CIENCIAS NATURALES E INGENIERÍA

Álgebra lineal, con ejercicios en Python, aplicaciones y notas históricas

JUAN MANUEL
ROMERO SANPEDRO

MATEMÁTICAS APLICADAS Y SISTEMAS

COLECCIÓN LIBRO DE TEXTO

Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

Esta obra es una de las ganadoras de la “Convocatoria para libros de texto como apoyo en la impartición de los programas de estudios”, 2023. Fue evaluada para su publicación por el Consejo Editorial de la UAM Unidad Cuajimalpa, con base en los dictámenes solicitados a pares académicos mediante un esquema que preserva el anonimato mutuo. Estos dictámenes resultaron favorables.

Diseño de portada: Aldo Juárez Herrera

ISBN: 978-607-28-3141-4
ISBN: 978-607-28-3020-2 (Colección)

EDITORIAL

UNIVERSIDAD AUTÓNOMA METROPOLITANA
UNIDAD CUAJIMALPA

DEPARTAMENTO DE MATEMÁTICAS APLICADAS
Y SISTEMAS

Casa abierta al tiempo

Álgebra Lineal, con Ejercicios en Python, Aplicaciones y Notas Históricas.

JUAN MANUEL ROMERO SANPEDRO
jromero@cua.uam.mx

Índice general

Dedicatoria

Antes de la llegada de los españoles al territorio que hoy llamamos México, existían diversas culturas que desarrollaron áreas de las matemáticas como la geometría, el álgebra, la teoría de números, etc. Mucho de ese desarrollo se perdió durante la conquista y el período colonial. En la historia de México, los pueblos indígenas han sido carne de cañón de diferentes luchas como la independencia, la intervención norteamericana, la intervención francesa y la revolución. Si México es independiente de las potencias extrajeras es en gran medida gracias a que los pueblos indígenas metieron las manos al fuego por la libertad de este país. No obstante, en la actualidad los indígenas son la clase más discriminada de México. Así, desde hace varios siglos la gran preocupación de los indígenas se centra en sobrevivir y evitar su desaparición, más que regresar su mirada al desarollo de las ciencias y las matemáticas.

Este libro está dedicado a los pueblos indígenas de México.

Prólogo

La humanidad ha estado interesada en los temas que aborda el álgebra lineal desde tiempos lejanos. Por ejemplo, uno de los problemas fundamentales de esta disciplina es resolver sistemas de ecuaciones lineales y, notablemente, desde hace más de dos mil años los babilonios sabían resolver sistemas de dos ecuaciones lineales con dos incógnitas. De hecho, el llamado método de Gauss-Jordan que se usa para resolver sistemas de m ecuaciones lineales con n incógnitas ya era conocido por los chinos al menos 200 años antes de la era cristiana.

En la actualidad, gracias a las aportaciones de diferentes matemáticos, físicos, ingenieros, astrónomos, etc., el álgebra lineal es una de las áreas de las matemáticas más elegantes en cuanto a su estructura, sus teoremas y demostraciones. Además, nos abre las puertas a otras áreas de las matemáticas como la teoría de grupos, las ecuaciones diferenciales, la geometría diferencial, etc.

Asimismo el álgebra lineal no solo es un cúmulo de teoremas y demostraciones elegantes, también tiene una amplia gama de aplicaciones en áreas como la mecánica cuántica, probabilidad, finanzas, economía, cálculo, física de partículas, ecuaciones diferenciales, inteligencia artificial, etc. Por ejemplo, el surgimiento de la mecánica cuántica no hubiera sido posible sin la existencia previa de los resultados fundamentales del álgebra lineal.

Por estas razones, para estudiantes de matemática puras y aplicadas, de física y de diversas ingenierías, es importantes entender y dominar los aspectos puros y aplicados del álgebra lineal.

El objetivo de este libro es introducir al estudiante en el álgebra lineal y sus diversas aplicaciones. En el escrito se encuentra la justificación de los resultados fundamentales del álgebra lineal, se realizan varios ejercicios para mostrar cómo se usa esta disciplina, se muestran diversas aplicaciones y se explica cómo emplear código de Python para resover diferentes problemas.

El material de este libro ha sido utilizado en cursos de Álgebra Lineal I y II para estudiantes de las carreras de Matemáticas Aplicadas e Ingeniería en Computación de la Universidad Autónoma Metropolitana-Unidad Cuajimalpa. Normalmente los alumnos muestran más interés en las aplicaciones. No obstante, en el texto se buscó fundamentar los resultados más importantes del álgebra lineal para que los alumnos obtengan bases sólidas y al mismo tiempo sean capaces de aplicar los conocimientos adquiridos. Además, se trató de incluir una nota histórica sobre las personas que realizaron aportaciones a los temas estudiados en cada capítulo. Con la idea de no saturar a los estudiantes con muchos temas en pocos capítulos, el texto se divide en varios capítulos con pocos temas.

El autor es originario del pueblo Hñahñu del Valle del Mezquital del Estado de Hidalgo. Los antepasadso de este pueblo habitaron el territorio de México varios milenios antes de la era cristiana. En español, al pueblo *Hñahñu* se le llama Otomí, que es una variante de la palabra náhuatl *otómitl*, que significa *quien camina con flechas* o *flechador de pájaros*. Es doctor en Física por la Universidad Nacional Autónoma de México y Profesor de Tiempo Completo de la Universidad Autónoma Metropolitana Unidad Cuajimalpa. Ha realizado investigación en temas como Altas Energías, Biofísica y Finanzas Cuánticas. Sobre estos temas ha publicado diversos artículos en revistas internacionales. Ha sido miembro del Sistema Nacional de Investigadores con Nivel II de México.

Capítulo 1

Espacios vectoriales

En este capítulo daremos la definición de espacio vectorial y veremos algunas de sus implicaciones.

1.1. Definición

Un espacio vectorial se define con un conjunto $\mathbf{V}$, un campo $\mathbf{K}$ y dos operaciones

$$\boxed{\begin{aligned} + &: \mathbf{V} \times \mathbf{V} \to \mathbf{V}, \\ \mu &: \mathbf{K} \times \mathbf{V} \to \mathbf{V}. \end{aligned}}$$

Esta operaciones deben cumplir que si u, v pertenecen a $\mathbf{V}$, entonces $u + v$ pertenece a $\mathbf{V}$ y si α pertenece a $\mathbf{K}$, entonces $\mu(\alpha, v) = \alpha \cdot v$ pertenece a $\mathbf{V}$. Además, se debe cumplir

$$1) \quad \boxed{\forall u, v \in \mathbf{V}, \qquad u + v = v + u,} \tag{1.1}$$

$$2) \quad \boxed{\forall u, v, w \in \mathbf{V}, \qquad (u + v) + w = u + (v + w),} \tag{1.2}$$

$$3) \quad \boxed{\forall u, v \in \mathbf{V}, \forall \alpha, \in \mathbf{K}, \quad \alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v,} \tag{1.3}$$

$$4) \quad \boxed{\forall v \in \mathbf{V}, \forall \alpha, \beta \in \mathbf{K}, \qquad (\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v,} \tag{1.4}$$

$$5) \quad \boxed{\forall v \in \mathbf{V}, \forall \alpha, \beta \in \mathbf{K}, \qquad (\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v),} \tag{1.5}$$

$$6) \quad \boxed{\exists \quad 0 \in \mathbf{V} \quad \text{tal que} \quad \forall v \in \mathbf{V}, \qquad 0 + v = v,} \tag{1.6}$$

$$7) \quad \boxed{\forall v \in \mathbf{V}, \quad \exists - v \in \mathbf{V}, \quad \text{tal que} \quad v + (-v) = 0,} \tag{1.7}$$

$$8) \quad \boxed{\forall v \in \mathbf{V}, \quad e \cdot v = v,} \tag{1.8}$$

aquí e representa el neutro multiplicativo de $\mathbf{K}$.

1.2. Ejemplos

Ahora, veremos algunos ejemplos de espacios vectoriales. El lector puede verificar que los siguientes espacios cumplen las reglas de espacios vectoriales.

1.2.1. $\mathbb{C}^n$

Supongamos que $\mathbb{C}$ es el conjunto de los números complejos. Un ejemplo de espacio vectorial son los arreglos de la forma

$$\mathbb{C}^n = \{(c_1, c_2, \cdots, c_n), c_i \in \mathbb{C}\}.$$

Si se tienen dos vectores de $\mathbb{C}^n$,

$$(c_1, c_2, \cdots, c_n), \quad (d_1, d_2, \cdots, d_n)$$

la suma se define como

$$(c_1 + d_1, c_2 + d_2, \cdots, c_n + d_n).$$

Mientras que si λ es un número complejo, el producto por un escalar se define como

$$\lambda(c_1, c_2, \cdots, c_n) = (\lambda c_1, \lambda c_2, \cdots, \lambda c_n).$$

1.2.2. Sucesiones

Una generalización de $\mathbb{C}^n$ es tomar el límite $n \to \infty$, que nos da el espacio de sucesiones

$$\{a_n\}_{n=0}^{\infty}, \qquad a_n \in \mathbb{C}.$$

Así, si se tienen dos sucesiones

$$\{a_n\}_{n=0}^{\infty}, \quad \{b_n\}_{n=0}^{\infty}, \qquad a_n, b_n \in \mathbb{C}$$

la suma se define como

$$\{a_n + b_n\}_{n=0}^{\infty}.$$

Mientras que si λ es un número complejo, el producto por un escalar se define como

$$\lambda\{a_n\}_{n=0}^{\infty} = \{\lambda a_n\}_{n=0}^{\infty}.$$

1.2.3. Matrices

Otra generalización de $\mathbb{C}^n$ es el espacio de matrices M_{nm} de entradas complejas

$$\boxed{M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1m} \\ M_{21} & M_{22} & \cdots & M_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nm} \end{pmatrix}, \quad M_{ij} \in \mathbb{C}, \quad i = 1, \cdots, n, j = 1, \cdots, m.}$$

Si se tienen dos matrices

$$\boxed{M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1m} \\ M_{21} & M_{22} & \cdots & M_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nm} \end{pmatrix}, \quad N = \begin{pmatrix} N_{11} & M_{12} & \cdots & N_{1m} \\ N_{21} & N_{22} & \cdots & N_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ N_{n1} & N_{n2} & \cdots & N_{nm} \end{pmatrix},}$$

la suma se define como

$$\boxed{M + N = \begin{pmatrix} M_{11}+N_{11} & M_{12}+N_{12} & \cdots & M_{1m}+N_{1m} \\ M_{21}+N_{21} & M_{22}+N_{22} & \cdots & M_{2m}+N_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1}+N_{n1} & M_{n2}+N_{n2} & \cdots & M_{nm}+N_{nm} \end{pmatrix}.}$$

Mientras que el producto por un escalar, $\lambda \in \mathbb{C}$, se define como

$$\boxed{\lambda M = \begin{pmatrix} \lambda M_{11} & \lambda M_{12} & \cdots & \lambda M_{1m} \\ \lambda M_{21} & \lambda M_{22} & \cdots & \lambda M_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda M_{n1} & \lambda M_{n2} & \cdots & \lambda M_{nm} \end{pmatrix}.}$$

1.2.4. Funciones

Otro ejemplo de espacio vectorial es el espacio de funciones $f : [a, b] \to \mathbb{C}$, donde $a, b \in \mathbb{R}, a < b$. Supongamos que tenemos dos funciones

$$\boxed{f : [a, b] \to \mathbb{C}, \qquad g : [a, b] \to \mathbb{C},}$$

la suma se define como

$$\boxed{f+g:[a,b]\to\mathbb{C},}$$

con la regla de correspondencia

$$(f+g)(x)=f(x)+g(x),\qquad x\in[a,b].$$

Mientras que el producto por un escalar, $\lambda\in\mathbb{C}$, se define como

$$\boxed{\lambda f:[a,b]\to\mathbb{C},}$$

con la regla de correspondencia

$$(\lambda f)(x)=\lambda f(x),\qquad x\in[a,b].$$

1.3. Unicidad del neutro aditivo

Si V es un espacio vectorial existe un elemento $0\in V$ tal que para todo $v\in V$ se cumple que

$$\boxed{v+0=v.} \tag{1.9}$$

Supongamos que existe otro elemento $0'\in V$ que tiene las mismas propiedades que 0, es decir que para todo $v\in V$ se cumple

$$v+0'=v. \tag{1.10}$$

Por lo que si tomamos $v=0$ en esta última ecuación se encuentra

$$0+0'=0.$$

Además, tomando $v=0'$ en la ecuación (1.9) se llega a

$$0'+0=0'.$$

Así, como

$$0'+0=0+0',$$

concluimos que

$$0'=0.$$

Este resultado nos indica que en cualquier espacio vectorial solo existe un elemento con las propiedades del neutro aditivo, es decir el neutro aditivo es único.

Otro resultado interesante sobre el neutro aditivo de un espacio vectorial es que para todo $v \in V$ se cumple

$$0 \cdot v = 0. \tag{1.11}$$

En efecto, note que para todo $v \in V$ se cumple

$$\begin{aligned} v + 0 \cdot v &= 1 \cdot v + 0 \cdot v \\ &= (1+0) \cdot v \\ &= 1 \cdot v \\ &= v, \end{aligned}$$

esto es

$$v + 0 \cdot v = v. \tag{1.12}$$

Ahora, sumando en cada término de la expresión (1.12) el inverso aditivo $-v$ se encuentra

$$(-v) + v + 0 \cdot v = v + (-v)$$

de donde

$$0 + 0 \cdot v = 0,$$

esto es

$$0 \cdot v = 0.$$

Que es lo que queríamos demostrar.

1.4. Unicidad del inverso aditivo

Si V es un espacio vectorial y $v \in V$, entonces existe un elemento $-v \in V$ tal que se cumple que

$$\boxed{v + (-v) = 0.} \tag{1.13}$$

Supongamos que existe otros elemento $v' \in V$ que tiene las mismas propiedades que $-v$, es decir, que para todo $v \in V$, se cumple

$$v + v' = 0. \tag{1.14}$$

Ahora, si en esta ecuación sumamos de ambos lados el vector $-v$ se llega a

$$v + v' + (-v) = (-v) + 0,$$

entonces, usando la propiedad del neutro aditivo y la asociatividad de la suma de vectores se llega a

$$\begin{aligned} -v &= v + (-v) + v' \\ &= 0 + v' \\ &= v', \end{aligned}$$

es decir,

$$-v = v'.$$

Por lo tanto, en cualquier espacio vectorial un vector tiene un único inverso aditivo.

Una implicación de la unicidad del inverso aditivo de un vector un espacio vectorial es que para todo $v \in V$ se cumple

$$\boxed{-1 \cdot v = -v.} \tag{1.15}$$

En efecto, observe que para todo $v \in V$ se cumple

$$\begin{aligned} v + (-1) \cdot v &= 1 \cdot v + (-1) \cdot v \\ &= (1 + (-1)) \cdot v \\ &= 0 \cdot v \\ &= 0, \end{aligned}$$

esto es

$$v + (-1) \cdot v = 0.$$

Como el inverso aditivo es único, al igualdad (1.15) es correcta.

La versión axiomática de los espacios vectoriales se deben a Giuseppe Peano, quien se basó en los trabajos de Hermann Grassmann. Grasmman desarrolló un espacio que llamó números hipercomplejos, que tiene las propiedades de lo que hoy llamamos espacio vectorial. Vale la pena mencionar que William R. Hamilton desarrolló una estructura matemática que llamó cuaterniones en donde formuló la noción de vector, incluso en la actualidad se sigue usando su notación de i, j, k para la base canónica del espacio tridimensional. Los axiomas de espacio vectorial se hicieron populares después de que en 1918 Hermann

Wey los retomó en su libro *Space, time, matter*, que es una introducción a la relatividad general de Einsten.

Hermann Grassmann (1809-1877) fue un matemático, físico y lingüista alemán. Realizó sus estudios universitarios en temas de Teología, Filosofía y Literatura en la Universidad de Berlín. A pesar de no tener estudios en matemáticas o física, realizó investigaciones por su propia cuenta en esos temas. Además de desarrollar los conceptos de los espacios vectoriales, fundamentó lo que en la actualidad se llaman álgebras y números de Grassman, que son empleados para describir partículas subatómicas. Su sueño profesional fue ser profesor universitario de matemáticas, pero nunca lo logró. Sus trabajos en lingüística le permitieron ser reconocido con un doctorado honoris causa por la Universidad de Tübingen de Alemania.

1.5. Operaciones con Python

Si tenemos los vectores

$$a = (-1, 9, 5, 2), \qquad b = (1, 3, -1, 7)$$

y queremos realizar la suma de ellos, es decir

$$a + b,$$

se puede emplear el siguiente código en Python

```python
import numpy as np
a=np.array([-1,9,5,2])
b=np.array([1,3,-1,7])
np.add(a,b)
```

Listing 1.1: Suma de vectores

que da como resultado

```
array([0,12,4,9])
```

Además, si tenemos los vectores

$$a = (-1, 9, 5, 2), \qquad b = (1, 3, -1, 7)$$

y queremos realizar la resta

$$a - b$$

se puede emplear el código de Python

```
import numpy as np
a = np.array([1,0,6,9])
b=np.array([-1,3,4,7])
np.subtract(a,b)
```

Listing 1.2: Resta de vectores

que da el resultado

```
array([2,-3,2,2])
```

También se puede ver que si se quieren sumar las matrices

$$a = \begin{pmatrix} 1 & 0 & 6 & 9 \\ -1 & 3 & 6 & 4 \\ 3 & -1 & 6 & 3 \\ 1 & 5 & 6 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 & 3 & 4 & 7 \\ -2 & 2 & 1 & 5 \\ 1 & -2 & 3 & 1 \\ 2 & 2 & 1 & -2 \end{pmatrix} \tag{1.16}$$

se puede emplear el código de Python

```
import numpy as np
a = np.array([[1,0,6,9], [-1,3,6,4],[3,-1,6,3],[1,5,6,-1]
             ])
b=np.array([[-1,3,4,7], [-2,2,1,5],[1,-2,3,1],[2,2,1,-2]
             ])
np.add(a,b)
```

Listing 1.3: Suma de matrices

que da como resultado

```
array([[0,3,10,16],
[-3,5,7,9],
[4,-3,9,4],
[3,7,7,-3]])
```

Para restar las matrices (1.16) se puede emplear el código de Python

```
import numpy as np
a = np.array([[1,0,6,9], [-1,3,6,4],[3,-1,6,3],[1,5,6,-1]
            ])
b=np.array([[-1,3,4,7], [-2,2,1,5],[1,-2,3,1],[2,2,1,-2]
            ])
np.subtract(b,a)
```

Listing 1.4: Suma de matrices

que da el resultado

```
array([[-2,3,-2,-2],
[-1,-1,-5,1],
[-2,-1,-3,-2],
[1,-3,-5,-1]])
```

1.6. Ejercicios

1.- Dé un ejemplo de espacio vectoria en la vida real.

2.- Muestre que con la suma usual y el producto de un escalar por una función, el conjunto de los polinomios con coefientes complejos de grado n

$$\{P(x) = a_0 + a_1 z + \cdots + a_n z^n, a_i \in \mathbb{R}, i = 0, 1, 2, \cdots, n\}$$

es un espacio vectorial.

3.- Con la suma usual y el producto de un escalar por una sucesión, muestre detalladamente que el conjunto suscesiones complejas convergentes forma un espacio vectorial.

4.- Muestre que con la suma usual y el producto de un escalar por una serie, que el conjunto de las series complejas convergentes es un espacio vectorial.

5.- Con la suma usual y el producto de un escalar por una función, muestre que el conjunto de funciones $\{F(x)\}$ que cumplen

$$\int_{-\infty}^{\infty} |F(x)|^2 dx < \infty$$

es un espacio vectorial.

Capítulo 2

Subespacios vectoriales

En este capítulo daremos la definición de subespacio vectorial, sus implicaciones, conceptos relacionados y algunos ejemplos.

2.1. Definición

Sea $(V, K, +, \cdot)$ espacio vectorial sobre el campo K. Si V_1 es subconjunto de V y $(V_1, K, +, \cdot)$ es espacio vectorial sobre K se dice que V_1 es subespacio vectorial de V.

Si V es espacio vectorial y $V_1 \subset V$, entonces V_1 es espacio vectorial si cumple las condiciones

$$\boxed{\begin{aligned} & 0 \in V_1, && (2.1)\\ \forall v, u \in V_1, \quad & u + v \in V_1, && (2.2)\\ \forall v \in V, \forall \alpha \in K, \quad & \alpha \cdot v \in V_1. && (2.3)\end{aligned}}$$

Para probar esta afirmación, note que la condición (2.2) indica que la suma de vectores en V_1 es cerrada en V_1 y la condición (2.3) indica que el producto de vectores con escalares de K es cerrada en V_1. Además, debido a que $V_1 \subset V$ y V es espacio vectorial, de manera automática se cumplen las condiciones (1.1)-(1.5) y (1.8). Ahora, debido a que V_1 satisface la condición (2.1), también satisface la condición (1.6). Adicionalmente, sabemos que para cualquier $v \in V$ se cumple

$$-1 \cdot v = -v,$$

en particular si $v \in V_1$. Así, como V_1 satisface la condición (2.3), si $v \in V_1$ entonces $-v \in V_1$ y se cumple la condición (1.7).

Por lo tanto, si V es espacio vectorial y V_1 es un subconjunto de V que satisface las condiciones (2.1)-(2.3), entonces V_1 es espacio vectorial y en consecuencia subespacio vectorial de V.

Existen varias formas de expresar las condiciones para que un subconjunto sea subespacio vectorial. Por ejemplo, V_1 es subespacio vectorial de V si y solo si $V_1 \subset V$ y

$$\boxed{\forall v_1, v_2 \in V_1, \forall \alpha_1, \alpha_2 \in K, \qquad \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 \in V_1.} \tag{2.4}$$

Para probar esta afirmación, notemos que si tomamos $\alpha_1 = \alpha_2 = 0$ la condición (2.4) implica la condición (2.1); si tomamos $\alpha_1 = \alpha_2 = 1$ la condición (2.4) implica la condición (2.2); si tomamos $\alpha_1 = 1, \alpha_2 = 0, v_1 = v$ la condición (2.4) implica la condición (2.3). Así, la condición (2.4) implica las condiciones (2.1)-(2.3).

Además, si $\alpha_1, \alpha_2 \in K$ y $v_1, v_2 \in V_1$ entonces la condición (2.3) implica

$$v_1' = \alpha_1 v_1 \in V_1, \qquad v_2' = \alpha_2 v_2 \in V_2,$$

mientras que la condición (2.2) implica

$$v_1' + v_2' \in V_2,$$

es decir

$$\alpha_1 v_1 + \alpha_2 v_2 \in V_2.$$

Así, las condiciones (2.1)-(2.3) implican la condición (2.4). Por lo tanto la condición (2.4) es equivalente a las condiciones (2.1)-(2.3).

2.1.1. Ejemplos de subespacios vectoriales

Por ejemplo, el conjunto de vectores de la forma

$$(x, y, 0), \quad x, y \in \mathbb{R}$$

es un subespacio vectorial del espacio vectorial $\mathbb{R}^3$.

Además, en el espacio vectorial de la funciones de la forma

$$f : [a, b] \to \mathbb{R},$$

el conjunto de las funciones pares, es decir que cumplen

$$f(-x) = f(x),$$

forman un subespacio vectorial.

En el espacio vectorial de las matrices reales de 2×2, el conjunto de las matrices de la forma

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

es un subespacio vectorial.

2.2. Intersección de espacios vectoriales

Sea V un espacio vectorial sobre un campo K. Si $V_1 \subset V, V_2 \subset V$, podemos definir el conjunto $V_1 \cap V_2$, observe que si V_1 y V_2 son subespacios vectoriales de V entonces $0 \in V_1 \cap V_2$. En general, se puede probar que cuando V_1 y V_2 son subespacios vectoriales de V, el conjunto $V_1 \cap V_2$ también es subespacio vectorial de V. Para probar esta aseveración supongamos que $v_1, v_2 \in V_1 \cap V_2$, entonces $v_1, v_2 \in V_1$ y $v_1, v_2 \in V_2$. Ahora, si $\alpha_1, \alpha_2 \in K$, como V_1 y V_2 son subespacios vectoriales se satisface

$$\alpha_1 v_1 + \alpha_2 v_2 \in V_1, \qquad \alpha_1 v_1 + \alpha_2 v_2 \in V_2,$$

entonces

$$\alpha_1 v_1 + \alpha_2 v_2 \in V_1 \cap V_2.$$

Por lo tanto, si V_1, V_2 son subespacios vectoriales de V, entonces $V_1 \cap V_2$ es subespacio vectorial de V.

2.3. Suma de subespacios vectoriales

Sea V_1 y V_2 subespacios vectoriales del espacio vectorial V, se define la suma de subespacios vectoriales como

$$\boxed{V_1 + V_2 = \{v \in V | \exists v_1 \in V_1, \exists v_2 \in V_2, \quad v = v_1 + v_2\}.}$$

El conjunto $V_1 + V_2$ es subespacio vectorial de V. En efecto, si u_1, u_2 son elementos de $V_1 + V_2$ entonces existen v_1, v_1' en V_1 y v_2, v_2' en V_2 de tal forma que

$$u_1 = v_1 + v_2 \qquad \text{y} \qquad u_2 = v_1' + v_2'.$$

Entonces, si α_1, α_2 son escalares de K, se obtiene

$$\begin{aligned} \alpha_1 u_1 + \alpha_2 u_2 &= \alpha_1(v_1 + v_2) + \alpha_2(v_1' + v_2') \\ &= (\alpha_1 v_1 + \alpha_2 v_1') + (\alpha_1 v_2 + \alpha_2 v_2'). \end{aligned}$$

Además como V_1 y V_2 son subespacios vectoriales tenemos que

$$v_1'' = \alpha_1 v_1 + \alpha_2 v_1' \in V_1, \qquad \text{y} \qquad v_2'' = \alpha_1 v_2 + \alpha_2 v_2' \in V_2.$$

Así, existe $v_1'' \in V_1$ y $v_2'' \in V_2$ de tal forma que

$$\alpha_1 u_1 + \alpha_2 u_2 = v_1'' + v_2''.$$

Por lo tanto,

$$\alpha_1 u_1 + \alpha_2 u_2 \in V_1 + V_2$$

y $V_1 + V_2$ es un subespacio vectorial de V

Cuando

$$\boxed{V_1 + V_2 = V}$$

y para cada

$$v \in V, \exists! v_1 \in V_1, \quad \text{y} \quad \exists! v_2 \in V_2$$

de tal forma que

$$v = v_1 + v_2$$

dice que V es la suma directa de V_1 y V_2 y se escribe

$$V = V_1 \oplus V_2.$$

2.4. Espacio generado

Sea V un espacio vectorial y

$$\boxed{S = \{v_1, v_2, \cdots, v_n\} \subset V.}$$

Sea $\langle S\rangle$ el conjunto todas las combinaciones lineales de los elementos de S, es decir

$$\boxed{\langle S\rangle = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n\}}$$

con

$$\alpha_1, \alpha_2, \cdots, \alpha_n \in K.$$

El conjunto $\langle S\rangle$ es subespacio vectorial de V. En efecto, supongamos que $\alpha, \beta \in K$ y $v, w \in \langle S\rangle$, entonces existen

$$\alpha_1, \alpha_2, \cdots, \alpha_n \in K, \qquad \beta_1, \beta_2, \cdots, \beta_n \in K$$

tal que

$$\begin{aligned} v &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, \\ w &= \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n. \end{aligned}$$

Por lo tanto

$$\begin{aligned} \alpha v + \beta w &= \alpha(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) + \beta(\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n) \\ &= (\alpha\alpha_1 + \beta\beta_1)v_1 + (\alpha\alpha_2 + \beta\beta_2)v_2 + \cdots + (\alpha\alpha_n + \beta\beta_n)v_n, \end{aligned}$$

que es una combinación lineal del conjunto S, entonces

$$\alpha v + \beta w \in \langle S\rangle .$$

Por lo tanto $\langle S\rangle$ es un subespacio vectorial de V.

Cuando $\langle S\rangle = V$ se dice que el conjunto S genera a V.

2.5. Independencia lineal

Sean

$$\boxed{v_1, v_2, \cdots, v_n} \tag{2.5}$$

elementos de un espacio vectorial V. Se dice que el conjunto de vectores (2.5) es linealmente independiente si

$$\boxed{\alpha_1, \alpha_2, \cdots, \alpha_n \in K}$$

y

$$\boxed{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0}$$

implica

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Se dice que el conjunto (2.5) es linealmente dependiente si existen escalares

$$\alpha_1, \alpha_2, \cdots, \alpha_n \in K$$

no todos nulos tales que

$$\boxed{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0.} \tag{2.6}$$

Por ejemplo, en $\mathbb{R}^2$ los vectores

$$(1, 0), \quad (0, 1)$$

son linealmente independientes.

Además, en el espacio vectorial de las matrices reales de 2×2, las matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & -4 \end{pmatrix}$$

son linealmente independientes.

2.6. Dependencia lineal

Podemos afirmar que un conjunto de vectores

$$\{v_1, v_2, \cdots, v_n\}$$

es linealmente dependiente si alguno de ellos se puede escribir como combinación lineal de los restantes. En efecto, sin pérdida de generalidad, supongamos que se cumple (2.6) y que $\alpha_1 \neq 0$. Por lo tanto, de (2.6) se obtiene

$$\alpha_1 v_1 = -\alpha_2 v_2 - \cdots - \alpha_n v_n,$$

de donde

$$v_1 = -\frac{1}{\alpha_1}(\alpha_2 v_2 + \cdots + \alpha_n v_n).$$

Si v_1 se puede escribir como combinación de los vectores

$$v_2, \cdots, v_n, \tag{2.7}$$

se dice que v_1 depende linealmente de los vectores (2.7).

Por ejemplo, en $\mathbb{R}^2$ los vectores

$$(1,0), \quad (0,1), \quad (1,2)$$

son linealmente dependientes.

Además, en el espacio vectorial de las matrices reales de 2×2, las matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

son linealmente dependientes.

2.7. Base y dimensión

Sea V un espacio vectorial. Se dice que el conjunto de vectores

$$\boxed{\{v_1, v_2, \cdots, v_n\} \subset V} \tag{2.8}$$

es base de V si es linealmente independiente y es generador de V. En dicho caso, se dice que la dimensión de V es la cardinalidad de dicho conjunto.

Si el espacio vectorial V tiene dimensión n y el conjunto de vectores de V

$$\boxed{\{v_1, v_2, \cdots, v_r\} \quad \text{con} \quad r < n} \tag{2.9}$$

es linealmente independiente, este último conjunto se puede completar para formar una base. En efecto, si el conjunto (2.9) no genera a todo el espacio vectorial V, entonces existe un vector v_{r+1} en V que no puede ser generado por los vectores del conjunto (2.9), entonces el conjunto

$$\{v_1, v_2, \cdots, v_r, v_{r+1}\} \tag{2.10}$$

es linealmente independiente. Si el conjunto (2.10) no genera a todo el espacio vectorial V, entonces existe un vector v_{r+2} en V que no puede ser generado por dicho conjunto, así el conjunto

$$\{v_1, v_2, \cdots, v_r, v_{r+1}, v_{r+2}\}$$

es linealmente independiente. Si este nuevo conjunto de vectores no genera al espacio vectorial V podemos seguir con el mismo procedimiento, el cual es finito si la dimensión de V es finita.

Usando otra notación, los conceptos de combinación lineal, subespacio vectorial, dimensión y espacio generado fueron establecidos por Hermann Gunther Grassmann en las dos versiones de su libro Cálculo de Extensión (1884 y 1862).

2.8. Ejercicios

1.- En el espacio vectorial de la funciones de la forma

$$f : [a, b] \to \mathbb{R},$$

mostrar que el conjunto de las funciones impares, es decir, que cumplen

$$f(-x) = -f(x),$$

forman un subespacio vectorial.

2.- En el espacio vectorial de las matrices reales de 2×2, el conjunto de las matrices antisimétricas, es decir, la matrices de la forma

$$M = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix},$$

es un subespacio vectorial.

3.- Determinar si el siguiente conjunto de polinomios es linealmente dependiente o independiente

$$\{P_0 = 2x - 3, \quad P_1 = 5x^2 + 4, \quad P_2 = 7\}.$$

4.- Determinar si el conjunto de polinomios es linealmente dependientes o independiente

$$\{P_0 = 2x, \quad P_1 = x^2 + 4x, \quad P_2 = x^2\}.$$

5.- Determinar si los siguientes vectores son linealmente dependientes o independientes

$$(1,1,0,1), \quad (1,-1,1,1), \quad (2,2,1,2), \quad (0,1,0,0).$$

6.- Determinar si los siguientes vectores son linealmente dependientes o independientes

$$(1,0,0,1), \quad (0,1,0,1), \quad (0,0,1,0), \quad (1,1,1,1).$$

7.- Determinar si las siguientes matrices son linealmente dependientes o independientes

$$M_1 = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix}, \quad M_3 = \begin{pmatrix} -1 & 4 \\ 4 & -3 \end{pmatrix}.$$

8.- Encontrar una base para el conjunto de matrices reales de 3×3

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

9.- Obtener una base para los polinomios de coeficientes reales de grado 4.

10.- Obtener el espacio generado por las matrices

$$M_1 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Capítulo 3

Producto de matrices y sus propiedades

En este capítulo veremos algunas operaciones con matrices como el producto, la inversa y la traza.

3.1. Definición

Supongamos que A es una matriz de $n \times m$ y B es una matriz de $m \times l$, ambas matrices con entradas complejas,

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1l} \\ B_{21} & B_{22} & \cdots & B_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{n2} & \cdots & B_{ml} \end{pmatrix},$$

con

$$A_{ij}, B_{kr} \in \mathbb{C}, \qquad i = 1, 2, \cdots, n, \qquad k, j = 1, 2, \cdots, m, \qquad r = 1, 2, \cdots, l.$$

Entonces definimos el producto de matrices como

$$
\begin{aligned}
AB &= \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1l} \\ B_{21} & B_{22} & \cdots & B_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{n2} & \cdots & B_{ml} \end{pmatrix} \\
&= \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} + \cdots + A_{1m}B_{m1} & \cdots & A_{11}B_{1l} + A_{12}B_{2l} + \cdots + A_{1m}B_{ml} \\ A_{21}B_{11} + A_{22}B_{21} + \cdots + A_{2m}B_{m1} & \cdots & A_{21}B_{1l} + A_{22}B_{2l} + \cdots + A_{2m}B_{ml} \\ \vdots & \ddots & \vdots \\ A_{n1}B_{11} + A_{n2}B_{21} + \cdots + A_{nm}B_{m1} & \cdots & A_{n1}B_{1l} + A_{n2}B_{2l} + \cdots + A_{nm}B_{ml} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{k=1}^m A_{1k}B_{k1} & \cdots & \sum_{k=1}^m A_{1k}B_{kl} \\ \sum_{k=1}^m A_{2k}B_{k1} & \cdots & \sum_{k=1}^m A_{2k}B_{kl} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^m A_{nk}B_{k1} & \cdots & \sum_{k=1}^m A_{nk}B_{kl} \end{pmatrix} \\
&= \begin{pmatrix} A_{1k}B_{k1} & \cdots & A_{1k}B_{kl} \\ A_{2k}B_{k1} & \cdots & A_{2k}B_{kl} \\ \vdots & \ddots & \vdots \\ A_{nk}B_{k1} & \cdots & A_{nk}B_{kl} \end{pmatrix}, \quad k = 1, 2, \cdots, m.
\end{aligned}
$$

Por lo tanto

$$\boxed{(AB)_{ir} = \sum_{k=1}^{m} A_{ik}B_{kr}, \qquad i = 1, 2, \cdots, n, \qquad r = 1, 2, \cdots, l.} \tag{3.1}$$

3.2. Convención de Einstein

Antes de continuar notemos que por simplicidad, en una suma del tipo

$$\boxed{\sum_{i=1}^{n} V^i W^i}$$

es común escribir

$$\boxed{\sum_{i=1}^{n} V^i W^i = V^i W^i = V^1 W^1 + V^2 W^2 + \cdots + V^n W^n.}$$

En donde se entiende que los índices repetidos se suman y se omite el símbolo de suma Σ. A esta regla se le llama convención de Einstein.

