{e}_{j_1}, \dots, \vec{u}_s = u^{j_s} \vec{e}_{j_s}$ en V . Por linealidad,

$$\begin{aligned} S(\tilde{\omega}^1, \dots, \tilde{\omega}^r; \vec{u}_1, \dots, \vec{u}_s) &= \omega_{i_1} \dots \omega_{i_r} u^{j_1} \dots u^{j_s} S(\tilde{e}^{i_1}, \dots, \tilde{e}^{i_r}; \vec{e}_{j_1}, \dots, \vec{e}_{j_s}) \\ &= \omega_{i_1} \dots \omega_{i_r} u^{j_1} \dots u^{j_s} S_{j_1 \dots j_s}^{i_1 \dots i_r}. \end{aligned} \quad (1.76)$$

Por otro lado, dados los productos tensoriales

$$\vec{e}_{i_1} \otimes \dots \otimes \vec{e}_{i_r} \otimes \tilde{e}^{j_1} \otimes \dots \otimes \tilde{e}^{j_s} \quad (1.77)$$

puede verificarse fácilmente que

$$S = \vec{e}_{i_1} \otimes \dots \otimes \vec{e}_{i_r} \otimes \tilde{e}^{j_1} \otimes \dots \otimes \tilde{e}^{j_s} S_{j_1 \dots j_s}^{i_1 \dots i_r} \in V_r^s, \quad (1.78)$$

es decir que S así definido es un tensor de tipo (r, s) . Esta forma de escribir un tensor es completamente análoga a la manera en que escribimos un vector $\vec{u} = u^i \vec{e}_i$ o una forma $\tilde{\omega} = \omega_i \tilde{e}^i$ como combinaciones lineales de los elementos de una base, donde los coeficientes de la combinación son las respectivas componentes. Puede mostrarse que los productos tensoriales (1.77) forman una *base* del espacio tensorial V_r^s .

Ejemplo 1.9.3. Sea $\mathcal{A} : V \rightarrow V$ un operador lineal sobre V , con componentes A^i_j en la base $\{\vec{e}_i\}$ de V . Entonces

$$\mathcal{A} = A^i_j \vec{e}_i \otimes \tilde{e}^j,$$

con lo cual si $\vec{x} = x^i \vec{e}_i \in V$,

$$\mathcal{A}\vec{x} = A^i_j \vec{e}_i \otimes \tilde{e}^j x^l \vec{e}_l = A^i_j x^l \vec{e}_i \otimes \tilde{e}^j (\vec{e}_l) = A^i_j x^l \vec{e}_i \delta^j_l = (A^i_j x^j) \vec{e}_i \in V,$$

y análogamente, si $\tilde{\omega} = \omega_i \tilde{e}^i \in V^*$,

$$\tilde{\omega}\mathcal{A} = (\omega_i A^i_j) \tilde{e}^i \in V^*.$$

Ejemplo 1.9.4. Podemos pensar al operador identidad $\mathcal{I} : V \rightarrow V$ como un tensor $I \in V_1^1$ tal que

$$I(\tilde{\omega}, \vec{u}) = \tilde{\omega}\mathcal{I}\vec{u} \equiv \tilde{\omega}\vec{u} \quad \forall \tilde{\omega} \in V^*, \vec{u} \in V.$$

Entonces

$$I = \delta^i_j \vec{e}_i \otimes \tilde{e}^j,$$

donde δ^i_j es la delta de Kronecker.

Ejercicio 1.9.1. Muestre que los productos tensoriales (1.77) forman una base del espacio tensorial V_r^s , y que (1.78) es un tensor de tipo (r, s) .

Ejercicio 1.9.2. Muestre que $I = \delta_j^i \vec{e}_i \otimes \vec{e}^j$.

1.9.5. Cambio de base

El comportamiento de las componentes de un tensor ante un cambio de base en los espacios V y V^* puede deducirse de la misma forma en que lo hicimos para las componentes de vectores, formas y operadores lineales. La idea central es, nuevamente, que un tensor tiene existencia independiente de las bases en que expresamos sus componentes.

Así, si $\{\vec{e}_i\}_{i=1}^n$ y $\{\vec{e}'_j\}_{j=1}^n$ son bases de V con

$$\vec{e}'_j = \vec{e}_i A^i{}_j,$$

y $\{\vec{e}^i\}_{i=1}^n$, $\{\vec{e}^j\}_{j=1}^n$ son sus respectivas bases duales (bases de V^*), el mismo tensor S podrá expresarse en cualquiera de estas bases, como

$$S = \vec{e}_{i_1} \otimes \cdots \otimes \vec{e}_{i_r} \otimes \vec{e}^{j_1} \otimes \cdots \otimes \vec{e}^{j_s} S_{j_1 \dots j_s}^{i_1 \dots i_r} \quad (1.79)$$

$$= \vec{e}'_{k_1} \otimes \cdots \otimes \vec{e}'_{k_r} \otimes \vec{e}^{l_1} \otimes \cdots \otimes \vec{e}^{l_s} S_{l_1 \dots l_s}^{k_1 \dots k_r}. \quad (1.80)$$

De aquí resulta

$$S_{l_1 \dots l_s}^{k_1 \dots k_r} = [\mathbf{A}^{-1}]_{i_1}^{k_1} \dots [\mathbf{A}^{-1}]_{i_r}^{k_r} A^{j_1}{}_{l_1} \dots A^{j_s}{}_{l_s} S_{j_1 \dots j_s}^{i_1 \dots i_r}, \quad (1.81)$$

mostrando que S es s veces covariante y r veces contravariante.

Esto nos permite introducir una definición alternativa de tensor: Cualquier objeto con $r + s$ índices, cada uno que corre de 1 a $\dim(V)$, y que ante un cambio de base $\vec{e}'_j = \vec{e}_i A^i{}_j$ en V transforme de acuerdo a (1.81), lo llamaremos tensor de rango $r+s$, s veces covariante (índices abajo) y r veces contravariante (índices arriba). Esta definición práctica es la que usaremos en adelante.

Ejemplo 1.9.5. Un vector $\vec{u} = u^j \vec{e}_j$ con componentes u^j , es un tensor de rango 1 una vez contravariante, también llamado vector contravariante (o simplemente *vector*).

Ejemplo 1.9.6. Una 1-forma $\bar{\omega} = \omega_i \vec{e}^i$ con componentes ω_i , es un tensor de rango 1 una vez covariante, también llamado vector covariante o *covector*.

Ejemplo 1.9.7. Un operador lineal $\mathcal{A} = A^i{}_j \vec{e}_i \otimes \vec{e}^j$ con componentes $A^i{}_j$, es un tensor de rango 2 *mixto*, una vez covariante (j) y una vez contravariante (i) .

1.9.6. Contracción de índices

Ya hemos introducido el producto tensorial, que nos permite construir un tensor de rango mayor a partir de dos o más tensores de rangos menores. La

operación contraria, que nos permite obtener un tensor de rango menor a partir de otro de rango mayor, es la *contracción de índices*.