Por ejemplo, con la notación de Einstein el producto de matrices se escribe como

$$\boxed{\begin{aligned} AB &= \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1l} \\ B_{21} & B_{22} & \cdots & B_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{n2} & \cdots & B_{ml} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=1}^{m} A_{1k} B_{k1} & \cdots & \sum_{k=1}^{m} A_{1k} B_{kl} \\ \sum_{k=1}^{m} A_{2k} B_{k1} & \cdots & \sum_{k=1}^{m} A_{2k} B_{kl} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} A_{nk} B_{k1} & \cdots & \sum_{k=1}^{m} A_{nk} B_{kl} \end{pmatrix} \\ &= \begin{pmatrix} A_{1k} B_{k1} & \cdots & A_{1k} B_{kl} \\ A_{2k} B_{k1} & \cdots & A_{2k} B_{kl} \\ \vdots & \ddots & \vdots \\ A_{nk} B_{k1} & \cdots & A_{nk} B_{kl} \end{pmatrix}, \quad k = 1, 2, \cdots, m. \end{aligned}}$$

En componentes se tiene

$$\boxed{(AB)_{ir} = A_{ik} B_{kr} = \sum_{k=1}^{m} A_{ik} B_{kr}, \;\; i = 1, 2, \cdots, n, \quad r = 1, 2, \cdots, l.} \tag{3.2}$$

Se puede observar que cuando se trabaja con componentes es más conveniente ocupar la notación de Einstein.

$\boxed{\text{Albert Einstein (1879-1955)}}$ fue un físico alemán de origen judío. En su adolecencia asistió a una escuela progresista basada en la filosofía pedagógica de Pestalozzi, en donde aprendió a ser autodidácta y a cuestionar todo tipo de autoridad y conocimiento. Su estadía en dicha escuela fue una de las épocas más felices de su vida. Si bien Einstein tenía una habilidad especial para las matemáticas, como pensador universal también se interesó por la música, literatura, filosofía, política, antropología, geología, etc. Se introdujo por su propia cuenta en los trabajos de Boltzman, Hertz, Lorentz, Kant, Darwin entre otros. A pesar de tener un amplio conocimiento de la física de su época, cuando terminó sus estudios universitarios era poco conocido y solo logró conseguir un empleo en unas oficinas de patentes. En esas oficinas empezó a resolver varios problemas importantes de la física y en 1905 publicó cinco artículos que cambiaron para siempre la física y nuestra percepción de la naturaleza. En uno de esos artículos dio los fundamentos de lo que hoy conocemos como Relatividad Especial, la cual cambió nuestro concepto de espacio-tiempo. En otro de los artículos explicó el movimiento Browiniano, en donde fundamenta el concepto de mólecula de Boltzman y cambió nuestro concepto de lo microscópico. En otro de los trabajos explicó el efecto fotoeléctrico. En el aspecto teórico, este último artículo dio un impulso importante a una naciente Mecánica Cuántica y desde el punto de vista tecnológico permitió el desarollo de las celdas solares. 1905 es recordado como uno de los años más importantes de la física moderna. Después de esos cinco artículos logró encontrar un trabajo de profesor y empredió un largo viaje intelectual en solitario para formular la llamada Relatividad General en 1915, la cual nos permite entender el origen de universo, las ondas gravitacionales, los hoyos negros, etc. Por ser judío, tuvo que dejar la Alemania de Hitler para emigrar a Estados Unidos, en donde finalmente murió un 18 de Abril de 1955. Después de I. Newton, quizá Einstein sea el físico más relevante de la historia.

3.3. Matriz identidad

A la matriz de $n \times n$ que en la diagonal tiene unos y fuera de ella tienen ceros

$$\boxed{I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}}$$

se le llama matriz identidad.

Note que si A es una matriz de $n \times n$ se tiene

$$
\begin{aligned}
AI &= \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \\
&= \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix},
\end{aligned}
$$

es decir

$$
\boxed{AI = IA = A.}
$$

Otra forma de definir la matriz identidad de $n \times n$ es mediante la llamada delta de Kronecker, que se define como el símbolo tal que

$$
\boxed{\delta_{ij} = \begin{cases} 1 & \text{si} \quad i = j, \\ 0 & \text{si} \quad i \neq j, \end{cases} \qquad i, j = 1, 2, \cdots, n.}
$$

Por ejemplo, para el caso $n = 3$ se tiene

$$
\boxed{\delta_{ij} = \begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.}
$$

Observe que usando la delta de Kronecker y la convención de Einstein, el producto de una matriz A por la matriz identidad I se escribe de la forma

$$\boxed{\begin{aligned} AI &= \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nn} \end{pmatrix} \\ &= \begin{pmatrix} A_{1k}\delta_{k1} & \cdots & A_{1k}\delta_{kn} \\ A_{2k}\delta_{k1} & \cdots & A_{2k}\delta_{kn} \\ \vdots & \ddots & \vdots \\ A_{nk}\delta_{k1} & \cdots & A_{nk}\delta_{kn} \end{pmatrix} \qquad (k = 1, 2, \cdots, n) \\ &= \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}. \end{aligned}}$$

En componentes esta igualdad es simplemente

$$\boxed{(AI)_{ij} = A_{ik}\delta_{kj} = A_{ij}.}$$

Así, cuando se trabaja con componentes, es más conveniente usar la delta de Kronecker.

$\boxed{\text{Leopold Kronecker (1823-1891)}}$ fue un matemático alemán que aportó resultados relevantes en áreas como teoría de números, álgebra y lógica. Descendiente de comerciantes y banqueros judíos, su educación inicial la tomó en su casa con profesores particulares. Posteriormente asistió al Gymnasium de Liegnitz, en donde se interesó por diferentes disciplinas como la filosofía, química, astronomía y las matemáticas. En 1845 realizó su trabajo doctoral en teoría de números bajo la supervisión de Peter Gustav Dirichlet, en la Universidad de Berlín. Después de terminar su doctorado, se dedicó a manejar los negocios de su familia y tomó a las matemáticas solo como un pasatiempo. Debido a su talento, recibió ofertas para trabajar como profesor universitario, pero las rechazó. Durante más de 30 años no fue propiamente un profesor universitario, pero escribió artículos notables. Por ejemplo, realizó extenciones de algunos resultados de Évariste Galois en la teoría de ecuaciones usando teoría de grupos. Además, sostenía grandes debates sobre los fundamentos de las matemáticas con personajes como Cantor, Bolzano, Weierstrass y Dedekind. En 1883 aceptó una plaza de profesor en la Universidad de Berlín. Como profesor dirigió la tesis doctoral de varios alumnos y continuó su debate con respecto al los fundamentos de las matemáticas. Él defendía las demostraciones constructivas en un número finitos de pasos. Con respecto a ese debate, es recordado por su frase *Dios hizo los números enteros; el resto es obra del hombre.* La muerte de su mujer fue un duro golpe para él y murió pocos meses después que ella.

3.4. Inversa de matrices

Supongamos que A es una matriz de $n \times n$. La matriz A^{-1} es la inversa de A si

$$\boxed{AA^{-1} = A^{-1}A = I.}$$

Por ejemplo, la matriz

$$M = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

tiene como inversa la matriz

$$M^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

La inversa de una matriz es única. En efecto, supongamos que existe otra matriz A' tal que

$$A'A = I.$$

Entonces, multiplicamos a esta ecuación por la derecha por A^{-1} obtenemos

$$A'AA^{-1} = A^{-1},$$

como $AA^{-1} = I$, llegamos a

$$A' = A^{-1}.$$

Por lo que la matriz inversa es única.

3.5. Transpuesta de matrices

Supongamos que A es una matriz de $n \times m$

$$\boxed{A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix}.}$$

Se define la matriz transpuesta de A como la matriz A^T de $m \times n$ que tiene como columnas los renglones de A y como renglones las columnas de A, es decir

$$\boxed{A^T = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1m} & A_{2m} & \cdots & A_{nm} \end{pmatrix},}$$

por lo que

$$\boxed{A^T_{ij} = A_{ji}.} \tag{3.3}$$

Por ejemplo, consideremos la matriz

$$M = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 4 & -9 \\ 5 & -3 & 8 \end{pmatrix},$$

entonces su transpuesta es

$$M^T = \begin{pmatrix} 2 & 3 & 5 \\ 0 & 4 & -3 \\ 1 & -9 & 8 \end{pmatrix}.$$

3.5.1. Transpuesta de un producto de matrices

Supongamos que A es una matriz de $n \times m$ y B es una matriz de $m \times l$, entonces se cumple la igualdad

$$\boxed{(AB)^T = B^T A^T.} \tag{3.4}$$

Para probar esta afirmación ocuparemos el producto de dos matrices (3.2) y la definición de transpuesta (3.3), entonces

$$\begin{aligned} (AB)^T_{ij} &= (AB)_{ji} \\ &= A_{jk}B_{ki} \\ &= B_{ki}A_{jk} \\ &= B^T_{ik}A^T_{kj} \\ &= \left(B^T A^T\right)_{ij} \end{aligned}$$

por lo que

$$(AB)^T_{ij} = \left(B^T A^T\right)_{ij},$$

que es lo que se quería demostrar.

Por ejemplo, consideremos las matrices

$$M_1 = \begin{pmatrix} 3 & 0 & -1 \\ 1 & 4 & -5 \\ 2 & 3 & 6 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} -1 & 3 & 2 \\ 7 & 4 & -1 \\ -3 & 2 & 5 \end{pmatrix},$$

cuyas transpuestas son

$$M_1^T = \begin{pmatrix} 3 & 1 & 2 \\ 0 & 4 & 3 \\ -1 & -5 & 6 \end{pmatrix}, \qquad M_2^T = \begin{pmatrix} -1 & 7 & -3 \\ 3 & 4 & 2 \\ 2 & -1 & 5 \end{pmatrix},$$

de donde

$$M_1M_2 = \begin{pmatrix} 0 & 7 & 1 \\ 42 & 9 & -27 \\ 1 & 30 & 31 \end{pmatrix}, \qquad M_2^T M_1^T = \begin{pmatrix} 0 & 42 & 1 \\ 7 & 9 & 30 \\ 1 & -27 & 31 \end{pmatrix},$$

así

$$(M_1M_2)^T = M_2^T M_1^T = \begin{pmatrix} 0 & 42 & 1 \\ 7 & 9 & 30 \\ 1 & -27 & 31 \end{pmatrix}.$$

3.5.2. Inversa de la matriz transpuesta

Si la matriz inversa de la matriz A es A^{-1}, entonces la inversa de la matriz transpuesta A^T es la transpuesta de la matriz inversa de A, esto es

$$\boxed{\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.} \tag{3.5}$$

Para mostrar este resultado notemos que

$$AA^{-1} = I,$$

entonces

$$\begin{aligned} I &= I^T \\ &= \left(AA^{-1}\right)^T \\ &= (A^{-1})^T A^T, \end{aligned}$$

así

$$(A^{-1})^T A^T = I,$$

multiplicando esta igualdad por la derecha por $(A^{-1})^T$, se encuentra

$$\begin{aligned} \left(A^T\right)^{-1} A^T (A^{-1})^T &= I(A^{-1})^T \\ &= (A^{-1})^T, \end{aligned}$$

de donde

$$\left(A^T\right)^{-1} = (A^{-1})^T$$

Esta igualdad es equivalente a la expresión (3.5).

Por ejemplo, consideremos la matriz

$$M = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 3 & 2 \\ 0 & 1 & 0 \end{pmatrix},$$

cuya inversa es

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

por lo que

$$M^T = \begin{pmatrix} 2 & 4 & 0 \\ 3 & 3 & 1 \\ 4 & 2 & 0 \end{pmatrix}, \qquad \left(M^{-1}\right)^T = \frac{1}{6}\begin{pmatrix} -1 & 0 & 2 \\ 2 & 0 & -1 \\ -3 & 6 & -3 \end{pmatrix},$$

de donde

$$M^T\left(M^{-1}\right)^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

3.6. Traza de una matriz

Supongamos que M es una matriz de $n \times n$ de entradas complejas

$$M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{pmatrix},$$

entonces la traza de M se define como la suma de las componentes de la diagonal, es decir

$$\boxed{Tr(M) = M_{11} + M_{22} + \cdots + M_{nn} = \sum_{i}^{n} M_{ii} = M_{ii}.}$$

Por ejemplo, consideremos la matriz

$$M = \begin{pmatrix} 2 & 0 & -1 \\ 5 & 6 & -3 \\ 2 & 3 & 9 \end{pmatrix},$$

así

$$Tr(M) = 2 + 6 + 9 = 17.$$

De la definición de traza se deducen los resultados

$$\boxed{\begin{aligned} Tr\left(M^T\right) &= Tr\left(M\right), \\ Tr(M+N) &= Tr(M) + Tr(N), \\ Tr(NM) &= Tr(MN). \end{aligned}}$$

En efecto,

$$
\begin{aligned}
Tr\left(M^T\right) &= \sum_{i=1}^{n} \left(M^T\right)_{ii} \\
&= \sum_{i=1}^{n} M_{ii} \\
&= Tr\left(M\right), \\
Tr(M+N) &= \sum_{i=1}^{n} \left(M+N\right)_{ii} \\
&= \sum_{i=1}^{n} \left(M_{ii}+N_{ii}\right) \\
&= \sum_{i=1}^{n} M_{ii} + \sum_{i=1}^{n} N_{ii} \\
&= Tr(M)+Tr(N), \\
Tr\left(MN\right) &= \sum_{i=1}^{n} (MN)_{ii} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} N_{ji} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} N_{ji} M_{ij} \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} N_{ji} M_{ij} \right) \\
&= \sum_{j=1}^{n} \left(NM\right)_{jj} \\
&= Tr(NM).
\end{aligned}
$$

Más adelante veremos algunas implicaciones de estos resultados.

Para ver un ejemplo consideremos la matriz

$$
M = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 3 & 2 \\ 0 & 1 & 0 \end{pmatrix},
$$

cuya matriz transpuesta es

$$M^T = \begin{pmatrix} 2 & 4 & 0 \\ 3 & 3 & 1 \\ 4 & 2 & 0 \end{pmatrix},$$

de donde

$$Tr(M) = 6, \qquad Tr\left(M^T\right) = 6.$$

Además, sean las matrices

$$M = \begin{pmatrix} 1 & 3 & 0 \\ -1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix}, \qquad N = \begin{pmatrix} 2 & 3 & 0 \\ 3 & 1 & 1 \\ 1 & 2 & -8 \end{pmatrix},$$

de donde

$$M + N = \begin{pmatrix} 3 & 6 & 0 \\ 2 & 3 & 4 \\ 1 & 3 & -8 \end{pmatrix},$$

por lo que

$$\begin{aligned} &T(M+N) = -2, \\ &Tr(M) + Tr(N) = 3 - 5 = -2, \end{aligned}$$

por lo que

$$T(M+N) = Tr(M) + Tr(N) = -2.$$

Además, sean las matrices

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 0 \\ 1 & 2 & 3 \end{pmatrix}, \\ B &= \begin{pmatrix} 0 & 2 & 1 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}, \end{aligned}$$

que implican

$$
\begin{aligned}
AB &= \begin{pmatrix} -5 & 4 & 0 \\ -3 & 3 & 0 \\ 7 & 4 & 4 \end{pmatrix}, \\
BA &= \begin{pmatrix} 1 & 8 & 3 \\ -1 & 1 & 1 \\ 4 & 8 & 0 \end{pmatrix},
\end{aligned}
$$

de donde

$$Tr\,(AB) = Tr(BA) = 2.$$

3.7. Matriz diagonal

Una matriz A de entradas complejas de $n \times n$ es diagonal si todas sus componentes fuera de la diagonal son cero, por ejemplo

$$\boxed{A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.}$$

Note que usando la delta de Kronecker, las componentes de una matriz diagonal se pueden escribir como

$$\boxed{A_{ij} = a_i \delta_{ij}, \qquad i, j = 1, 2 \cdots, n.}$$

Cuando una matriz es diagonal, la mayoría de las operaciones para matrices son muy fácil de hacer. Por ejemplo, la transpuesta es ella misma

$$A^T = A,$$

la m-ésima potencia es

$$\boxed{A^m = \begin{pmatrix} a_1^m & 0 & \cdots & 0 \\ 0 & a_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^m \end{pmatrix}.}$$

Además, si $a_i \neq 0$ $(i = 1, 2, \cdots, n)$, la inversa es

$$\boxed{A^{-1} = \begin{pmatrix} a_1^{-1} & 0 & \cdots & 0 \\ 0 & a_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{-1} \end{pmatrix}.}$$

Más adelante, veremos que cierto tipo de matrices se pueden relacionar con matrices diagonales y que en esas matrices se pueden simplificar muchas operaciones.

La idea de matriz y sus propiedades se encuentra en los trabajos de diferentes matemáticos como Gauss, Cauchy, Hermite y Eiseintein. James Joseph Sylvester introdujó el término de matriz en 1850. Sin embargo, fue Arthur Caley quien introdujó de forma general las operaciones de las matrices como suma, producto, potencia e inversa. Además, probó la asociatividad del producto de matrices, extendió la idea de determinante para más de dos dimensiones y desarrolló la noción de matrices simétricas y antisimétricas.

Arthur Cayley (1821-1895) fue un abogado y matemático inglés que ralizó diferentes aportaciones relevantes en las matemáticas puras, principalmente en álgebra y geometría. Además cultivó amistades con grandes científicos como William Hamilton, George Boole, William Thomson, entre otros. Notablemente, por sus diferentes aportaciones a la ciencia británica el físico Clerk Maxwell le escribió un poema. Arthur Cayley murió por múltiples enfermedades causadas por su avanzada edad.

3.8. Operaciones con Python

Supongamos que tenemos las matrices

$$A = \begin{pmatrix} 1 & 0 & 6 & 1 \\ -1 & 3 & 6 & 4 \\ 3 & -1 & 6 & 3 \\ 1 & 0 & 6 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 3 & 4 & -1 \\ 0 & 1 & -2 & 3 \\ -5 & 3 & 4 & 8 \\ 1 & -9 & 3 & 8 \end{pmatrix}$$

entonces para hacer el producto de ellas se puede emplear el siguiente código en Python

```
import numpy as np
a = np.array([[1,0,6,1], [-1,3,6,4],[3,-1,6,3],[1,0,6,-1]
             ])
b=np.array([[2,3,4,-1],[0,1,-2,3],[-5,3,4,8],[1,-9,3,8]])
np.dot(a,b)
```

Listing 3.1: Producto de dos matrices

que da el resultado

```
array([[-27,12,31,55],
[-28,-18,26,90],
[-21,-1,47,66],
[-29,30,25,39]])
```

Además, si para la matriz

$$A = \begin{pmatrix} 1 & 7 & 8 & 9 \\ -1 & 3 & 6 & 5 \\ 2 & 1 & 0 & -3 \\ -1 & 2 & 1 & 8 \end{pmatrix}$$

queremos sacar la transpuesta podemos emplear el código en Python

```
import numpy as np
A = np.array([[1,7,8,9],[-1,3,6,5],[2,1,0,-3],[-1,2,1,8]
             ])
np.transpose(A)
```

Listing 3.2: Transpuesta de una matriz

el cual da el resultado

```
array([[1,-1,2,-1],
[7,3,1,2],
[8,6,0,1],
[9,5,-3,8]])
```

Ahora, supongamos que queremos sacar la traza de la matriz

$$A = \begin{pmatrix} 1 & 0 & 6 & 9 \\ -1 & 3 & 6 & 4 \\ 3 & -1 & 6 & 3 \\ 1 & 5 & 6 & -1 \end{pmatrix}$$

se tiene el código en Python

```
import numpy as np
from numpy.linalg import eig
A= np.array([[1,0,6,9], [-1,3,6,4],[3,-1,6,3],[1,5,6,-1]
            ])
np.trace(A)
```

Listing 3.3: Traza de una matriz

que da como resultado

9

3.9. Ejercicios

1.- Obtener el producto de las matrices

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 3 & 3 & 5 \\ 1 & 2 & 7 \end{pmatrix}, \qquad B = \begin{pmatrix} -11 & 2 & 8 \\ 3 & -4 & 5 \\ 6 & -3 & 1 \end{pmatrix}.$$

2.- Para las matrices

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 5 & 6 \\ 3 & 2 & 9 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} -\frac{33}{10} & \frac{8}{5} & -\frac{7}{10} \\ \frac{9}{5} & -\frac{3}{5} & \frac{1}{5} \\ \frac{7}{10} & -\frac{2}{5} & \frac{3}{10} \end{pmatrix},$$

mostrar que se cumple

$$A^{-1}A = AA^{-1} = I.$$

3.- Obtener la traza de matriz

$$A = \begin{pmatrix} 4 & -5 & 6 \\ 1 & -2 & 7 \\ 5 & 2 & 9 \end{pmatrix}.$$

4.- Obtener la transpuesta de la matriz

$$A = \begin{pmatrix} 1 & 7 & -2 \\ 11 & 12 & -4 \\ 4 & -2 & 6 \end{pmatrix}.$$

5.- Considere las matrices

$$A = \begin{pmatrix} 3 & -1 & 0 \\ 1 & -3 & 5 \\ 3 & 2 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -3 & 7 \\ 2 & -4 & 3 \\ 1 & -6 & 2 \end{pmatrix}.$$

a) Obtener $A^T B^T$.

b) Obtener $Tr(AB)$.

c) Obtener $Tr(BA)$.

d) Obtener $Tr(A+B)$.

e) Obtener $Tr\left((AB)^T\right)$.

Capítulo 4

Determinante de una matriz

En este capítulo daremos la definición determinante de una matriz y alguna de sus implicaciones. Gauss nombró a esta operación determinante porque determina varias propiedades de una matrix, en particular si una matriz tiene inversa.

4.1. Inversa de una matriz de 2×2

Consideremos la matriz de 2×2 de entradas complejas

$$\boxed{M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},}$$

si queremos buscar su matriz inversa podemos proponer la matriz

$$\boxed{M^{-1} = \begin{pmatrix} e & f \\ g & h \end{pmatrix},}$$

la cual debe cumplir

$$\boxed{MM^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.}$$

Esta condición implica

$$MM^{-1} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Resolver este problema implica encontrar los valores e, f, g, h que cumplen

$$\begin{aligned} ae + bg &= 1, &(4.1)\\ af + bh &= 0, &(4.2)\\ ce + dg &= 0, &(4.3)\\ cf + dh &= 1. &(4.4) \end{aligned}$$

De la ecuación (4.3) se obtiene

$$e = -\frac{dg}{c}, \tag{4.5}$$

usando este resultado en la ecuación (4.1), encontramos

$$g = -\frac{c}{ad - bc},$$

además ocupando este último resultado en la expresión (4.5), llegamos a

$$e = \frac{d}{ad - bc}.$$

Mientras que, de la ecuación (4.2) se obtiene

$$f = -\frac{bh}{a}, \tag{4.6}$$

sustituyendo este resultado en la ecuación (4.4) encontramos

$$h = \frac{a}{ad - bc}$$

y considerando este resultado en (4.6), se obtiene

$$f = -\frac{b}{ad - bc}.$$

En resumen, tenemos los valores

$$\begin{aligned} e &= \frac{d}{ad - bc},\\ f &= -\frac{b}{ad - bc},\\ g &= -\frac{c}{ad - bc},\\ h &= \frac{a}{ad - bc}, \end{aligned}$$

que implican la matriz

$$\boxed{M^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.}$$

Observe que esta matriz existe solo si

$$\boxed{ad - bc \neq 0,} \tag{4.7}$$

a esta cantidad se le llama determinante. Note que el determinante es importante para conocer la inversa de una matriz de 2×2.

4.2. Tensor de Levi-Civita

Si bien el determinante (4.7) fue obtenido para encontrar la inversa de una matriz de 2×2, esta cantidad es interesante por sí misma como objeto matemático y tiene diversas aplicaciones. Antes de generalizar el determinante para una matriz de $n \times n$, expresaremos el determinante de una matriz de 2×2 de una manera más sugerente. Consideremos la matriz

$$\boxed{A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},}$$

entonces el determinante (4.7) tiene la forma

$$\boxed{\det A = A_{11}A_{22} - A_{12}A_{21}.} \tag{4.8}$$

Ahora definamos la cantidad

$$\boxed{\epsilon_{ij}, \quad i, j = 1, 2} \tag{4.9}$$

que cumple las propiedades

$$\boxed{\begin{aligned} \epsilon_{12} &= 1, \\ \epsilon_{ij} &= -\epsilon_{ij}. \end{aligned}}$$

Observe que esta propiedades implican

$$\boxed{\begin{aligned}\epsilon_{12} =& -1,\\ \epsilon_{11} =& \epsilon_{22} = 0.\end{aligned}}$$

Entonces, usando el objeto (4.9) el determinante (4.8) se puede escribir como

$$\boxed{\det A = \epsilon_{i_1 i_2} A_{1i_1} A_{2i_2}.} \tag{4.10}$$

Ahora consideremos un conjunto de dos objetos,

$$\{1, 2\},$$

de este conjunto tenemos las permutaciones

$$\sigma(12) = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad \sigma(21) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Definiremos la paridad de una permutación como

$$P(\sigma) = (-1)^P$$

donde P es el número de movimientos que se deben hacer para pasar de la permutación σ a la permutación identidad. Por lo que

$$P(\sigma(12)) = 1, \qquad P(\sigma(21)) = -1.$$

Se puede ver que se cumple

$$P(\sigma(12)) = \epsilon_{12}, \qquad P(\sigma(21)) = \epsilon_{21}.$$

Así, la cantidad ϵ_{ij} es un objeto totalmente antisimétrico que nos da la paridad de la permutación $\sigma(ij)$.

La versión de ϵ_{ij} para n dimensiones nos ayuradará definir el determinante para una matriz de $n \times n$.

Ahora definamos la cantidad

$$\boxed{\epsilon_{ijk}, \qquad ij, k = 1, 2, 3}$$

totalmente antisimétrica que satisface

$$\boxed{\epsilon_{123} = 1.}$$

Debido a que ϵ_{ijk} es una cantidad antisimétrica se tiene

$$\epsilon_{112} = -\epsilon_{112} = 0.$$

En general se tiene

$$\epsilon_{iij} = -\epsilon_{iij} = 0, \quad \epsilon_{ijj} = -\epsilon_{jij} = 0, \quad \epsilon_{iji} = -\epsilon_{iij} = 0, \quad \epsilon_{iii} = 0,$$

es decir, cada que vez que dos índices se repiten la cantidad ϵ_{ijk} es cero. Además, usando de nuevo que ϵ_{ijk} es una cantidad antisimétrica, se encuentra

$$\epsilon_{123} = 1, \quad \epsilon_{132} = -1, \quad \epsilon_{213} = 1, \quad \epsilon_{231} = -1, \quad \epsilon_{312} = 1, \quad \epsilon_{321} = -1.$$

Ahora veamos el conjunto de tres objetos

$$\{1, 2, 3\},$$

de este conjunto tenemos las siguientes permutaciones

$$\begin{aligned}
\sigma(123) &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma(132) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma(213) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \\
\sigma(231) &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma(312) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \sigma(321) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.
\end{aligned}$$

La paridad de estas permutaciones es

$$\begin{aligned}
P(\sigma(123)) &= 1, \quad P(\sigma(132)) = -1, \quad P(\sigma(213)) = -1, \\
P(\sigma(231)) &= 1, \quad P(\sigma(312)) = 1, \quad P(\sigma(321)) = -1,
\end{aligned}$$

note que se cumple

$$\begin{aligned}
P(\sigma(123)) &= \epsilon_{123}, \quad P(\sigma(132)) = \epsilon_{132}, \quad P(\sigma(213)) = \epsilon_{213}, \\
P(\sigma(231)) &= \epsilon_{231}, \quad P(\sigma(312)) = \epsilon_{312}, \quad P(\sigma(321)) = \epsilon_{321}.
\end{aligned}$$

Así, la cantidad ϵ_{ijk} es un objeto totalmente antisimétrico que nos da la paridad de la permutación $\sigma(ijk)$.

En general definiremos

$$\boxed{\epsilon_{i_1 i_2 \cdots i_n} \qquad i_j, j = 1, 2, \cdots n}$$

como el objeto totalmente antisimétrico que satisface

$$\boxed{\epsilon_{123\cdots n} = 1,}$$

a este objeto se le llama tensor de Levi-Civita, símbolo de Levi-Civita o símbolo de permutaciones.

4.3. Propiedades del tensor de Levi-Civita

Antes de definir el determinante, veremos algunas propiedades del tensor de Levi-Civita.

4.3.1. Intercambio de dos índices

Primero probaremos que si intercambiamos dos índices cualquiera i_p y i_q, se requiere un número impar de permutaciones. En efecto, supongamos que tenemos la permutación

$$\begin{pmatrix} 1 & 2 & \cdots & p & p+1 & \cdots & q-1 & q\cdots & n \\ i_1 & i_2 & \cdots & i_p & i_{p+1} & \cdots & i_{q-1} & i_q\cdots & i_n \end{pmatrix},$$

con $q - p = k$. Entonces, si queremos pasar el índice i_q del lugar p al lugar q se deben hacer k movimientos

$$\begin{pmatrix} 1 & 2 & \cdots & p & p+1 & \cdots & q-1 & q\cdots & n \\ i_1 & i_2 & \cdots & i_q & i_p & \cdots & i_{q-2} & i_{q-1}\cdots & i_n \end{pmatrix},$$

ahora para pasar al índice p del lugar $p + 1$ al lugar q debemos hacer $k - 1$ movimientos

$$\begin{pmatrix} 1 & 2 & \cdots & p & p+1 & \cdots & q-1 & q\cdots & n \\ i_1 & i_2 & \cdots & i_q & i_{p+1} & \cdots & i_{q-1} & i_p\cdots & i_n \end{pmatrix}.$$

Por lo tanto, para intercambiar dos índices debemos hacer $2k-1$ movimientos, es decir debemos hacer una permutación impar. Así, debido a que el tensor de Levi-Civita es totalmente antisimétrico, al intercambiar dos de sus índices tendremos un signo menos, esto es,

$$\epsilon_{i_1 i_2 \cdots i_p \cdots i_q \cdots i_n} = (-1)^{2k-1} \epsilon_{i_1 i_2 \cdots i_q \cdots i_q \cdots i_n} = -\epsilon_{i_1 i_2 \cdots i_q \cdots i_p \cdots i_n},$$

de donde

$$\boxed{\epsilon_{i_1 i_2 \cdots i_p \cdots i_q \cdots i_n} = -\epsilon_{i_1 i_2 \cdots i_q \cdots i_p \cdots i_n}.} \tag{4.11}$$

4.3.2. Índices iguales

También podemos observar que si el tensor de Levi-Civita tiene dos índices iguales, entonces es cero. En efecto, al cambiar los dos índices iguales debemos hacer $2k-1$ permutaciones y en ese cambio el tensor de Levi-Civita tendrá la misma forma, pero ahora tendremos un número negativo, es decir

$$\epsilon_{i_1 i_2 \cdots i_p \cdots i_p \cdots i_n} = -\epsilon_{i_1 i_2 \cdots i_p \cdots i_p \cdots i_n} = 0.$$

por lo que

$$\boxed{\epsilon_{i_1 i_2 \cdots i_p \cdots i_p \cdots i_n} = 0.}$$

4.3.3. Suma de permutaciones

Ahora recordemos que si tenemos un conjunto de n elementos, podemos hacer $n!$ permutaciones. Esto quiere decir que el tensor de Levi-Civita tiene $n!$ términos diferentes de cero. Además, debido a que se cumple que $\epsilon_{123\cdots n} = 1$, los términos diferentes de cero del tensor de Levi-Civita tiene el valor 1 o -1. Por lo tanto, el cuadrado de cualquier componente diferente de cero del tensor de Levi-Civita siempre es 1. Esta última afirmación nos permite concluir que la suma de los cuadrados de todos los términos diferentes de cero del tensor de Levi-Civita es $n!$, esto es

$$\boxed{\epsilon_{i_1 i_2 \cdots i_p \cdots i_n} \epsilon_{i_1 i_2 \cdots i_p \cdots i_n} = n!.}$$

Tullio Levi-Civita (1873-1941) fue un matemático italiano que realizó diversas contribuciones a las matemáticas y física. Realizó sus estudios en la Universidad de Padua y a partir de 1918 fue profesor de la Universidad de Roma. Desarrolló lo que hoy se conoce como cálculo tensorial, que es la base para la geometría diferencial. También realizó aportaciones en el problema de los tres cuerpos de la mecánica celeste y en la hidrodinámica. Los trabajos de Levi-Civita fueron fundamentales para el desarrollo de la Relatividad General de Einstein. De hecho, Albert Einstein mantuvo una fructífera correspondencia con él. Debido a sus orígenes judíos, fue despedido de su trabajo en 1938, lo cual lo hizo caer depresión. No obstante, luchó para que varios matemáticos italianos pudieran escapar de la Italia fascista de Mussolini.

4.4. Definición general del determinante

Para el caso de una matriz de 2×2 el determinante está definido en la ecuación (4.10). Si tenemos una matriz de $n \times n$

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \quad A_{ij} \in \mathbb{C}, \quad i, j = 1, \cdots n,$$

el determinante se define como

$$\boxed{\det A = \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}.} \tag{4.12}$$

Por ejemplo, para una matriz de 3×3

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix},$$

tenemos

$$\boxed{\det A = \epsilon_{i_1 i_2 i_3} A_{1i_1} A_{2i_2} \cdots A_{3i_n}.}$$

Considerando las propiedades del tensor de Levi-Civita, encontramos

$$\begin{aligned}
\det A &= \epsilon_{1i_2i_3}A_{11}A_{2i_2}A_{i3_n} + \epsilon_{2i_2i_3}A_{12}A_{2i_2}A_{i3_n} + \epsilon_{3i_2i_3}A_{13}A_{2i_2}A_{i3_n} \\
&= A_{11}(\epsilon_{1i_2i_3}A_{2i_2}A_{i3_n}) + A_{12}(\epsilon_{2i_2i_3}A_{2i_2}A_{i3_n}) + A_{13}(\epsilon_{3i_2i_3}A_{13}A_{2i_2}A_{i3_n}) \\
&= A_{11}(\epsilon_{123}A_{22}A_{33} + \epsilon_{132}A_{23}A_{32}) \\
&\quad +A_{12}(\epsilon_{213}A_{21}A_{33} + \epsilon_{231}A_{23}A_{31}) \\
&\quad +A_{13}(\epsilon_{312}A_{21}A_{32} + \epsilon_{321}A_{22}A_{31}) \\
&= A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{12}(A_{21}A_{33} - A_{23}A_{31}) + A_{13}(A_{21}A_{32} - A_{22}A_{31}).
\end{aligned}$$

esto es

$$\det A = A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{12}(A_{21}A_{33} - A_{23}A_{31}) + A_{13}(A_{21}A_{32} - A_{22}A_{31}).$$

4.5. Propiedades del determinante

Antes de continuar veremos algunas propiedades del determinante

4.5.1. Suma de dos renglones

Supongamos que tenemos la siguiente matriz de $n \times n$ con entradas complejas

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ A_{\mathbf{1}p1} + A_{\mathbf{2}p1} & \cdots & A_{\mathbf{1}pn} + A_{\mathbf{2}pn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix},$$

note que de esta matriz se puede obtener las dos matrices

$$A_1 = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ A_{\mathbf{1}p1} & \cdots & A_{\mathbf{1}pn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \quad A_2 = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ A_{\mathbf{2}p1} & \cdots & A_{\mathbf{2}pn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix},$$

Probaremos que se cumple la igualdad

$$\boxed{\det A = \det A_1 + \det A_2.}$$

Para realizar esta prueba, observe que por definición se cumple

$$\begin{aligned} \det A_1 &= \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{\mathbf{1}pi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n}, \\ \det A_2 &= \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{\mathbf{2}pi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n}, \end{aligned}$$

además,

$$\begin{aligned} \det A &= \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} (A_{\mathbf{1}pi_p} + A_{\mathbf{2}pi_p}) A_{(p+1)i_{p+1}} \cdots A_{ni_n} \\ &= \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} (A_{\mathbf{1}pi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n} + A_{\mathbf{2}pi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n}) \\ &= \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{\mathbf{1}pi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n} + \\ &\quad + \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{\mathbf{2}pi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n} \\ &= \det A_1 + \det A_2, \end{aligned}$$

esto es,

$$\det A = \det A_1 + \det A_2,$$

que es lo que queríamos demostrar.

Por ejemplo, de la matriz

$$A = \begin{pmatrix} 4 & 3 \\ 5 & 2 \end{pmatrix},$$

podemos extraer las matrices

$$A_1 = \begin{pmatrix} 2 & 1 \\ 5 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 2 \\ 5 & 2 \end{pmatrix}.$$

Se puede observar que las suma del primer renglón de A_1 con el primer renglón de A_2 nos da el primer renglón de A. Así mismo se cumple

$$\begin{aligned} \det A &= -7, \\ \det A_1 &= -1, \\ \det A_2 &= -6, \end{aligned}$$

por lo cual se satisface

$$\det A = \det A_1 + \det A_2.$$

4.5.2. Escalar por un renglón

Ahora, supongamos que A y A' son dos matrices de $n\times n$ donde los renglones de la matriz A' es son iguales a los de A, salvo el renglón p de A' que es múltiplo de renglón de p de A. Es decir, las matrices A y A' tienen la forma

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pn} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \quad A' = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ \lambda A_{p1} & \cdots & \lambda A_{pn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix},$$

por lo que las componentes de la matriz A' se pueden definir como

$$A'_{ij} = A_{ij}, \quad i \neq p, \quad A'_{pj} = \lambda A_{pj}.$$

En este caso se cumple

$$\boxed{\det A' = \lambda \det A.}$$

Para probar esta afirmación usaremos la definición de determinante (4.12), de donde

$$\begin{aligned} \det A' &= \epsilon_{i_1 i_2 \cdots i_n} A'_{1i_1} A'_{2i_2} \cdots A'_{ni_n} \\ &= \epsilon_{i_1 i_2 \cdots i_{p-1} i_p i_{p+1} \cdots i_n} A'_{1i_1} A'_{2i_2} \cdots A'_{(p-1)i_{p-1}} A'_{pi_p} A'_{(p+1)p} \cdots A'_{ni_n} \\ &= \epsilon_{i_1 i_2 \cdots i_{p-1} i_p i_{p+1} \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} \lambda A_{qi_p} A_{(p+1)p} \cdots A_{ni_n} \\ &= \lambda \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots \cdots A_{ni_n} \\ &= \lambda \det A, \end{aligned}$$

que es lo que queríamos demostrar.

Note que este último resultado implica que si A tiene un renglón igual a cero, entonces el determinante de A es cero.

Por ejemplo consideremos la matriz

$$A' = \begin{pmatrix} 6 & 3 \\ 1 & 2 \end{pmatrix},$$

se puede observar que el primer reglón de esta matriz es multiplo de 3, esto es

$$A' = \begin{pmatrix} 3 \cdot 2 & 3 \cdot 1 \\ 1 & 2 \end{pmatrix}.$$

Así podemos construir la matriz

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

Además se cumple

$$\begin{aligned} \det A' &= 9, \\ \det A &= 3, \end{aligned}$$

por lo cual se satisface

$$\det A' = 9 = 3 \cdot \det A = 3 \cdot 3.$$

4.5.3. Escalar por una matriz

Sea A un matriz de $n \times n$ y λ un escalar, entonces se cumple

$$\boxed{\det(\lambda A) = \lambda^n \det A.} \tag{4.13}$$

Para probar este resultado definamos la matriz $A' = \lambda A$, observe que esta definición implica $A'_{ij} = \lambda A_{ij}$. Por lo cual

$$\begin{aligned} \det(\lambda A) &= \det(A') \\ &= \epsilon_{i_1 i_2 \cdots i_n} A'_{1i_1} A'_{2i_2} \cdots A'_{ni_n} \\ &= \epsilon_{i_1 i_2 \cdots i_n} (\lambda A_{1i_1})(\lambda A_{2i_2}) \cdots (\lambda A_{ni_n}) \\ &= \lambda^n \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n} \\ &= \lambda^n \det A, \end{aligned}$$

esto es

$$\det(\lambda A) = \lambda^n \det A,$$

por lo que el resultado (4.13) es correcto.

Por ejemplo consideremos la matriz

$$A = \begin{pmatrix} 21 & 3 \\ 12 & 9 \end{pmatrix},$$

que se puede escribir como

$$A = 3 \begin{pmatrix} 7 & 1 \\ 4 & 3 \end{pmatrix}.$$

De donde podemos ver que se cumple

$$\det A = \det \begin{pmatrix} 21 & 3 \\ 12 & 9 \end{pmatrix} = 153,$$

$$\det \begin{pmatrix} 7 & 1 \\ 4 & 3 \end{pmatrix} = 17,$$

por lo cual

$$\det A = 3^2 \det \begin{pmatrix} 7 & 1 \\ 4 & 3 \end{pmatrix} = 3^2 \cdot 17 = 153.$$

4.5.4. Combinación lineal en un regón

Ahora veremos que los dos resultados anteriores se pueden unir y generalizar.

Supongamos que tenemos la matriz

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ \lambda_1 A_{1p1} + \lambda_2 A_{2p1} + \cdots + \lambda_l A_{lp1} & \cdots & \lambda_1 A_{1pn} + \lambda_2 A_{2pn} + \cdots + \lambda_l A_{lpn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix},$$

es decir

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ \lambda_k A_{kp1} & \cdots & \lambda_k A_{kpn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \qquad k = 1, 2, \cdots, l.$$

Note que de esta matriz se pueden obtener las siguientes l matrices

$$
\begin{aligned}
A_1 &= \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ A_{1p1} & \cdots & A_{1pn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \\
A_2 &= \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ A_{2p1} & \cdots & A_{2pn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \\
&\vdots \\
A_l &= \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ A_{lp1} & \cdots & A_{lpn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}.
\end{aligned}
$$

Estas matrices satisfacen

$$\boxed{\det A = \lambda_k \det A_k.} \tag{4.14}$$

Para realizar esta prueba, primero observemos que

$$\det A_k = \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{kpi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n},$$

además

$$
\begin{aligned}
\det A &= \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} (\lambda_k A_{kpi_p}) A_{(p+1)i_{p+1}} \cdots A_{ni_n} \\
&= \lambda_k \epsilon_{i_1 i_2 \cdots i_n} A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{kpi_p} A_{(p+1)i_{p+1}} \cdots A_{ni_n} \\
&= \lambda_k \det A_k,
\end{aligned}
$$

que es lo que queríamos demostrar.

4.5.5. Renglones como vectores

Supongamos que tenemos la siguiente matriz de $n\times n$ de entradas complejas

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & \ddots & \vdots\\ A_{n1} & \cdots & A_{nn}\end{pmatrix},\quad A_{ij}\in\mathbb{C},\quad i,j=1,\cdots n,$$

con los renglones de esta matriz se pueden formar los siguientes n vectores

$$\begin{aligned} R_1 &= (A_{11},A_{12},\cdots,A_{1n}),\\ R_2 &= (A_{21},A_{22},\cdots,A_{2n}),\\ R_3 &= (A_{31},A_{32},\cdots,A_{3n}),\\ &\vdots\\ R_n &= (A_{n1},A_{n2},\cdots,A_{nn}),\end{aligned}$$

Ahora, supongamos que A' es una matriz de $n\times n$ que se obtiene al intercambiar dos renglones de A, entonces se cumple

$$\boxed{\det A' = -\det A.}$$

Note que en este caso las matrices tienen la forma

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & & \vdots\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{p1} & \cdots & A_{pn}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ \vdots & & \vdots\\ A_{(q-1)1} & \cdots & A_{(q-1)n}\\ A_{q1} & \cdots & A_{qn}\\ A_{(q+1)1} & \cdots & A_{(q+1)n}\\ \vdots & & \vdots\\ A_{n1} & \cdots & A_{nn}\end{pmatrix},\quad A'=\begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & & \vdots\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{q1} & \cdots & A_{qn}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ \vdots & & \vdots\\ A_{(q-1)1} & \cdots & A_{(q-1)n}\\ A_{p1} & \cdots & A_{pn}\\ A_{(q+1)1} & \cdots & A_{(q+1)n}\\ \vdots & & \vdots\\ A_{n1} & \cdots & A_{nn}\end{pmatrix},$$

es decir, las componentes de la matriz A' se pueden definir como

$$A'_{ij}=A_{ij},\quad i\neq p, i\neq q,\quad A'_{pj}=A_{qj},\quad A'_{qj}=A_{pj}.$$

Por lo que, usando las propiedades del tensor de Levi-Civita (4.11), se obtiene

$$
\begin{aligned}
& \det A' = \epsilon_{i_1 i_2 \cdots i_n} A'_{1i_1} A'_{2i_2} \cdots A'_{ni_n} \\
= \ & \epsilon_{i_1 i_2 \cdots i_{p-1} i_p i_{p+1} \cdots i_{q-1} i_q i_{q+1} \cdots i_n} \\
& A'_{1i_1} A'_{2i_2} \cdots A'_{(p-1)i_{p-1}} A'_{pi_p} A'_{(p+1)i_{p+1}} \cdots A'_{(q-1)i_{q-1}} A'_{qi_q} A'_{(q+1)i_{q+1}} \cdots A'_{ni_n} \\
= \ & \epsilon_{i_1 i_2 \cdots i_{p-1} i_p i_{p+1} \cdots i_{q-1} i_q i_{q+1} \cdots i_n} \\
& A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{qi_p} A_{(p+1)i_{p+1}} \cdots A_{(q-1)i_{q-1}} A_{pi_q} A_{(q+1)i_{q+1}} \cdots A_{ni_n} \\
= \ & (-)\epsilon_{i_1 i_2 \cdots i_{p-1} i_q i_{p+1} \cdots i_{q-1} i_p i_{q+1} \cdots i_n} \\
& A_{1i_1} A_{2i_2} \cdots A_{(p-1)i_{p-1}} A_{pi_q} A_{(p+1)i_{p+1}} \cdots A_{(q-1)i_{q-1}} A_{qi_p} A_{(q+1)i_{q+1}} \cdots A_{ni_n} \\
= \ & -\det A,
\end{aligned}
$$

que es lo que queríamos demostrar.

Observe que este último resultado implica que si A tiene dos renglones iguales, entonces su determinante es cero.

Por ejemplo, consideremos la matriz

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix},$$

al intercambiar los dos renglones tenemos

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

Además, se cumple

$$
\begin{aligned}
\det A &= 4 - 6 = -2, \\
\det A' &= 6 - 4 = 2,
\end{aligned}
$$

por lo cual

$$\det A = -\det A'.$$

También podemos ver que la matriz

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 4 & 7 \\ 3 & 1 & 5 \end{pmatrix},$$

tiene dos renglones iguales y se cumple

$$\det M = 0.$$

4.5.6. Renglones linealmente dependientes

Sea la matriz A de la forma

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \quad A_{ij} \in \mathbb{C}, \quad i, j = 1, \cdots n,$$

cuyos renglones son los n vectores

$$R_i = (A_{i1}, A_{i2}, \cdots, A_{in}).$$

Mostraremos que si uno de estos vectores es linealmente dependiente, el determinante de A es cero.

Sin pérdida de generalidad, supongamos que el renglón p es linealmente dependiente, es decir,

$$R_p = \sum_{k=1,k\neq p}^{n} \lambda_k R_k = \left(\sum_{k=1,k\neq p}^{n} \lambda_k A_{k1}, \sum_{k=1,k\neq p}^{n} \lambda_k A_{k2}, \cdots, \sum_{k=1,k\neq p}^{n} \lambda_k A_{kn} \right),$$

note que, en este caso la matriz A tienen la forma

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ \sum_{k=1,k\neq p}^{n} \lambda_k A_{k1} & \cdots & \sum_{k=1,k\neq p}^{n} \lambda_k A_{kn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \tag{4.15}$$

es decir la matriz A se puede escribir como

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{(p-1)1} & \cdots & A_{(p-1)n} \\ A_{p1} & \cdots & A_{pn} \\ A_{(p+1)1} & \cdots & A_{(p+1)n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix},$$

con

$$
\begin{aligned}
A_{p1} &= \sum_{k=1,k\neq p}^{n} \lambda_k A_{k1} = \lambda_1 A_{11} + \cdots + \lambda_{p-1} A_{(p-1)1} + \lambda_{p+1} A_{(p+1)1} + \cdots + \lambda_n A_{n1},\\
A_{p2} &= \sum_{k=1,k\neq p}^{n} \lambda_k A_{k2} = \lambda_1 A_{12} + \cdots + \lambda_{p-1} A_{(p-1)2} + \lambda_{p+1} A_{(p+1)2} + \cdots + \lambda_n A_{n2},\\
&\vdots\\
A_{pn} &= \sum_{k=1,k\neq p}^{n} \lambda_k A_{kn} = \lambda_1 A_{1n} + \cdots + \lambda_{p-1} A_{(p-1)n} + \lambda_{p+1} A_{(p+1)n} + \cdots + \lambda_n A_{nn}.
\end{aligned}
$$

Observe que de la matriz (4.15) se obtiene las siguientes $n-1$ matrices

$$
\begin{aligned}
A_1 &= \begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & & \vdots\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{11} & \cdots & A_{1n}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ \vdots & & \vdots\\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, A_2 = \begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & & \vdots\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{21} & \cdots & A_{2n}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ \vdots & & \vdots\\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \cdots,\\
A_{(p-1)} &= \begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & & \vdots\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ \vdots & & \vdots\\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, A_{(p+1)} = \begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & & \vdots\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ \vdots & & \vdots\\ A_{n1} & \cdots & A_{nn} \end{pmatrix}, \cdots,\\
A_n &= \begin{pmatrix} A_{11} & \cdots & A_{1n}\\ \vdots & & \vdots\\ A_{(p-1)1} & \cdots & A_{(p-1)n}\\ A_{n1} & \cdots & A_{nn}\\ A_{(p+1)1} & \cdots & A_{(p+1)n}\\ \vdots & & \vdots\\ A_{n1} & \cdots & A_{nn} \end{pmatrix}.
\end{aligned}
$$

También, note que todas estas matrices tienen dos renglones iguales y por lo tanto su determinante es igual a cero. Además, usado el resultado (4.14), se obtiene

$$\det A = \sum_{k=1, k\neq p}^{n} \lambda_{\mathbf{k}} \det A_{\mathbf{k}} = 0.$$

Por lo tanto, si un rengón de la matriz A es linealmente dependiente, entonces

$$\boxed{\det A = 0.} \tag{4.16}$$

Así, si una matriz A satisface

$$\boxed{\det A \neq 0,} \tag{4.17}$$

ninguno de sus renglones es linealmente dependiente. En otras palabras si se cumple (4.17), entonces todos los renglones de la matriz A son linealmente independientes.

Por ejemplo, consideremos la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 4 & 3 \end{pmatrix}.$$

Se puede observa que se cumple

$$(5, 4, 3) = -3(1, 2, 3) + 2(4, 5, 6),$$

por lo que el tercer renglón es linealmente dependiente. De donde

$$\det A = 0.$$

Ahora veamos la matriz

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

En este caso los vectores

$$(1, 0, 0), \quad (1, 1, 0), \quad (1, 1, 1),$$

son linealmente independientes y se satisface

$$\det M = 1 \neq 0.$$

4.6. Versión alternativa del determinante

Ahora veremos una versión alternativa del determinante que nos permitirá obtener resultados de una forma más directa y clara.

Para iniciar, consideremos la expresión

$$a = \epsilon_{i_1 i_2 i_3 \cdots i_n} A_{2i_1} A_{1i_2} A_{3i_3} \cdots A_{ni_n}. \tag{4.18}$$

Note que esta a no es el determinante, pues los términos A_{1i_2} y A_{2i_1} están intercambiados. Al intercambiar estos términos, obtenemos

$$a = \epsilon_{i_1 i_2 i_3 \cdots i_n} A_{1i_2} A_{2i_1} A_{3i_3} \cdots A_{ni_n},$$

Aún no tenemos el determinante, pues los índices del tensor de Levi-Civita están intercambiados. Entonces al intercambiar los dos primeros índices del tensor de Levi-Civita se tiene

$$a = -\epsilon_{i_2 i_1 i_3 \cdots i_n} A_{1i_2} A_{2i_1} A_{3i_3} \cdots A_{ni_n}.$$

Adicionalmente, definiendo

$$i_1' = i_2, \quad i_2' = i_1, \quad i_j' = i_j, \quad j > 2,$$

se encuentra

$$\begin{aligned} a &= -\epsilon_{i_1' i_2' \cdots i_n'} A_{1i_1'} A_{2i_2'} A_{3i_3'} \cdots A_{ni_n'} \\ &= -\det A, \end{aligned}$$

esto es

$$\det A = -a = -\epsilon_{i_1 i_2 i_3 \cdots i_n} A_{2i_1} A_{1i_2} A_{3i_3} \cdots A_{ni_n},$$

o sea

$$\det A = -\epsilon_{i_1 i_2 i_3 \cdots i_n} A_{2i_1} A_{1i_2} A_{3i_3} \cdots A_{ni_n}. \tag{4.19}$$

Además, considerando que se cumple

$$\epsilon_{213\cdots n} = -\epsilon_{123\cdots n} = -1$$

la expresión (4.19) se puede escribir como

$$\det A = \epsilon_{i_1 i_2 i_3 \cdots i_n} \epsilon_{213\cdots n} A_{2i_1} A_{1i_2} A_{3i_3} \cdots A_{ni_n}. \tag{4.20}$$

El resultado (4.20) se consiguió permutando dos índices del tensor de Levi-Civita y cambiando de lugar dos términos del producto de los coeficientes A_{ij}.

Ahora veremos que se puede obtener un resultado semejante con cualquier permutación. Para el caso general, consideremos la permutación

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}, \tag{4.21}$$

con paridad $(-)^P$. Note que en ese caso se tiene

$$\epsilon_{\sigma(1)\sigma(2)\cdots\sigma(n)} = (-)^P. \tag{4.22}$$

Entonces, definamos la cantidad

$$b = \epsilon_{i_1 i_2 \cdots i_n} A_{\sigma(1)i_1} A_{\sigma(2)i_2} \cdots A_{\sigma(n)i_n} \tag{4.23}$$

y definamos los números

$$k_1, k_2, \cdots, k_n$$

tales que, bajo la permutación (4.21), se satisface

$$\sigma(k_1) = 1, \quad \sigma(k_2) = 2, \cdots, \sigma(k_n) = n. \tag{4.24}$$

Entonces, reordenando los términos del producto de los coeficientes A_{ij} en la expresión (4.23), se tiene

$$b = \epsilon_{i_1 i_2 \cdots i_n} A_{1 i_{k_1}} A_{2 i_{k_2}} \cdots A_{n i_{k_n}}. \tag{4.25}$$

Ahora debemos ordenar los índices del tensor de Levi-Civita. Para realizar esta ordenación, debemos pasar el índice k_1 a lugar 1, el índice k_2 a lugar 2, el índice k_3 a lugar 3, etc. En otras palabras, en los índices del tensor de Levi-Civita debemos hacer la permutación (4.21). Se puede observar que al hacer dicha permutación se obtendrá un factor de

$$(-1)^P.$$

Por lo cual, la expresión (4.25) se puede escribir de la forma

$$b = (-1)^P \epsilon_{i_{k_1} i_{k_2} \cdots i_{k_n}} A_{1 i_{k_1}} A_{2 i_{k_2}} \cdots A_{n i_{k_n}}.$$

Ahora, definiendo

$$i'_1 = i_{k_1}, i'_2 = i_{k_2}, \cdots, i'_n = i_{k_n},$$

tenemos

$$\begin{aligned} b &= (-1)^P \epsilon_{i'_1 i'_2 \cdots i'_n} A_{1i'_1} A_{2i'_2} \cdots A_{ni'_n} \\ &= (-1)^P \det A, \end{aligned}$$

esto es

$$b = (-)^P \det A.$$

Este resultado implica la igualdad

$$\det A = (-)^P b = (-)^P \epsilon_{i_1 i_2 \cdots i_n} A_{\sigma(1) i_1} A_{\sigma(2) i_2} \cdots A_{\sigma(n) i_n},$$

de donde

$$\det A = (-)^P \epsilon_{i_1 i_2 \cdots i_n} A_{\sigma(1) i_1} A_{\sigma(2) i_2} \cdots A_{\sigma(n) i_n}.$$

Además, considerando la igualdad (4.22) se encuentra

$$\begin{aligned} \det A &= (-)^P \epsilon_{i_1 i_2 \cdots i_n} A_{\sigma(1) i_1} A_{\sigma(2) i_2} \cdots A_{\sigma(n) i_n} \\ &= \epsilon_{\sigma(1)\sigma(2)\cdots\sigma(n)} \epsilon_{i_1 i_2 \cdots i_n} A_{\sigma(1) i_1} A_{\sigma(2) i_2} \cdots A_{\sigma(n) i_n} \end{aligned}$$

por lo cual

$$\det A = \epsilon_{\sigma(1)\sigma(2)\cdots\sigma(n)} \epsilon_{i_1 i_2 \cdots i_n} A_{\sigma(1) i_1} A_{\sigma(2) i_2} \cdots A_{\sigma(n) i_n}. \tag{4.26}$$

Ahora, por cada permutación tenemos una expresión de la forma (4.26), entonces como existen $n!$ permutaciones, se encuentra

$$\epsilon_{i_1 i_2 \cdots i_n} \epsilon_{j_1 j_2 \cdots j_n} A_{i_1 j_1} A_{i_2 j_2} \cdots A_{i_i j_n} = n! \det A$$

es decir,

$$\boxed{\det A = \frac{1}{n!} \epsilon_{i_1 i_2 \cdots i_n} \epsilon_{j_1 j_2 \cdots j_n} A_{i_1 j_1} A_{i_2 j_2} \cdots A_{i_i j_n}.} \tag{4.27}$$

Esta versión del determinante es muy útil para realizar diversas demostraciones.

4.7. Determinante de la transpuesta de una matriz

Supongamos que si A es una matriz de $n\times n$, entonces se cumple la igualdad

$$\boxed{\det A^T = \det A.} \tag{4.28}$$

Para probar esta afirmación, observe que usando la ecuación (4.27), tenemos

$$\begin{aligned}
\det A^T &= \frac{1}{n!}\epsilon_{i_1 i_2\cdots i_n}\epsilon_{j_1 j_2\cdots j_n}(A)^T_{i_1 j_1}(A)^T_{i_2 j_2}\cdots(A)^T_{i_n j_n}\\
&= \frac{1}{n!}\epsilon_{i_1 i_2\cdots i_n}\epsilon_{j_1 j_2\cdots j_n}A_{j_1 i_1}A_{j_2 i_2}\cdots A_{j_n i_n}\\
&= \frac{1}{n!}\epsilon_{j_1 j_2\cdots j_n}\epsilon_{i_1 i_2\cdots i_n}A_{j_1 i_1}A_{j_2 i_2}\cdots A_{j_n i_n},
\end{aligned}$$

esto es,

$$\det A^T = \frac{1}{n!}\epsilon_{j_1 j_2\cdots j_n}\epsilon_{i_1 i_2\cdots i_n}A_{j_1 i_1}A_{j_2 i_2}\cdots A_{j_n i_n}.$$

Ahora, definiendo

$$i'_k = j_k, \qquad j'_k = i_k$$

tenemos

$$\begin{aligned}
\det A^T &= \frac{1}{n!}\epsilon_{j_1 j_2\cdots j_n}\epsilon_{i_1 i_2\cdots i_n}A_{j_1 i_1}A_{j_2 i_2}\cdots A_{j_n i_n}\\
&= \frac{1}{n!}\epsilon_{i'_1 i'_2\cdots i'_n}\epsilon_{j'_1 j'_2\cdots j'_n}A_{i'_1 j'_1}A_{i'_2 j'_2}\cdots A_{i'_n j'_n},\\
&= \det A,
\end{aligned}$$

de donde

$$\det A^T = \det A.$$

Así, hemos probado el resultado (4.28).

Particularmente, si A es antisimétrica, es decir, si cumple

$$A^T = -A,$$

se obtiene

$$\det A^T = \det(-A) = (-)^n \det A = \det A.$$

es decir

$$\det A = (-)^n \det A.$$

Por lo que si n es impar y A es una matriz antisimétrica se encuentra

$$\boxed{\det A = 0.}$$

4.7.1. Ejemplos

Consideremos la matriz

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & -1 & 4 \\ 2 & 1 & -2 \end{pmatrix},$$

en este caso el determinate es

$$\det A = -6.$$

Además se puede observar que la transpuesta de A es

$$A^T = \begin{pmatrix} 1 & 3 & 2 \\ -1 & -1 & 1 \\ 2 & 4 & -2 \end{pmatrix},$$

que tiene el determinante

$$\det A^T = -6.$$

Por lo que se cumple

$$\det A = \det A^T = -6.$$

También se puede observar que la siguiente matriz de 3×3

$$M = \begin{pmatrix} 0 & 3 & -2 \\ -3 & 0 & 4 \\ 2 & -4 & 0 \end{pmatrix},$$

es antisimétrica y se cumple

$$\det M = 0.$$

4.8. Determinante de un producto de matrices

Supongamos que A y B son matrices de $n \times n$, probaremos que se cumple

$$\boxed{\det(AB) = \det A \det B.} \tag{4.29}$$

Se puede observar que de la ecuación (4.27) se tiene

$$\begin{aligned}
\det(AB) &= \frac{1}{n!}\epsilon_{i_1 i_2 \cdots i_n}\epsilon_{j_1 j_2 \cdots j_n}(AB)_{i_1 j_1}(AB)_{i_2 j_2}\cdots(AB)_{i_n j_n} \\
&= \frac{1}{n!}\epsilon_{i_1 i_2 \cdots i_n}\epsilon_{j_1 j_2 \cdots j_n}A_{i_1 k_1}B_{k_1 j_1}A_{i_2 k_2}B_{k_2 j_2}\cdots A_{i_n k_n}B_{k_n j_n} \\
&= \frac{1}{n!}\left(\epsilon_{i_1 i_2 \cdots i_n}A_{i_1 k_1}A_{i_2 k_2}\cdots A_{i_n k_n}\right)\left(\epsilon_{j_1 j_2 \cdots j_n}B_{k_1 j_1}B_{k_2 j_2}\cdots B_{k_n j_n}\right).
\end{aligned}$$

Antes de continuar analicemos la expresión

$$\epsilon_{i_1 i_2 \cdots i_n}A_{i_1 k_1}A_{i_2 k_2}\cdots A_{i_n k_n},$$

observe que usando la definición de matriz transpuesta, obtenemos

$$\epsilon_{i_1 i_2 \cdots i_n}A_{i_1 k_1}A_{i_2 k_2}\cdots A_{i_n k_n} = \epsilon_{i_1 i_2 \cdots i_n}A^T_{k_1 i_1}A^T_{k_2 i_2}\cdots A^T_{k_n i_n}. \tag{4.30}$$

También recordemos que según la ecuación (4.12), el determinante de la matriz transpuesta se puede escribir como

$$\det A^T = \epsilon_{i_1 i_2 \cdots i_n}A^T_{1 i_1}A^T_{2 i_2}\cdots A^T_{n i_n}. \tag{4.31}$$

Notemos que para pasar de la expresión (4.30) a la expresión (4.31) debemos colocar el índice 1 en lugar del índice k_1, el índice 2 en lugar k_2 y así sucesivamente hasta colocar n en el lugar del índice k_n. En otras palabras, para pasar de la expresión (4.30) a la expresión (4.31) debemos realizar la permutación

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}, \tag{4.32}$$

que supondremos que tiene paridad P. Por lo que las expresiones (4.30) y (4.31) difieren en un término de la forma $(-1)^P$. Entonces

$$\begin{aligned}
\epsilon_{i_1 i_2 \cdots i_n}A_{i_1 k_1}A_{i_2 k_2}\cdots A_{i_n k_n} &= \epsilon_{i_1 i_2 \cdots i_n}A^T_{k_1 i_1}A^T_{k_2 i_2}\cdots A^T_{k_n i_n} \\
&= (-)^P \det A^T \\
&= (-)^P \det A.
\end{aligned}$$

esto es

$$\epsilon_{i_1 i_2 \cdots i_n} A_{i_1 k_1} A_{i_2 k_2} \cdots A_{i_n k_n} = (-)^P \det A. \tag{4.33}$$

De la misma forma se cumple

$$\epsilon_{j_1 j_2 \cdots j_n} B_{k_1 j_1} B_{k_2 j_2} \cdots B_{k_n j_n} = (-)^P \det B. \tag{4.34}$$

Considerando los resultados (4.33) y (4.34), si los índices $k_1, k_2, \cdots, k_n$ están fijos, se encuentra

$$\epsilon_{i_1 i_2 \cdots i_n} A_{i_1 k_1} A_{i_2 k_2} \cdots A_{i_n k_n} \epsilon_{j_1 j_2 \cdots j_n} B_{k_1 j_1} B_{k_2 j_2} \cdots B_{k_n j_n} = \det A \det B.$$

Así, debido a que existen $n!$ permutaciones, si los índices $k_1, k_2, \cdots, k_n$ no están fijos y se suman, tenemos

$$\left(\epsilon_{i_1 i_2 \cdots i_n} A_{i_1 k_1} A_{i_2 k_2} \cdots A_{i_n k_n}\right)\left(\epsilon_{j_1 j_2 \cdots j_n} B_{k_1 j_1} B_{k_2 j_2} \cdots B_{k_n j_n}\right) = n! \det A \det B.$$

Usando este resultado en la ecuación (4.30) llegamos a

$$\begin{aligned} \det(AB) &= \frac{1}{n!}\left(\epsilon_{i_1 i_2 \cdots i_n} A_{i_1 k_1} A_{i_2 k_2} \cdots A_{i_n k_n}\right)\left(\epsilon_{j_1 j_2 \cdots j_n} B_{k_1 j_1} B_{k_2 j_2} \cdots B_{k_n j_n}\right). \\ &= \frac{1}{n!} n! \det A \det B = \det A \det B. \end{aligned}$$

es decir

$$\det(AB) = \det A \det B.$$

Por lo tanto la expresión (4.29) es correcta.

En especial si

$$A^{-1} A = I,$$

entonces

$$1 = \det I = \det\left(A^{-1} A\right) = \det\left(A^{-1}\right) \det(A).$$

Por lo tanto si $\det(A) \neq 0$, entonces

$$\boxed{\det\left(A^{-1}\right) = \frac{1}{\det A}.}$$

4.8.1. Ejemplos

Como un ejemplo consideremos las matrices

$$A = \begin{pmatrix} 2 & -3 & 2 \\ 1 & -1 & 0 \\ 2 & 4 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} -2 & 1 & 5 \\ 3 & 0 & 1 \\ 6 & 1 & 1 \end{pmatrix},$$

cuyo producto es

$$AB = \begin{pmatrix} -1 & -4 & 9 \\ -5 & 1 & 4 \\ -4 & 0 & 12 \end{pmatrix}.$$

Además, los determinantes son

$$\begin{aligned} \det A &= 10, \\ \det B &= 20, \\ \det(AB) &= 200, \end{aligned}$$

por lo que se cumple

$$\det(AB) = \det(A)\det(B).$$

Como otro ejemplo, consideremos la matriz

$$M = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

cuya inversa es

$$M^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

En este caso se tienen los determinantes

$$\begin{aligned} \det M &= 2, \\ \det M^{-1} &= \frac{1}{2}, \end{aligned}$$

por lo cual se cumple

$$\det M^{-1} = \frac{1}{\det M}.$$

En 1812, Augustin-Louis Cauchy introduce la versión moderna de determinante. Sin embargo, en China el uso de determinantes para resolver sistemas de ecuaciones lineales data de dos siglos antes de Cristo. En la idea de determinante convergen diferentes áreas de las matemáticas como la combinatoria, el álgebra y la geometría euclidiana. Por ejemplo, en 1773 Lagrange muestra que los determinantes de 3×3 están relacionados con el volumen de un paralelépipedo. De la misma forma, el determinante es importante para el cálculo de varias variables y la geometría diferencial.

Augustin-Louis Cauchy (1789-1857) fue un matemático francés que realizó aportaciones fundamentales en áreas como el análisis matemático, variable compleja, geometría y mecánica analítica. Publicó un poco menos de 800 artículos y su nombre está asociado a teoremas de diferentes áreas de las matemáticas. Además, a Cauchy se le debe la fundamentación rigurosa de diversos conceptos matemáticos, como los de límite y convergencia. Cauchy nació un mes antes de que iniciara la revolución francesa. El padre de Cauchy fue funcionario de la monarquía francesa, pero perdió su cargo cuando la monarquía fue derrotada. El padre de Cauchy pudo acomodarse en el nuevo régimen republicano francés y cultivó amistad con matemáticos republicanos como Lagrange y Laplace. En particular, Lagrange influyó en la educación de Cauchy hijo y lo recomendó para sus estudios universitarios. Como estudiante, Cauchy destaca en humanidades, pero se decide por las matemáticas y la física. Al terminar sus estudios universitarios trabaja como ingeniero en la naval de Napoleón, pero su sueño era ser profesor de matemáticas y realizar investigación. Mientras trabajaba como ingeniero, dedicó su tiempo libre a la investigación matemática y logró algunos resultados importantes. Después de que Napoleón es derrotado y se reinstala la monarquía en Francia, se reestructura la Academia de Ciencias de Francia y varios acadámicos como Carnot son despedidos. Sin embargo, la monarquía le ofrece Cauchy un puesto en la Academia de Ciencias y él acepta, por lo que se le consideró como un traidor por parte de sus colegas republicanos. Cuando la Monarquía es derrotada de nuevo, Cauchy tiene que huir de Francia a Suiza y después a varios países de Europa. Cuando se estabiliza Francia, Cauchy pudo retomar su trabajo y terminó su vida como profesor de matemáticas. Cauchy es reconocido como uno de los matemáticos más importantes de Francia y su nombre se encuentra gravado en la Torre Eiffel.

4.9. Operaciones con Python

Si queremos calcular el determinante de una matriz de la forma

$$A = \begin{pmatrix} 1 & 7 & 8 & 9 \\ -1 & 3 & 6 & 5 \\ 2 & 1 & 0 & -3 \\ -1 & 2 & 1 & 8 \end{pmatrix}$$

podemos emplear el siguiente código en Python

```python
import numpy as np
A = np.array([[1,7,8,9],[-1,3,6,5],[2,1,0,-3],[-1,2,1,8]
             ])
np.linalg.det(A)
```

Listing 4.1: Determinante de una matriz

el cual da el resultado

```
-4
```

4.10. Ejercicios

1.- Considere las matrices

$$A = \begin{pmatrix} 1 & -12 & 7 \\ 3 & -11 & 4 \\ 2 & 1 & -5 \end{pmatrix}$$

a) Obtener $\det A$.

b) Obtener A^T y $\det A^T$.

c) Obtener $\det(7A)$.

2.- Considere las matrices

$$A = \begin{pmatrix} 11 & -3 & 1 \\ 0 & 2 & 5 \\ 3 & 1 & -4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -3 & 6 \\ 2 & 4 & 3 \\ 7 & 5 & 1 \end{pmatrix}.$$

a) Obtener AB.

b) Obtener $\det A$ y $\det B$

c) Obtener $\det(B)\det(A)$.

d) Obtener $\det(BA)$.

Capítulo 5

Matriz adjunta e inversa de una matriz

En este capítulo veremos la definición de matriz adjunta. Esta definición es importante porque nos ayuda a obtener la matriz inversa de una matriz.

5.1. Definición

Hemos mostrado que el determinante de una matriz A se puede escribir como

$$\boxed{\det A = \frac{1}{n!}\epsilon_{i_1 i_2\cdots i_n}\epsilon_{j_1 j_2\cdots j_n}A_{i_1 j_1}A_{i_2 j_2}\cdots A_{i_n j_n}.}$$

Observe que dado un número fijo r tenemos

$$\begin{aligned}\det A &= \frac{1}{n!}\Big(\epsilon_{r i_2\cdots i_n}\epsilon_{i j_2\cdots j_n}A_{ri}A_{i_2 j_2}\cdots A_{i_n j_n} + \\ &\quad +\epsilon_{i_1 r i_3\cdots i_n}\epsilon_{j_1 i j_3\cdots j_n}A_{i_1 j_1}A_{ri}A_{i_3 j_3}\cdots A_{i_n j_n} + \\ &\quad +\cdots+\epsilon_{i_1 i_2\cdots i_{n-1} r}\epsilon_{j_1 i j_3\cdots j_{n-1} i}A_{i_1 j_1}A_{i_3 j_3}\cdots A_{i_{n-1} j_{n-1}}A_{ri}\Big),\end{aligned}$$

note que en esta expresión el índice r es fijo, mientras que el índice i se suma. Además, factorizando el término A_{ri} llegamos a

$$\begin{aligned}\det A &= \frac{1}{n!}A_{ri}\Big(\epsilon_{r i_2\cdots i_n}\epsilon_{i j_2\cdots j_n}A_{i_2 j_2}\cdots A_{i_n j_n} + \\ &\quad +\epsilon_{i_1 r i_3\cdots i_n}\epsilon_{j_1 i j_3\cdots j_n}A_{i_1 j_1}A_{i_3 j_3}\cdots A_{i_n j_n} + \\ &\quad +\cdots+\epsilon_{i_1 i_2\cdots i_{n-1} r}\epsilon_{j_1 i j_3\cdots j_{n-1} i}A_{i_1 j_1}A_{i_2 j_2}\cdots A_{i_{n-1} j_{n-1}}\Big),\end{aligned}$$

ahora pasado al inicio de cada tensor de Levi-Civita los índices r, i obtenemos

$$\begin{aligned}
\det A &= \frac{n}{n!} A_{ri} \epsilon_{k i_2 \cdots i_n} \epsilon_{i j_2 \cdots j_n} A_{i_2 j_2} \cdots A_{i_n j_n} \\
&= \frac{1}{(n-1)!} A_{ri} \epsilon_{k i_2 \cdots i_n} \epsilon_{i j_2 \cdots j_n} A_{i_2 j_2} \cdots A_{i_n j_n}.
\end{aligned}$$

esto es

$$\det A = \frac{1}{(n-1)!} A_{ri} \epsilon_{k i_2 \cdots i_n} \epsilon_{i j_2 \cdots j_n} A_{i_2 j_2} \cdots A_{i_n j_n}. \tag{5.1}$$

A la matriz

$$\boxed{(Adj A)_{ri} = \frac{1}{(n-1)!} \epsilon_{r i_2 \cdots i_n} \epsilon_{i j_2 \cdots j_n} A_{i_2 j_2} \cdots A_{i_n j_n}.} \tag{5.2}$$

le llamaremos matriz adjunta de la matriz A. Observe que en la matriz adjunta $Adj A$ están todos los renglones de A excepto el renglón r.