Si $S \in V_r^s$ es un tensor de tipo (r, s) con componentes $S_{j_1 \dots j_s}^{i_1 \dots i_r}$, definimos su *contraído* respecto a los índices i_n y j_m como

$$S_{j_1 \dots j_{m-1} \textcolor{red}{i}_n j_{m+1} \dots j_s}^{i_1 \dots i_{n-1} \textcolor{red}{i}_n i_{n+1} \dots i_r} = S_{j_1 \dots j_s}^{i_1 \dots i_r} \delta_{j_m}^{j_m} {}_{i_n}.$$

De acuerdo a la ley de transformación (1.81), éste será un tensor de tipo $(r - 1, s - 1)$ y rango $r + s - 2$.

Ejemplo 1.9.8. Sea $T = T_j^i \vec{e}_i \otimes \vec{e}^j$ un tensor mixto de rango 2. Su *traza*, definida como

$$\mathrm{Tr}(T) := T_i^i,$$

es un *escalar* (tensor de rango 0).

Es fundamental aclarar que el contraído de un tensor será a su vez un tensor, si y sólo si los índices que se contraen son tomados de a pares, uno covariante y el otro contravariante. Por ejemplo, si S_{ij} son las componentes de un tensor covariante de rango 2, S_{ii} no es un tensor (escalar, en este caso), ya que no resulta invariante ante cambios de base.

Ejercicio 1.9.3. Verifique que S_{ii} no es un escalar, aunque S_{ij} sea un tensor.

1.9.7. Simetría

Una propiedad importante que pueden tener los tensores, y que ya hemos conocido a través de las matrices, es la simetría o antisimetría.

Sea $S \in V_r^s$ un tensor de tipo (r, s) con componentes $S_{j_1 \dots j_s}^{i_1 \dots i_r}$. Diremos que S es *simétrico* respecto a los índices i_n, i_m si

$$S_{j_1 \dots j_s}^{i_1 \dots \textcolor{blue}{i}_m \dots \textcolor{red}{i}_n \dots i_r} \equiv S_{j_1 \dots j_s}^{i_1 \dots i_n \dots \textcolor{blue}{i}_m \dots i_r},$$

es decir si sus componentes son invariantes respecto del intercambio de los índices i_n, i_m . De igual modo, le llamaremos simétrico respecto a los índices j_n, j_m si

$$S_{j_1 \dots \textcolor{blue}{j}_m \dots \textcolor{red}{j}_n \dots j_s}^{i_1 \dots i_r} \equiv S_{j_1 \dots \textcolor{red}{j}_n \dots \textcolor{blue}{j}_m \dots j_s}^{i_1 \dots i_r}.$$

Si un tensor es simétrico respecto a cualquier par de índices, le llamaremos simplemente *simétrico*.

Diremos que S es *antisimétrico* respecto a los índices i_n, i_m si

$$S_{j_1 \dots j_s}^{i_1 \dots \textcolor{blue}{i}_m \dots \textcolor{red}{i}_n \dots i_r} \equiv -S_{j_1 \dots j_s}^{i_1 \dots i_n \dots \textcolor{blue}{i}_m \dots i_r},$$

es decir si sus componentes cambian de signo ante el intercambio de los índices i_n, i_m . De igual modo, le llamaremos antisimétrico respecto a los índices j_n, j_m si

$$S_{j_1 \dots \textcolor{blue}{j}_m \dots \textcolor{red}{j}_n \dots j_s}^{i_1 \dots i_r} \equiv -S_{j_1 \dots \textcolor{red}{j}_n \dots \textcolor{blue}{j}_m \dots j_s}^{i_1 \dots i_r}.$$

Si $S \in V_r^0$, le llamaremos *totalmente antisimétrico* si

$$S^{\Pi(i_1, \dots, i_r)} = \text{sgn}(\Pi) S^{i_1 \dots i_r},$$

donde $\Pi(i_1, \dots, i_r)$ es una permutación cualquiera del conjunto ordenado de índices $\{i_1, \dots, i_r\}$, y $\text{sgn}(\Pi)$ es su signo (+ si es una permutación par y - si es impar). De igual modo, si $S \in V_0^s$, le llamaremos *totalmente antisimétrico* si

$$S_{\Pi(j_1, \dots, j_s)} = \text{sgn}(\Pi) S_{j_1 \dots j_s}.$$

Notemos que definimos la simetría o antisimetría de un tensor solamente respecto al intercambio de pares de índices ambos covariantes o ambos contravariantes. Ello es así porque de este modo la simetría o antisimetría se preseva ante cambios de base. Si la definiéramos respecto de índices uno covariante y el otro contravariante, cambiaría con la base en que representáramos el tensor, y no podríamos considerarla una propiedad del tensor, sino sólo de su representación en alguna base particular.

Ejercicio 1.9.4. Verifique que $T_j^i = T_i^j$ no implica $T'^i_j = T'^j_i$.

Dos operaciones relacionadas con las definiciones anteriores son la simetrización y antisimetrización de tensores. Dado un tensor $T \in V_r^0$ de tipo $(r, 0)$ con componentes $T^{i_1 \dots i_r}$, definimos su *parte simétrica* como el tensor

$$\mathcal{S}T \in V_r^0$$

con componentes

$$T^{\{i_1 \dots i_r\}} := \frac{1}{r!} \sum_{\Pi} T^{\Pi(i_1, \dots, i_r)},$$

donde la suma corre sobre todas las permutaciones de los índices. Asimismo, definimos su *parte antisimétrica* como el tensor

$$\mathcal{A}T \in V_r^0$$

con componentes

$$T^{[i_1 \dots i_r]} := \frac{1}{r!} \sum_{\Pi} \text{sgn}(\Pi) T^{\Pi(i_1, \dots, i_r)}$$

Definiciones y notación completamente análogas valen para tensores $T \in V_0^s$ de tipo $(0, s)$.

Las operaciones de simetrización y antisimetrización pueden extenderse a conjuntos arbitrarios de índices, mientras sean todos del mismo tipo (covariantes o contravariantes). Así por ejemplo

$$T_k^{[ij]} = \frac{1}{2} [T_k^{ij} - T_k^{ji}], \quad S_{j\{kl\}}^i = \frac{1}{2} [S_{jkl}^i + S_{jlk}^i], \quad \text{etc.}$$

1.9.8. Producto exterior

Cuando tratamos con vectores, nos es familiar la noción de producto vectorial entre dos vectores, que es otro vector³⁰. Un concepto relacionado a él, y hasta cierto punto similar, es el de producto exterior entre dos tensores.