También observe que usando la igualdad (5.1) tenemos que

$$\det A = A_{ri} (Adj A)_{ri} = A_{ri} (Adj A)^T_{ir} = (A (Adj A)^T)_{rr},$$

es decir

$$\boxed{\det A = \left(A (Adj A)^T\right)_{rr}.} \tag{5.3}$$

Como este resultado es válido para cualquier número r, entonces las componentes de la diagonal de la matriz $A\,(Adj A)^T$ son todas iguales al determinante de A, $\det A$.

Ahora veamos como son las componentes de la matriz $A(Adj A)^T$ fuera de la diagonal. Consideremos $r \neq s$ y que

$$\left(A (Adj A)^T\right)_{rs} = A_{rk} (Adj A)^T_{ks} = A_{rk} (Adj A)_{sk}, \tag{5.4}$$

note que en la definición de las componentes de la matriz $(Adj A)_{sk}$ están todos los renglones de la matriz A excepto el renglón s. En particular, está el renglón r. Ahora, usando la definición de la matriz adjunta en la expresión (5.4), obtenemos

$$\begin{aligned}
\left(A (Adj A)^T\right)_{rs} &= A_{rk} \frac{1}{(n-1)!} \epsilon_{s i_2 \cdots i_n} \epsilon_{k j_2 \cdots j_n} A_{i_2 j_2} \cdots A_{i_n j_n} \\
&= \frac{1}{(n-1)!} \epsilon_{s i_2 \cdots i_n} \epsilon_{k j_2 \cdots j_n} A_{rk} A_{i_2 j_2} \cdots A_{i_n j_n}. \qquad (5.5)
\end{aligned}$$

Debido a que el reglón r ya está considerado en la definición de la matriz adjunta (5.2), la ecuación (5.5) representa al determinante de una matriz con dos renglones que se repiten, entonces se debe cumplir que

$$\left(A(AdjA)^T\right)_{rs} = 0, \qquad r \neq s.$$

Por lo tanto,

$$\left(A(AdjA)^T\right)_{rs} = \det A\delta_{rs},$$

es decir,

$$A(AdjA)^T = (\det A)I. \tag{5.6}$$

de donde

$$A\frac{(AdjA)^T}{\det A} = I.$$

De este resultado se puede deducir que si

$$\boxed{\det A \neq 0,}$$

entonces la inversa de la matriz A es

$$\boxed{A^{-1} = \frac{1}{\det A}(AdjA)^T.}$$

Por lo tanto, una matriz A tiene inversa siempre y cuando sus renglones (y por lo tanto sus columnas) sean linealmente independientes.

5.2. Matriz de 2×2

Como un ejemplo veamos la matriz de 2×2

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

podemos ver que en este caso la matriz adjunta (5.2) es

$$\boxed{(Adj)_{ab} = \epsilon_{ai}\epsilon_{bj}A_{ij}, \quad a, b, i, j = 1, 2.}$$

De donde se obtiene

$$\begin{aligned}
(Adj)_{11} &= \epsilon_{1i}\epsilon_{1j}A_{ij} = \epsilon_{12}\epsilon_{12}A_{22} = A_{22}, \\
(Adj)_{12} &= \epsilon_{1i}\epsilon_{2j}A_{ij} = \epsilon_{12}\epsilon_{21}A_{21} = -A_{21}, \\
(Adj)_{21} &= \epsilon_{2i}\epsilon_{1j}A_{ij} = \epsilon_{21}\epsilon_{12}A_{12} = A_{12}, \\
(Adj)_{22} &= \epsilon_{2i}\epsilon_{2j}A_{ij} = \epsilon_{21}\epsilon_{21}A_{11} = A_{11},
\end{aligned}$$

así tenemos la matriz

$$Adj(A) = \begin{pmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{pmatrix},$$

por lo que la matriz inversa es

$$A^{-1} = \frac{1}{\det A} Adj(A)^T = \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}.$$

5.3. Matriz de 3×3

Como otro ejemplo veamos la matriz de 3×3

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}.$$

Podemos ver que en este caso la matriz adjunta (5.2) es

$$\boxed{(Adj)_{ab} = \frac{1}{2!}\epsilon_{ai_1i_2}\epsilon_{bj_1j_2}A_{i_1j_1}A_{i_2j_2}, \quad a, b = 1, 2, 3,}$$

De donde se obtiene

$$\begin{aligned}
(Adj)_{11} &= \frac{1}{2}\epsilon_{1i_1i_2}\epsilon_{1j_1j_2}A_{i_1j_1}A_{i_2j_2} \\
&= \frac{1}{2}\left(\epsilon_{123}\epsilon_{1j_1j_3}A_{2j_1}A_{3j_3} + \epsilon_{132}\epsilon_{1j_1j_2}A_{3j_1}A_{2j_2}\right) \\
&= \frac{1}{2}\epsilon_{1j_1j_3}\left(A_{2j_1}A_{3j_3} - A_{3j_1}A_{2j_2}\right) \\
&= \frac{1}{2}\epsilon_{123}\left(A_{22}A_{33} - A_{32}A_{23}\right) + \frac{1}{2}\epsilon_{132}\left(A_{23}A_{32} - A_{33}A_{22}\right) \\
&= \left(A_{22}A_{33} - A_{32}A_{23}\right) \\
&= \det \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} = \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix},
\end{aligned}$$

esto es

$$(Adj)_{11} = (A_{22}A_{33} - A_{32}A_{23}) = \det\begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} = \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix}.$$

De la misma forma, se encuentran

$$\begin{aligned}
(Adj)_{12} &= A_{23}A_{31} - A_{21}A_{33} = -\det\begin{pmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{pmatrix} = -\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix}, \\
(Adj)_{13} &= A_{21}A_{32} - A_{22}A_{31} = \det\begin{pmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix} = \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix}, \\
(Adj)_{21} &= A_{13}A_{32} - A_{12}A_{33} = -\det\begin{pmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{pmatrix} = -\begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix}, \\
(Adj)_{22} &= A_{11}A_{33} - A_{13}A_{31} = \det\begin{pmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{pmatrix} = \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix}, \\
(Adj)_{23} &= A_{31}A_{12} - A_{32}A_{11} = -\det\begin{pmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{pmatrix} = -\begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix}, \\
(Adj)_{31} &= A_{12}A_{23} - A_{13}A_{22} = \det\begin{pmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{pmatrix} = \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix}, \\
(Adj)_{32} &= A_{13}A_{21} - A_{11}A_{23} = -\det\begin{pmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{pmatrix} = -\begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix}, \\
(Adj)_{33} &= A_{11}A_{22} - A_{12}A_{21} = \det\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}.
\end{aligned}$$

Por lo tanto,

$$\boxed{Adj(A) = \begin{pmatrix} +\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} & +\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} \\ -\begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} & +\begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} \\ +\begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix} & +\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \end{pmatrix}.}$$

Así,

$$
\boxed{
\begin{aligned}
A^{-1} &= \frac{1}{\det A} Adj(A)^T \\
&= \frac{1}{\det A}\begin{pmatrix}
+\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} & +\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} \\
-\begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} & +\begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} \\
+\begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix} & +\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}
\end{pmatrix}^T \\
&= \frac{1}{\det A}\begin{pmatrix}
+\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} & +\begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} \\
-\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} & +\begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix} \\
+\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} & -\begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} & +\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}
\end{pmatrix}.
\end{aligned}
}
$$

De forma análoga, se puede obtener la matriz inversa de matrices de otras dimensiones.

5.4. Operaciones con Python

Si queremos obtener la inversa de la matriz

$$
A = \begin{pmatrix} 1 & 0 & 6 & 9 \\ -1 & 3 & 6 & 4 \\ 3 & -1 & 6 & 3 \\ 1 & 5 & 6 & -1 \end{pmatrix}
$$

podemos emplear el siguiente código en Python

```python
import numpy as np
A = np.array([[1,0,6,9], [-1,3,6,4],[3,-1,6,3],[1,5,6,-1]
            ])
np.linalg.inv(A)
```

Listing 5.1: Inversa de una matriz

el cual da el resultado

```
array([[0.31428571,-0.57142857,-0.07142857,0.32857143],
[0.25714286,-0.28571429,-0.28571429,0.31428571],
[-0.22857143,0.30952381,0.22619048,-0.14047619],
[0.22857143,-0.14285714,-0.14285714,0.05714286]])
```

5.5. Ejercicios

1.- Considere la matriz

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

a) Obtener la matriz adjunta de A.

b) Obtener la matriz inversa de A.

Capítulo 6

Sistema de ecuaciones lineales de $n \times n$, método de Cramer

En este capítulo estudiaremos el llamado método de Cramer para obtener la solución de un sistema de n ecuaciones lineales con n incógnitas.

6.1. Caso general

Supongamos que tenemos n incógnitas

$$x_1, x_2, \cdots, x_n$$

que satisfacen n ecuaciones lineales

$$\boxed{\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= k_1, & (6.1)\\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= k_2, & (6.2)\\ &\vdots \\ A_{n1}x_1 + A_{n2}x_2 + \cdots + A_{nn}x_n &= k_n, & (6.3) \end{aligned}}$$

donde $A_{ij}, k_i (i, j = 1, \cdots, n)$ son constantes dadas. Observe que definiendo la matriz y los vectores

$$\boxed{A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad K = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix},}$$

el sistema de ecuaciones (6.1)-(6.3) se puede escribir como

$$\boxed{AX = K.}$$

Note que cuando $\det A \neq 0$, la matriz A es invertible y se cumple

$$\boxed{X = A^{-1}K.} \tag{6.4}$$

Si la matriz A no es invertible se debe a que no todos sus renglones son linealmente independientes, esto nos indica que no todas las ecuaciones son linealmente independientes. Cuando esto ocurre habrá más incógnitas que ecuaciones y tendremos un número infinito de soluciones. Pero si la matriz A es invertible, la solución es única y está dada por la ecuación (6.4).

Como podemos observar, que utilizando las herramientas que desarrollamos en el capítulo anterior es posible resolver un sistema de n ecuaciones con n incógnitas.

6.2. Caso 2×2

Supongamos que tenemos el siguiente sistema de ecuaciones

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 &= k_1, \\ A_{21}x_1 + A_{22}x_2 &= k_2, \end{aligned}$$

Por lo que, con la matriz y los vectores

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad K = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix},$$

este sistema de ecuaciones se puede escribir como

$$AX = K.$$

Si $\det A \neq 0$, la matriz A es invertible y se cumple

$$X = A^{-1}K.$$

En este caso la matriz inversa es

$$A^{-1} = \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix},$$

Así tenemos la solución

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix},$$

de donde se obtienen los valores

$$\begin{aligned} x_1 &= \frac{A_{22}k_1 - A_{12}k_2}{A_{11}A_{22} - A_{12}A_{21}}, \\ x_2 &= \frac{-A_{21}k_1 + A_{11}k_2}{A_{11}A_{22} - A_{12}A_{21}}. \end{aligned}$$

6.3. Caso 3×3

Supongamos que tenemos el siguiente sistema de ecuaciones

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + A_{13}x_3 &= k_1, \\ A_{21}x_1 + A_{22}x_2 + A_{23}x_3 &= k_2, \\ A_{31}x_1 + A_{32}x_2 + A_{33}x_3 &= k_3. \end{aligned}$$

Por lo que, usando la matriz y los vectores

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad K = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix},$$

este sistema de ecuaciones se puede escribir como

$$AX = K.$$

Si $\det A \neq 0$, la matriz A es invertible y se cumple

$$X = A^{-1}K.$$

esta ecuación tiene la forma concreta

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} + \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} & - \begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} & + \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} \\ - \begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} & + \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} & - \begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix} \\ + \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} & - \begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} & + \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix},$$

es decir

$$
\begin{aligned}
x_1 &= \frac{1}{\det A}\left(\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} k_1 - \begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} k_2 + \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} k_3\right), \\
x_2 &= \frac{1}{\det A}\left(-\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} k_1 + \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} k_2 - \begin{vmatrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{vmatrix} k_3\right), \\
x_3 &= \frac{1}{\det A}\left(\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} k_1 - \begin{vmatrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{vmatrix} k_2 + \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} k_3\right).
\end{aligned}
$$

A este método de solución de sistemas ecuaciones lineales se le llama método de Cramer.

Gabriel Cramer (1704-1752) fue un matemático suizo que trabajó en mecánica celeste, curvas algebraicas y sistemas de ecuaciones. En particular, demostró que los sistemas de ecuaciones lineales se pueden resolver mediante determinantes. Fue un matemático precoz que se doctoró a los 18 años y a los 20 años fue profesor de la Universidad de Ginebra. Murió debido a las complicaciones de una caída de un carruaje.

6.4. Ejemplo concreto

Ahora consideremos el sistema de ecuaciones

$$
\begin{aligned}
x + 2z &= 1, && (6.5)\\
x + 3y + 2z &= 4, && (6.6)\\
3y + 2z &= 3. && (6.7)
\end{aligned}
$$

En este caso tenemos la matriz asociada y los vectores

$$
A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 3 & 2 \\ 0 & 3 & 2 \end{pmatrix}, \qquad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad K = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}.
$$

Podemos ver que

$$
\det A = 6,
$$

por lo que la matriz A tiene inversa y el sistema de ecuaciones tiene solución. Además,

$$\begin{aligned} x &= \frac{1}{6}\left(\begin{vmatrix} 3 & 2 \\ 3 & 2 \end{vmatrix} - \begin{vmatrix} 0 & 2 \\ 3 & 3 \end{vmatrix} 4 + \begin{vmatrix} 0 & 2 \\ 3 & 2 \end{vmatrix} 3\right), \\ y &= \frac{1}{6}\left(-\begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} 4 - \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} 3\right), \\ z &= \frac{1}{6}\left(\begin{vmatrix} 1 & 3 \\ 0 & 3 \end{vmatrix} - \begin{vmatrix} 1 & 0 \\ 0 & 3 \end{vmatrix} 4 + \begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} 3\right), \end{aligned}$$

por lo cual la solución al sistema de ecuaciones lineales (6.5)-(6.7) es

$$x = 1, \qquad y = 1, \qquad z = 0.$$

El método de Cramer es sistemático y nos indica cuando el sistema de ecuaciones tiene solución. Sin embargo, en algunos casos puede ser muy laborioso. En el próximo capítulo veremos un método de solución diferente al de Cramer.

6.5. Ejercicios

1.- Usando el método de Cramer, resolver el sistema de ecuaciones lineales

$$\begin{aligned} x + 2y + 3z &= 9, \\ 4x + 5y + 6z &= -5, \\ 7x + 8y - z &= 3. \end{aligned}$$

2.- Usando el método de Cramer, resolver el sistema de ecuaciones lineales

$$\begin{aligned} 8x + y - 3z &= -1, \\ x + 3y + 2z &= 2, \\ 6x + y + 5z &= -7. \end{aligned}$$

Capítulo 7

Sistema de ecuaciones lineales de $n \times n$, método de Gauss-Jordan

En este capítulo veremos el método de Gauss-Jordan para resolver sistemas de ecuaciones lineales con n ecuaciones y n variables.

Mostraremos este método mediante ejemplos y después daremos una descripción del algoritmo general.

7.1. Ecuación de una variable

Una ecuación lineal de una variable tiene la forma

$$\boxed{ax + b = 0,} \tag{7.1}$$

donde se supone que a y b son constantes conocidas, con $a \neq 0$. En este problema debemos determinar el valor de x para que se cumpla la ecuación (7.1).

Sumando de ambos lados de la ecuación (7.1) el inverso aditivo de b, se obtiene

$$ax + b + (-b) = 0 + (-b),$$

es decir

$$ax = -b. \tag{7.2}$$

Además, multiplicando en ambos lados de la ecuación (7.2) por el inverso multiplicativo de a, se obtiene

$$a^{-1}ax = -a^{-1}b,$$

es decir

$$\boxed{x = -\frac{b}{a}.}$$

Esta es la solución de la ecuación (7.1).

7.2. Sistemas ecuaciones lineales de 2×2

Ahora veamos el caso de 2×2.

Un sistema de ecuaciones lineales de 2×2 tiene la forma

$$\boxed{\begin{aligned} ax + by =& c, \\ dx + ey =& f, \end{aligned}} \qquad \begin{aligned} (7.3) \\ (7.4) \end{aligned}$$

donde se supone que a, b, c, d, e, f son constantes conocidas.

En este caso, podemos ocupar que sabemos resolver una ecuación lineal de una incógnita. En efecto, multipliquemos por d a la ecuación (7.3) y por a a la ecuación (7.4), así obtenemos

$$\begin{aligned} dax + dby &= dc, \\ adx + aey &= af, \end{aligned}$$

al restar estas dos últimas ecuaciones, llegamos a

$$dax - adx + (db - ae)y = dc - af,$$

es decir,

$$(db - ae)y = dc - af.$$

Así, tenemos el sistema de ecuaciones

$$ax + by = c, \tag{7.5}$$
$$(db - ae)y = dc - af. \tag{7.6}$$

Observe que la ecuación (7.6) es una ecuación lineal de una incóngnita, de hecho tenemos

$$\begin{aligned} ax + by &= c, \\ y &= \frac{dc - af}{db - ae} = \frac{af - dc}{ae - db}, \end{aligned}$$

esto es

$$\begin{aligned} ax + by &= c, && (7.7) \\ y &= \frac{af - dc}{ae - db}. && (7.8) \end{aligned}$$

Ahora al multiplicar por b a la ecuación (7.8) y restar el resultado a la ecuación (7.7) obtenemos el sistema

$$\begin{aligned} ax + by - by &= c - b\frac{af - dc}{ae - db}, \\ y &= \frac{af - dc}{ae - db}, \end{aligned}$$

que se puede escribir de la forma

$$\begin{aligned} ax &= a\frac{ce - fb}{ae - db}, \\ y &= \frac{af - dc}{ae - db}. \end{aligned}$$

Por lo tanto, la solución al sistema de ecuaciones que estamos estudiando es

$$\boxed{\begin{aligned} x &= \frac{ce - bf}{ae - db}, && (7.9) \\ y &= \frac{af - dc}{ae - db}. && (7.10) \end{aligned}}$$

7.2.1. Ejemplo

Consideremos el sistemas de ecuaciones

$$\begin{aligned} 2x + 3y &= 6, \\ 4x - 5y &= 2. \end{aligned}$$

Entonces, multiplicando por 2 a la primera ecuación y restándola a la segunda ecuación, llegamos a

$$\begin{aligned} 2x + 3y &= 6, \\ -11y &= -10. \end{aligned}$$

Ahora, en este nuevo sistema, multiplicando por $-\frac{1}{11}$ a la segunda ecuación tenemos

$$\begin{aligned} 2x + 3y &= 6, \\ y &= \frac{10}{11}. \end{aligned}$$

Además, de este último sistema multiplicando por -3 a la segunda ecuación y sumándola a la primera ecuación tenemos

$$\begin{aligned} 2x &= \frac{36}{11}, \\ y &= \frac{10}{11}, \end{aligned}$$

que implica la solución

$$\begin{aligned} x &= \frac{18}{11}, \\ y &= \frac{10}{11}. \end{aligned}$$

7.3. Sistemas ecuaciones lineales de 3×3

Ahora veremos el caso de un sistema de ecuaciones de 3×3 Un sistema de ecuaciones lineales de 3×3 tiene la forma

$$\boxed{\begin{aligned} a_1x + a_2y + a_3z &= a_4, && (7.11) \\ b_1x + b_2y + b_3z &= b_4, && (7.12) \\ c_1x + c_2y + c_3z &= c_4. && (7.13) \end{aligned}}$$

Para resolver este problema ocuparemos el mismo método desarrollado para resolver el sistema de 2×2.

Primero multiplicamos a la ecuación (7.13) por b_1 y a la ecuación (7.12) por c_1, posteriormente restamos el resultado de estas dos operaciones y colocamos el resultado en el lugar de la tercera ecuación. Por lo que se obtienen el sistema

$$\begin{array}{rcl} a_1x + a_2y + a_3z &=& a_4, \\ b_1x + b_2y + b_3z &=& b_4, \\ (c_1b_1 - b_1c_1)x + (c_2b_1 - c_1b_2)y + (c_3b_1 - c_1b_3)z &=& (c_4b_1 - c_1b_4), \end{array}$$

esto es

$$\begin{array}{rcll} a_1x + a_2y + a_3z &=& a_4, & (7.14) \\ b_1x + b_2y + b_3z &=& b_4, & (7.15) \\ (c_2b_1 - c_1b_2)y + (c_3b_1 - c_1b_3)z &=& (c_4b_1 - c_1b_4). & (7.16) \end{array}$$

Ahora, multiplicamos por a_1 a la ecuación (7.15) y por b_1 a la ecuación (7.14), posteriomente restamos el resultado de estas dos operaciones y colocamos el resultado en el lugar de la segunda ecuación. Así se obtienen el sistema

$$\begin{array}{rcl} a_1x + a_2y + a_3z &=& a_4, \\ (b_1a_1 - a_1b_1)x + (b_2a_1 - b_1a_2)y + (b_3a_1 - b_1a_3)z &=& (b_4a_1 - b_1a_4), \\ (c_2b_1 - c_1b_2)y + (c_3b_1 - c_1b_3)z &=& (c_4b_1 - c_1b_4), \end{array}$$

de donde

$$\begin{array}{rcll} a_1x + a_2y + a_3z &=& a_4, & (7.17) \\ (b_2a_1 - b_1a_2)y + (b_3a_1 - b_1a_3)z &=& (b_4a_1 - b_1a_4), & (7.18) \\ (c_2b_1 - c_1b_2)y + (c_3b_1 - c_1b_3)z &=& (c_4b_1 - c_1b_4). & (7.19) \end{array}$$

Note, que definiendo las constantes

$$\begin{array}{rcl} \tilde{b}_2 &=& (b_2a_1 - b_1a_2), \quad \tilde{b}_3 = (b_3a_1 - b_1a_3), \quad \tilde{b}_4 = (b_4a_1 - b_1a_4), \\ \tilde{c}_2 &=& (c_2b_1 - c_1b_2), \quad \tilde{c}_3 = (c_3b_1 - c_1b_3), \quad \tilde{c}_4 = (c_4b_1 - c_1b_4), \end{array}$$

el sistema (7.17)-(7.19) se puede escribir como

$$\begin{array}{rcll} a_1x + a_2y + a_3z &=& a_4, & (7.20) \\ \tilde{b}_2y + \tilde{b}_3z &=& \tilde{b}_4, & (7.21) \\ \tilde{c}_2y + \tilde{c}_3z &=& \tilde{c}_4. & (7.22) \end{array}$$

Observe que en las ecuaciones (7.21)-(7.22) no aparece la incógnita x, por lo que estas dos ecuaciones forman un sistema de ecuaciones lineales de 2×2.

Así, usando el método desarrollado anteriormente, llegamos a

$$a_1x + a_2y + a_3z = a_4, \tag{7.23}$$

$$y = B_4 = \frac{\tilde{b}_4\tilde{c}_3 - \tilde{b}_3\tilde{c}_4}{\tilde{b}_2\tilde{c}_3 - \tilde{c}_2\tilde{b}_3}, \tag{7.24}$$

$$z = C_4 = \frac{\tilde{b}_2\tilde{c}_4 - \tilde{c}_2\tilde{b}_4}{\tilde{b}_2\tilde{c}_3 - \tilde{c}_2\tilde{b}_3}. \tag{7.25}$$

Adicionalmente, multiplicando la ecuación (7.24) por $-a_2$ y sumando el resultado a la ecuación (7.23), se obtiene el sistema

$$a_1x + a_3z = a_4 - B_4a_2, \tag{7.26}$$

$$y = B_4, \tag{7.27}$$

$$z = C_4. \tag{7.28}$$

Además, multiplicando la ecuación (7.28) por $-a_3$ y sumando el resultado a la ecuación (7.26), se obtiene el sistema

$$\begin{aligned} a_1x &= a_4 - B_4a_2 - C_4a_3, \\ y &= B_4, \\ z &= C_4. \end{aligned}$$

Por lo que finalmente se encuentra la solución

$$\boxed{\begin{aligned} x &= A_4, \qquad A_4 = \frac{a_4 - B_4a_2 - C_4a_3}{a_1}, \\ y &= B_4, \\ z &= C_4. \end{aligned}}$$

7.4. Ejemplos

Consideremos el siguiente sistema de ecuaciones

$$3x - 5y + 6z = 9, \tag{7.29}$$

$$5x - y + 2z = 2, \tag{7.30}$$

$$x - 2y + z = -1. \tag{7.31}$$

Para resolver este sistema ecuaciones primero pasamos la tercera ecuación al inicio y obtenemos

$$\begin{aligned} x - 2y + z &= -1, \\ 3x - 5y + 6z &= 9, \\ 5x - y + 2z &= 2. \end{aligned}$$

Ahora, de este sistema multiplicamos por -3 a la primera ecuación y la sumamos a la segunda ecuación, el resultado lo ponemos en la segunda ecuación. Además multiplicamos por -5 a la primera ecuación y la sumamos a la tercera ecuación, el resultado lo ponemos en la tercera ecuación. Así, obtenemos

$$\begin{aligned} x - 2y + z &= -1, \\ y + 3z &= 12, \\ 9y - 3z &= 7. \end{aligned}$$

Para continuar, de este nuevos sistema multiplicamos por -9 a la segunda ecuación y la sumamos a la primera ecuación, el resultado de este procedimiento lo ponemos en la tercera ecuación. Entonces llegamos al sistema

$$\begin{aligned} x - 2y + z &= -1, \\ y + 3z &= 12, \\ -30z &= -101 \end{aligned}$$

que implica

$$\begin{aligned} x - 2y + z &= -1, \\ y + 3z &= 12, \\ z &= \frac{101}{30}. \end{aligned}$$

Adicionalmente, de este nuevo sistema multiplicamos por -3 a la tercera ecuación y la sumamos a la segunda ecuación, el resultado de este procedimiento lo ponemos en la segunda ecuación y llegamos a

$$\begin{aligned} x - 2y + z &= -1, \\ y &= \frac{19}{10}, \\ z &= \frac{101}{30}. \end{aligned}$$

Además, en el último sistema que hemos obtenido, restando la tercera ecuación a la primera y colocando el resultado de este procedimiento en la primera ecuación obtenemos

$$\begin{aligned} x - 2y &= -\frac{131}{30}, \\ y &= \frac{19}{10}, \\ z &= \frac{101}{30}. \end{aligned}$$

Finamente, multiplicando la segunda ecuación por -2 y sumándola a la primera ecuación llegamos a

$$\begin{aligned} x &= -\frac{17}{30}, \\ y &= \frac{19}{10}, \\ z &= \frac{101}{30}, \end{aligned}$$

que es la solución del sistema de ecuaciones (7.29)-(7.31).

7.5. Método de Gauss-Jordan

El método que hemos usado para resolver sistemas de ecuaciones lineales se basa en los siguientes pasos:

1.- Podemos cambiar de lugar las ecuaciones o variables. Note que al hacer esto no cambia en nada al sistema.

2.- Multiplicar por un escalar a una ecuación y colocar el resultado en el lugar de dicha ecuación.

3.- Realizar combinaciones lineales de ecuaciones y colocar el resultado en el lugar de una de las ecuacione involucradas en la combinación lineal.

Todas estas operaciones se realizan con el objetivo de dejar una sola variable en cada renglón.

Wilhelm Jordan (1842-1899) fue un ingeniero alemán. Realizó sus estudios en el Instituto Politécnico de Stuttgart. Sus libros de Geodesia son considerados clásicos en su área. Especialmente su libro *Handbuch der Vermessungskunde* ha sido usado como texto en diferentes partes del mundo, aún después de más de un siglo de haber sido escrito. En dicha obra, Jordan introduce lo que hoy se conoce como método de Gauss-Jordan para resolver sistemas de ecuaciones lineales. Además de la ingeniería, tenía pasión por los viajes y un espíritu aventurero que lo llevó a realizar expediciones a África. Terminó su vida como profesor en la Universidad Técnica de Hannover.

Carl Friedrich Gauss (1777-1855) fue un matemático alemán que contribuyó enormemente en diversas áreas como la teoría de números, geometría diferencial, probabilidad, análisis matemático, electrostática, etc. Sus padres fueron campesinos alemanes que no sabían leer ni escribir, pero gracias a su trabajo y talento hoy se le conoce como el Príncipe de las Matemáticas. No sembró en campos de cultivos como sus padres, pero sembró resultados importantes en las matemáticas. Por ejemplo, la llamada distribución gaussiana se usa para modelar fenómenos naturales, sociales, económicos, financieros, etc.; la llamada ley de Gauss se usa para describir fenómenos eléctricos y gravitacionales; el teorema de Gauss del cálculo de varias variables es una generalización del teorema fundamental del cálculo; la curvatura gaussiana fue el primer paso para lo que hoy se conoce como geometría diferencial, que permitió la construcción de la Relatividad General de Einstein; su ley de mínimos cuadrados también tiene diversas aplicaciones en diferentes disciplinas. Cuando apenas era un niño, un profesor rural le ayudó a mejorar su alemán y descubrió su enorme talento para las matemáticas. Dicho profesor lo recomendó para que pudiera continuar sus estudios, posteriormente conoció al matemático Martin Barteles con quien estudió los trabajos clásicos de Newton, Euler y Lagrange. Realizó sus estudios doctorales en la universidad de Götingen en donde fue profesor. Gauss murió de un ataque al corazón.

Cabe hacer notar que el método de Gauss-Jordan se le atribuye a Carl Friedrich Gauss y Wilhelm Jordan. Sin embargo, este método se puede ver en los libros del matemático chino Jiuzhang Suanshu, los cuales se escribieron 150 años antes de Cristo.

7.6. Ejercicios con Python

Si tenemos el sistema de ecuaciones

$$\begin{aligned}
x - 2y + 6z + 9w &= 5, \\
-x + 3y + 6z + 4w &= 7, \\
3x - y + 6z + 3w &= 3, \\
x + 5y + 6z - w &= 1,
\end{aligned}$$

donde la matriz

$$\begin{pmatrix} 1 & -2 & 6 & 9 \\ -1 & 3 & 6 & 4 \\ 3 & -1 & 6 & 3 \\ x & 5 & 6 & -1 \end{pmatrix}$$

es invertible, entonces para resolver el sistema de ecuaciones se puede ocupar el siguiente código de Python

```
import numpy as np
a= np.array([[1,-2,6,9], [-1,3,6,4],[3,-1,6,3],[1,5,6,-1]
              ])
b=np.array([5,7,3,1])
c=np.linalg.inv(a)
np.dot(c,b)
```

Listing 7.1: Sistema de ecuaciones lineales

que arroja el resultado

```
array([-3.94117647,-2.58823529,2.74509804,-1.41176471])
```

7.7. Ejercicios

1.- Usando el método de Gauss-Jordan, resolver el sistema de ecuaciones lineales

$$\begin{aligned} -x + y + 4z &= 8, \\ x + 5y + 2z &= -5, \\ 3x + y - 5z &= -7. \end{aligned}$$

Comparar sus resultado con el obtenido con Python.

2.- Usando el método de Gauss-Jordan, resolver el sistema de ecuaciones lineales

$$\begin{aligned} 4x - 4y + z &= 5, \\ 2x + y + z &= -2, \\ 3x + 2y + 9z &= 9. \end{aligned}$$

Comparar sus resultado con el obtenido con Python.

Capítulo 8

Método de Gauss-Jordan usando matrices

En este capítulo veremos el método de Gauss-Jornan usando matrices.

8.1. Caso 2×2

En el método de Gauss-Jornan para resolver sistemas de ecuaciones lineales las incógnitas no juegan un papel relevante, por lo que es posible olvidarlas mientras se resuelve el sistema. Para mostrar este hecho, recordemos un sistema de ecuaciones de 2×2 tiene la forma

$$\boxed{\begin{aligned} ax + by &= c, \\ dx + ey &= f, \end{aligned}}$$

y su solución es

$$\begin{aligned} x &= \frac{ce - bf}{ae - db}, \\ y &= \frac{af - dc}{ae - db}. \end{aligned}$$

Ahora, este sistema de ecuaciones se puede escribir de la forma

$$\boxed{\begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ f \end{pmatrix}.} \tag{8.1}$$

Además, usando esta notación su solución se puede escribir como

$$\boxed{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{ce-bf}{ae-db} \\ \frac{af-dc}{ae-db} \end{pmatrix}.} \tag{8.2}$$

Desde el punto matricial, para obtener la solución un sistema de ecuaciones de 2×2 debemos pasar de la matriz

$$\boxed{\begin{pmatrix} a & b \\ d & e \end{pmatrix}}$$

a la matriz identidad

$$\boxed{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.}$$

Ahora, note que para tener todas las constantes importantes del sistema de ecuaciones en una matriz podemos proponer la llamada **Matriz extendida**

$$\boxed{\left(\begin{array}{cc||c} a & b & c \\ d & e & f \end{array}\right),} \tag{8.3}$$

que es una representación del sistema de ecuaciones (8.1). Con esta representación la solución del sistema de ecuaciones (8.2) se puede escribir como

$$\boxed{\left(\begin{array}{cc||c} 1 & 0 & \frac{ce-bf}{ae-db} \\ 0 & 1 & \frac{af-dc}{ae-db} \end{array}\right),} \tag{8.4}$$

Por lo tanto, resolver el sistema de ecuaciones (8.1) equivale a pasar de la matriz extendida (8.3) a la matriz extendida (8.4) usando los pasos permitidos del método del Gauss-Jordan.

En este caso el Método de Gauss-Jordan consiste en:

1.- Intercambiar renglones o columnas.

2.- Multiplicar por un escalar a un renglón y colocar el resultado en el lugar de dicho renglón.

3.- Realizar combinaciones lineales de renglones y colocar el resultado en el lugar de uno de los renglones involucrados en la combinación lineal.

Todas estas operaciones se realizan con el objetivo de dejar una matriz diagonal de lado izquierdo de la matriz extendida.

8.2. Ejemplos

8.2.1. Sistema 2×2

Consideremos el siguiente sistema

$$\begin{aligned} 2x - y &= 6, \qquad (8.5)\\ 3x + 5y &= 7. \qquad (8.6) \end{aligned}$$

En este caso, la matriz extendida es

$$\left(\begin{array}{cc||c} 2 & -1 & 6 \\ 3 & 5 & 7 \end{array} \right).$$

Observe que en este caso tenemos los renglones

$$\begin{aligned} R_1 &= (2 \quad -1 \quad || \, 6) \, , \\ R_2 &= (3 \quad 5 \quad || \, 7) \, , \end{aligned}$$

Ahora, en el segundo regón de esta matriz colocaremos la operación

$$-3R_1 + 2R_2,$$

por lo que obtenemos

$$\left(\begin{array}{cc||c} 2 & -1 & 6 \\ 0 & 13 & -4 \end{array} \right).$$

Además, de esta última matriz extendida, al multiplicar por el segundo rengón por $\frac{1}{13}$ se encuentra la matriz

$$\left(\begin{array}{cc||c} 2 & -1 & 6 \\ 0 & 1 & -\frac{4}{13} \end{array}\right).$$

Adicionalmente, colocando en el primer renglón la suma de los dos renglones de esta última matriz extendida se llega a

$$\left(\begin{array}{cc||c} 2 & 0 & \frac{74}{13} \\ 0 & 1 & -\frac{1}{4} \end{array}\right).$$

Finalmente, multiplicando por $\frac{1}{2}$ al primer renglón de la nueva matriz extendida se encuentra

$$\left(\begin{array}{cc||c} 1 & 0 & \frac{37}{13} \\ 0 & 1 & -\frac{4}{13} \end{array}\right).$$

Así, la solución del sistema de ecuaciones (8.5)-(8.6) es

$$\begin{aligned} x &= \frac{37}{13}, \\ y &= -\frac{4}{13}. \end{aligned}$$

8.2.2. Sistema 3×3

Consideremos el sistema de ecuaciones

$$\begin{aligned} 2y + 4z &= 2, &(8.7)\\ 2x + 4y + 2z &= 3, &(8.8)\\ 3x + 3y + z &= 1. &(8.9) \end{aligned}$$

En este caso, tenemos la matriz extendida

$$\left(\begin{array}{ccc||c} 0 & 2 & 4 & 2 \\ 2 & 4 & 2 & 3 \\ 3 & 3 & 1 & 1 \end{array}\right),$$

Al intercambiar el primer renglón por el segundo renglón encontramos la matriz extendida

$$\left(\begin{array}{ccc||c} 2 & 4 & 2 & 3 \\ 0 & 2 & 4 & 2 \\ 3 & 3 & 1 & 1 \end{array}\right),$$

además, multiplicando por $\frac{1}{2}$ al primer renglón de esta matriz, se llega a

$$\left(\begin{array}{ccc||c} 1 & 2 & 1 & \frac{3}{2} \\ 0 & 1 & 2 & 1 \\ 3 & 3 & 1 & 1 \end{array}\right).$$

Por lo que multiplicando por -3 al primer renglón y sumando el resultado al tercer renglón de esta nueva matriz extendida se encuentra

$$\left(\begin{array}{ccc||c} 1 & 2 & 1 & \frac{3}{2} \\ 0 & 1 & 2 & 1 \\ 0 & -3 & -2 & -\frac{7}{2} \end{array}\right).$$

Ahora, multiplicando por 3 al segundo renglón y sumando el resultado al tercer renglón de esta última matriz extendida se obtiene

$$\left(\begin{array}{ccc||c} 1 & 2 & 1 & \frac{3}{2} \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 4 & -\frac{1}{2} \end{array}\right),$$

así multiplicando por $\frac{1}{4}$ al tercer renglón de esta nueva matriz extendida tenemos

$$\left(\begin{array}{ccc||c} 1 & 2 & 1 & \frac{3}{2} \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -\frac{1}{8} \end{array}\right).$$

Además, si en la última matriz extendida multiplicamos por -2 al tercer renglón y sumamos el resultado al segundo renglón se llega a

$$\left(\begin{array}{ccc||c} 1 & 2 & 1 & \frac{3}{2} \\ 0 & 1 & 0 & \frac{5}{4} \\ 0 & 0 & 1 & -\frac{1}{8} \end{array}\right).$$

Para continuar con el procedimiento, al primer renglón de la nueva matriz extendida le restamos su tercer renglón y se obtiene

$$\left(\begin{array}{ccc||c} 1 & 2 & 0 & \frac{13}{8} \\ 0 & 1 & 0 & \frac{5}{4} \\ 0 & 0 & 1 & -\frac{1}{8} \end{array}\right).$$

Finalmente, si en la última matriz extendida multiplicamos por -2 al segundo renglón y sumamos el resultado al primer renglón llegamos a

$$\left(\begin{array}{ccc||c} 1 & 0 & 0 & -\frac{7}{8} \\ 0 & 1 & 0 & \frac{5}{4} \\ 0 & 0 & 1 & -\frac{1}{8} \end{array}\right).$$

Así, la solución al sistema de ecuaciones (8.7)-(8.9) es

$$\begin{aligned} x &= -\frac{7}{8}, \\ y &= \frac{5}{4}, \\ z &= -\frac{1}{8}. \end{aligned}$$

8.3. Notación de matriz similar

Cuando una matriz se obtiene de otra mediante un paso del método de Gauss-Jordan se dice que son similares mediante Gauss-Jordan y se escribe

$$\sim .$$

Por ejemplo, supongamos que tenemos el sistema de ecuaciones

$$\begin{aligned} 2x + y + z &= 2, &(8.10)\\ x + 4y + 2z &= 1, &(8.11)\\ 5x + y + 3z &= 0. &(8.12) \end{aligned}$$

En este caso tenemos la matriz extendida

$$\left(\begin{array}{ccc||c} 2 & 1 & 1 & 2 \\ 1 & 4 & 2 & 1 \\ 5 & 1 & 3 & 0 \end{array}\right),$$

por lo que usando el método de Gauss-Jordan se tiene

$$\begin{aligned} &\left(\begin{array}{ccc||c} 2 & 1 & 1 & 2 \\ 1 & 4 & 2 & 1 \\ 5 & 1 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc||c} 1 & 4 & 2 & 1 \\ 2 & 1 & 1 & 2 \\ 5 & 1 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc||c} 1 & 4 & 2 & 1 \\ 0 & -7 & -3 & 0 \\ 0 & -19 & -7 & -5 \end{array}\right) \\ \sim &\left(\begin{array}{ccc||c} 1 & 4 & 2 & 1 \\ 0 & 1 & \frac{3}{7} & 0 \\ 0 & 1 & \frac{7}{19} & \frac{5}{19} \end{array}\right) \sim \left(\begin{array}{ccc||c} 1 & 4 & 2 & 1 \\ 0 & 1 & \frac{3}{7} & 0 \\ 0 & 0 & -\frac{8}{133} & \frac{5}{19} \end{array}\right) \sim \left(\begin{array}{ccc||c} 1 & 4 & 2 & 1 \\ 0 & 1 & \frac{3}{7} & 0 \\ 0 & 0 & 1 & -\frac{35}{8} \end{array}\right) \\ \sim &\left(\begin{array}{ccc||c} 1 & 4 & 2 & 1 \\ 0 & 1 & 0 & \frac{15}{8} \\ 0 & 0 & 1 & -\frac{35}{8} \end{array}\right) \sim \left(\begin{array}{ccc||c} 1 & 4 & 0 & \frac{78}{8} \\ 0 & 1 & 0 & \frac{15}{8} \\ 0 & 0 & 1 & -\frac{35}{8} \end{array}\right) \sim \left(\begin{array}{ccc||c} 1 & 0 & 0 & \frac{18}{8} \\ 0 & 1 & 0 & \frac{15}{8} \\ 0 & 0 & 1 & -\frac{35}{8} \end{array}\right). \end{aligned}$$

Por lo que la solución al sistema de ecuaciones (8.10)-(8.12) es

$$\begin{aligned} x &= \frac{18}{8} \\ y &= \frac{15}{8}, \\ z &= -\frac{35}{8}. \end{aligned}$$

8.4. Ejercicios

1.- Usando el método desarrollado en este capítulo, resolver el sistema de ecuaciones lineales

$$\begin{aligned} x + 2y + 3z &= -2, \\ 2x + y + 2 &= -1, \\ -2x + y - z &= 8. \end{aligned}$$

2.- Usando el método desarrollado en este capítulo, resolver el sistema de ecuaciones lineales

$$\begin{aligned} 3x + 2y - 2z + 10w &= 1, \\ 3x + y + z + 2w &= -3, \\ -2x + 2y + 3z + 4w &= -2, \\ x + y + 5z + 2w &= -3. \end{aligned}$$

Capítulo 9

Método de Gauss-Jordan y la inversa de una matriz

En este capítulo veremos que usando el método de Gauss-Jordan es posible obtener la matriz inversa de una matriz.

9.1. Caso 2×2

Sabemos que cualquier sistema de ecuaciones de 2×2 se puede escribir de la forma

$$\boxed{\begin{aligned} ax + by =& c, \\ dx + ey =& f, \end{aligned}}$$

se puede escribir de la forma

$$\boxed{MX = X_0,} \tag{9.1}$$

con

$$\boxed{M = \begin{pmatrix} a & b \\ d & e \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad X_0 = \begin{pmatrix} c \\ f \end{pmatrix}.}$$

Por lo tanto, usando la matriz inversa de la matriz M, de la ecuación (9.1) se obtiene

$$\boxed{X = M^{-1}X_0.}$$

Así, si existe la matriz inversa de M, es posible resolver un sistema de 2 ecuaciones lineales con 2 incógnitas. Claramente este resultado se puede generalizar a un sistema de n ecuaciones con n variables indeterminadas. Por lo que podemos afirmar que al resolver un sistema de n ecuaciones lineales con n incógnitas con el método de Gauss-Jordan, de forma indirecta estamos obteniendo la inversa de una matriz.

De hecho, con el método de Gauss-Jordan es posible obtener la inversa de una matriz. Por simplicidad, solo probaremos el caso de 2×2.

Supongamos que tenemos la matriz de 2×2

$$\boxed{M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},} \tag{9.2}$$

a esta matriz le asociaremos la matriz extendida

$$\boxed{\left(\begin{array}{cc||cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right)}$$

de donde

$$\boxed{\begin{array}{l} \left(\begin{array}{cc||cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc||cc} a & b & 1 & 0 \\ 0 & da-bc & -c & a \end{array}\right) \sim \\ \sim \left(\begin{array}{cc||cc} a & b & 1 & 0 \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array}\right) \sim \left(\begin{array}{cc||cc} a & 0 & 1+\frac{cb}{da-bc} & -\frac{ab}{da-bc} \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array}\right) \sim \\ \sim \left(\begin{array}{cc||cc} a & 0 & \frac{da}{da-bc} & -\frac{ab}{da-bc} \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array}\right) \sim \left(\begin{array}{cc||cc} 1 & 0 & \frac{d}{da-bc} & -\frac{b}{da-bc} \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array}\right), \end{array}}$$

esto es

$$\boxed{\left(\begin{array}{cc||cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc||cc} 1 & 0 & \frac{d}{da-bc} & -\frac{b}{da-bc} \\ 0 & 1 & -\frac{c}{da-bc} & \frac{a}{da-bc} \end{array}\right).}$$

Note que del lado derecho de esta última matriz extendida se obtiene la matriz

$$\boxed{\frac{1}{da-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1},}$$

que es la matriz inversa de A definida en (9.2).

Por lo tanto, usando el método de Gauss-Jordan es posible obtener la inversa de una matriz de 2×2. Lo mismo ocurre con matrices de $n \times n$. No probaremos esta última afirmación, pero veremos algunos ejemplos.

9.2. Ejemplo de 3×3

Consideremos la matriz

$$M = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 1 & 6 \\ 3 & 4 & 0 \end{pmatrix}. \tag{9.3}$$

En este caso tenemos la matriz aumentada

$$\left(\begin{array}{ccc||ccc} 1 & 0 & 5 & 1 & 0 & 0 \\ 2 & 1 & 6 & 0 & 1 & 0 \\ 3 & 4 & 0 & 0 & 0 & 1 \end{array}\right),$$

además usando el método de Gauss-Jordan tenemos

$$\left(\begin{array}{ccc||ccc} 1 & 0 & 5 & 1 & 0 & 0 \\ 2 & 1 & 6 & 0 & 1 & 0 \\ 3 & 4 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc||ccc} 1 & 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & -4 & -2 & 1 & 0 \\ 0 & 4 & -15 & -3 & 0 & 1 \end{array}\right)$$
$$\sim \left(\begin{array}{ccc||ccc} 1 & 0 & 5 & 1 & 0 & \\ 0 & 1 & -4 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -4 & 1 \end{array}\right) \sim \left(\begin{array}{ccc||ccc} 1 & 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & 0 & 18 & -15 & 4 \\ 0 & 0 & 1 & 5 & -4 & 1 \end{array}\right)$$
$$\sim \left(\begin{array}{ccc||ccc} 1 & 0 & 0 & -24 & 20 & -5 \\ 0 & 1 & 0 & 18 & -15 & 4 \\ 0 & 0 & 1 & 5 & -4 & 1 \end{array}\right).$$

Observe que la matriz

$$\begin{pmatrix} -24 & 20 & -5 \\ 18 & -15 & 4 \\ 5 & -4 & 1 \end{pmatrix}$$

es la inversa de la matriz M definida en (9.3).

9.3. Ejemplo de 4×4

Consideremos la matriz

$$A = \begin{pmatrix} 3 & 0 & 2 & -1 \\ 1 & 2 & 0 & -2 \\ 4 & 0 & 6 & -3 \\ 5 & 0 & 2 & 0 \end{pmatrix}, \tag{9.4}$$

para la cual tenemos la matriz aumentada

$$\left(\begin{array}{cccc||cccc} 3 & 0 & 2 & -1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & -2 & 0 & 1 & 0 & 0 \\ 4 & 0 & 6 & -3 & 0 & 0 & 1 & 0 \\ 5 & 0 & 2 & 0 & 0 & 0 & 0 & 1 \end{array} \right).$$

Así, usando el método de Gauss-Jordan tenemos

$$
\left(\begin{array}{cccc||cccc}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
1 & 2 & 0 & -2 & 0 & 1 & 0 & 0\\
4 & 0 & 6 & -3 & 0 & 0 & 1 & 0\\
5 & 0 & 2 & 0 & 0 & 0 & 0 & 1
\end{array}\right)
$$

$$
\sim \left(\begin{array}{cccc||cccc}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 2 & 5 & 1 & -3 & 0 & 0\\
0 & 0 & -10 & 5 & 4 & 0 & -3 & 0\\
0 & 0 & 4 & -5 & 5 & 0 & 0 & -3
\end{array}\right)
$$

$$
\sim \left(\begin{array}{cccc||cccc}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 2 & 5 & 1 & -3 & 0 & 0\\
0 & 0 & -10 & 5 & 4 & 0 & -3 & 0\\
0 & 0 & 0 & -30 & 66 & 0 & -12 & -30
\end{array}\right)
$$

$$
\sim \left(\begin{array}{cccc||cccc}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 2 & 5 & 1 & -3 & 0 & 0\\
0 & 0 & -10 & 5 & 4 & 0 & -3 & 0\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim \left(\begin{array}{cccc||cccc}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 2 & 5 & 1 & -3 & 0 & 0\\
0 & 0 & -10 & 0 & 15 & 0 & -5 & -5\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim \left(\begin{array}{cccc||cccc}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 2 & 5 & 1 & -3 & 0 & 0\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim \left(\begin{array}{cccc||cccc}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 2 & 0 & 12 & -3 & -2 & -5\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

además,

$$
\sim\left(\begin{array}{rrrr||rrrr}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 2 & 0 & 12 & -3 & -2 & -5\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim\left(\begin{array}{rrrr||rrrr}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & -6 & 0 & 0 & 15 & -3 & -3 & -6\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim\left(\begin{array}{rrrr||rrrr}
3 & 0 & 2 & -1 & 1 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} & 1\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim\left(\begin{array}{rrrr||rrrr}
3 & 0 & 2 & 0 & -\frac{6}{5} & 0 & \frac{2}{5} & 1\\
0 & 1 & 0 & 0 & -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} & 1\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim\left(\begin{array}{rrrr||rrrr}
3 & 0 & 0 & 0 & \frac{9}{5} & 0 & -\frac{3}{5} & 0\\
0 & 1 & 0 & 0 & -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} & 1\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

$$
\sim\left(\begin{array}{rrrr||rrrr}
1 & 0 & 0 & 0 & \frac{3}{5} & 0 & -\frac{1}{5} & 0\\
0 & 1 & 0 & 0 & -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} & 1\\
0 & 0 & 1 & 0 & -\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
0 & 0 & 0 & 1 & -\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right).
$$

Por que lo tenemos la matriz

$$
\left(\begin{array}{rrrr}
\frac{3}{5} & 0 & -\frac{1}{5} & 0\\
-\frac{5}{2} & \frac{1}{2} & \frac{1}{2} & 1\\
-\frac{3}{2} & 0 & \frac{1}{2} & \frac{1}{2}\\
-\frac{11}{5} & 0 & \frac{2}{5} & 1
\end{array}\right)
$$

que es la inversa de la matriz A definida en (9.4).

9.4. Ejercicios

1.- Usando el método de Gauss-Jordan, obtener la inversa de la matriz

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

2.- Usando el método de Gauss-Jordan, obtener la inversa de la matriz

$$\begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & -1 & 4 \\ -4 & 5 & -10 & -6 \\ 3 & -2 & 10 & -1 \end{pmatrix}.$$

Capítulo 10

Sistemas de ecuaciones lineales de $n \times m$

Ahora veremos como resolver un sistema de m ecuaciones con n variables y algunas propiedades de la soluciones.

10.1. Caso general

Supongamos que tenemos n incógnitas

$$x_1, x_2, \cdots, x_n$$

y m ecuaciones lineales

$$\boxed{\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n =& J_1, \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n =& J_2, \\ & \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n =& J_m, \end{aligned}}$$

donde $A_{ij}, J_i (i = 1, 2, \cdots m, j = 1, \cdots, n)$ son constantes dadas. A este sistema de ecuaciones se puede asociar a la matriz extendida

$$\boxed{\left(\begin{array}{cccc||c} A_{11} & A_{12} & \cdots & A_{1n} & J_1 \\ A_{21} & A_{22} & \cdots & A_{2n} & J_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} & J_m \end{array} \right)}$$

y para resolver el sistema de ecuaciones podemos ocupar el método de Gauss-Jordan.

Después de aplicar el método de Gauss-Jordan a la matriz extendida, tenemos tres posibles resultados. El primero es que el número de renglones linealmente independientes de la matriz extendida sea mayor al número de incógnitas. En este caso no es posible tener soluciones. En efecto, cada renglón de la matriz extendida representa una ecuación, de la cual se puede despejar una incógnita, pero si tenemos menos incógnitas que ecuaciones no es posible satisfacer todas la ecuaciones.

El segundo caso posible es que el número de renglones linealmente independientes de la matriz extendida sea igual al número de incógnitas. Este caso es equivalente a tener un sistema de ecuaciones con el mismo número de ecuaciones que el de incógnitas. Este caso ya fue estudiado anteriormente.

El tercer caso posible es en el cual el número de incógnitas sea mayor al número de renglones linealmente independientes de la matriz extendida. Por ejemplo, supongamos que después de aplicar el método de Gauss-Jordan obtenemos l ecuaciones linealmente independientes, con $l < n$. Por lo que tendremos un sistema de ecuaciones de la forma

$$\boxed{\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1l}x_l + a_{1(l+1)}x_{l+1} + \cdots + a_{1n}x_n = \tilde{J}_1, \\ &a_{21}x_1 + a_{22}x_2 + \cdots + a_{2l}x_l + a_{2(l+1)}x_{l+1} + \cdots + a_{2n}x_n = \tilde{J}_2, \\ &\qquad\vdots \\ &a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ll}x_l + a_{l(l+1)}x_{l+1} + \cdots + a_{ln}x_n = \tilde{J}_l. \end{aligned}}$$

En este caso podemos pasar los términos que tengan las incógnitas

$$\boxed{x_{l+1}, x_{l+2}, \cdots, x_n}$$

al lado derecho del sistema de ecuaciones, de donde se obtiene

$$\boxed{\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1l}x_l =& \tilde{J}_1 - a_{1(l+1)}x_{l+1} - \cdots - a_{1n}x_n, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2l}x_l =& \tilde{J}_2 - a_{2(l+1)}x_{l+1} - \cdots - a_{2n}x_n, \\ &\vdots \\ a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ll}x_l =& \tilde{J}_l - a_{l(l+1)}x_{l+1} - \cdots - a_{ln}x_n. \end{aligned}}$$

Se puede observar que si suponemos conocidas la últimas $n-(l+1)$ incógnitas entonces tendremos un sistema de l ecuaciones con l incógnitas, el cual se puede resolver con el método de Gauss-Jordan obteniendo la inversa de la matriz asociada del sistema de ecuaciones reducidas. Nótese que en este caso la soluciones están dadas en términos de las incógnitas $x_{l+1}, x_{l+2}, \cdots, x_n$ que pueden tomar cualquier valor, por lo que tendremos un número infinito de soluciones.

Por ejemplo, consideremos el sistema de ecuaciones

$$\begin{aligned} 3x + y - z &= 1, & (10.1) \\ x + y + 3z &= 3, & (10.2) \end{aligned}$$

el cual tiene la matriz extendida

$$\left(\begin{array}{ccc||c} 3 & 1 & -1 & 1 \\ 1 & 1 & 3 & 3 \end{array} \right).$$

Aplicando el método de Gauss-Jordan obtenemos

$$\left(\begin{array}{ccc||c} 3 & 1 & -1 & 1 \\ 1 & 1 & 3 & 3 \end{array} \right) \sim \left(\begin{array}{ccc||c} 3 & 1 & -1 & 2 \\ 0 & -2 & -10 & -8 \end{array} \right) \sim \left(\begin{array}{ccc||c} 3 & 1 & -1 & 1 \\ 0 & 1 & 5 & 4 \end{array} \right),$$

que implica el sistema de ecuaciones equivalente

$$\begin{aligned} 3x + y - z &= 1, \\ y + 5z &= 4. \end{aligned}$$

Pasando los términos con la incógnita z al lado derecho de cada ecuación, se obtienen

$$\begin{aligned} 3x + y &= 1 + z, \\ y &= 4 - 5z. \end{aligned}$$

A este sistema lo trataremos como un sistema de dos ecuaciones con dos incógnitas y puede ser asociado a la matriz extendida

$$\left(\begin{array}{cc||c} 3 & 1 & 1+z \\ 0 & 1 & 4-5z \end{array}\right),$$

aplicando de nuevo el método de Gauss-Jordan se obtiene

$$\begin{aligned} &\left(\begin{array}{cc||r} 3 & 1 & 1+z \\ 0 & 1 & 4-5z \end{array}\right) \\ \sim &\left(\begin{array}{cc||r} 3 & 0 & -3+6z \\ 0 & 1 & 4-5z \end{array}\right) \\ \sim &\left(\begin{array}{cc||r} 1 & 0 & -1+2z \\ 0 & 1 & 4-5z \end{array}\right). \end{aligned}$$

Por lo que la solución al sistema de ecuaciones (10.1)-(10.2) es

$$\begin{aligned} x &= -1+2z, \\ y &= 4-5z, \end{aligned}$$

la cual se puede escribir como

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1+2z \\ 4-5z \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \end{pmatrix} + z\begin{pmatrix} 2 \\ -5 \end{pmatrix},$$

de forma tridimensional esta solución se puede escribir como

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix} + z\begin{pmatrix} 2 \\ -5 \\ 1 \end{pmatrix}.$$

Note que en este caso tenemos un número infinito de soluciones, pues z puede tener cualquier valor.

Para tener desde un inicio un sistema de l ecuaciones con l incógnitas, en algunos casos es conveniente pasar directamente al lado derecho del sistema de ecuaciones un número de incógnitas conveniente. Por ejemplo, consideremos el sistema de ecuaciones

$$\begin{aligned} x+2y+3z+w &= 1, \\ 4x+5y+6z-w &= 3, \\ 7x+2y+9z-2w &= 2, \end{aligned}$$

al pasar los términos que tienen w al lado derecho de cada ecuación obtenemos

$$\begin{aligned} x+2y+3z &= 1-w, \\ 4x+5y+6z &= 3+w, \\ 7x+2y+9z &= 2+2w, \end{aligned}$$

que se puede escribir como

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 2 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1-w \\ 3+w \\ 2+2w \end{pmatrix}.$$

Ahora, debido a que

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 2 & 9 \end{pmatrix}^{-1} = \frac{1}{12} \begin{pmatrix} -11 & 4 & 1 \\ -2 & 4 & -2 \\ 9 & -4 & 1 \end{pmatrix},$$

se obtiene la solución

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{12} \begin{pmatrix} -11 & 4 & 1 \\ -2 & 4 & -2 \\ 9 & -4 & 1 \end{pmatrix} \begin{pmatrix} 1-w \\ 3+w \\ 2+2w \end{pmatrix},$$

es decir

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 3 \\ 6 \\ -1 \end{pmatrix} + \frac{w}{12} \begin{pmatrix} 17 \\ 2 \\ -11 \end{pmatrix},$$

que de forma cuadridimensional se puede expresar como

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 3 \\ 6 \\ -1 \\ 0 \end{pmatrix} + \frac{w}{12} \begin{pmatrix} 17 \\ 2 \\ -11 \\ 12 \end{pmatrix}.$$

Note que tenemos un número infinito de soluciones pues w puede tomar cualquier valor.

10.2. Propiedades de las soluciones

Hemos visto cómo resolver un sistema de m ecuaciones lineales con n incógnitas. Vale la pena analizar diferentes propiedades de las soluciones.

De nuevo supongamos que tenemos el sistema de ecuaciones

$$\boxed{\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n =& J_1, \qquad (10.3)\\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n =& J_2, \qquad (10.4)\\ &\vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n =& J_m, \qquad (10.5) \end{aligned}}$$

donde $A_{ij}, J_i\,(i = 1, 2, \cdots m, j = 1, \cdots, n)$ son constantes dadas. Nótese que definiendo la matriz y los vectores

$$\boxed{A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad J = \begin{pmatrix} J_1 \\ J_2 \\ \vdots \\ J_m \end{pmatrix},}$$

el sistema de ecuaciones (10.3)-(10.5) se puede escribir como

$$\boxed{AX = J.} \qquad (10.6)$$

Al sistema de ecuaciones (10.6) se le puede asociar un sistema de ecuaciones homogéneo de la forma

$$\boxed{AX = 0.} \qquad (10.7)$$

El conjunto de soluciones de este sistema de ecuaciones homogéneo forma un espacio vectorial. En efecto, supongamos que X_1 y X_2 satisfacen (10.7) entonces si λ_1 y λ_2 son constantes la combinación lineal

$$\boxed{\lambda_1 X_1 + \lambda_2 X_2}$$

cumple que

$$A(\lambda_1 X_1 + \lambda_2 X_2) = A\lambda_1 X_1 + A\lambda_2 X_2 = \lambda_1 A X_1 + \lambda_2 A X_2 = 0 + 0 = 0,$$

es decir,

$$A(\lambda_1 X_1 + \lambda_2 X_2) = 0.$$

Así, cualquier combinación lineal de las soluciones del sistema homogéneo (10.7) también es solución de dicho sistema homogéneo. Así, el conjunto de las soluciones del sistema homogéneo (10.7) forman un espacio vectorial.

Ahora, supongamos que X_p es un vector particular que satisface el sistema de ecuaciones (10.6), es decir que cumple

$$\boxed{AX_p = J,}$$

también supongamos que $X_{H1}, X_{H2}, \cdots, X_{Hs}$ es la base del espacio vectorial que satisface el sistema de ecuaciones homogéneo (10.7),

$$\boxed{AX_{Hi} = 0, \qquad i = 1, 2, \cdots, s,}$$

entonces el vector

$$\boxed{X = \lambda_i X_{Hi} + X_p, \qquad \lambda_i \in \mathbb{R}} \tag{10.8}$$

también satisface el sistema de ecuaciones (10.6). Esta afirmación es correcta, pues

$$AX = A(\lambda_i X_{Hi} + X_p) = \lambda_i A X_{Hi} + AX_p = 0 + J = J,$$

en otras palabras

$$AX = J.$$

Como podemos ver un sistema de ecuaciones puede tener muchas soluciones. Esto depende del sistema de ecuaciones homogéneas.

10.3. Rango de una matriz

Antes de continuar definiremos como rango r de la matriz A como el número de renglones linealmente independientes que se obtienen después aplicar el método de Gauss-Jordan a la matriz A. Nótese que el rango de la matriz A también nos da el número de ecuaciones linealmente independientes del sistema de ecuaciones (10.3)-(10.5). Por lo tanto, si tenemos n incógnitas y el rango del matriz A es r, el número de incógnitas libres es

$$n - r$$

y este es la dimensión del espacio de soluciones del sistema homogéneo.

10.4. Ejemplos

Consideremos el sistema de ecuaciones

$$x + 2y - z + 3w + 2u = 1, \tag{10.9}$$

$$-2x + 4y + 2z + 2w - u = 3, \tag{10.10}$$

$$x + 2y + 3z + w + u = -2, \tag{10.11}$$

al pasar los términos que tienen w y u al lado derecho de cada ecuación obtenemos

$$\begin{aligned} x + 2y - z &= 1 - 3w - 2u, \\ -2x + 4y + 2z &= 3 - 2w + u, \\ x + 2y + 3z &= -2 - w - u, \end{aligned}$$

que se puede escribir como

$$\begin{pmatrix} 1 & 2 & -1 \\ -2 & 4 & 2 \\ 2 & 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 - 3w - 2u \\ 3 - 2w + u \\ -2 - w - u \end{pmatrix}.$$

Además, se puede mostrar que

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 2 & 9 \end{pmatrix}^{-1} = \frac{1}{8} \begin{pmatrix} 2 & -2 & 2 \\ 2 & 1 & 0 \\ -2 & 0 & 2 \end{pmatrix},$$

de donde

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 2 & -2 & 2 \\ 2 & 1 & 0 \\ -2 & 0 & 2 \end{pmatrix}\begin{pmatrix} 1-3w-2u \\ 3-2w+u \\ -2-w-u \end{pmatrix},$$

es decir,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{8}\begin{pmatrix} -8 \\ 5 \\ -6 \end{pmatrix} + \frac{w}{8}\begin{pmatrix} -4 \\ -8 \\ 4 \end{pmatrix} + \frac{u}{8}\begin{pmatrix} -8 \\ -3 \\ 2 \end{pmatrix}.$$

Por lo que, usando las cinco dimensiones del problema se obtiene

$$\begin{pmatrix} x \\ y \\ z \\ w \\ u \end{pmatrix} = \frac{1}{8}\begin{pmatrix} -8 \\ 5 \\ -6 \\ 0 \\ 0 \end{pmatrix} + \frac{w}{8}\begin{pmatrix} -4 \\ -8 \\ 4 \\ 8 \\ 0 \end{pmatrix} + \frac{u}{8}\begin{pmatrix} -8 \\ -3 \\ 2 \\ 0 \\ 8 \end{pmatrix}.$$

Note que tenemos un número infinito de soluciones pues w y u puede tomar cualquier valor. También observe que el vector

$$X_p = \frac{1}{8}\begin{pmatrix} -8 \\ 5 \\ -6 \\ 0 \\ 0 \end{pmatrix}$$

es un solución particular del sistema de ecuaciones (10.9)-(10.11), mientras que los vectores

$$X_1 = \begin{pmatrix} -4 \\ -8 \\ 4 \\ 8 \\ 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} -8 \\ -3 \\ 2 \\ 0 \\ 8 \end{pmatrix}.$$

son soluciones al sistema de ecuaciones homogéneo asociado al sistema de ecuaciones no homogéneo (10.9)-(10.11).

10.5. Ejercicios

1.- Resolver el sistema de ecuaciones

$$\begin{aligned} 5x - y + 2z + w &= 7, \\ 2x + y + 4z - 2w &= 1, \\ x - 3y - 6z + 5w &= 0. \end{aligned}$$

2.- Resolver el sistema de ecuaciones

$$\begin{aligned} x + y + 3z - 2w - q &= 1, \\ 5x - 2y + 3z + 7w + 8q &= 3, \\ -3x - y + 2z + 7w + 5q &= 2, \\ 5x + 3y + z - 2w - 7q &= 3. \end{aligned}$$

Capítulo 11

Matriz de cambio de base

En este capítulo veremos la llamada matriz de cambio de base.

11.1. Definición

Sea V un espacio vectorial de dimensión n sobre el campo K y

$$\boxed{\{v_1, v_2, \cdots, v_n\}} \tag{11.1}$$

una bases de V. Entonces, cualquier vector de V se puede escribir como combinación lineal de la base (11.1).

Además, supongamos que

$$\boxed{\{v_1', v_2', \cdots, v_n'\}} \tag{11.2}$$

es otra base de V. Por lo cual también cualquier vector de V se puede escribir como combinación lineal de la base (11.2).

En particular, para los vectores de la base (11.1) existen n^2 escalares $\beta_{ij} \in K, j = 1, 2, \cdots, n$ tales que

$$\boxed{v_i' = \beta_{ij} v_j,} \tag{11.3}$$

observe que esta igualdad se puede escribir como

$$\boxed{\begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{pmatrix} = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.}$$

Así, el conjunto de números β_{ij} define una matriz.

De la misma forma, usando la base (11.1), para los vectores de la base (11.2) existen n^2 escalares $\beta'_{ij} \in K$ tales que

$$v_i = \beta'_{ij} v'_j. \tag{11.4}$$

Usando esta última expresión en la ecuación (11.3), encontramos

$$v_i = \beta'_{ij} \beta_{jk} v_k, \qquad k = 1, 2, \cdots, n$$

es decir

$$v_i = (\beta' \beta)_{ik} v_k. \tag{11.5}$$

Ahora, debido a que

$$v_i = \delta_{ik} v_k,$$

la igualdad (11.5) se puede escribir como

$$v_k \delta_{ik} = (\beta' \beta)_{ik} v_k,$$

por lo que

$$(\delta_{ik} - (\beta' \beta)_{ik}) v_k = 0. \tag{11.6}$$

Como el conjunto de vectores v_k forman una base de V, son linealmente independientes. Por lo que la ecuación (11.6) implica la igualdad

$$\delta_{ik} - (\beta' \beta)_{ik} = 0$$

de donde

$$\delta_{ik} = (\beta' \beta)_{ik}$$

lo que es lo mismo

$$\beta'\beta = I.$$

Por lo tanto, la matriz β es invertible y su inversa es β'.

Ahora, si $v \in V$ existe un conjunto de escalares

$$\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \qquad \lambda_i \in K, i = 1, 2, \cdots, n$$

tales que

$$\boxed{v = \lambda_i v_i} \tag{11.7}$$

Pero también existen un conjunto de escalares $\lambda'_i \in K$ tal que

$$\boxed{v = \lambda'_i v'_i.} \tag{11.8}$$

Tomando en cuenta que las ecuaciones (11.7) y (11.8) representan al mismo vector en diferentes bases, veamos como se relacionan los escalares λ_i con los escalares λ'_i.

Se puede observar que sustituyendo la igualdad (11.4) en la ecuación (11.7) encontramos

$$v = \lambda_i v_i = \lambda_i \beta'_{ij} v'_j,$$

además, considerando la ecuación (11.8) se encuentra la expresión

$$v = \lambda_i v_i = \lambda_i \beta'_{ij} v'_j = \lambda'_j v'_j,$$

de donde se deduce el resultado

$$0 = \lambda_i \beta'_{ij} v'_j - \lambda'_j v'_j = \left(\lambda_i \beta'_{ij} - \lambda'_j\right) v'_j.$$

Es decir,

$$\left(\lambda_i \beta'_{ij} - \lambda'_j\right) v'_j = 0. \tag{11.9}$$

Además, como los vectores v'_j forman una base, son linealmente independientes y la ecuación (11.9) implica la igualdad

$$\lambda_i \beta'_{ij} - \lambda'_j = 0,$$

de donde se encuentra la expresión

$$\lambda'_j = \lambda_i \beta'_{ij} = \beta'^{T}_{ji} \lambda_i. \tag{11.10}$$

A la matriz

$$\boxed{N' = \beta'^{T},}$$

le llamamos matriz de cambio de base. Entonces, la ecuación (11.10) toma la forma

$$\boxed{\lambda'_j = N'_{ji} \lambda_i = N^{-1}_{ji} \lambda_i,}$$

de donde

$$\boxed{\lambda_j = N_{ji} \lambda'_i,}$$

por lo que

$$\boxed{\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} N_{11} & N_{12} & \cdots & N_{1n} \\ N_{21} & N_{22} & \cdots & N_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ N_{n1} & N_{n2} & \cdots & N_{nn} \end{pmatrix} \begin{pmatrix} \lambda'_1 \\ \lambda'_2 \\ \vdots \\ \lambda'_n \end{pmatrix}.}$$

11.2. Ejemplo

Como un ejemplo consideremos la base canónica

$$\{v_1 = (1,0,0), v_2 = (0,1,0), v_3 = (0,1,0)\}$$

y la base

$$\{v'_1 = (1,0,0), v'_2 = (1,1,0), v'_3 = (1,1,1)\}$$

en este caso tenemos

$$\begin{aligned} (1,0,0) &= 1(1,0,0) + 0(0,1,0) + 0(0,0,1), \\ (1,1,0) &= 1(1,0,0) + 1(0,1,0) + 0(0,0,1), \\ (1,1,1) &= 1(1,0,0) + 1(0,1,0) + 1(0,0,1), \end{aligned}$$

es decir

$$\begin{aligned} v_1' &= 1v_1 + 0v_2 + 0v_3, \\ v_2' &= 1v_1 + 1v_2 + 0v_3, \\ v_3' &= 1v_1 + 1v_2 + 1v_3, \end{aligned}$$

de donde

$$\beta = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

y la matriz asociada al cambio de base es

$$N = (\beta)^T = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note que también se cumple

$$\beta^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \qquad N^{-1} = \left(\beta^{-1}\right)^T = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

las cuales se obtienen al hacer el cambio de base inverso.