Dado un tensor $S \in V_0^s$ con componentes $S_{j_1 \dots j_s}$, totalmente antisimétrico, y otro tensor $T \in V_0^t$ con componentes $T_{j_1 \dots j_t}$, también totalmente antisimétrico, su *producto exterior*

$$S \wedge T := \frac{(s+t)!}{s!t!} \mathcal{A}(S \otimes T)$$

es un nuevo tensor $\in V_0^{s+t}$ totalmente antisimétrico, con componentes

$$S_{[j_1 \dots j_s} T_{l_1 \dots l_t]}.$$

Una definición completamente equivalente vale para tensores contravariantes. El producto exterior posee las siguientes propiedades:

1. $S \wedge (T_1 + T_2) = S \wedge T_1 + S \wedge T_2$.
2. $S \wedge (T \wedge R) = (S \wedge T) \wedge R = S \wedge T \wedge R$.
3. $S \wedge T = (-1)^{st} T \wedge S$.

Ejercicio 1.9.5. Demuestre las propiedades precedentes.

Ejemplo 1.9.9. El producto exterior entre dos vectores \vec{u} y \vec{v} de \mathbb{R}^3 será el tensor con componentes

$$u^{[i} v^{j]} = \frac{1}{2} [u^i v^j - u^j v^i],$$

que podemos representar por la matriz antisimétrica

$$\frac{1}{2} \begin{pmatrix} 0 & u^1 v^2 - u^2 v^1 & u^1 v^3 - u^3 v^1 \\ u^2 v^1 - u^1 v^2 & 0 & u^2 v^3 - u^3 v^2 \\ u^3 v^1 - u^1 v^3 & u^3 v^2 - u^2 v^3 & 0 \end{pmatrix}.$$

Notemos que $\vec{u} \wedge \vec{u} = 0$.

1.9.9. Densidades tensoriales

Existen muchos objetos que ante cambios de base no transforman exactamente como tensores, sino de una forma estrechamente relacionada, y que tienen significado físico. El producto vectorial entre dos vectores es un ejemplo; las densidades de masa y carga son otro.

Todo objeto con $r+s$ índices, cada uno que corre de 1 a $\dim(V)$, y que ante un cambio de base $\vec{e}'_j = \vec{e}_i A^i_j$ en V transforme de acuerdo a

$$S'^{k_1 \dots k_r}_{l_1 \dots l_s} = \det(\mathbf{A})^p [\mathbf{A}^{-1}]^{k_1}_{i_1} \dots [\mathbf{A}^{-1}]^{k_r}_{i_r} A^{j_1}_{l_1} \dots A^{j_1}_{l_1} S^{i_1 \dots i_r}_{j_1 \dots j_s}, \quad (1.82)$$

se llama *densidad tensorial* o pseudotensor de peso p .

³⁰En realidad, como veremos en seguida, es un pseudovector

Ejemplo 1.9.10. Las densidades de carga eléctrica y de masa no son escalares, sino pseudoescalares de peso 1. Análogamente, el diferencial de volumen es un pseudoescalar de peso -1 . Estos resultados serán discutidos en mas detalle en la Sección siguiente.

Ejemplo 1.9.11. El tensor completamente antisimétrico de Levi-Civitta se define como

$$\epsilon_{j_1 \dots j_s} = \begin{cases} 1 & \text{si } \{j_1 \dots j_s\} \text{ es permutación par de } \{1, \dots, s\}, \\ -1 & \text{si } \{j_1 \dots j_s\} \text{ es permutación impar de } \{1, \dots, s\}, \\ 0 & \text{para todo otro caso,} \end{cases}$$

y es una densidad tensorial (contravariante) de peso -1 y rango $s = \dim(V)$. Alternativamente podemos definir $\epsilon^{i_1 \dots i_r}$, cuyas componentes coinciden numéricamente con las del anterior, pero que es una densidad tensorial (contravariante) de peso 1.

Demostraremos la afirmación precedente para $s = 2$. En ese caso $\epsilon_{12} = 1$, $\epsilon_{21} = -1$, y $\epsilon_{11} = \epsilon_{22} = 0$. Luego

$$\epsilon_{kl} A^k{}_i A^l{}_j = A^1{}_i A^2{}_j - A^2{}_i A^1{}_j = \begin{cases} \det(\mathbf{A}) & \text{si } i = 1, j = 2, \\ -\det(\mathbf{A}) & \text{si } i = 2, j = 1, \\ 0 & \text{si } i = j, \end{cases}$$

de donde es inmediato que

$$\det(\mathbf{A})^{-1} \epsilon_{kl} A^k{}_i A^l{}_j = \begin{cases} 1 & \text{si } \{i, j\} \text{ es permutación par de } \{1, 2\}, \\ -1 & \text{si } \{i, j\} \text{ es permutación impar de } \{1, 2\}, \\ 0 & \text{para todo otro caso.} \end{cases}$$

Ejercicio 1.9.6. Demuestre que $\epsilon'_{j_1 \dots j_s} = \det(\mathbf{A}) \epsilon_{j_1 \dots j_s}$ y $\epsilon'^{i_1 \dots i_s} = \det(\mathbf{A})^{-1} \epsilon^{i_1 \dots i_s}$.

Ejercicio 1.9.7. Demuestre que el producto tensorial de dos densidades tensoriales de pesos p_1 y p_2 , es otra densidad tensorial de peso $p_1 + p_2$.

Ejercicio 1.9.8. Demuestre que la contracción de índices no altera el peso de una densidad tensorial, sino sólo su rango.

1.9.10. Tensor adjunto

Un concepto íntimamente relacionado a los de densidades tensoriales y producto tensorial, y que frecuentemente aparece en conjunción con ellos, es el de tensor adjunto o dual.

Dado $\dim(V) = n$ y un tensor completamente antisimétrico $T^{i_1 \dots i_r}$ de rango $r \leq n$, definimos el tensor *adjunto*³¹ (o dual) como

$$\bar{T}_{i_1 \dots i_{n-r}} := \epsilon_{i_1 \dots i_{n-r} j_1 \dots j_r} T^{j_1 \dots j_r}, \quad (1.83)$$

³¹No confundir con la adjunta Hermitiana.

es decir, contraemos los índices de $T^{i_1 \dots i_r}$ con los últimos r índices del tensor de Levi-Civitta. El tensor adjunto es también totalmente antisimétrico, y su rango es $n - r$. Sin embargo, si $T^{i_1 \dots i_r}$ era un tensor, su adjunto será una densidad tensorial.

Definiciones completamente análogas valen para tensores covariantes e incluso para densidades tensoriales.

Ejemplo 1.9.12. El producto vectorial usual es el adjunto del producto exterior (ver Ejemplo 1.9.9) de dos vectores \vec{u} y \vec{v} de \mathbb{R}^3 ,

$$\epsilon_{ijk} u^{[j} v^{k]} = \begin{cases} u^2 v^3 - u^3 v^2 & \text{si } i = 1, \\ u^3 v^1 - u^1 v^3 & \text{si } i = 2, \\ u^1 v^2 - u^2 v^1 & \text{si } i = 3, \end{cases}$$

y es una densidad tensorial de peso -1 o *pseudovector*.