11.3. Ejercicios

1.- Considere las bases

$$\{v_1 = (1,0,0), v_2 = (0,1,0), v_3 = (0,0,1)\}$$

y

$$\{v_1' = (1,2,0), v_2' = (-2,1,0), v_3' = (0,0,1)\},$$

obtener la matriz de cambio de base entre ellas.

2.- Considere las bases

$$\{v_1 = (1,0,0,0), v_2 = (0,1,0,0), v_3 = (0,0,1,0), v_4 = (0,0,0,1\}$$

y

$$\{v_1' = (4,-1,-3,6), v_2' = (-2,3,1,4), v_3' = (0,5,0,0), v_4' = (1,2,3,-1)\},$$

obtener la matriz de cambio de base entre ellas.

Capítulo 12

Transformaciones lineales

En este capítulo veremos la definición de transformación lineal y algunas implicaciones importantes de dicha definición.

12.1. Definición

Sean dos espacios vectoriales $(V, K, +, \cdot)$ y $(W, K, +, \cdot)$ sobre el mismo campo K. Se dice que la función

$$T : V \longrightarrow W$$

es lineal si satisface las propiedades

$$\boxed{\begin{aligned} &T(v_1 + v_2) = T(v_1) + T(v_2), \qquad \forall v_1, v_2 \in V, &(12.1)\\ &T(\lambda v) = \lambda T(v), \qquad \forall \lambda \in K, \forall v \in V. &(12.2)\end{aligned}}$$

Es posible probar que una función T es lineal si y solo si satisface la propiedad

$$\boxed{T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2), \quad \forall \lambda_1, \lambda_2 \in K, \forall v_1, v_2 \in V.} \tag{12.3}$$

En efecto, supongamos que se cumple la propiedad (12.3), en particular se cumple si $\lambda_1 = \lambda_2 = e$, con e la identidad del producto del campo K. Entonces

$$T(ev_1 + ev_2) \;\; = \;\; eT(v_1) + eT(v_2),$$

por lo que, para cualquier par de vectores v_1, v_2 en V se cumple

$$T(v_1 + v_2) \;\; = \;\; T(v_1) + T(v_2),$$

que es la propiedad (12.1). Ahora, si tomamos $\lambda_1 = \lambda, v_1 = v$ y $\lambda_2 = 0$, la ecuación (12.3) toma la forma

$$T(\lambda v) \quad = \quad \lambda T(v),$$

que es la propiedad (12.2). Por lo tanto, si una función T satisface la propiedad (12.3), entonces es lineal.

Ahora supongamos T es lineal y que $v_1, v_2 \in V, \lambda_1, \lambda_2 \in K$. Usando las propiedades (12.1)-(12.2) se cumple

$$T(\lambda_1 v_1 + \lambda_2 v_2) = T(\lambda_1 v_1) + T(\lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2),$$

que es la propiedad (12.3). Por lo tanto, si una función T es lineal, cumple la propiedad (12.3).

En conclusión, una función T es lineal si y solo si satisface la propiedad (12.3).

12.2. Transformaciones lineales en $\mathbb{R}^n$

Si V es un espacio vectorial y $\lambda \in K$, podemos definir la función

$$T : V \longrightarrow V$$

tal que

$$T(v) = \lambda v, \qquad v \in V,$$

la cual es un ejemplo de función lineal. En efecto, si $\alpha, \beta \in K$ y $v_1, v_2 \in V$, se cumple

$$\begin{aligned} T(\alpha v_1 + \beta v_2) &= \lambda(\alpha v_1 + \beta v_2) \\ &= \lambda \alpha v_1 + \lambda \beta v_2 \\ &= \alpha \lambda v_1 + \beta \lambda v_2 \\ &= \alpha T(v_1) + \beta T(v_2). \end{aligned}$$

esto es

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2).$$

Para los espacios vectoriales $\mathbb{R}^2$ y $\mathbb{R}^2$, la función

$$T(x, y) = (x + 3y, -2x + 5y) \tag{12.4}$$

es lineal. En efecto, si $\lambda_1, \lambda_2 \in \mathbb{R}$ se cumple

$$\begin{aligned}
&T(\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2)) = T(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2) \\
&= (\lambda_1 x_1 + \lambda_2 x_2 + 3(\lambda_1 y_1 + \lambda_2 y_2), -2(\lambda_1 x_1 + \lambda_2 x_2) + 5(\lambda_1 y_1 + \lambda_2 y_2)) \\
&= (\lambda_1(x_1 + 3y_1) + \lambda_2(x_2 + 3y_2), \lambda_1(-2x_1 + 5y_1) + \lambda_2(-2x_2 + 5y_2)) \\
&= (\lambda_1(x_1 + 3y_1), \lambda_1(-2x_1 + 5y_1)) + (\lambda_2(x_2 + 3y_2), \lambda_2(-2x_2 + 5y_2)) \\
&= \lambda_1(x_1 + 3y_1, -2x_1 + 5y_1) + \lambda_2(x_2 + 3y_2, -2x_2 + 5y_2) \\
&= \lambda_1 T(x_1, y_1) + \lambda_2 T(x_1, y_1),
\end{aligned}$$

por lo cual

$$T(\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2)) = \lambda_1 T(x_1, y_1) + \lambda_2 T(x_1, y_1).$$

En general, se puede probar que si tenemos el vector $(x_1, x_2, \cdots, x_m)$ donde $x_k \in \mathbb{R}$ y $k = 1, 2, \cdots, m$, entonces dados los números $\alpha_{lk} \in \mathbb{R}$ con $l = 1, 2, \cdots n$, la función

$$T(x_1, x_2, \cdots, x_m) = (\alpha_{1k} x_k, \alpha_{2k} x_k, \cdots, \alpha_{nk} x_k) \tag{12.5}$$

que es una función lineal de $\mathbb{R}^m$ en $\mathbb{R}^n$. En este caso se puede obsevar que definiendo la matriz

$$M = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nm} \end{pmatrix}, \quad \alpha_{ij} \in \mathbb{R}, \quad i = 1, \cdots n, j = 1, \cdots, m.$$

se tiene

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \alpha_{1k} x_k \\ \alpha_{2k} x_k \\ \vdots \\ \alpha_{nk} x_k \end{pmatrix}.$$

Por lo que la transformación lineal (12.5) se puede asociar con una matriz. En particular, la transformación lineal (12.4) se puede asociar con la matriz

$$\begin{pmatrix} 1 & 3 \\ -2 & 5 \end{pmatrix},$$

y se cumple

$$\begin{pmatrix} 1 & 3 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 3y \\ -2x + 5y \end{pmatrix}.$$

12.3. Transformaciones entre funciones

12.3.1. Derivada

En el espacio de las funciones suaves, la derivada cumple con las propiedades de transformación lineal. En efecto, si definimos

$$\boxed{T = \frac{\partial}{\partial x}}$$

tenemos

$$\begin{aligned} T\left(\alpha f_1(x) + \beta f_2(x)\right) &= \frac{\partial}{\partial x}\left(\alpha f_1(x) + \beta f_2(x)\right) \\ &= \alpha \frac{\partial}{\partial x} f_1(x) + \beta \frac{\partial}{\partial x} f_2(x) \\ &= \alpha T\left(f_1(x)\right) + \beta T\left(f_2(x)\right). \end{aligned}$$

12.3.2. Producto por una función fija

Dado una función $f(x)$, usando el producto por un escalar en un espacio vectorial V podemos definir una función lineal T de la siguiente forma

$$\boxed{T(v) = f(x)v,}$$

donde v es un vector en V. Para probar que T es una función lineal, consideremos $v_1, v_2 \in V$ y $\alpha, \beta \in K$, entonces

$$\begin{aligned} T(\alpha v_1 + \beta v_2) &= f(x)(\alpha v_1 + \beta v_2) = \alpha f(x) v_1 + f(x)\beta v_2 \\ &= \alpha T\left(v_1\right) + \beta T\left(v_2\right). \end{aligned}$$

Por lo que T es una función lineal.

12.3.3. Transformadas integrales

Dada una función $g(k,x)$ integrable se puede definir una transformación lineal entre funciones integrales. Si $f(x)$ es una función integrable definimos

$$\boxed{\tilde{f}(k) = T(f(x)) = \int_a^b dx g(k,x) f(x).} \tag{12.6}$$

Esta transformación es lineal, pues si $f_1(x), f_2(x)$ son funciones suaves y α, β escalares, se encuentra

$$\begin{aligned} T(\alpha f_1(x) + \beta f_2(x)) &= \int_a^b dx g(k,x) \left(\alpha f_1(x) + \beta f_2(x)\right) \\ &= \alpha \int_a^b dx g(k,x) f_1(x) + \beta \int_a^b dx g(k,x) f_2(x) \\ &= \alpha T(f_1(x)) + \beta T(f_2(x)). \end{aligned}$$

A la transformación (12.6) se le llama transformada integral en base $g(k,x)$.

Por ejemplo, con $g(k,x) = e^{-ikx}$ se define la transformada de Fourier

$$\tilde{f}(k) = \int_{-\infty}^{\infty} dx e^{-ikx} f(x).$$

Para cada función, $g(k,x)$, bien portada (derivable y suave) se puede definir una transformada integral.

12.4. Núcleo

Sean dos espacios vectoriales $(V, K, +, \cdot)$ y $(W, K, +, \cdot)$ sobre el mismo campo K y

$$T : V \longrightarrow W$$

una función lineal.

Sabemos que si 0 es el neutro aditivo del campo K y v es un vector del espacio vectorial V, entonces su producto también es el neutro aditivo de V, es decir,

$$0v = 0.$$

Por lo que si T es una función lineal, se cumple

$$T(0) = T(0v) = 0T(v) = 0,$$

o lo que es lo mismo

$$T(0) = 0.$$

Así, si T es una función lineal, T manda el cero de V al cero de W.

Ahora definiremos el conjunto de todos los vectores de V tal que van al cero de W

$$\boxed{Ker(T) = \{v \in V | \quad T(v) = 0\}.}$$

A este conjunto lo llamamos *kernel* o núcleo de T. Note que $0 \in Ker(T)$ por lo que el núcleo de T es un conjunto no vacío.

El conjunto $Ker(T)$ es un subespacio vectorial de V. En efecto, supongamos que

$$v_1, v_2 \in Ker(T), \qquad \alpha_1, \alpha_2 \in K.$$

Por lo que podemos formar el vector

$$v = \alpha_1 v_1 + \alpha_2 v_2$$

y como T es una función lineal, se satisface

$$\begin{aligned} T(v) &= T(\alpha_1 v_1 + \alpha_2 v_2) \\ &= \alpha_1 T(v_1) + \alpha_2 T(v_2) \\ &= \alpha_1 0 + \alpha_2 0 = 0, \end{aligned}$$

es decir

$$T(v) = 0.$$

Este resultado nos indica que

$$v = \alpha_1 v_1 + \alpha_2 v_2 \in Ker(T).$$

En resumen hemos mostrado que si

$$v_1, v_2 \in Ker(T)$$

y $\alpha_1, \alpha_2 \in K$, se cumple que

$$\alpha_1 v_1 + \alpha_2 v_2 \in Ker(T).$$

Este resultado nos indica que $Ker(T)$ es un subespacio vectorial de V.

12.4.1. Núcleo y funciones inyectivas

El conjunto $Ker(T)$ contiene información de la función T. Por ejemplo, un resultado interesante es que si el núcleo de T solo contiene al vector cero, entonces T es una función invectiva. Para probar esta afirmación supongamos que

$$Ker(T) = \{0\}.$$

Ahora si

$$v_1, v_2 \in V$$

y se cumplen

$$T(v_1) = T(v_2),$$

entonces

$$0 = T(v_1) - T(v_2).$$

Además, como T es una función lineal, se encuentra

$$T(v_1 - v_2) = 0,$$

por lo que

$$v_1 - v_2 \in Ker(T) = \{0\},$$

en consecuencia

$$v_1 - v_2 = 0,$$

es decir

$$v_1 = v_2.$$

Así, hemos probado que si $Ker(T) = \{0\}$ y $T(v_1) = T(v_2)$ entonces $v_1 = v_2$. En otras palabras si $Ker(T) = \{0\}$, la función T es inyectiva.

12.4.2. Núcleo e independencia lineal

Otro resultado interesantes es que si el conjunto de vectores de V

$$v_1, v_2, \cdots, v_n$$

es linealmente independiente y $Ker(T) = \{0\}$, entonces el conjunto de vectores de W

$$T(v_1), T(v_2), \cdots, T(v_n)$$

es linealmente independiente. Para probar esta aseveración supongamos que existen escalares en K

$$\alpha_1, \alpha_2, \cdots, \alpha_n$$

tal que

$$\alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) = 0.$$

Como T es una función lineal, se llega a

$$0 = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) = T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n)$$

es decir

$$T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = 0,$$

lo que implica

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in Ker(T) = \{0\}.$$

Por lo que

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

y como el conjunto $v_1, v_2, \cdots, v_n$ es linealmente independiente, se obtiene

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

En resumen, si $v_1, v_2, \cdots, v_n$ es linealmente independiente y $Ker(T) = \{0\}$ entonces el conjunto de vectores $T(v_1), T(v_2), \cdots, T(v_n)$ es linealmente independiente.

12.5. Imagen

Sean dos espacios vectoriales $(V, K, +, \cdot)$ y $(W, K, +, \cdot)$ sobre el mismo campo K y

$$T : V \longrightarrow W$$

una función lineal. Definimos la **imagen** de T como el subconjunto de los vectores w en W tales que existe al menos un vector v en V que cumple

$$T(v) = w,$$

es decir,

$$\boxed{T(V) = Im(T) = \{w \in W| \quad \exists v \in V, \qquad T(v) = w\}.}$$

12.5.1. Imagen como subespacio del contradominio

El conjunto $Im(T)$ es un subespacio vectorial de W. En efecto si

$$w_1, w_2 \in Im(T),$$

entonces existe

$$v_1, v_2 \in V$$

tal que

$$T(v_1) = w_1, \qquad T(v_2) = w_2.$$

Ahora, sea $\alpha_1, \alpha_2 \in K$, con estos escalares podemos formar el vector

$$w = \alpha_1 w_1 + \alpha_2 w_2 \in W,$$

el cual se puede escribir como

$$\begin{aligned} w &= \alpha_1 w_1 + \alpha_2 w_2 \\ &= \alpha_1 T(v_1) + \alpha_2 T(v_2) \\ &= T(\alpha_1 v_1 + \alpha_2 v_2), \end{aligned}$$

por lo que existe

$$v = \alpha_1 v_1 + \alpha_2 v_2 \in V$$

tal que

$$T(v) = w = \alpha_1 w_1 + \alpha_2 w_2.$$

Así hemos probado que si $w_1, w_2 \in Im(T)$ y $\alpha_1, \alpha_2 \in K$, se cumple que $\alpha_1 w_1 + \alpha_2 w_2 \in Im(T)$. Lo cual nos indica que $Im(T)$ es un subespacio vectorial de W.

12.5.2. Imagen y subespacios en el dominio

Ahora supongamos que A es un subconjunto de V, entonces definiremos

$$\boxed{T(A) = \{w \in W| \quad \exists v \in A, \qquad T(v) = w\}.}$$

Observe que $T(A)$ es subconjunto de W, se dice que $T(A)$ es el conjunto imagen de A bajo la función T.

Si A es subespacio vectorial de V, entonces $T(A)$ es subespacio vectorial de W. Para probar esta afirmación, supongamos que

$$w_1, w_2 \in T(A),$$

entonces existe

$$v_1, v_2 \in A$$

tal que

$$T(v_1) = w_1, \qquad T(v_2) = w_2.$$

Ahora, sea $\alpha_1, \alpha_2 \in K$, con estos escalares podemos formar el vector

$$w = \alpha_1 w_1 + \alpha_2 w_2 \in W,$$

el cual se puede escribir como

$$w = \alpha_1 w_1 + \alpha_2 w_2 = \alpha_1 T(v_1) + \alpha_2 T(v_2) = T\left(\alpha_1 v_1 + \alpha_2 v_2\right).$$

Adicionalmente, debido a que A es un subespacio vectorial se cumple que

$$v = \alpha_1 v_1 + \alpha_2 v_2 \in A,$$

así $\exists v \in A$ tal que

$$T(v) = w = \alpha_1 w_1 + \alpha_2 w_2,$$

de donde

$$\alpha_1 w_1 + \alpha_2 w_2 \in T(A).$$

Como se puede notar, hemos probado que si $w_1, w_2 \in T(A)$ y $\alpha_1, \alpha_2 \in K$, se cumple que $\alpha_1 w_1 + \alpha_2 w_2 \in T(A)$, lo cual nos indica que $T(A)$ es un subespacio vectorial de W.

12.6. Imagen inversa

Ahora supongamos que B es un subconjunto de W, entonces definiremos el conjunto

$$\boxed{T^{-1}(B) = \{v \in V| \quad \exists w \in B, \qquad T(v) = w\}.}$$

Note que $T^{-1}(B)$ es subconjunto de V, se dice que este conjunto es la imagen inversa de B bajo la función T.

12.6.1. Imagen inversa como subespacio vectorial del dominio

Si B es un subespacio vectorial de W, entonces $T^{-1}(B)$ es subespacio vectorial de V.

Para probar este resultado supongamos que

$$v_1, v_2 \in T^{-1}(B),$$

entonces existe

$$w_1, w_2 \in B$$

tal que

$$T(v_1) = w_1, \qquad T(v_2) = w_2.$$

Ahora, sea $\alpha_1, \alpha_2 \in K$, con estos escalares podemos formar la combinación lineal

$$v = \alpha_1 v_1 + \alpha_2 v_2.$$

Considerando que T es un función lineal tenemos

$$\begin{aligned} T(v) &= T(\alpha_1 v_1 + \alpha_2 v_2) \\ &= \alpha_1 T(v_1) + \alpha_2 T(v_2) \\ &= \alpha_1 w_1 + \alpha_2 w_2, \end{aligned}$$

además como B es subespacio vectorial de W se cumple

$$T(v) = \alpha_1 w_1 + \alpha_2 w_2 \in W.$$

Por lo que, existe $w \in B$ tal que $T(v) = w$ entonces

$$v = \alpha_1 v_1 + \alpha_2 v_2 \in T^{-1}(B).$$

Así hemos probado que si $v_1, v_2 \in T^{-1}(B)$ y $\alpha_1, \alpha_2 \in K$, entonces $\alpha_1 v_1 + \alpha_2 v_2 \in T^{-1}(B)$. Lo cual dice que si B es subespacio vectorial de W, $T^{-1}(B)$ es un subespacio vectorial de V.

12.6.2. Imagen inversa e independecia lineal

Otro resultado interesante es que un conjunto de vectores linealmente independientes de $T(V)$ viene de un conjunto de vectores linealmente de V. En efecto, supongamos que los vectores

$$w_1, w_2, \cdots, w_l \tag{12.7}$$

son linealmente independientes y pertenecen a $T(V)$. Entonces existen vectores

$$v_1, v_2, \cdots, v_l \tag{12.8}$$

en V tales que

$$T(v_1) = w_1, T(v_2) = w_2, \cdots, T(v_l) = w_l. \tag{12.9}$$

Ahora supongamos que existen los escalares

$$\alpha_1, \alpha_2, \cdots, \alpha_l$$

en K tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_l v_l = 0.$$

Entonces, como T es una función lineal, se tiene

$$\begin{aligned} 0 &= T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_l v_l) \\ &= \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_l T(v_l) \\ &= \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_l w_l, \end{aligned}$$

es decir

$$\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_l w_l = 0.$$

Además, como los vectores (12.7) son linealmente independientes, se encuentra que

$$\alpha_1 = \alpha_2 = \cdots = \alpha_l = 0.$$

Por lo tanto si $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_l v_l = 0$, entonces $\alpha_1 = \alpha_2 = \cdots = \alpha_l = 0$. Es decir, el conjunto de vectores (12.8) es linealmente independiente.

Así, podemos afirmar que si los vectores $w_1, w_2, \cdots, w_l$ pertenecen a $T(V)$ y son linealmente independientes, entonces existe un conjunto de vectores $v_1, v_2, \cdots, v_l$ en V que son linealmente independientes y satisfacen $T(v_1) = w_1, T(v_2) = w_2, \cdots, T(v_l) = w_l$.

12.7. Teorema de la dimensión

Sean dos espacios vectoriales $(V, K, +, \cdot)$ y $(W, K, +, \cdot)$ de dimensión finita sobre el mismo campo K y

$$T : V \longrightarrow W$$

una función lineal. Mostraremos que la dimensión de $Ker(T)$ más la dimensión de $Im(T)$ es la dimensión del espacio vectorial V.

Sin pérdida de generalidad, podemos suponer que la dimensión del espacio vectorial V es n y que los vectores

$$v_1, v_2, \cdots, v_r, \qquad r \leq n \tag{12.10}$$

son una base del espacio $Ker(T)$. Entonces, existen $n - r$ vectores linealmente independientes

$$v_{r+1}, v_{r+2}, \cdots, v_n \tag{12.11}$$

de tal forma que los vectores

$$v_1, v_2, \cdots v_r, v_{r+1}, \cdots, v_n \tag{12.12}$$

formen una base de V.

Observe que si

$$T(v_j) = 0, \qquad j = r+1, r+2, \cdots, n$$

entonces

$$v_j \in Ker(T)$$

y como los vectores (12.10) forman una base del conjunto $Ker(T)$, entonces existen r escalares $\alpha_1, \alpha_2, \cdots, \alpha_r$ tales que

$$v_j = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r.$$

Esta última ecuación contradice el hecho de que los vectores (12.12) sean linealmente independientes, y por lo tanto, que sean base de V. Entonce, deben existir $n-r$ vectores no nulos en W de tal forma que

$$T(v_{r+1}) = w_{r+1}, T(v_{r+2}) = w_{r+2}, \cdots, T(v_n) = w_n. \tag{12.13}$$

Estos vectores son linealmente independientes. En efecto, supongamos que existen $n-r$ escalares

$$\lambda_{r+1}, \lambda_{r+2}, \cdots, \lambda_n$$

de tal forma que

$$\lambda_{r+1} w_{r+1} + \lambda_{r+2} w_{r+2} + \cdots + \lambda_n w_n = 0,$$

es decir,

$$\lambda_{r+1} T(v_{r+1}) + \lambda_{r+2} T(v_{r+2}) + \cdots + \lambda_n T(v_n) = 0.$$

Por lo que

$$T(\lambda_{r+1} v_{r+1} + \lambda_{r+2} v_{r+2} + \cdots + \lambda_n v_n) = 0,$$

así,

$$\lambda_{r+1} v_{r+1} + \lambda_{r+2} v_{r+2} + \cdots + \lambda_n v_n \in Ker(T).$$

Considerando que el conjunto de vectores (12.10) es una base del conjunto $Ker(T)$, existen r escalares $\mu_1, \mu_2, \cdots, \mu_r$ tales que

$$\lambda_{r+1} v_{r+1} + \lambda_{r+2} v_{r+2} + \cdots + \lambda_n v_n = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_r v_r$$

o sea

$$\lambda_{r+1} v_{r+1} + \lambda_{r+2} v_{r+2} + \cdots + \lambda_n v_n - \mu_1 v_1 - \mu_2 v_2 - \cdots - \mu_r v_r = 0.$$

Esta expresión es una combinación lineal nula de los vectores (12.12), los cuales son una base de V, por lo tanto

$$\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = \mu_1 = \mu_2 = \cdots = \mu_r = 0.$$

En consecuencia, hemos mostrado que si $\lambda_{r+1} w_{r+1} + \lambda_{r+2} w_{r+2} + \cdots + \lambda_n w_n = 0$, entonces $\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$. Es decir, los $n-r$ vectores (12.13) son linealmente independientes.

Ahora, supongamos que $w \in Im(T)$, entonces existe al menos un vector $v \in V$ de tal forma que

$$T(v) = w.$$

Además, debido a que el conjunto de vectores (12.12) son una base de V, existen n escalares $\zeta_1, \zeta_2, \cdots \zeta_n$ tales que

$$v = \zeta_1 v_1 + \zeta_2 + \cdots \zeta_r v_r + \zeta_{r+1} v_{r+1} + \cdots + \zeta_n v_n.$$

Esta igualdad implica

$$\begin{aligned} w &= T(v) = T(\zeta_1 v_1 + \zeta_2 v_2 + \cdots \zeta_r v_r + \zeta_{r+1} v_{r+1} + \cdots + \zeta_n v_n) \\ &= \zeta_1 T(v_1) + \zeta_2 T(v_2) + \cdots \zeta_r T(v_r) + \zeta_{r+1} T(v_{r+1}) + \cdots + \zeta_n T(v_n) \\ &= \zeta_{r+1} T(v_{r+1}) + \cdots + \zeta_n T(v_n) \\ &= \zeta_{r+1} w_{r+1} + \cdots + \zeta_n w_n, \end{aligned}$$

dicho de otro modo

$$w = \zeta_{r+1} w_{r+1} + \cdots + \zeta_n w_n.$$

Así, hemos mostrado que cualquier vector del conjunto $Im(T)$ se puede escribir como combinación lineal de los $n - r$ vectores (12.13), por lo que dichos vectores son generadores del $Im(T)$.

En resumen, el conjunto de vectores (12.13) es base del espacio vectorial $Im(T)$, lo cual implica que la dimensión de $Im(T)$ es $n - r$. Observe estos resultados implican

$$\boxed{dim\,(Im(T)) + dim\,(Ker(T)) = dimV.}$$

La idea de función lineal está de forma implícita en la obra de Grassmann. Sin embargo, este concepto es formalizado en los trabajo de Peano, quien fue de los pocos matemáticos que logran apreciar la importancia de los trabajos de Grassmann.

Giuseppe Peano (1858-1932) fue un matemático italiano que realizó contribuciones relevantes en los fundamentos de las matemáticas y la lógica. Fue hijo de dos granjeros del Piamonte italiano. De niño tenía que caminar cinco kilómetros al día para asistir a la escuela. Realizó sus estudios doctorales bajo la dirección de Enrico D' Ovidio, en la Universidad de Turín. Se interesó en construir un lenguaje universal escrito para las matemáticas que todos los matemáticos de Mundo pudieran entender y usar. En este sentido, inventó su propia notación que la actualidad se sigue usando en la lógica. También se interesó en dar fundamentos sólidos a las matemáticas, por lo que incursionó en la lógica y en la teoría de conjuntos. En estos temas sus trabajos influyeron a importantes filósofos como Bertrand Russell. Por estas razones es considerado uno de los padres de la Lógica Matemática. Siendo muy joven inició el proyecto llamado *Formulario*, una obra en la cual buscó tener todas las demostraciones y fórmulas matemáticas conocidas hasta su tiempo. Debido a que no le gustaba la calidad de las impresiones de los libros de matemáticas, fundó su propia imprenta. También realizó contribuciones al cálculo y a la geometría. Por ejemplo, construyó la llamada curva de Peano, que tiene propiedades fractales. Peano estaba interesado en la enseñanza de las matemáticas por lo que escribió diversos libros que influyeron a estudiantes italianos de todos los niveles educativos. Murió de un ataque al corazón en 1932.

12.8. Ejercicios

1.- Suponiendo que $g(x)$ es una función continua, determinar si la siguiente transformación es lineal

$$T\left[f(x)\right] = \left(\frac{\partial^2}{\partial x^2} + g(x)\right) f(x).$$

2.- Suponiendo que $g(x)$ es una función continua, determinar si la siguiente transformación es lineal

$$T\left[f(x)\right] = \frac{\partial^2 f(x)}{\partial x^2} + g(x)\left(f(x)\right)^2.$$

3.- Determinar si la siguiente transformación es lineal

$$T\left[f(x)\right] = \ln\left(f(x)\right).$$

4.- Determinar si la siguiente transformación es lineal

$$F(x,y) = (2x - y, x + 4y).$$

5.- Determinar si la siguiente transformación es lineal

$$F(x,y) = (2x - y, x^2 + 4y).$$

6.- Determinar si la siguiente transformación es lineal

$$T\left[f(x)\right] = \int_0^\infty e^{-sx} \frac{1}{f(x)}.$$

7.- En el espacio de las funciones derivables de una variable real, sea el operador lineal

$$\hat{A} = m\frac{d^2}{dx^2} + m\omega^2, \quad m, \omega \in \mathbb{R}, m > 0.$$

Encontrar una función no trivial que pertenezca al conjunto $Ker\left(\hat{A}\right)$.

8.- Sea la matriz

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

Encontrar un vector no trivial que pertenezca al conjunto $Ker(A)$.

9.- Sea el operador lineal

$$\hat{A} = \frac{d^2}{dx^2} + \lambda\frac{d}{dx} + \gamma.$$

y la función

$$Y(x) = \frac{b}{\alpha^2 + \lambda\alpha + \gamma} e^{\alpha x}.$$

Encontrar la función imagen de $Y(x)$ bajo $\hat{A}$.

10.- La Transformada de Laplace de una función $f(x)$ se define como

$$F(s) = L\left[f(x)\right] = \int_0^\infty e^{-sx} f(x) dx.$$

Encontrar la imagen de $f(x) = e^{ax}$ bajo la transformada de Laplace.

Capítulo 13

Combinación lineal y composición de transformaciones lineales

En este capítulo mostraremos que al hacer combinaciones lineales de transformaciones lineales o composición de ellas se obtiene otra transformación lineal.

13.1. Combinación lineal

Sea V un espacio vectorial de dimensión n sobre el campo K y W un espacio vectorial de dimensión m sobre el mismo campo y

$$\begin{aligned} T_1 &: V \longrightarrow W, \\ T_2 &: V \longrightarrow W, \end{aligned}$$

son dos transformaciones lineales .

Si $a, b \in K$ podemos definir otra transformación

$$T : V \longrightarrow W,$$

de la forma

$$\boxed{T = aT_1 + bT_2.}$$

En este caso, si $v \in V$, se define

$$T(v) = (aT_1 + bT_2)(v) = aT_1(v) + bT_2(v).$$

Note que esta nueva transformación cumple

$$\begin{aligned} T(\alpha v_1 + \beta v_2) &= (aT_1 + bT_2)(\alpha v_1 + \beta v_2) \\ &= aT_1(\alpha v_1 + \beta v_2) + bT_2(\alpha v_1 + \beta v_2) \\ &= a(\alpha T_1(v_1) + \beta T_1(v_2)) + b(\alpha T_2(v_1) + \beta T_2(v_2)) \\ &= \alpha(aT_1(v_1) + bT_2(v_1)) + \beta(aT_1(v_2) + bT_2(v_2)) \\ &= \alpha(aT_1 + bT_2)(v_1) + \beta(aT_1 + bT_2)(v_2) \\ &= \alpha T(v_1) + \beta T(v_2). \end{aligned}$$

Así, cualquier combinación lineal de dos operadores lineales nos da otro operador lineal.

13.2. Composición de funciones lineales

Supongamos que V_1, V_2, V_3 son espacios vectoriales sobre el campo K y que las siguientes funciones

$$\begin{aligned} F_1 &: V_1 \longrightarrow V_2, \\ F_2 &: V_2 \longrightarrow V_3, \end{aligned}$$

son lineales. Entonces podemos realizar la composición de funciones lineales

$$\boxed{F = F_2 \circ F_1 : V_1 \longrightarrow V_3,}$$

Probaremos que F es una función lineal. Para realizar esta prueba supongamos que $v_1, v_2 \in V_1$ y $\alpha, \beta \in K$, de donde

$$\begin{aligned} F(\alpha v_1 + \beta v_2) &= (F_1 \circ F_2)(\alpha v_1 + \beta v_2) \\ &= F_1(F_2(\alpha v_1 + \beta v_2)) \\ &= F_1(\alpha F_2(v_1) + \beta F_2(v_2)) \\ &= \alpha F_1(F_2(v_1)) + \beta F_1(F_2(v_2)) \\ &= \alpha(F_1 \circ F_2)(v_1) + \beta(F_1 \circ F_2)(v_2) \\ &= \alpha F(v_1) + \beta F(v_2), \end{aligned}$$

esto es

$$F(\alpha v_1 + \beta v_2) = \alpha F(v_1) + \beta F(v_2).$$

Por lo tanto, la composición de dos funciones lineales también da como resultado otro operador lineal.

13.3. Ejemplos

Por ejemplo, sabemos que los operadores

$$\frac{\partial}{\partial x}, \qquad \frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial z}$$

son lineales, entonces también lo son

$$\frac{\partial^2}{\partial x^2}, \qquad \frac{\partial^2}{\partial y^2}, \qquad \frac{\partial^2}{\partial z^2}.$$

Esto implica que el operador Laplaciano

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

es lineal. Recordemos que cualquier función $V(x, y, z)$ que cumpla el papel de operador, es lineal. Entonces el operador Hamiltoniano

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(x, y, z)$$

es lineal. Esto se debe a que es combinación lineal de opeadores lineales. Además las variables

$$x, \qquad y, \qquad z$$

como operadores son lineales. Entonces los operadores

$$L_x = -i\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \quad L_y = -i\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \quad L_z = -i\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right),$$

son lineales, puesto que son combinaciones lineales de productos de operadores lineales. Por la misma razón, el operador

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

es lineal.

13.4. Ejercicios

1.- Sin c es una constante, con los siguientes operadores lineales del espaciotiempo

$$x, \quad y, \quad z, \quad ct,$$
$$\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial z}, \quad \frac{\partial}{\partial ct}$$

obtener todos los operadores lineales que se pueden obtener hasta una potencia 2.

Capítulo 14

Matriz asociada a una transformación lineal

En este capítulo mostraremos que las transformaciones lineales que actúan en espacios vectoriales de dimesión finita tienen asociada una matriz y veremos algunas propiedades de dicha matriz.

14.1. Matriz

Sea V un espacio vectorial de dimensión n sobre el campo K y sea W un espacio vectorial de dimensión m sobre el mismo campo y

$$T : V \longrightarrow W$$

una transformación lineal. Además, supongamos que

$$\{v_1, v_2, \cdots, v_n\} \tag{14.1}$$

es una base del espacio vectorial V y

$$\{w_1, w_2, \cdots, w_m\} \tag{14.2}$$

es una base del espacio vectorial W.

Observe que si $v \in V$, entonces $T(v) \in W$. Por lo que, debido a que el conjunto (14.2) es base de W, existen m escalares $\tilde{\lambda}_j \in K, j = 1, 2, \cdots, m$ tales que

$$\boxed{T(v) = \tilde{\lambda}_j w_j.} \tag{14.3}$$

Ahora, como $T(v_i) \in W, i = 1, 2, \cdots, n$, existen nm escalares en K tales que

$$\boxed{T(v_i) = \alpha_{ij} w_j.} \tag{14.4}$$

Además, como el conjunto (14.1) es base de V, existen n escalares $\lambda_i \in K$ tales que

$$v = \lambda_i v_i$$

y como T es una función lineal obtenemos

$$\begin{aligned} T(v) &= T(\lambda_i v_i) \\ &= \lambda_i T(v_i) \\ &= \lambda_i \alpha_{ij} w_j, \end{aligned}$$

es decir

$$T(v) = \lambda_i \alpha_{ij} w_j.$$

Igualando este último resultado con la expresión (14.3) encontramos

$$\lambda_i \alpha_{ij} w_j = \tilde{\lambda}_j w_j,$$

que induce la igualdad

$$\begin{aligned} 0 &= \lambda_i \alpha_{ij} w_j - \tilde{\lambda}_j w_j \\ &= \left(\lambda_i \alpha_{ij} - \tilde{\lambda}_j\right) w_j, \end{aligned}$$

esto es

$$\left(\lambda_i \alpha_{ij} - \tilde{\lambda}_j\right) w_j = 0.$$

Entonces, como el conjunto (14.2) es linealmente independiente llegamos a

$$\lambda_i \alpha_{ij} - \tilde{\lambda}_j = 0,$$

que implica

$$\begin{aligned} \tilde{\lambda}_j &= \lambda_i \alpha_{ij} \\ &= \alpha^T_{ji} \lambda_i, \end{aligned}$$

esto es

$$\tilde{\lambda}_j = \alpha^T_{ji} \lambda_i.$$

A la matriz

$$\boxed{M = \alpha^T}$$

se le llama matriz asociada a la transformación lineal T.

14.2. Ejemplo

Por ejemplo, consideremos la transformación lineal

$$T(x, y, z) = (3x - y + 5z, 5x - 6y + 7z, x - 7y + 8z), \tag{14.5}$$

ocupando la base canónica

$$(1, 0, 0), \quad (0, 1, 0) \quad (0, 0, 1) \tag{14.6}$$

se encuentra

$$\begin{aligned}
T(1, 0, 0) &= (3, 5, 1) = 3(1, 0, 0) + 5(0, 1, 0) + 1(0, 0, 1) \\
T(0, 1, 0) &= (-1, -6, -7) = -1(1, 0, 0) - 6(0, 1, 0) - 7(0, 0, 1), \\
T(0, 0, 1) &= (5, 7, 8) = 5(1, 0, 0) + 7(0, 1, 0) + 8(0, 0, 1),
\end{aligned}$$

de donde se obtiene

$$\alpha = \begin{pmatrix} 3 & 5 & 1 \\ -1 & -6 & -7 \\ 5 & 7 & 8 \end{pmatrix},$$

por lo tanto, usando la base canónica (14.6), la matriz asociada a la transformación lineal (14.5) es

$$M = \alpha^T = \begin{pmatrix} 3 & -1 & 5 \\ 5 & -6 & 7 \\ 1 & -7 & 8 \end{pmatrix}.$$

Note que

$$M = \alpha^T = \begin{pmatrix} 3 & -1 & 5 \\ 5 & -6 & 7 \\ 1 & -7 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x - y + 5 \\ 5x - 6y + 7z \\ x - 7y + 8z \end{pmatrix},$$

así la transformación lineal (14.5) se recupera como el producto de una matriz por un vector columna.

14.3. Matriz asociada a composición de funciones lineales

Sean V, W_1, W_2 espacios vectoriales de dimensión finita sobre un campo K y

$$\begin{aligned} T_1 &: V \longrightarrow W_1, \\ T_2 &: W_1 \longrightarrow W_2 \end{aligned}$$

funciones lineales. Supongamos que

$$\{v_1, v_2, \cdots, v_n\}$$

es base de V, que

$$\{w_1, w_2, \cdots, w_{m_1}\}$$

es base de W_1 y

$$\{w'_1, w'_2, \cdots, w'_{m_2}\}$$

es base de W_2. Además, supongamos que con dichas bases M_1 es la matriz asociada a T_1 y M_2 es la matriz asociada a T_2. Entonces, la matriz asociada a la composición de las funciones lineales

$$\boxed{T = (T_2 \circ T_1) : V \longrightarrow W_2}$$

es el producto de matrices

$$\boxed{M = M_2 M_1.}$$

Para demostrar esta afirmación, recordemos que debido a que los conjuntos $\{v_i\}(i = 1, 2, \cdots, n), \{w_j\}(j = 1, 2, \cdots, m_1), \{w'_k\}(k = 1, 2, \cdots, m_2)$ son bases de V, W_1, W_2, respectivamente, entonces si $v \in V$ existen n números λ_i tales que

$$v = \lambda_i v_i;$$

como $T_1(v_i) \in W_1$ existen nm_1 números α_{ij} tales que

$$T_1(v_i) = \alpha_{ij} w_j;$$

como $T_2(w_j) \in W_2$ existen $m_1 m_2$ números α'_{jk} tales que

$$T_2(w_j) = \alpha'_{jk} w'_k;$$

como $T(v) = (T_2 \circ T_1)(v) \in W_2$ existen nm_2 números $\tilde{\lambda}_k$ tales que

$$T(v) = \tilde{\lambda}_k w'_k. \tag{14.7}$$

Por lo tanto,

$$\begin{aligned}
T(v) &= (T_2 \circ T_1)(v) \\
&= T_2(T_1(v)) \\
&= T_2(T_1(\lambda_i v_i)) \\
&= T_2(\lambda_i T_1(v_i)) \\
&= T_2(\lambda_i \alpha_{ij} w_j) \\
&= \lambda_i \alpha_{ij} T_2(w_j) \\
&= \lambda_i \alpha_{ij} \alpha'_{jk} w'_k \\
&= \lambda_i (\alpha\alpha')_{ik} w'_k,
\end{aligned}$$

es decir

$$T(v) = \lambda_i (\alpha\alpha')_{ik} w'_k.$$

Igualando este resultado con la expresión (14.7), tenemos

$$T(v) = \lambda_i (\alpha\alpha')_{ik} w'_k = \tilde{\lambda}_k w'_k,$$

que implica

$$\lambda_i (\alpha\alpha')_{ik} w'_k - \tilde{\lambda}_k w'_k = (\lambda_i (\alpha\alpha')_{ik} - \tilde{\lambda}_k) w'_k = 0,$$

esto es

$$(\lambda_i (\alpha\alpha')_{ik} - \tilde{\lambda}_k) w'_k = 0.$$

Ahora, como el conjunto $\{w'_k\}$ es linealmente independiente, el último resultado induce la igualdad

$$\lambda_i (\alpha\alpha')_{ik} - \tilde{\lambda}_k = 0$$

de donde

$$\tilde{\lambda}_k = \lambda_i (\alpha\alpha')_{ik} = (\alpha\alpha')^T_{ki} \lambda_i,$$

así obtenemos

$$\tilde{\lambda}_k = (\alpha\alpha')^T_{ki} \lambda_i.$$

Por lo tanto, la matriz asociada a T es

$$\begin{aligned} M &= (\alpha\alpha')^T \\ &= \alpha'^T\alpha^T \\ &= M_2M_1, \end{aligned}$$

esto es,

$$M = M_2M_1,$$

que es lo que se quería demostrar.

14.4. Ejemplo

Por ejemplo, consideremos las transformaciones lineal

$$\begin{aligned} T_1(x,y,z) &= (x-y+z, x+y-z, 2x+y+z), \\ T_2(x,y,z) &= (-x+2y+z, -x-y+z, -x+y+z), \end{aligned}$$

por lo que

$$T(x,y,z) = (T_2 \circ T_1)(x,y,z) = (3x+4y-2z, x+z, 2x+4y-z).$$

Además, usando la base canónica

$$(1,0,0), \quad (0,1,0) \quad (0,0,1)$$

se obtienen

$$\begin{aligned} T_1(1,0,0) &= (1,1,2) = 1(1,0,1) + 1(0,1,0) + 2(0,0,1), \\ T_1(0,1,0) &= (-1,1,1) = -1(1,0,0) + 1(0,1,0) + 1(0,0,1), \\ T_1(0,0,1) &= (1,-1,1) = 1(1,0,0) - 1(0,1,0) + 1(0,0,1), \end{aligned}$$

de donde se obtiene la matriz

$$\alpha_1 = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix},$$

por lo cual la matriz asociada a T_1 es

$$M_1 = \alpha_1^T = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & 1 \end{pmatrix}.$$

También utilizando la base canónica se llega a

$$
\begin{aligned}
T_2(1,0,0) &= (-1,-1,-1) = -1(1,0,0) - 1(0,1,0) - 1(0,0,1),\\
T_2(0,1,0) &= (2,-1,1) = 2(1,0,0) - 1(0,1,0) + 1(0,0,1),\\
T_2(0,0,1) &= (1,1,1) = 1(1,0,0) + 1(0,1,0) + 1(0,0,1),
\end{aligned}
$$

de donde

$$\alpha_2 = \begin{pmatrix} -1 & -1 & -1 \\ 2 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

por lo que la matriz asociada a T_2 es

$$M_2 = \alpha_2^T = \begin{pmatrix} -1 & 2 & 1 \\ -1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Además, con la misma base canónica, se encuentra

$$
\begin{aligned}
T(1,0,0) &= (3,0,2) = 3(1,0,1) + 0(0,1,0) + 2(0,0,1),\\
T(0,1,0) &= (4,1,3) = 4(1,0,0) + 1(0,1,0) + 3(0,0,1),\\
T(0,0,1) &= (-2,1,-1) = -2(1,0,0) + 1(0,1,0) - 1(0,0,1),
\end{aligned}
$$

de donde

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

así la matriz asociada a T es

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

Se puede observar que se cumple $M = M_2 M_1$, es decir

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

14.5. Cambio de base y matriz asociada

Sea V un espacio vectorial de dimensión finita y

$$T : V \longrightarrow V$$

una transformación lineal, además sean

$$\{v_1, v_2, \cdots v_n\}, \qquad \{v'_1, v'_2, \cdots v'_n\}$$

dos bases diferentes de V. Supongamos que N es la matriz de cambio de base, M la matriz asociada a T con la base $\{v_i\}(i = 1, 2 \cdots , n)$ y M' la matriz asociada a T con la base $\{v'_i\}$. Entonces, se cumple

$$\boxed{M' = N^{-1}MN.}$$

Para mostrar esta afirmación, primero recordemos que debido a que $\{v_i\}$ y $\{v'_i\}$ son bases de V y N es la matriz de cambio de base se tiene

$$v'_i = N^T_{ij} v_j, \quad j = 1, 2, \cdots , n.$$

También recordemos que como M es la matriz asociada a T con la base $\{v_i\}$ y M' es la matriz asociada a T con la base $\{v'_i\}$ se cumple

$$\begin{aligned} T(v_i) &= M^T_{ij} v_j, \\ T(v'_i) &= M'^T_{ij} v'_j. \end{aligned}$$

Por lo tanto

$$\begin{aligned} T(v'_i) &= T\left(N^T_{ij} v_j\right) \\ &= N^T_{ij} T(v_j) \\ &= N^T_{ij} M^T_{jk} v_k \\ &= (MN)^T_{ik} v_k, \end{aligned}$$

además

$$\begin{aligned} T(v'_i) &= M'^T_{ij} v'_j \\ &= M'^T_{ij} N^T_{jk} v_k \\ &= (NM')^T_{ik} v_k, \end{aligned}$$

de donde

$$(MN)^T_{ik} v_k = (NM')^T_{ik} v_k,$$

es decir

$$\left[(MN)^T_{ik} - (NM')^T_{ik}\right] v_k = 0.$$

Como el conjunto de vectores $\{v'\}$ es una base de V, se encuentra

$$(MN)^T_{ik} = (NM')^T_{ik},$$

que implica

$$(MN)^T = (NM')^T.$$

Esta igualdad se puede escribir como

$$MN = NM',$$

así obtenemos

$$M' = N^{-1}MN,$$

que es lo que se quería demostrar.

14.6. Ejemplo

Por ejemplo, consideremos la transformación lineal

$$T(x, y, z) = (2x + y + z, -x - y + 3z, 2x - y + 3z),$$

con la base canónica

$$(1, 0, 0), \quad (0, 1, 0), \quad (0, 0, 1)$$

se obtienen

$$\begin{aligned}
T(1, 0, 0) &= (2, -1, 2) = 2(1, 0, 0) - 1(0, 1, 0) + 2(0, 0, 1),\\
T(0, 1, 0) &= (1, -1, -1) = 1(1, 0, 0) - 1(0, 1, 0) - 1(0, 0, 1),\\
T(0, 0, 1) &= (1, 3, 3) = 1(1, 0, 0) + 3(0, 1, 0) + 3(0, 0, 1),
\end{aligned}$$

por lo que se encuentra

$$\alpha = \begin{pmatrix} 2 & -1 & 2 \\ 1 & -1 & -1 \\ 1 & 3 & 3 \end{pmatrix},$$

entonces la matriz asociada con esta base es

$$M = \alpha^T = \begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 3 \\ 2 & -1 & 3 \end{pmatrix}.$$

Mientras que con la base

$$(1,1,0), \quad (1,-1,0), \quad (1,1,1)$$

se encuentra

$$\begin{aligned} T(1,1,0) &= (3,-2,1) = -\frac{1}{2}(1,1,0) + \frac{5}{2}(1,-1,0) + 1(1,1,1) \\ T(1,-1,0) &= (1,0,3) = -\frac{5}{2}(1,1,0) + \frac{1}{2}(1,-1,0) + 3(1,1,1), \\ T(1,1,1) &= (4,1,4) = -\frac{3}{2}(1,1,0) + \frac{3}{2}(1,-1,0) + 4(1,1,1), \end{aligned}$$

que implica la matriz

$$\alpha' = \begin{pmatrix} -\frac{1}{2} & \frac{5}{2} & 1 \\ -\frac{5}{2} & \frac{1}{2} & 3 \\ -\frac{3}{2} & \frac{3}{2} & 4 \end{pmatrix},$$

de donde la matriz asociada con esta base es

$$M' = \alpha^T = \frac{1}{2}\begin{pmatrix} -1 & -5 & -3 \\ 5 & 1 & 3 \\ 2 & 6 & 8 \end{pmatrix}.$$

Además podemos ver que

$$\begin{aligned} (1,1,0) &= 1(1,0,0) + 1(0,1,0) + 0(0,0,1), \\ (1,-1,0) &= 1(1,0,0) - 1(0,1,0) + 0(0,0,1), \\ (1,1,1) &= 1(1,0,0) + 1(0,1,0) + 1(0,0,1), \end{aligned}$$

que induce la matriz

$$\beta = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix},$$

así la matriz de cambio de base y su inversa son

$$N = \beta^T = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad N^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -2 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Con estas matrices se puede ver que se satisface $M' = NMN^{-1}$, es decir

$$\frac{1}{2}\begin{pmatrix} -1 & -5 & -3 \\ 5 & 1 & 3 \\ 2 & 6 & 8 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -2 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 3 \\ 2 & -1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

14.7. Ejercicios

1.- Con la base canónica, obtener la matriz asociada a la transformación lineal

$$T(x, y, z) = (2x + 3y - z, x + y + z, x - 2y + 3z).$$

2.- Suponiendo que se usa la base canónica, encontrar la transformación lineal está asociada a la matriz

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

3.- Usando la base canónica y la matriz asociada, obtener la función inversa de transformación lineal

$$T(x, y, z) = (3x - y + z, x + y + z, y - z).$$

Capítulo 15

Producto escalar

En este capítulo veremos la definición de producto escalar en un espacio vectorial. Además veremos ejemplos de espacios vectoriales con producto escalar.

El concepto de producto escalar para vectores en $\mathbb{R}^n$ se encuentra en los trabajos fundamentales de Hermann Gunther Grassmann, que datan de 1884. En la actualidad este concepto importante para entender temas de diversas áreas como geometría diferencial, mecánica cuántica, análisis funcional, entre otras.

15.1. Definición

Una operación importante entre vectores es el **producto escalar**. Este producto manda dos vectores a un número complejo

$$\boxed{\langle | \rangle : \mathbf{V} \times \mathbf{V} \to \mathbb{C}}$$

y debe satisfacer los axiomas:

$$\begin{aligned}
\cdot) &\quad \forall v \in \mathbf{V} \qquad \langle v|v\rangle \geq 0, \qquad \langle v|v\rangle = 0 \Longleftrightarrow v = 0, && (15.1)\\
\cdot\cdot) &\quad \forall v, u, w \in \mathbf{V} \qquad \langle v + u|w\rangle = \langle v|w\rangle + \langle u|w\rangle, && (15.2)\\
\cdots) &\quad \forall v, u \in \mathbf{V}, \lambda \in \mathbb{C} \qquad \langle v|\lambda u\rangle = \lambda \langle v|u\rangle, && (15.3)\\
\cdot\cdot\cdot\cdot) &\quad \forall v, u \in \mathbf{V}, \qquad \langle v|u\rangle = (\langle u|v\rangle)^{*}. && (15.4)
\end{aligned}$$

Existen diferentes implicaciones de estos axiomas. Por ejemplo, para cualquier vector v se cumple

$$\boxed{\langle v|0\rangle = 0.}$$

En efecto sabemos que $v - v = 0$, entonces

$$\begin{aligned} \langle v|0\rangle &= \langle v|v - v\rangle \\ &= \langle v|v| - \langle v|v\rangle = 0, \end{aligned}$$

que es lo que se quería demostrar

Otra implicación es que si λ es un número complejo y v_1, v_2 dos vectores, entonces se cumple

$$\boxed{\langle \lambda v_1|v_2\rangle = \lambda^* \langle v_1|v_2\rangle .}$$

Esta igualdad es correcta, pues considerando que los número complejos cumplen que

$$(z_1 z_2)^* = z_1^* z_2^*, \quad (z^*)^* = z$$

y los axiomas (15.3)-(15.4) se encuentra

$$\begin{aligned} \langle \lambda v_1|v_2\rangle &= (\langle \lambda v_2|v_1\rangle)^* \\ &= (\lambda \langle v_2|v_1\rangle)^* \\ &= \lambda^* (\langle v_2|v_1\rangle)^* \\ &= \lambda^* ((\langle v_1|v_2\rangle)^*)^* \\ &= \lambda^* \langle v_1|v_2\rangle . \end{aligned}$$

Antes de ver otras propiedades del producto escalar, veremos algunos ejemplos de ellos.

15.2. Ejemplos de producto escalar

15.2.1. Producto escalar en $\mathbb{C}^n$

Si tenemos dos vectores en $\mathbb{C}^n$,

$$v = (v_1, v_2, \cdots, v_n) \quad \text{y} \quad w = (w_1, w_2, \cdots, w_n),$$

el producto escalar se define como

$$\boxed{\langle v|w\rangle = v_1^* w_1 + v_2^* w_2 + \cdots + v_n^* w_n = \sum_{i=1}^{n} v_i^* w_i.}$$

Note que a los vectores v y w se les puede asignar las matrices columna

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \qquad w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix},$$

por lo que

$$\boxed{\begin{aligned} v^{*T} w &= \begin{pmatrix} v_1^* & v_2^* & \cdots & v_n^* \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \\ &= v_1^* w_1 + v_2^* w_2 + \cdots + v_n^* w_n \\ &= \sum_{i=1}^{n} v_i^* w_i \\ &= \langle v|w\rangle , \end{aligned}}$$

así el producto escalar en el espacio $\mathbb{C}^n$ se puede escribir como

$$\boxed{\langle v|w\rangle = v^{*T} w.}$$

Ahora podemos ver que

$$\begin{aligned} \langle v|v\rangle &= v^{*T} v \\ &= v_1^* v_1 + v_2^* v_2 + \cdots + v_n^* v_n \\ &= |v_1|^2 + |v_2|^2 + \cdots + |v_n|^2 \geq 0, \end{aligned}$$

además si

$$|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2 = 0,$$

entonces

$$v_1 = 0, v_2 = 0, \cdots, v_n = 0,$$

por lo que

$$v = (0, 0, \cdots, 0).$$

Así esta definción de producto escalar cumple el axioma (15.1).

Además, tenemos que

$$\begin{aligned}
\langle v|(w_1 + w_2)\rangle &= v^{*T}(w_1 + w_2) \\
&= v^{*T}w_1 + v^{*T}w_2 \\
&= \langle v|w_1\rangle + \langle v|w_2\rangle ,
\end{aligned}$$

entonces esta definición de producto escalar respeta la suma de vectores y se cumple el axioma (15.2).

También se puede ver que si λ es un número complejo se tiene que

$$\begin{aligned}
\langle v|\lambda w\rangle &= v^{*T}\lambda w \\
&= \lambda v^{*T}w \\
&= \lambda \langle v|w\rangle ,
\end{aligned}$$

en este caso podemos concluir este producto escalar cumple el axioma (15.3).

Ahora, recordemos que el complejo conjugado respeta la suma y que cualquier número complejo satisface $(z^*)^* = z$, entonces se puede observar que

$$\begin{aligned}
(\langle v|w\rangle)^* &= (v_1^*w_1 + v_2^*w_2 + \cdots + v_n^*w_n)^* \\
&= (v_1^*)^* w_1^* + (v_2^*)^* w_2^* + \cdots + (v_n^*)^* w_n^* \\
&= v_1w_1^* + v_2w_2^* + \cdots + v_nw_n^* \\
&= w_1^*v_1 + w_2^*v_2 + \cdots + w_n^*v_n \\
&= w^{*T}v \\
&= \langle w|v\rangle,
\end{aligned}$$

es decir,

$$\langle w|v\rangle = \langle v|w\rangle^*$$

de donde se cumple el axioma (15.4).

15.2.2. Funciones

Si $q(x)$ es una función real, continua y positiva en el intervalo (a, b), para el espacio vectorial de las funciones continuas, $\{f\}$, que va del intervalo $[a, b]$ a los complejos, tales que

$$\boxed{\int_a^b dxq(x)f^*(x)f(x) < \infty,}$$

se puede definir el producto escalar como

$$\boxed{\langle f|g\rangle = \int_a^b dxq(x)f^*(x)g(x).} \tag{15.5}$$

Considerando las propiedades de $q(x)$ y $f(x)$, el primer axioma de producto escalar se cumple pues

$$\begin{aligned} \langle f|f\rangle &= \int_a^b dxq(x)f^*(x)f(x) \\ &= \int_a^b dxq(x)|f(x)|^2 \geq 0. \end{aligned}$$

Además, como f es una función continua, se tiene

$$\langle f|f\rangle = \int_a^b dxq(x)|f(x)|^2 = 0 \Longleftrightarrow f(x) = 0.$$

El segundo axioma de producto escalar se cumple, en efecto

$$\begin{aligned} \langle f|g_1 + g_2\rangle &= \int_a^b dxq(x)f^*(x)\left(g_1(x) + g_2(x)\right) \\ &= \int_a^b dx\left(q(x)f^*(x)g_1(x) + q(x)f^*(x)g_2(x)\right) \\ &= \int_a^b dxq(x)f^*(x)g_1(x) + \int_a^b dxq(x)f^*(x)g_2(x) \\ &= \langle f|g_1\rangle + \langle f|g_2\rangle . \end{aligned}$$

Los axiomas restantes también se cumplen, para probar esta afirmación supongamos que λ es un número complejo, entonces

$$\begin{aligned} \langle f|g\rangle &= \int_a^b dxq(x)f^*(x)\lambda g(x) \\ &= \lambda\int_a^b dxq(x)f^*(x)g(x) = \lambda\langle f|g\rangle\,, \end{aligned}$$

además

$$\begin{aligned} \langle f|g\rangle^* &= \left(\int_a^b dxq(x)f^*(x)g(x)\right)^* \\ &= \int_a^b dxq(x)\left(f^*(x)\right)^* g^*(x) \\ &= \int_a^b dxq(x)f(x)g^*(x) \\ &= \int_a^b dxq(x)g^*(x)f(x) \\ &= \langle g|f\rangle\,. \end{aligned}$$

Por lo tanto, la ecuaciíon (15.5) es un producto escalar para el espacio vectorial de las funciones.

15.2.3. Matrices

Supongamos que M y N pertenecen al espacio vectorial de las matrices de entradas complejas de $n \times n$, definiremos el producto escalar entre estas matrices como

$$\boxed{\langle M|N\rangle = Tr\left(M^{*T}N\right).} \tag{15.6}$$

El primer axioma se cumple, pues

$$
\begin{aligned}
\langle M|M\rangle &= Tr\left(M^{*T}M\right)\\
&= \sum_{i=i}^{n}\left(M^{*T}M\right)_{ii}\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n}\left(M^{*T}\right)_{ik}M_{ki}\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n}M^{*}_{ki}M_{ki}\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n}|M_{ik}|^{2}\geq 0,
\end{aligned}
$$

esto es

$$\langle M|M\rangle = \sum_{i=1}^{n}\sum_{k=1}^{n}|M_{ik}|^{2}\geq 0.$$

De donde, si $\langle M|M\rangle = 0$, entonces $M_{ik} = 0$, es decir $M = 0$.

Además, si N_1 y N_2 son matrices de $n\times n$,

$$
\begin{aligned}
\langle M|N_1+N_2\rangle &= Tr\left(M^{*T}\left(N_1+N_2\right)\right)\\
&= Tr\left(M^{*T}N_1+M^{*T}N_2\right)\\
&= Tr\left(M^{*T}N_1\right)+Tr\left(M^{*T}N_2\right)\\
&= \langle M|N_1\rangle+\langle M|N_2\rangle\,.
\end{aligned}
$$

Por lo que se cumple el segundo axioma de producto escalar.

También se puede observar que si λ es un número complejo, entonces

$$
\begin{aligned}
\langle M|\lambda N\rangle &= Tr\left(M^{*T}\lambda N\right)\\
&= \sum_{i=i}^{n}\left(M^{*T}\lambda N\right)_{ii}\\
&= \lambda\sum_{i=i}^{n}\left(M^{*T}N\right)_{ii}\\
&= \lambda Tr\left(M^{*T}N\right)\\
&= \lambda\left\langle M|N\right\rangle.
\end{aligned}
$$

Adicionalmente, se encuentra

$$\begin{aligned}\langle M|N\rangle &= Tr\left(M^{*T}N\right)\\ &= Tr\left(\left(M^{*T}N\right)^T\right)\\ &= Tr\left(N^TM^*\right)\\ &= \left(\left[Tr\left(N^TM^*\right)\right]^*\right)^*\\ &= \left(Tr\left(N^{*T}M\right)\right)^*\\ &= \left(\langle N|M\rangle\right)^*.\end{aligned}$$

Por lo tanto, (15.6) es un producto escalar para el espacio vectorial de las matrices de $n \times n$.

Por ejemplo, consideremos las matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

que tienen diferentes aplicaciones en mecánica cuántica. Suponiendo que I es matriz identidad, se puede probar que estas matrices satisfacen

$$\begin{aligned}\langle I|\sigma_i\rangle &= Tr(\sigma_i) = 0,\\ \langle \sigma_i|\sigma_j\rangle &= Tr\left((\sigma_i)^{*T}\sigma_j\right) = Tr(\sigma_i\sigma_j) = 2\delta_{ij}.\end{aligned}$$

15.2.4. Sucesiones

Si tenemos dos sucesiones $s_1 = \{a_n\}_{n=0}^{\infty}$ y $s_2 = \{b_n\}_{n=0}^{\infty}$ donde $a_n, b_n \in \mathbb{C}$ y $\sum_{n\geq 0}|a_n|^2 < \infty, \sum_{n\geq 0}|b_n|^2 < \infty$, se puede definir el producto escalar como

$$\langle s_1|s_2\rangle = \sum_{n\geq 0} a_n^* b_n.$$

Observe que éste es una generalización del producto escalar entre vectores.

15.3. Vectores ortogonales

Se dice que los vectores v_1 y v_2 son ortogonales si

$$\boxed{\langle v_1 | v_2 \rangle = 0.}$$

Por ejemplo, los vectores

$$v_1 = (1, -1, 1), \qquad v_2 = (2, 2, 0)$$

son ortogonales, pues

$$\begin{aligned} \langle v_1 | v_2 \rangle &= (1, -1, 1) \cdot (2, 2, 0) \\ &= 1 \cdot 2 + (-1) \cdot 2 + 1 \cdot 0 \\ &= 2 - 2 + 0 = 0. \end{aligned}$$

De la misma forma, tomando $q(x) = 1$ y el intervalo $[-1, 1]$, la funciones

$$f(x) = x, \qquad g(x) = x^2$$

son ortogonales, pues

$$\begin{aligned} \langle f(x) | g(x) \rangle &= \int_{-1}^{1} x \cdot x^2 \\ &= \int_{-1}^{1} x^3 \\ &= \left. \frac{x^4}{4} \right|_{-1}^{1} = 0. \end{aligned}$$

Además, las matrices

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

son ortogonales, pues

$$\begin{aligned}
\langle M_1|M_2\rangle &= Tr\left(M_1^T M_2\right)\\
&= Tr\left(\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}^T \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}\right)\\
&= Tr\left(\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}\right)\\
&= Tr\begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}\\
&= 1-1=0.
\end{aligned}$$

15.4. Vectores normales o unitarios

Se dice que un vector v es normal o unitario si cumple que

$$\boxed{\langle v|v\rangle = 1.}$$

Por ejemplo, el vector

$$v=\frac{1}{\sqrt{6}}(1,-1,2)$$

es normal, pues

$$\begin{aligned}
\langle v|v\rangle &= \frac{1}{\sqrt{6}}(1,-1,2)\cdot\frac{1}{\sqrt{6}}(1,-1,2)\\
&= \frac{1\cdot 1+(-1)(-1)+2\cdot 2}{6}\\
&= \frac{6}{6}=1.
\end{aligned}$$

De la misma forma, tomando $q(x)=1$ y el intervalo $[-1,1]$, la función

$$f(x)=\sqrt{\frac{3}{2}}x,$$

es normal, pues

$$\begin{aligned}
\langle f(x)|f(x)\rangle &= \int_{-1}^{1}\sqrt{\frac{3}{2}}x\cdot\sqrt{\frac{3}{2}}x \\
&= \frac{3}{2}\int_{-1}^{1}x^2 \\
&= \frac{3}{2}\frac{x^3}{3}\bigg|_{-1}^{1} = 1.
\end{aligned}$$

Además, la matriz

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

cumple que

$$\begin{aligned}
\langle M|M\rangle &= Tr\left(M^T M\right) \\
&= Tr\left(\frac{1}{\sqrt{6}}\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}^T \frac{1}{\sqrt{6}}\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}\right) \\
&= \frac{1}{6}Tr\left(\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}\right) \\
&= \frac{1}{6}Tr\begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \\
&= \frac{4+2}{6} = 1.
\end{aligned}$$

Así, M es una matriz normal.

En general, si v es un vector diferente de cero, siempre podemos construir un vector unitario. De hecho, en este caso se cumple que $\langle v|v\rangle > 0$ y se puede definir el vector

$$\boxed{u = \frac{v}{\sqrt{\langle v|v\rangle}},}$$

el cual es unitario. En efecto, se satisface

$$\begin{aligned}
\langle u|u\rangle &= \langle \frac{v}{\sqrt{\langle v|v\rangle}} | \frac{v}{\sqrt{\langle v|v\rangle}} \rangle \\
&= \left(\frac{1}{\sqrt{\langle v|v\rangle}} \right)^2 \langle v|v\rangle \\
&= \frac{1}{\langle v|v\rangle} \langle v|v\rangle = 1,
\end{aligned}$$

es decir

$$\langle u|u\rangle = 1,$$

por lo que u es un vector unitario.

15.5. Ortonormalidad e independencia lineal

Se dice que un conjunto de vectores,

$$\{v_i\}_{i=1}^n = \{v_1, v_2, \cdots, v_n\},$$

es ortonormal si

$$\boxed{\langle v_i|v_j\rangle = \delta_{ij}, \qquad i, j = 1, 2, \cdots, n.}$$

En otras palabras, un conjunto de vectores, $\{v_i\}_{i=1}^n$, es ortonormal si cada vector es unitario y es ortogonal al resto de los vectores del conjunto.

Si un conjunto de vectores es ortonomal, entonces es linealmente independiente. Para probar esta afirmación, recordemos que un conjunto de vectores $\{w_i\}_{i=1}^n$ es linealmente independiente si cualquier combinación lineal de la forma

$$0 = \sum_{i=1}^n b_i w_i,$$

implica $b_i = 0$.

Ahora, supongamos que

$$\{a_1, a_2, \cdots, a_n\}$$

son un conjunto de escalares tales que

$$0 = \sum_{i=1}^{n} a_i v_i.$$

Entonces, como los vectores del conjunto $\{v_i\}_{i=1}^{n}$ son ortonormales, si el vector v_j es un vector de ese conjunto se encuentra que

$$\begin{aligned} 0 &= \langle v_j | 0 \rangle = \left\langle v_j \middle| \sum_{i=1}^{n} a_i v_i \right\rangle \\ &= \sum_{i=1}^{n} \langle v_j | a_i v_i \rangle \\ &= \sum_{i=1}^{n} a_i \langle v_j | v_i \rangle \\ &= \sum_{i=1}^{n} a_i \delta_{ij} = a_j, \end{aligned}$$

esto es $a_j = 0$, lo que completa la prueba.

Un conjunto de vectores ortonormal, $\{v_i\}_{i=1}^{n}$, tiene varias propiedades interesantes. Por ejemplo, si v es una combinación lineal de estos vectores, es decir, si

$$\boxed{v = \sum_{i=1}^{n} a_i v_i,}$$

se cumple

$$\boxed{\langle v | v \rangle = \sum_{i=1}^{n} |a_i|^2.}$$

En efecto,

$$
\begin{aligned}
\langle v|v\rangle &= \langle\sum_{i=1}^{n} a_i v_i \Big| \sum_{j=1}^{n} a_j v_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \langle a_i v_i | a_j v_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i^* a_j \langle v_i | v_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i^* a_j \delta_{ij} \\
&= \sum_{i=1}^{n} |a_i|^2,
\end{aligned}
$$

que es lo que queríamos demostrar.

15.6. Ejemplos de conjuntos de vectores ortonormales

En esta sección veremos diferentes conjuntos de vectores ortonormales.

15.6.1. Exponencial compleja

Sea el conjunto de funciones

$$\boxed{\{\Phi_n(\varphi)\} = \left\{\frac{e^{in\varphi}}{\sqrt{2\pi}}\right\}, \quad n \in \mathbb{Z}} \tag{15.7}$$

definidas en el intervalo $[0, 2\pi]$. Este conjunto de funciones es ortonormal con el producto escalar

$$\langle f|g\rangle = \int_0^{2\pi} dx f^*(x) g(x).$$

De hecho

$$\begin{aligned}
\langle \Phi_n(\varphi)|\Phi_m(\varphi)\rangle &= \int_0^{2\pi} d\varphi \left(\Phi_n(\varphi)\right)^* \Phi_m(\varphi) \\
&= \frac{1}{2\pi}\int_0^{2\pi} d\varphi e^{i(m-n)\varphi}.
\end{aligned}$$

Si $m = n$, es claro que

$$\frac{1}{2\pi}\int_0^{2\pi} d\varphi e^{i(n-n)\varphi} = \frac{1}{2\pi}\int_0^{2\pi} d\varphi = 1.$$

Adicionalmente, si $n \neq m$ se tiene

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} d\varphi e^{i(m-n)\varphi} &= \frac{1}{2\pi}\frac{1}{i(n-m)} e^{i(m-n)\varphi}\Big|_0^{2\pi} \\
&= \frac{1}{2\pi}\frac{1}{i(m-n)}\left((-1)^{2(m-n)} - 1\right) = 0.
\end{aligned}$$

Por lo tanto,

$$\boxed{\langle \Phi_n(\varphi)|\Phi_m(\varphi)\rangle = \delta_{mn},} \tag{15.8}$$

es decir, el conjunto (15.7) es ortonormal y en consecuencia linealmente independiente.