1.9.11. Ley del cociente

Este es un procedimiento rápido y práctico para reconocer si un objeto dado es un tensor. Lo ilustraremos con un ejemplo.

Supongamos que sabemos que \vec{u} es un vector con componentes u^i , y determinamos que $T^i_j u^j$ es un vector $\forall \vec{u}$. Entonces T es un tensor, en este caso de rango 2 y mixto.

Un procedimiento análogo vale para tensores de rango superior, y para densidades tensoriales.

1.10. Coordenadas curvilíneas

Hasta ahora hemos venido considerando cambios de coordenadas de la forma $\vec{e}'_j = \vec{e}_i A^i_j$ con A^i_j constantes, es decir cambios de coordenadas *globales*. Estos incluyen rotaciones, reflexiones y permutaciones de la base, y cambios entre sistemas Cartesianos no necesariamente ortogonales. Pero todos estos cambios de coordenadas comparten una característica: si los vectores de la vieja base eran paralelos a sí mismos en cada punto del espacio (es decir, si era una base Cartesiana), los de la nueva base también lo serán.

Sin embargo, en Física estamos acostumbrados a usar sistemas de coordenadas no Cartesianos, las llamadas *coordenadas curvilíneas*, que muchas veces resultan más adecuados a la geometría del problema que estamos estudiando. Estos sistemas se caracterizan por el hecho de que los vectores base (que en general se toman como versores tangentes a las curvas coordenadas) cambian de dirección de un punto a otro del espacio. Por lo tanto, cualquier cambio de coordenadas que involucre coordenadas curvilíneas será *local*, es decir, los elementos A^i_j de la matriz de transformación serán funciones de punto.

En esta Sección consideraremos las consecuencias de realizar cambios de coordenadas locales, es decir de la forma

$$\vec{e}'_j = \vec{e}_i A^i_j(\vec{x}).$$

1.10.1. Cambios de coordenadas locales

Un cambio de coordenadas cualquiera en V [$\dim(V) = n$] queda completamente caracterizado dando, por ejemplo, la expresión de las nuevas coordenadas x'^i como funciones de las viejas coordenadas x^j , es decir

$$x'^i = x'^i(x^1, \dots, x^n), \quad i = 1, \dots, n. \quad (1.84)$$

En el caso de cambios de coordenadas globales, las funciones $x'^i(x^1, \dots, x^n)$ eran lineales en todos sus argumentos, con una matriz de coeficientes no singular. En el caso de cambios de coordenadas locales, consideraremos admisible cualquier forma funcional para las x'^i , siempre y cuando podamos invertir estas funciones para despejar

$$x^j = x^j(x'^1, \dots, x'^n), \quad j = 1, \dots, n. \quad (1.85)$$

La condición necesaria y suficiente para ésto es que el determinante Jacobiano de la transformación no se anule,

$$J := \left| \frac{\partial(x^1, \dots, x^n)}{\partial(x'^1, \dots, x'^n)} \right| = \left| \frac{\partial x^i}{\partial x'^j} \right| \neq 0. \quad (1.86)$$

Si esto ocurre, diremos que la transformación es *no singular* (las nuevas coordenadas están “bien definidas” en términos de las viejas).

Ejemplo 1.10.1. Sean x^1 y x^2 coordenadas Cartesianas usuales en \mathbb{R}^2 , y sean x'^1 y x'^2 coordenadas polares (en notación usual, $x^1 = x$, $x^2 = y$, $x'^1 = \rho$ y $x'^2 = \theta$). Entonces

$$x'^1 = \sqrt{(x^1)^2 + (x^2)^2}, \quad x'^2 = \arctan \frac{x^2}{x^1},$$

y

$$x^1 = x'^1 \cos x'^2, \quad x^2 = x'^1 \sin x'^2.$$

Notemos que el Jacobiano de la transformación,

$$J = \left| \frac{\partial(x^1, x^2)}{\partial(x'^1, x'^2)} \right| = \det \begin{pmatrix} \cos x'^2 & \sin x'^2 \\ -x'^1 \sin x'^2 & x'^1 \cos x'^2 \end{pmatrix} = x'^1,$$

sólo se anula para $x'^1 = 0$, hecho que conocemos como “el sistema de coordenadas polares es singular en el origen”.

Las ecuaciones (1.84) y (1.85) pueden ser altamente no lineales, lo cual generalmente dificulta trabajar con ellas. Sin embargo, la relación entre los *diferenciales* de las coordenadas viejas y nuevas,

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j, \quad (1.87)$$

es *siempre* lineal y homogénea, aunque los coeficientes puedan depender de las coordenadas de manera no lineal. Notando que si las relaciones (1.84) son lineales y tienen la forma

$$x'^i = a^i_j x^j$$

la relación entre diferenciales se reduce a

$$dx'^i = a^i_j dx^j,$$

vemos que la generalización a coordenadas curvilíneas de las definiciones de tensores, densidades, etc., dadas en la Sección precedente, puede hacerse simplemente reemplazando la matriz de transformación \mathbf{A} de los cambios de coordenadas globales, por la matriz de derivadas parciales (matriz Jacobiana) \mathbf{J} , es decir haciendo los reemplazos ³²

$$[\mathbf{A}^{-1}]^i_j \leftrightarrow \frac{\partial x'^i}{\partial x^j} =: \mathbf{J}^{-1}, \quad A^i_j \leftrightarrow \frac{\partial x^i}{\partial x'^j} =: \mathbf{J}, \quad \det(\mathbf{A}) \leftrightarrow \det(\mathbf{J}) =: J. \quad (1.88)$$

Esta identificación, junto con la linealidad de la relación entre los diferenciales, simplificarán notablemente los cálculos subsiguientes.

1.10.2. Vectores contravariantes

Definiremos como vector *contravariante* a cualquier objeto con componentes u^i que, ante cambios de coordenadas, transformen de acuerdo a

$$u'^i = \frac{\partial x'^i}{\partial x^j} u^j, \quad (1.89)$$

que es la ley de transformación de los diferenciales de coordenadas.

Ejemplo 1.10.2. Son vectores contravariantes: las componentes del vector desplazamiento $d\vec{x}$ (por definición); las del vector velocidad $\frac{d\vec{x}}{dt}$; etc.

1.10.3. Vectores covariantes

Sea $\varphi(\vec{x})$ una función de punto (escalar); entonces por la regla de la cadena

$$\frac{\partial \varphi}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i} \frac{\partial \varphi}{\partial x^j}.$$

Definiremos como vector *covariante* a cualquier objeto con componentes u_i que, ante cambios de coordenadas, transformen de acuerdo a

$$u'_i = \frac{\partial x^j}{\partial x'^i} u_j, \quad (1.90)$$

que es la inversa de la ley de transformación de los diferenciales de coordenadas.

Ejemplo 1.10.3. El gradiente de una función escalar de punto, $\nabla \varphi$, es un vector covariante con componentes

$$\varphi_{,i} := \frac{\partial \varphi}{\partial x^i}.$$

Son vectores covariantes: fuerza, campo eléctrico, y todo otro que se obtenga como el gradiente de un potencial escalar.

³²La justificación rigurosa de esta identificación pasa por *definir* los vectores de la base curvilínea como tangentes a las curvas coordenadas. Geométricamente los identificamos con los vectores desplazamiento $d\vec{x}$ obtenidos variando cada coordenada con todas las demás fijas; analíticamente los identificamos con las formas diferenciales $\frac{\partial}{\partial x^i}$.

1.10.4. Tensores

Análogamente a las definiciones anteriores, un objeto con componentes $S_{j_1 \dots j_s}^{i_1 \dots i_r}$ que ante cambios de coordenadas transforme de acuerdo a

$$S'_{l_1 \dots l_s}^{k_1 \dots k_r} = \frac{\partial x'^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial x'^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial x'^{l_1}} \cdots \frac{\partial x^{j_s}}{\partial x'^{l_s}} S_{j_1 \dots j_s}^{i_1 \dots i_r}, \quad (1.91)$$

es un *tensor* de rango $r+s$, r veces contravariante (índices arriba) y s veces covariante (índices abajo).

Ejemplo 1.10.4. La delta de Kronecker puede definirse como

$$\delta^i_j := \frac{\partial x^i}{\partial x^j},$$

y es un tensor mixto de rango 2

Ejemplo 1.10.5. Sean u^i las componentes de un vector contravariante, y v_j las de un vector covariante. Entonces $u^i v_j$ son las componentes de un tensor mixto de rango 2 (su producto tensorial), y el contraído $u^i v_i$ es un tensor de rango 0 (escalar), su producto interno. Notemos que ni $u^i v^i$ ni $u_i v_i$ son escalares (dependen de la base). Por ejemplo con fuerzas conservativas, la velocidad de un móvil es un vector contravariante con componentes v^i , y la fuerza aplicada es un vector covariante (ya que es el gradiente de un potencial escalar) con componentes f_i , de modo que la potencia

$$W := f_i v^i$$

es un escalar.

1.10.5. Densidades tensoriales

Un objeto con componentes $S_{j_1 \dots j_s}^{i_1 \dots i_r}$ que ante cambios de coordenadas transforme de acuerdo a

$$S'_{l_1 \dots l_s}^{k_1 \dots k_r} = J^p \frac{\partial x'^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial x'^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial x'^{l_1}} \cdots \frac{\partial x^{j_s}}{\partial x'^{l_s}} S_{j_1 \dots j_s}^{i_1 \dots i_r}, \quad (1.92)$$

es una *densidad tensorial* de peso p y rango $r+s$, r veces contravariante (índices arriba) y s veces covariante (índices abajo).

Ejemplo 1.10.6. El tensor completamente antisimétrico de Levi-Civitta $\epsilon_{i_1 \dots i_n}$ es una densidad tensorial (covariante) de peso -1 , y $\epsilon^{i_1 \dots i_n}$ es una densidad tensorial (contravariante) de peso 1 .

Ejemplo 1.10.7. El producto vectorial $\epsilon_{ijk} u^{[j} v^{k]}$ entre dos vectores contravariantes en \mathbb{R}^3 es una densidad tensorial de rango 1 y peso -1 , o pseudovector (covariante).

1.10.6. Integrales multidimensionales

En \mathbb{R}^3 y en coordenadas Cartesianas ortonormales estamos acostumbrados a escribir una integral de volumen como

$$\int dV = \int dx dy dz,$$

escribiendo el diferencial de volumen dV simplemente como $dx dy dz$. Sin embargo esta expresión es engañosa: debemos tener presente que geométricamente dV representa el volumen (infinitesimal) del paralelepípedo generado por tres vectores desplazamiento LI, y bajo cambios de coordenadas transformará de manera muy diferente al producto de los diferenciales dx , dy y dz .

La definición de diferencial de volumen que toma en cuenta adecuadamente estas consideraciones se basa en tomar n [= dim(V)] vectores desplazamiento $d\vec{x}_{(1)}, \dots, d\vec{x}_{(n)}$ LI, formar su producto exterior y tomar el tensor adjunto,

$$dV := \overline{d\vec{x}_{(1)} \wedge \dots \wedge d\vec{x}_{(n)}} \equiv \epsilon_{i_1 \dots i_n} dx_{(1)}^{[i_1} \dots dx_{(n)}^{i_n]} \quad (1.93)$$

donde $dx_{(k)}^{i_j}$, $i_j = 1, \dots, n$, son las componentes (contravariantes) de cada vector desplazamiento $d\vec{x}_{(k)}$.

Puede verificarse que, si construimos una matriz cuadrada $n \times n$ cuyas columnas sean las componentes de los $d\vec{x}_{(k)}$, la definición anterior hace que

$$dV = \det \begin{pmatrix} d\vec{x}_{(1)} & \cdots & d\vec{x}_{(n)} \\ \downarrow & & \downarrow \end{pmatrix},$$

que es justamente el volumen del paralelepípedo subtendido por $d\vec{x}_{(1)}, \dots, d\vec{x}_{(n)}$. Podemos interpretar así la definición (1.93) como una “receta” algebraica para realizar esta construcción geométrica: El producto tensorial de los $d\vec{x}_{(k)}$ genera un tensor de rango n cuyas componentes son todos los productos de la forma $dx_{(1)}^{i_1} \dots dx_{(n)}^{i_n}$, con $i_1, \dots, i_n = 1, \dots, n$. La antisimetrización selecciona de entre éstos los que corresponden a i_1, \dots, i_n todos distintos, les asigna signos alternantes y divide por $n!$. Finalmente, la contracción con $\epsilon_{i_1 \dots i_n}$ produce el determinante en cuestión.

Ejemplo 1.10.8. Construyamos el diferencial de volumen en \mathbb{R}^2 (diferencial de área): El producto tensorial de $d\vec{x}_{(1)}$ y $d\vec{x}_{(2)}$ tiene componentes $dx_{(1)}^i dx_{(2)}^j$ con $i, j = 1, 2$, que podemos ordenar en la matriz

$$\begin{pmatrix} dx_{(1)}^1 dx_{(2)}^1 & dx_{(1)}^1 dx_{(2)}^2 \\ dx_{(1)}^2 dx_{(2)}^1 & dx_{(1)}^2 dx_{(2)}^2 \end{pmatrix}.$$

Antisimetrizando, podemos ordenar las componentes del tensor producto exterior $d\vec{x}_{(1)} \wedge d\vec{x}_{(2)}$ en la matriz

$$\frac{1}{2} \begin{pmatrix} 0 & dx_{(1)}^1 dx_{(2)}^2 - dx_{(1)}^2 dx_{(2)}^1 \\ dx_{(1)}^2 dx_{(2)}^1 - dx_{(1)}^1 dx_{(2)}^2 & 0 \end{pmatrix}.$$

Finalmente, contrayendo con el tensor de Levi-Civitta de rango 2 obtenemos

$$dV = dx_{(1)}^1 dx_{(2)}^2 - dx_{(1)}^2 dx_{(2)}^1,$$

que es justamente el área del paralelogramo subtendido por $d\vec{x}_{(1)}$ y $d\vec{x}_{(2)}$.

Nótese que la expresión obtenida es justamente el determinante de

$$\begin{pmatrix} dx_{(1)}^1 & dx_{(1)}^2 \\ dx_{(2)}^1 & dx_{(2)}^2 \end{pmatrix}.$$

Nótese también que si $d\vec{x}_{(1)} = dx\vec{e}_1$ y $d\vec{x}_{(2)} = dy\vec{e}_2$, con $\{\vec{e}_1, \vec{e}_2\}$ una base ortonormal, esta expresión se reduce a $dV = dx dy$.

Ejercicio 1.10.1. Construya la expresión equivalente para el diferencial de volumen en \mathbb{R}^3 .

De lo anterior resulta claro que el diferencial de volumen no será un escalar, sino una densidad escalar de peso -1 .

Ejercicio 1.10.2. Demuestre la afirmación precedente.

Ejemplo 1.10.9. Sea ρ la densidad de carga eléctrica; si la carga total Q contenida en un volumen dado debe ser un escalar (invariante ante transformaciones de coordenadas), es decir si pedimos

$$Q = \int \rho dV = \int \rho' dV',$$

vemos que ρ debe ser un pseudotensor de rango 0 y peso 1, es decir una densidad escalar de peso 1.

El razonamiento por el que construimos el diferencial de volumen también nos permite decidir cómo construir otros diferenciales multidimensionales. Por ejemplo, el diferencial de área orientado en \mathbb{R}^3 debe ser un vector normal a la superficie y de módulo igual al área del paralelepípedo subtendido por dos vectores desplazamiento LI tangentes a ella. Esto nos sugiere definirlo como el producto vectorial de dichos vectores desplazamiento,

$$dA_i = \left[\overline{d\vec{x}_{(1)} \wedge d\vec{x}_{(2)}} \right]_i = \epsilon_{ijk} dx_{(1)}^{[j]} dx_{(2)}^{k]},$$

que no es un vector sino un pseudovector. El flujo de un vector contravariante a través de una superficie vendrá dado entonces por

$$\int u^i dA_i,$$

y si debe ser invariante ante cambios de coordenadas, en realidad u^i debe ser una densidad vectorial de peso 1.

Generalizando, si $\dim(V) = n$, el diferencial de “área” sobre una “superficie” de dimensión $n - 1$ será el pseudovector de componentes

$$dA_j = \left[\overline{d\vec{x}_{(1)} \wedge \cdots \wedge d\vec{x}_{(n-1)}} \right]_j = \epsilon_{j i_1 \dots i_{n-1}} dx_{(1)}^{[i_1} \dots dx_{(n-1)}^{i_{n-1}]}, \quad (1.94)$$

donde los vectores desplazamiento son LI, y nos permitirá calcular el “flujo” de una densidad vectorial de peso 1 como antes. Análogamente, el diferencial de “área” sobre una “superficie” de dimensión $n - r$ será el pseudotensor de peso -1 y componentes

$$dA_{j_1 \dots j_r} = \left[\overline{d\vec{x}_{(1)} \wedge \cdots \wedge d\vec{x}_{(n-r)}} \right]_{j_1 \dots j_r} = \epsilon_{j_1 \dots j_r i_1 \dots i_{n-r}} dx_{(1)}^{[i_1} \dots dx_{(n-r)}^{i_{n-r}]},$$

y deberemos integrarlo con una densidad tensorial $u^{j_1 \dots j_r}$ de peso 1.

El lector atento habrá notado que no hemos intentado definir ningún “ente” como, por ejemplo en \mathbb{R}^3 , la integral de una densidad (escalar) de superficie, con un diferencial de área “escalar” (no orientado). Ello es porque cualquier definición consistente del diferencial de área, donde pidamos que ante cambios de coordenadas transforme como un tensor (o escalar, o vector) o como una densidad tensorial, *siempre* nos llevará a la definición (1.94). Si queremos definir, por ejemplo, la integral de una densidad superficial (de carga, masa, etc.) ρ_s sobre una superficie en \mathbb{R}^3 definida por la ecuación $\sigma(\vec{x}) = 0$, primero deberemos convertirla en la densidad *volumétrica*

$$\rho = \rho_s \delta(\sigma(\vec{x})),$$

donde δ es la delta de Dirac,³³ y luego integrar en el volumen,

$$\int \rho_s \delta(\sigma(\vec{x})) dV.$$

El hecho de que, cuando $\sigma(\vec{x}) = 0$ es una superficie coordinada, la expresión pueda reducirse a una integral de la forma

$$\int \rho_s du dv,$$

donde u y v son las restantes coordenadas (ortogonales) sobre la superficie, es tan engañoso como lo era el que, a veces, pudiéramos escribir $\int dV = \int dx dy dz$.

Ejemplo 1.10.10. Integremos la densidad superficial de carga $\rho_s = Q/4\pi r_0^2$ sobre la superficie definida en coordenadas esféricas por $\sigma(\vec{x}) = r - r_0$. En este caso tendremos

$$\int \rho_s \delta(\sigma(\vec{x})) dV = \int \frac{Q}{4\pi r_0^2} \delta(r - r_0) r^2 \sin \theta dr d\theta d\phi = \frac{Q}{4\pi} \iint \sin \theta d\theta d\phi = Q.$$

³³La delta de Dirac será tratada en detalle más adelante, en el contexto de distribuciones.

1.10.7. Tensor métrico

Ya hemos introducido anteriormente la *métrica* a través del producto interno. Ahora veremos cómo ese concepto se adapta al entorno de las coordenadas curvilíneas. La principal diferencia estriba en que ahora la métrica puede no ser constante, sino función de punto.

Sea ds el *diferencial de arco* (o diferencial de longitud), y supongamos que, en las coordenadas que estamos usando,

$$ds^2 = g_{ij}dx^i dx^j, \quad (1.95)$$

con g_{ij} funciones continuas y diferenciables de las coordenadas x^i . Asumiremos $g_{ij} = g_{ji}$ y $\det[g_{ij}] \neq 0$ y definiremos

$$g := \det[g_{ij}], \quad g^{ij} := \frac{\text{cof } g_{ij}}{g}. \quad (1.96)$$

Entonces

$$g^{ij}g_{jk} = \delta_k^i. \quad (1.97)$$

Como ds debe ser un invariante (escalar), y dx^i es un vector contravariante, la ley del cociente nos asegura que g_{ij} es un tensor covariante de rango 2, conocido como *tensor métrico* o simplemente *métrica*. Asimismo, g^{ij} es un tensor contravariante de rango 2, y g es un pseudoescalar de peso 2: $g' = J^2 g$.

1.10.8. Ascenso y descenso de índices

Vimos anteriormente cómo el producto interno puede utilizarse para crear un mapa entre un EV V y su dual V^* . Utilizando la métrica podemos explicitar esa correspondencia.

Sea un vector contravariante con componentes u^i ; sus componentes covariantes se definen como

$$u_i := g_{ij}u^j.$$

Análogamente, dado un vector covariante con componentes v_i , sus componentes contravariantes se definen como

$$v^i := g^{ij}v_j.$$

En general, dado un tensor T^{ij} , definimos $T_j^i := g_{jk}T^{ik}$, $T_{ij} := g_{ik}T^{kl}g_{lj}$, etc.

Decimos así que usamos la métrica para “subir” o “bajar” índices, o sea para “traducir” de componentes covariantes a contravariantes y viceversa.

1.10.9. Producto escalar y norma

Utilizando la métrica para subir o bajar índices, podemos definir ahora el *producto escalar* entre dos vectores covariantes, o entre dos contravariantes, como

$$\vec{u} \cdot \vec{v} := u^i v_i = g^{ij}u_i v_j = g_{ij}u^i v^j. \quad (1.98)$$

La *norma* de un vector vendrá dada por

$$\|\vec{u}\|^2 := \vec{u} \cdot \vec{u} = u^i u_i = g^{ij} u_i u_j = g_{ij} u^i u^j. \quad (1.99)$$

De modo análogo, utilizando la métrica para subir y bajar índices, podemos ahora contraer pares de índices ambos covariantes o contravariantes, y obtener un tensor: por ejemplo, si T^{ij}_k es un tensor,

$$S_k = T^{ij}_k g_{ij}$$

será un vector covariante, etc.

Los sistemas de *coordenadas ortogonales* se definen como aquellos donde la métrica es diagonal: $g_{ij} = 0$ para $i \neq j$; y en ellos

$$\vec{u} \cdot \vec{v} := u^i v_i = g^{ii} u_i v_i = g_{ii} u^i v^i,$$

es decir, no aparecen términos cruzados. Un sistema de coordenadas será *ortogonal* si en él $g_{ij} = \delta_{ij}$, y allí

$$\vec{u} \cdot \vec{v} := u^i v_i = u_i v_i = u^i v^i,$$

es decir, la diferencia entre componentes covariantes y contravariantes se desvanece.

1.10.10. Símbolos de Christoffel

A partir de la métrica se pueden construir diversos entes con propiedades útiles. En particular, definimos el *símbolo de Christoffel de primera especie* como

$$[ij, k] := \frac{1}{2}(g_{ki,j} + g_{jk,i} - g_{ij,k}), \quad (1.100)$$

que *no* es un tensor. El *símbolo de Christoffel de segunda especie* se define como

$$\left\{ \begin{matrix} k \\ ij \end{matrix} \right\} := g^{kl} [ij, l], \quad (1.101)$$

y *tampoco* es un tensor.

Ejercicio 1.10.3. Demuestre que los símbolos de Christoffel no son tensores.

1.10.11. Derivación covariante

Consideremos un vector contravariante con componentes u^i . Sabemos que, ante una transformación de coordenadas,

$$u'^i = \frac{\partial x'^i}{\partial x^k} u^k.$$

Sin embargo, es fácil (aunque algo engoroso) ver que las derivadas del vector transformarán según

$$u'^i_{,j} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^l}{\partial x'^j} u^k_{,l} + \frac{\partial^2 x'^i}{\partial x^l \partial x^k} \frac{\partial x^l}{\partial x'^j} u^k. \quad (1.102)$$

Es evidente, entonces, que las derivadas $u'^i_{,j}$ no son las componentes de un tensor. Esto puede parecer una sorpresa desagradable, ya que frecuentemente tendremos que trabajar con las derivadas de vectores, tensores, etc. Sin embargo, mostraremos en seguida que en realidad esto no debería sorprendernos, y cómo esquivar la dificultad.

El inconveniente proviene de que, en coordenadas curvilíneas, los vectores coordinados cambian de dirección de un punto a otro; por lo tanto, si derivamos un vector escrito como $\vec{u} = u^i \vec{e}_i$, debemos derivar no sólo sus componentes u^i , sino también los vectores \vec{e}_i . Una “receta” para derivar vectores en coordenadas curvilíneas, por tanto, podría ser la siguiente: tome las componentes u^i , multiplíquelas por los vectores base \vec{e}_i y sume; derive la combinación lineal $u^i \vec{e}_i$ (*todos* los términos); y descomponga el resultado nuevamente en la base $\{\vec{e}_i\}$ para obtener las componentes de la derivada del vector. De este complicado procedimiento al menos algo quedaría claro: en general, en coordenadas curvilíneas *las componentes de la derivada no son las derivadas de las componentes*.

Evidentemente sería mucho más cómodo obtener, de una vez y para siempre, expresiones para las componentes de la derivada de un vector en términos de las componentes del vector. Para ello, consideremos de nuevo la derivada de $\vec{u} = u^i \vec{e}_i$, que escribiríamos

$$\vec{u}_{,k} = u^i_{,k} \vec{e}_i + u^j (\vec{e}_j)_{,k}.$$

Ahora bien, $(\vec{e}_j)_{,k}$ debe poder descomponerse en la base \vec{e}_i , es decir que deben existir coeficientes Γ^i_{jk} tales que

$$(\vec{e}_j)_{,k} = \Gamma^i_{jk} \vec{e}_i.$$

Por lo tanto, podremos poner

$$\vec{u}_{,k} = (u^i_{,k} + \Gamma^i_{jk} u^j) \vec{e}_i.$$

Puede mostrarse que los coeficientes Γ^i_{jk} , que son conocidos como los *elementos de la conexión afín*, satisfacen

$$\Gamma^i_{jk} \equiv \left\{ \begin{matrix} i \\ jk \end{matrix} \right\}, \quad (1.103)$$

y que si u^i es un vector contravariante, la cantidad

$$u^i_{,k} := u^i_{,k} + \Gamma^i_{jk} u^j \quad (1.104)$$

es un tensor (mixto), su *derivada covariante*. Un cálculo bastante largo permite ver que Γ^i_{jk} transforma como

$$\Gamma'_{kj}^i = \frac{\partial x'^i}{\partial x^m} \frac{\partial x^l}{\partial x'^k} \frac{\partial x^h}{\partial x'^j} \Gamma_{hl}^m - \frac{\partial^2 x'^i}{\partial x^l \partial x^h} \frac{\partial x^l}{\partial x'^k} \frac{\partial x^h}{\partial x'^j}, \quad (1.105)$$

de donde resulta obvio que Γ_{jk}^i no es un tensor.

Ejercicio 1.10.4. Demuestre que $u^i_{;k}$ es un tensor.

De forma análoga pueden definirse las derivadas covariantes de tensores de rango arbitrario:

- h escalar $\implies h_{;i} \equiv h_{,i}$
- v_i covariante $\implies v_{i;j} = v_{i,j} - \Gamma_{ij}^k v_k$
- u^i contravariante $\implies u^i_{;j} = u^i_{,j} + \Gamma_{kj}^i u^k$
- t_{ij} covariante $\implies t_{ij;k} = t_{ij,k} - \Gamma_{ik}^h t_{hj} - \Gamma_{jk}^h t_{ih}$
- t^{ij} contravariante $\implies t^{ij}_{;k} = t^{ij}_{,k} + \Gamma_{hk}^i t^{hj} + \Gamma_{hk}^j t^{ih}$
- t^i_j mixto $\implies t^i_{j;k} = t^i_{j,k} + \Gamma_{hk}^i t^h_j - \Gamma_{jk}^h t^i_h$
- etc.

Es de destacar que la derivada covariante y la derivada usual de un escalar coinciden. Por otro lado, si una cantidad H es una densidad escalar (pseudoscalar de peso 1), su derivada covariante es

$$H_{;k} = H_{,k} - \Gamma_{hk}^h H, \quad (1.106)$$

que es una densidad vectorial (pseudovector de peso 1). Pueden deducirse fórmulas análogas para las derivadas covariantes de pseudotensores de rango y peso arbitrarios.

También es de destacar que la derivada covariante se defina enteramente en términos de la métrica, aunque si recordamos que la métrica justamente *mide* (permite construir longitudes, productos escalares, etc.), y la derivación involucra el límite de diferencias entre puntos cercanos, no debería extrañarnos.

Por último, de la definición de la conexión afín puede deducirse el importante resultado

$$g_{ij;k} \equiv 0, \quad (1.107)$$

es decir, la métrica es “constante” (*invariante*) bajo derivación covariante (¡aunque en general las g_{ij} sean funciones de punto, no constantes!)

1.10.12. Operadores diferenciales

Podemos ahora introducir las definiciones correctas, y completamente generales, para los operadores diferenciales de uso más común en Física.

Gradiente: Definimos el gradiente de un escalar φ como el vector covariante de componentes

$$\nabla_i \varphi := \varphi_{;i} \equiv \varphi_{,i},$$

que es un vector covariante. Un ejemplo es el campo eléctrico, gradiente del potencial (escalar) en electrostática.

Si φ es un pseudoscalar, aparecen dos diferencias: por un lado, obtendremos un pseudovector; por otro lado, y más importante, la derivada covariante ya no será equivalente a la derivada usual.

Rotor: En un EV de dimensión arbitraria, el rotor de un vector covariante u_i se define como

$$r_{ij} := u_{[i;j]} = u_{i;j} - u_{j;i} = u_{i,j} - u_{j,i},$$

y el de un vector u^i contravariante se define como

$$r_{ij} := (u^k g_{k[i]})_{;j} = (u^k g_{ki})_{;j} - (u^k g_{kj})_{;i} = (u^k g_{ki})_{,j} - (u^k g_{kj})_{,i},$$

y en ambos casos es un tensor covariante antisimétrico de rango 2.

En un espacio tridimensional, podemos obtener un vector (que es la definición de rotor a la que estamos habituados) tomando el tensor adjunto de r_{ij} , es decir definiendo

$$R^k := \frac{1}{2} r_{ij} \epsilon^{jik} \equiv u_{i,j} \epsilon^{jik},$$

y resulta ser una densidad vectorial (pseudovector de peso 1). Un ejemplo es el campo magnético, rotor de un potencial vectorial.

Divergencia: Definimos la divergencia de un vector contravariante u^i como

$$\operatorname{div} \vec{u} := u^i_{;i} = \frac{1}{\sqrt{g}} (\sqrt{g} u^i)_{,i},$$

que es un escalar. Para un vector covariante u_i , primero subimos el índice con la métrica (pasándolo a un vector contravariante) y después derivamos:

$$\operatorname{div} \vec{u} := (g^{ij} u_j)_{;i} = \frac{1}{\sqrt{g}} (\sqrt{g} g^{ij} u_j)_{,i},$$

que también es un escalar.

Laplaciano: Definimos el Laplaciano de un escalar φ como

$$\nabla^2 \varphi := \operatorname{div}(\nabla \varphi) = \frac{1}{\sqrt{g}} (\sqrt{g} g^{ij} \varphi_{,ij})_{,i},$$

que es otro escalar.

1.10.13. Componentes físicas de un vector

Quizá el lector se descubra algo confuso al advertir que, una vez que podemos subir y bajar índices con la métrica, un vector contravariante como ser la velocidad, con componentes v^i , puede ser convertido en un vector covariante con componentes $v_i = g_{ij}v^j$. Es evidente que, salvo en coordenadas Cartesianas ortonormales, donde $g_{ij} \equiv \delta_{ij}$, los valores numéricos de v^i y v_i pueden ser muy diferentes. Si ahora se le pide que *mida* la componente de la velocidad del móvil en la dirección de \vec{e}_i , ¿qué contestaría?

La aparente ambigüedad proviene del hecho que ni las componentes covariantes ni las contravariantes de un vector son lo que usualmente identificamos como sus componentes *geométricas* o *físicas*; estas últimas son las proyecciones (geométricas) del vector sobre las direcciones tangentes a las curvas coordinadas, que son las direcciones de los vectores base \vec{e}_i . Podemos aprovechar esta idea si notamos que la norma de un vector, que usualmente identificamos como su “largo” (el módulo de la velocidad, por ejemplo), *sí* tiene un valor independiente de si es covariante o contravariante:

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v^i v_i} = \sqrt{g^{ij} v_i v_j} = \sqrt{g_{ij} v^i v^j}.$$

Una breve reflexión nos muestra que lo que llamariamos componentes físicas de un vector vienen dadas por

$$U^i := \sqrt{g_{ii}} u^i, \quad V_i := \sqrt{g^{ii}} v_i.$$

Estas cantidades *no son* componentes de vectores covariantes ni contravariantes.

En términos de las componentes físicas, las expresiones de los operadores diferenciales usuales son:

Gradiente:

$$\nabla_i V := \sqrt{g^{ii}} V_i.$$

Rotor:

$$\bar{R}^k := \frac{\sqrt{g^{kk}}}{\sqrt{g}} \left(\frac{V_i}{\sqrt{g^{ii}}} \right)_{,j} \epsilon^{jik}.$$

Divergencia:

$$\operatorname{div} \vec{U} := \frac{1}{\sqrt{g}} \left(\frac{\sqrt{g}}{\sqrt{g_{ii}}} U^i \right)_{,i}.$$

Laplaciano:

$$\nabla^2 U := \operatorname{div}(\nabla U) = \frac{1}{\sqrt{g}} (\sqrt{g} g^{ij} U_{,j})_{,i}$$

(igual que antes).