15.6.2. Funciones trigonométricas

Consideremos el conjunto de funciones trigonométricas

$$\boxed{\{\psi(x)\} = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}\right\}, \quad n \in \mathbb{N}} \tag{15.9}$$

Este conjunto forman un conjunto de funciones ortonormales en el intervalo $[-\pi, \pi]$ con el producto escalar

$$\langle f|g\rangle = \int_{-\pi}^{\pi} dx f(x) g(x). \tag{15.10}$$

Para mostrar esta afirmación ocuparemos las identidades

$$\begin{aligned}
\cos(\alpha)\cos(\beta) &= \frac{1}{2}\left[\cos(\alpha-\beta)+\cos(\alpha+\beta)\right], && (15.11)\\
\sin(\alpha)\sin(\beta) &= \frac{1}{2}\left[\cos(\alpha-\beta)-\cos(\alpha+\beta)\right], && (15.12)\\
\sin(\alpha)\cos(\beta) &= \frac{1}{2}\left[\sin(\alpha+\beta)+\sin(\alpha-\beta)\right]. && (15.13)
\end{aligned}$$

Ahora, obsérvese que se cumple

$$\begin{aligned}
\langle\frac{1}{\sqrt{2\pi}}\Big|\frac{1}{\sqrt{2\pi}}\rangle &= \int_{-\pi}^{\pi}dx\frac{1}{2\pi}=1,\\
\langle\frac{1}{\sqrt{2\pi}}\Big|\frac{\sin nx}{\sqrt{\pi}}\rangle &= \int_{-\pi}^{\pi}dx\frac{1}{\sqrt{2\pi}}\frac{\sin nx}{\sqrt{\pi}}=0,\\
\langle\frac{1}{\sqrt{2\pi}}\Big|\frac{\cos nx}{\sqrt{\pi}}\rangle &= \int_{-\pi}^{\pi}dx\frac{1}{\sqrt{2\pi}}\frac{\cos nx}{\sqrt{\pi}}=0.
\end{aligned}$$

Esto nos indica que la función constante

$$\frac{1}{\sqrt{2\pi}}$$

es ortogonal a las funciones

$$\{\cos nx, \sin nx\}.$$

Ahora, definamos $l = n - m$ y $k = n + m$, entonces considerando las identidades (15.11)-(15.13) se llega a

$$\begin{aligned}
\langle\frac{\sin mx}{\sqrt{\pi}}\Big|\frac{\cos nx}{\sqrt{\pi}}\rangle &= \int_{-\pi}^{\pi}dx\frac{\sin mx}{\sqrt{\pi}}\frac{\cos nx}{\sqrt{\pi}}\\
&= \frac{1}{\pi}\int_{-\pi}^{\pi}dx\sin mx\cos nx\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}dx\Big(\sin kx+\sin lx\Big)=0.
\end{aligned}$$

Para el caso en que $n \neq m$, usando de nuevo las identidades (15.11)-(15.13) se encuentra

$$\begin{aligned}
\langle \frac{\cos nx}{\sqrt{\pi}} \Big| \frac{\cos mx}{\sqrt{\pi}} \rangle &= \int_{-\pi}^{\pi} dx \frac{\cos nx}{\sqrt{\pi}} \frac{\cos mx}{\sqrt{\pi}} \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} dx \cos mx \cos nx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \Big(\cos(n-m)x + \cos(n+m)x \Big) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \Big(\cos lx + \cos kx \Big) = 0, \\
\langle \frac{\sin nx}{\sqrt{\pi}} \Big| \frac{\sin mx}{\sqrt{\pi}} \rangle &= \int_{-\pi}^{\pi} dx \frac{\sin nx}{\sqrt{\pi}} \frac{\sin mx}{\sqrt{\pi}} \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} dx \sin mx \sin nx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \Big(\cos(n-m)x - \cos(n+m)x \Big) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \Big(\cos lx - \cos kx \Big) = 0.
\end{aligned}$$

Además, se cumplen

$$\begin{aligned}
\langle \frac{\cos nx}{\sqrt{\pi}} \Big| \frac{\cos nx}{\sqrt{\pi}} \rangle &= \int_{-\pi}^{\pi} dx \frac{\cos nx}{\sqrt{\pi}} \frac{\cos nx}{\sqrt{\pi}} \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} dx \cos nx \cos nx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \Big(\cos(n-n)x + \cos(2n)x \Big) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx + \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \cos(2n)x = 1, \\
\langle \frac{\sin nx}{\sqrt{\pi}} \Big| \frac{\sin nx}{\sqrt{\pi}} \rangle &= \int_{-\pi}^{\pi} dx \frac{\sin nx}{\sqrt{\pi}} \frac{\sin nx}{\sqrt{\pi}} \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} dx \sin nx \sin nx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \Big(\cos(n-n)x - \cos(2n)x \Big) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx - \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \cos(2n)x = 1.
\end{aligned}$$

En resumen tenemos

$$\boxed{\begin{aligned}
\left\langle \frac{\sin mx}{\sqrt{\pi}} \middle| \frac{\sin nx}{\sqrt{\pi}} \right\rangle =& \delta_{nm}, && (15.14)\\
\left\langle \frac{\cos mx}{\sqrt{\pi}} \middle| \frac{\cos nx}{\sqrt{\pi}} \right\rangle =& \delta_{nm}, && (15.15)\\
\left\langle \frac{1}{\sqrt{2\pi}} \middle| \frac{1}{\sqrt{2\pi}} \right\rangle =& 1, && (15.16)\\
\left\langle \frac{\sin mx}{\sqrt{\pi}} \middle| \frac{\cos nx}{\sqrt{\pi}} \right\rangle =& 0, && (15.17)\\
\left\langle \frac{1}{\sqrt{2\pi}} \middle| \frac{\cos nx}{\sqrt{\pi}} \right\rangle =& 0, && (15.18)\\
\left\langle \frac{1}{\sqrt{2\pi}} \middle| \frac{\sin nx}{\sqrt{\pi}} \right\rangle =& 0. && (15.19)
\end{aligned}}$$

Por lo tanto, el conjunto de funciones (15.9) es un conjunto de funciones ortonormales en el intervalo $[-\pi, \pi]$ con el producto escalar (15.10).

15.6.3. Matrices de Dirac

Como otro ejemplo, consideremos las llamadas matrices de Dirac

$$\boxed{\begin{array}{cc}
\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, & \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \\
\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, & \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},
\end{array}}$$

que se usan en mecánica cuántica relativista. Se puede mostrar que las matrices de Dirac cumplen

$$\begin{aligned}
\langle I|\gamma^\mu\rangle &= Tr\left(\gamma^\mu\right) = 0, \\
\langle \gamma^\mu|\gamma^\nu\rangle &= Tr\left(\left(\gamma^\mu\right)^{*T} \gamma^\nu\right) = 4\delta^{\mu\nu}.
\end{aligned}$$

Así, el conjunto de matrices

$$\left\{\frac{1}{2}I, \frac{1}{2}\gamma^\mu\right\}, \qquad \mu = 0, 1, 2, 3,$$

forman un conjunto de matrices ortonormales.

15.7. Teorema de Pitágoras

Supongamos que v es un vector y $\{v_i\}_{i=1}^n$ es un conjunto de vectores ortonormales, mostraremos que v se puede descomponer como la suma de dos vectores ortogonales w_1 y w_2 y además que se cumple

$$\boxed{\langle v|v\rangle = \langle w_1|w_1\rangle + \langle w_2|w_2\rangle .}$$

En la figura (15.1) se muestra una representación geométrica de este resultado.

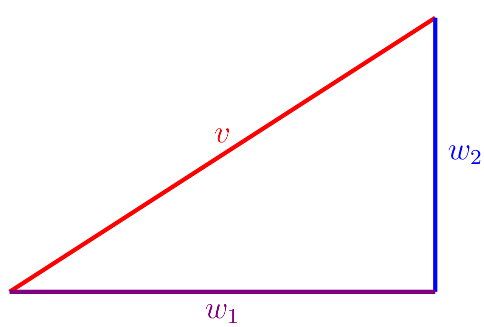

Figura 15.1: Representación geométrica de la descomposición de un vector v en dos vectores w_1, w_2 ortogonales.

Para demostrar esta afirmación primero notemos que se pueden formar los vectores

$$w_1 = \sum_{i=1}^{n} \langle v_i|v\rangle\, v_i$$

y

$$w_2 = v - \sum_{j=1}^{n} \langle v_j|v\rangle\, v_j,$$

los cuales satisfacen

$$\boxed{v = w_1 + w_2.}$$

Así, hemos descompuesto al vector v en los vectores w_1 y w_2.

Ahora mostraremos que los vectores w_1 y w_2 son ortogonales, para probar esta afirmación observemos que

$$\begin{aligned}
\langle w_1|w_2\rangle &= \left\langle \left(\sum_{i=1}^{n} \langle v_i|v\rangle\, v_i \right) \middle| \left(v - \sum_{j=1}^{n} \langle v_j|v\rangle\, v_j \right) \right\rangle \\
&= \left\langle \left(\sum_{i=1}^{n} \langle v_i|v\rangle\, v_i \right) \middle| v \right\rangle \\
&\quad - \left\langle \left(\sum_{i=1}^{n} \langle v_i|v\rangle\, v_i \right) \middle| \left(\sum_{j=1}^{n} \langle v_j|v\rangle\, v_j \right) \right\rangle \\
&= \sum_{i=1}^{n} \langle v_i \left(\langle v_i|v\rangle \right) |v\rangle \\
&\quad - \sum_{i=1}^{n} \sum_{j=1}^{n} \langle (\langle v_i|v\rangle\, v_i) \,|\, (\langle v_j|v\rangle\, v_j) \rangle \\
&= \sum_{i=1}^{n} \langle v_i|v\rangle^* \langle v_i|v\rangle - \sum_{i=1}^{n} \sum_{j=1}^{n} \langle v_i|v\rangle^* \langle v_j|v\rangle \langle v_i|v_j\rangle \\
&= \sum_{i=1}^{n} \langle v_i|v\rangle^* \langle v_i|v\rangle - \sum_{i=1}^{n} \sum_{j=1}^{n} \langle v_i|v\rangle^* \langle v_j|v\rangle\, \delta_{ij} \\
&= \sum_{i=1}^{n} \langle v_i|v\rangle^* \langle v_i|v\rangle - \sum_{i=1}^{n} \langle v_i|v\rangle^* \langle v_i|v\rangle = 0,
\end{aligned}$$

es decir

$$\boxed{\langle w_1|w_2\rangle = 0}$$

y los vectores w_1 y w_2 son ortogonales.

Además, recordando que $\{v_i\}_{i=1}^{n}$ es un conjunto de vectores ortonormales se cumple

$$\langle w_1|w_1\rangle = \sum_{i=1}^{n} |\, \langle v_i|v\rangle \,|^2.$$

Adicionalmente, considerando que w_1 y w_2 son vectores ortonormales, se puede observar que

$$\begin{aligned}
\langle v \big| v \rangle &= \Big\langle w_1 + w_2 \Big| w_1 + w_2 \Big\rangle \\
&= \Big\langle w_1 \Big| w_1 \Big\rangle + \Big\langle w_1 \Big| w_2 \Big\rangle + \Big\langle w_2 \Big| w_1 \Big\rangle + \Big\langle w_2 \Big| w_2 \Big\rangle \\
&= \langle w_1 | w_1 \rangle + \langle w_2 | w_2 \rangle ,
\end{aligned}$$

que es lo que queríamos demostrar.

Tomando en cuenta los resultados anteriores, claramente se cumple que si $\{v_n\}$ es un conjunto de vectores ortonormales, para cualquier vector v se tiene que

$$\boxed{\langle v|v\rangle = \sum_{i=1}^{n} |\,\langle v_i|v\rangle\,|^2 + \left\langle \left(v - \sum_{i=1}^{n} \langle v_i|v\rangle\, v_i \right) \middle| \left(v - \sum_{i=1}^{n} \langle v_i|v\rangle\, v_i \right) \right\rangle .}$$

Esta igualdad tiene diversas implicaciones importantes, algunas de ellas las veremos en las siguientes secciones.

15.7.1. Desigualdad de Bessel

Note que el teorema de Pitágoras implica la desigualdad

$$\sum_{i=1}^{n} |\,\langle v_i|v\rangle\,|^2 \leq \langle v|v\rangle\,, \tag{15.20}$$

que es conocida como desigualdad de Bessel.

$\boxed{\text{Friedrich Bessel (1784-1846)}}$ nació en Alemania y trabajó en diversas disciplinas como astronomía, física y matemáticas. Las aportaciones más importantes de Bessel se dieron en el campo de la astronomía, incluso fue director del observatorio de Königsberg y es considerado el mejor astrónomo de su época.

Una de sus contribuciones más relevantes en matemáticas son funciones que llevan su nombrel y que tienen aplicaciones en astronomía, electrodinámica, finanzas, etc. Además, la desigualdad que tiene su nombre es fundamental en análisis matemático y análisis funcional. Fue hijo de una sirvienta alemana. Por problemas económicos, de adolescente dejó la escuela para trabajar. En sus trabajos adquirió diferentes habilidades matemáticas. Primero trabajó con un comerciante, después en una compañía naval y posteriormente fue contratado en un observatorio en donde mostró su gran potencial para la astronomía y las matemáticas. No tenía una eduación universitaria, pero recibió un grado de Doctor honorario por parte de la Universidad de Göttingen, por recomendación de Gauss. Como profesor realizó contribuciones relevantes para reformar la educación universitaria de su época.

15.7.2. Desigualdad de Schwarz

Sea w un vector diferente de cero, claramente el conjunto formado por

$$\left\{\frac{w}{\sqrt{\langle w|w\rangle}}\right\}$$

es ortonormal. Entonces, de acuerdo con la desigualdad de Bessel (15.20), para cualquier vector v se cumple

$$\left|\langle \frac{w}{\sqrt{\langle w|w\rangle}}\middle|v\rangle\right|^2 \leq \langle v|v\rangle \quad \Longrightarrow \quad |\langle w|v\rangle|^2 \leq \langle w|w\rangle \langle v|v\rangle$$

que implica

$$\boxed{|\langle w|v\rangle| \leq \sqrt{\langle w|w\rangle}\sqrt{\langle v|v\rangle}.} \tag{15.21}$$

Esta eta desigualdad se atribuye H. Schwarz.

$\boxed{\text{Hermann Schwarz (1843-1921)}}$ fue un matemático y químico alemán que realizó contribuciones notables en áreas como análisis complejo, geometría diferencial y cálculo de variaciones. Realizó su tesis doctoral bajo la dirección de Karl Weierstrass en la Universidad de Berlín. Sus trabajos influenciaron a Emile Picard para mostrar la existencia de soluciones de ecuaciones diferenciales ordinarias. Fue profesor de la Universidad de Berlín y miembro de la Academia de Ciencias de Berlín. Como profesor influyó a varias generaciones de matemáticos. Fue director de 20 tesis doctorales y entre sus estudiantes se encuentran matemáticos importantes como L. Fejér, P. Koebe y E. Zermelo.

15.7.3. Desigualdad del triángulo

Sean v y w dos vectores, entonces

$$\begin{aligned}
\langle v+w|v+w\rangle &= \langle v|v+w\rangle + \langle w|v+w\rangle \\
&= \langle v|v\rangle + \langle v|w\rangle + \langle w|v\rangle + \langle w|w\rangle \\
&= \langle v|v\rangle + \langle w|w\rangle + \left(\langle w|v\rangle^* + \langle w|v\rangle\right) \\
&= \langle v|v\rangle + \langle w|w\rangle + 2\mathrm{Re}\left(\langle w|v\rangle\right) \\
&\leq \langle v|v\rangle + \langle w|w\rangle + 2|\langle w|v\rangle|.
\end{aligned}$$

De donde, recurriendo a la desigualdad de Schwarz (15.21) se tiene

$$\begin{aligned}
\langle v+w|v+w\rangle &\leq \langle v|v\rangle + \langle w|w\rangle + 2\sqrt{\langle w|w\rangle}\sqrt{\langle v|v\rangle} \\
&= \left(\sqrt{\langle w|w\rangle} + \sqrt{\langle v|v\rangle}\right)^2,
\end{aligned}$$

es decir,

$$\boxed{\langle v+w|v+w\rangle \leq \left(\sqrt{\langle w|w\rangle} + \sqrt{\langle v|v\rangle}\right)^2}$$

esta igualdad se nombra como del triángulo.

15.7.4. Igualdad del paralelogramo

Además, si v y w son dos vectores, entonces

$$\begin{aligned}
&\langle v+w|v+w\rangle + \langle v-w|v-w\rangle = \\
&= \langle v|v+w\rangle + \langle w|v+w\rangle + \langle v|v-w\rangle - \langle w|v-w\rangle \\
&= \langle v|v\rangle + \langle v|w\rangle + \langle w|v\rangle + \langle w|w\rangle + \langle v|v\rangle - \langle v|w\rangle \\
&- \langle w|v\rangle + \langle w|w\rangle \\
&= 2\left(\langle v|v\rangle + \langle w|w\rangle\right),
\end{aligned}$$

es decir,

$$\boxed{\langle v+w|v+w\rangle + \langle v-w|v-w\rangle = 2\left(\langle v|v\rangle + \langle w|w\rangle\right),}$$

que es la llamada igualdad del paralelogramo.

15.8. Espacios normados

Sea V un espacio vectorial y $|| \quad ||$ una función de V en los reales. Se dice que $(V, || \quad ||)$ es un espacio normado si

$$\boxed{\begin{array}{rlcl} I) & ||v|| \geq 0, & & \\ II) & ||v|| = 0, & \iff & v = 0, \\ III) & ||\alpha v|| = |\alpha| ||v||, & \alpha \in \mathbb{C}, & \\ IV) & ||v+w|| \leq ||v|| + ||w||. & & \end{array}}$$

Observe que si V es un espacio vectorial con producto escalar, entonces se puede definir un espacio normado con la norma dada por

$$\boxed{||v|| = \sqrt{\langle v|v\rangle}.}$$

Las propiedades I y II se cumplen, pues por los axiomas de producto escalar se tiene que $\langle v|v\rangle \geq 0$ y $\langle v|v\rangle = 0 \iff v = 0$. La propiedad III se cumple, pues si α es un escalar, se tiene

$$\begin{aligned} ||\alpha v|| &= \sqrt{\langle \alpha v|\alpha v\rangle} \\ &= \sqrt{\alpha^* \alpha \langle v|v\rangle} \\ &= \sqrt{|\alpha|^2 \langle v|v\rangle} \\ &= |\alpha| \sqrt{\langle v|v\rangle} \\ &= |\alpha| \, ||v||. \end{aligned}$$

La propiedad IV también se cumple, pues usando la desigualdad del triángulo se encuentra

$$\begin{aligned} ||v+w||^2 &= \langle v+w|v+w\rangle \\ &\leq \left(\sqrt{\langle w|w\rangle} + \sqrt{\langle v|v\rangle}\right)^2 \\ &= \left(||w|| + ||v||\right)^2, \end{aligned}$$

es decir,

$$||v+w|| \leq ||w|| + ||v||.$$

Por lo tanto, un espacio vectorial con producto escalar es un espacio normado con la norma dada por $||v|| = \sqrt{\langle v|v\rangle}$.

Note que utilizando la notación de espacios normados, la igualdad del paralelogramo toma la forma

$$||v+w||^2 + ||v-w||^2 = 2\left(||v||^2 + ||w||^2\right).$$

Mientras que, si $\{v_i\}_{i=1}^n$ es un conjunto de vectores ortonormales, el teorema de Pitágoras se puede escribir como

$$||v||^2 = \sum_{i=1}^{n} |\langle v_i|v\rangle|^2 + \left|\left| v - \sum_{i=1}^{n} \langle v_i|v\rangle\, v_i \right|\right|^2.$$

Además, la desigualdad de Bessel toma la forma

$$\sum_{i=1}^{n} |\langle v_i|v\rangle|^2 \leq ||v||^2$$

y la desigualdad de Schwarz es

$$|\langle w|v\rangle| \leq ||w|| \quad ||v||.$$

15.8.1. Espacios métricos

Sea un conjunto M y d una función de $M \times M$ en los reales. Se dice que (M, d) es un espacio métrico si satisface que $\forall x, y \in M$ se cumple

$$\boxed{\begin{array}{lll} A) & d(x,y) \geq 0, & \\ B) & d(x,y) = 0, \quad \Longleftrightarrow & x = y, \\ C) & d(x,y) = d(y,x), & \\ D) & d(x,z) \leq d(x,y) + d(y,z). & \end{array}}$$

A la función d se le llama distancia.

Si V es un espacio normado, entonces se tiene un espacio métrico con la distancia definida por

$$\boxed{d(v_1, v_2) = ||v_1 - v_2||.}$$

Los dos primeros axiomas de distancia se cumplen, pues

$$d(v_1, v_2) = ||v_1 - v_2|| \geq 0 \tag{15.22}$$

y

$$d(v_1, v_2) = ||v_1 - v_2|| = 0 \quad \Longleftrightarrow \quad v_1 - v_2 = 0,$$

es decir $v_1 = v_2$.

También se cumple,

$$\begin{aligned} d(v_1, v_2) &= ||v_1 - v_2|| \\ &= ||(-)(v_2 - v_1)|| \\ &= |(-)| \quad ||v_2 - v_1|| \\ &= ||v_2 - v_1|| \\ &= d(v_2, v_1), \end{aligned}$$

por lo tanto, se cumple el axioma C.

Además, por la desigualdad del triángulo, se tiene

$$\begin{aligned} d(v_1, v_3) &= ||v_1 - v_3|| \\ &= ||(v_1 - v_2) + (v_2 - v_3)|| \\ &\leq ||v_1 - v_2|| + ||v_2 - v_3|| \\ &= d(v_1, v_2) + d(v_2, v_3), \end{aligned}$$

de donde se cumple el axioma D.

Así, un espacio normado es métrico. Lo que quiere decir que cualquier espacio con producto escalar es normado y por lo tanto métrico.

15.9. Polinomios trigonométricos

Supongamos que $\{\psi_n(x)\}_{n=0}^{\infty}$ es un conjunto de funciones ortonormales en el intervalo $[a,b]$, es decir

$$\boxed{\langle \psi_n|\psi_m\rangle = \int_a^b dx q(x)\psi_n^*(x)\psi_m(x) = \delta_{nm},}$$

aquí $q(x)$ es una función de peso positiva en el intervalo (a,b). Con este conjunto de funciones podemos hacer las combinaciones lineales

$$\boxed{T_n(x) = \sum_{i=1}^{n} b_i\psi_i(x),}$$

las cuales llamaremos polinomios trigonométricos. Observe que debido a que $\{\psi_n(x)\}_{n=0}^{\infty}$ es un conjunto de funciones ortonormales, la norma de $T_n(x)$ es

$$||T_n||^2 = \sum_{i=1}^{n} |b_i|^2.$$

Sea $F(x)$ una función del intervalo (a,b) a los reales tal que

$$\boxed{\langle F|F\rangle = ||F||^2 = \int_a^b dx q(x)|F(x)|^2 < \infty,}$$

entonces definiremos los coeficientes de Fourier de F como

$$\boxed{a_n = \langle \psi_n|F\rangle = \int_a^b dx q(x)\psi_n^*(x)F(x).} \tag{15.23}$$

Ahora veremos que tanto se puede aproximar la función $F(x)$ con polinomios de la forma $T_n(x)$. El sentido de la distancia en este espacio está dada por la norma de la funciones. Así, el problema es encontrar los polinomios tales que

$$\boxed{d^2(F,T_n) = ||F-T_n||^2 = \int_a^b dx q(x)|F(x)-T_n(x)|^2} \tag{15.24}$$

es mínimo. Basicamente se trata de encontrar los coeficientes b_i que hacen mínimo (15.24). Podemos iniciar notando que

$$
\begin{aligned}
d^2(F,T_n) &= ||F-T_n||^2 = \langle F-T_n|F-T_n\rangle \\
&= \langle F|F\rangle - \langle F|T_n\rangle - \langle T_n|F\rangle + \langle T_n|T_n\rangle \\
&= ||F||^2 + ||T_n||^2 - \langle F|T_n\rangle - \langle T_n|F\rangle \\
&= ||F||^2 + ||T_n||^2 - \left\langle F\middle|\sum_{i=1}^{n} b_i\psi_i\right\rangle - \left\langle \sum_{i=1}^{n} b_i\psi_i\middle|F\right\rangle \\
&= ||F||^2 + ||T_n||^2 - \sum_{i=1}^{n} b_i\left\langle F\middle|\psi_i\right\rangle - \sum_{i=1}^{n} b_i^*\left\langle \psi_i\middle|F\right\rangle,
\end{aligned}
$$

considerando la norma de $T_n(x)$ y la definición de los coeficientes de Fourier se llega a

$$
||F-T_n||^2 = ||F||^2 + \sum_{i=1}^{n}\left(|b_i|^2 - b_i a_i^* - b_i^* a_i\right).
$$

Además, como

$$
\begin{aligned}
|b_i-a_i|^2 &= (b_i-a_i)(b_i-a_i)^* \\
&= |b_i|^2 + |a_i|^2 - (b_i a_i^* + b_i^* a_i),
\end{aligned}
$$

se tiene

$$
|b_i-a_i|^2 - |a_i|^2 = |b_i|^2 - (b_i a_i^* + b_i^* a_i),
$$

de donde,

$$
||F-T_n||^2 = ||F||^2 + \sum_{i=1}^{n}|b_i-a_i|^2 - \sum_{i=1}^{n}|a_i|^2.
$$

Claramente la distancia entre esta dos funciones es mínima cuando $b_i = a_i$, es decir, cuando el polinomio tienen los coeficientes de Fourier.

Si $b_i = a_i$ se encuentra

$$
\boxed{||F-T_n||^2 = ||F||^2 - \sum_{i=1}^{n}|a_i|^2 \geq 0.}
$$

Por lo que, para cualquier n

$$\boxed{||T_n||^2 = \sum_{i=1}^{n} |a_i|^2 \leq ||F||^2 < \infty.}$$

Esta es la desigualdad de Bessel, la cual implica que la sucesión $||T_n||^2 = \sum_{i=1}^{n} |a_i|^2$ está acotada.

Un resultado de cálculo diferencial nos dice que si una sucesión es monótona creciente y está acotada, entonces converge [8]. Note que la sucesión $||T_n||^2$ es monótona creciente y está acotada, por lo tanto converge. La pregunta es:¿ hacia donde converge? Sin demostrarlo, supondremos que converge a $||F||^2$, es decir,

$$\boxed{\lim_{n\to\infty} ||T_n||^2 = \lim_{n\to\infty} \sum_{i=1}^{n} |a_i|^2 = \sum_{n\geq 0} |a_n|^2 = ||F||^2.} \tag{15.25}$$

A esta igualdad se nombra de Parseval. Demostrar esta igualdad es un problema importante [9, 10], pero altamente no trivial y rebasa el propósito de este escrito, por lo que solo tocaremos este tema en casos particulares.

Un resultado de cálculo diferencial es que si $\sum_{n\geq 0} |a_n|^2$ converge, entonces $a_n \to 0$. Este resultado tiene diferentes implicaciones físicas. Por ejemplo, en mecánica cuántica significa que es más probable que el sistema esté en el estado base. Mientras que en electrostática, significa que en un sistema de cargas, los términos más importantes son el monopolo, dipolo, cuádrupolo.

Marc-Antoine Parseval (1755-1836) fue un matemático francés miembro de la aristocracia francesa. Debido a su origen aristócrata, durante el inicio de la revolución francesa fue tomado preso por un tiempo. Cuando Napoleón era emperador de Francia y dominaba a toda Europa, Perseval escribió un poema en su contra y tuvo que huir a otro país para no caer en la cárcel de nuevo.

Regresó a París cuando Napoleón fue derrotado. Intentó ser miembro de la Academia de Ciencia de Francia, pero fue rechazado cinco veces. No probó el teorema que lleva su nombre, pero mostró la importancia del teorema al usarlo para obtener soluciones a diferentes ecuaciones diferenciales. Escribió solo cinco artículos, pero le bastó para que su nombre pasara a la historia de las Matemáticas.

15.10. Espacios completos

Cuando se cumple la igualdad de Parseval se dice que el conjunto de funciones $\{\psi_n(x)\}_{n\geq 0}$ es completo. En este caso, cualquier función $F(x)$ con $||F|| < \infty$, se puede escribir como combinación lineal de $\{\psi_n(x)\}_{n\geq 0}$, es decir,

$$F(x) = \sum_{n\geq 0} a_n \psi_n(x),$$

con a_n el coeficiente de Fourier definido en (15.23).

En particular, cuando tenemos el conjunto ortonormal de funciones

$$\boxed{\{\psi(x)\} = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}\right\}, \quad n \in \mathbb{N}}$$

tenemos las series de Fourier

$$\boxed{F(x) = \frac{a_0}{\sqrt{2\pi}} + \sum_{n\geq 0}\left(a_n \frac{\cos nx}{\sqrt{\pi}} + b_n \frac{\sin nx}{\sqrt{\pi}}\right)}$$

con

$$\boxed{\begin{aligned} a_0 &= \left\langle \frac{1}{\sqrt{2\pi}} \middle| F(x) \right\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} dx F(x), \\ a_n &= \left\langle \frac{\cos nx}{\sqrt{\pi}} \middle| F(x) \right\rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} dx F(x) \cos x, \\ b_n &= \left\langle \frac{\sin nx}{\sqrt{\pi}} \middle| F(x) \right\rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} dx F(x) \sin x. \end{aligned}}$$

Las series de Fourier tienen aplicaciones en diversas disciplinas como electrónica, electrodinámica, finanzas, mecánica analítica, mecánica cuántica, etc.

Joseph Fourier (1768-1830) fue un matemático y físico francés cuyos resultados influyeron en diferentes áreas de las matemáticas. Fue el décimo hijo de una pareja de artesanos. A los 10 años quedó huérfano de ambos padres y seis de sus hermanos habían muerto. A esa edad fue adoptado por un músico que le enseñó a leer y escribir. A los 14 años ingresó a una academia de corte militar en donde destacó por su gusto por las matemáticas. Durante la revolución francesa participó en un *Comité Revolucionario* apoyando a las ideas republicanas. Después ingresó a la Escuela Normal Superior de París donde fue alumno de Lagrange y Laplace. Al término de sus estudios adquirió una plaza de profesor en la École Polytechnique. Fue parte del grupo que acompañó a Napoleón a Egipto. Como físico, trató de extender la ley de enfriamiento de Newton a casos más generales, por lo que se interesó en modelar la propagación del calor en los metales. En particular, obtuvo la ecuación de calor, que es una ecuación diferencial parcial que modela la difusión del calor. Sus trabajos fueron puestos en duda por físicos y matemáticos, incluidos Laplace y Lagrange. Sin embargo, los experimentos le daban la razón. En la física, sus trabajos ayudaron a fundamentar las leyes de la termodinámica. Desde el punto de vista matemático, sus contemporáneos cuestionaban sus métodos, pues tenía un número infinito de soluciones para la ecuación de calor. Sin embargo, a pesar de ser poco entendidos, los métodos de Fourier fueron ocupados para resolver con éxito otras ecuaciones diferenciales parciales. En realidad, Fourier estaba presentando a sus contemporáneos por primera vez lo que en la actualidad conocemos como un espacio vectorial de dimensión infinita, por lo que nadie era capaz de comprender las herramientas que él usaba de manera intuitiva. Para asimilar los resultados de Fourier, las matemáticas tuvieron que evolucionar en conceptos fundamentales como el de derivada, continuidad, función, serie de funciones, teoría de la medida, convergencia, etc. En su honor, las series de funciones y la transformada integral que se usan para estudiar la ecuación del calor llevan su nombre. Para Lord Kelvin, Fourier fue un poeta de las matemáticas. Fourier murió por problemas del corazón en 1830, su nombre está grabado en la Torre Eiffel.

15.11. Proceso de ornormalización de Gram-Schmidt

Sea V un espacio vectorial y supongamos que en este espacio el conjunto de vectores

$$\boxed{\{u_1, u_2, \cdots, u_n\}} \tag{15.26}$$

es ortormal, es decir, que se satisface

$$\boxed{\langle u_i | u_j \rangle = \delta_{ij}.}$$

Ahora, supongamos que $v \in V$, entonces se puede construir el vector

$$\boxed{\tilde{u} = v - \sum_{i=1}^{n} u_i \langle u_i | v \rangle .}$$

Este vector es ortogonal a todos los vectores del conjunto (15.26). En efecto, se cumple que

$$\begin{aligned}
\langle u_k | \tilde{u} \rangle &= \left\langle u_k \middle| v - \sum_{i=1}^{n} u_i \langle u_i | v \rangle \right\rangle \\
&= \langle u_k | v \rangle - \left\langle u_k \middle| \sum_{i=1}^{n} u_i \langle u_i | v \rangle \right\rangle \\
&= \langle u_k | v \rangle - \sum_{i=1}^{n} \langle u_k | u_i \rangle \langle u_i | v \rangle \\
&= \langle u_k | v \rangle - \sum_{i=1}^{n} \delta_{ki} \langle u_i | v \rangle \\
&= \langle u_k | v \rangle - \langle u_k | v \rangle \\
&= 0.
\end{aligned}$$

Además, si se satisface

$$\boxed{\tilde{u} = v - \sum_{i=1}^{n} u_i \langle u_i | v \rangle \neq 0.} \tag{15.27}$$

podemos formar el vector

$$\boxed{u = \frac{v - \sum_{i=1}^{n} u_i \langle u_i|v\rangle}{||v - \sum_{i=1}^{n} u_i \langle v|u_i\rangle\, ||}}$$

el cual es unitario.

Por lo tanto, si se satisface la condición (15.27), podemos construir el vector u el cual es unitario y ortogonal a todos los vectores del conjunto (15.26). Entonces el conjunto

$$\{u, u_1, u_2, \cdots, u_n\} \tag{15.28}$$

es ortonormal. El procedimiento que hemos mostrado nos ayudará a construir conjuntos de vectores ortonormales.

Supongamos que

$$\boxed{\{v_1, v_2, \cdots, v_n\}} \tag{15.29}$$

es un conjunto de vectores linealmente independiente. Usando el vector v_1 podemos construir el vector unitario

$$\boxed{u_1 = \frac{v_1}{||v_1||}.}$$

Observe que el conjunto

$$\boxed{\{u_1\}}$$

es ortonormal. Además, como v_2 es linealmente independiente de u_1 podemos formar el vector no nulo

$$\boxed{u_2 = \frac{v_2 - \langle u_1|v_2\rangle u_1}{||v_2 - \langle u_1|v_2\rangle u_1||},}$$

el cual es ortogonal a u_1. Por lo tanto el conjunto

$$\boxed{\{u_1, u_2\}}$$

es ortonormal. Además, como v_3 es linealmente independiente de u_1 y u_2 podemos forma el vector

$$\boxed{u_3 = \frac{v_3 - \langle u_2|v_3\rangle\, u_2 - \langle u_1|v_3\rangle\, u_1}{||v_3 - \langle u_2|v_3\rangle\, u_2 - \langle u_1|v_3\rangle\, u_1||},}$$

el cual es unitario y ortogonal a u_1 y u_2. Por lo tanto el conjunto

$$\boxed{\{u_1, u_2, u_3\}}$$

es ortonormal. Siguiendo este procedimiento, del conjunto linealmente independiente (15.29) se puede obtener un conjunto ortonormal. A este proceso se le llama proceso de ornormalización de Gram-Schmidt.

15.11.1. Ejemplos

Consideremos el conjunto de vectores

$$\{v_1 = (1, -1, 1)\,, v_2 = (1, 1, 1)\,, v_3 = (-1, -1, 1)\},$$

el cual es linealmente independiente.

En este caso tenemos

$$||v_1|| = \sqrt{3},$$

por lo que

$$u_1 = \frac{v_1}{||v_1||} = \frac{(1, -1, 1)}{\sqrt{3}}.$$

Además tenemos que

$$\langle u_1|v_2\rangle = \frac{1}{\sqrt{3}},$$

de donde

$$\begin{aligned}
\tilde{u}_2 &= v_2 - \langle u_1|v_2\rangle\, u_1 \\
&= (1, 1, 1) - \frac{1}{\sqrt{3}}\frac{(1, -1, 1)}{\sqrt{3}} \\
&= (1, 1, 1) - \frac{1}{3}(1, -1, 1) \\
&= \frac{2}{3}(1, 2, 1),
\end{aligned}$$

que implica

$$||\tilde{u}_2|| \;\; = \;\; \frac{2}{3}\sqrt{6},$$

así

$$u_2 = \frac{\tilde{u}_2}{||\tilde{u}_2||} \;\; = \;\; \frac{1}{\sqrt{6}}(1,2,1).$$

Adicionalmente, se encuentra que

$$\langle u_1|v_3\rangle = \frac{1}{\sqrt{3}}, \quad \langle u_2|v_3\rangle = -\frac{2}{\sqrt{6}},$$

entonces

$$\begin{aligned}
\tilde{u}_3 &= v_3 - \langle u_1|v_3\rangle\, u_1 - \langle u_2|v_3\rangle\, u_2 \\
&= (-1,-1,1) - \frac{1}{\sqrt{3}}\frac{(1,-1,1)}{\sqrt{3}} + \frac{2}{\sqrt{6}}\frac{1}{\sqrt{6}}(1,2,1) \\
&= (-1,-1,1) + \frac{-(1,-1,1)+(1,2,1)}{3}(1,-1,1) \\
&= (-1,0,1),
\end{aligned}$$

de donde

$$||\tilde{u}_3|| \;\; = \;\; \sqrt{2}$$

así

$$u_3 = \frac{\tilde{u}_3}{||u_3||} \;\; = \;\; \frac{(-1,0,1)}{\sqrt{2}}.$$

Por lo tanto el conjunto de vectores

$$\left\{u_1 = \frac{(1,-1,1)}{\sqrt{3}}, u_2 = \frac{(1,2,1)}{\sqrt{6}}, u_3 = \frac{(-1,0,1)}{\sqrt{2}}\right\}$$

es ortonormal.

Como otro ejemplo, estudiemos el conjunto de polinomios

$$\left\{Q_1(x) = 3, \quad Q_2(x) = 2x-1, \quad Q_3(x) = x - 2x^2\right\},$$

el cual es linealmente independiente. En este caso tomaremos el producto escalar

$$\langle f(x)|g(x)\rangle = \int_{-1}^{1} f(x)g(x).$$

Con este producto escalar obtenemos

$$\begin{aligned}||Q_1(x)||^2 &= \int_{-1}^{1} Q_1(x)Q_1(x)dx \\ &= \int_{-1}^{1} 3^2 dx \\ &= 3^2 \cdot 2,\end{aligned}$$

de donde

$$||Q_1(x)|| = 3\sqrt{2},$$

por lo que el polinomio

$$P_1(x) = \frac{Q_1(x)}{||Q_1(x)||} = \frac{1}{\sqrt{2}}$$

es unitario.

Además tenemos que

$$\langle P_1(x)|Q_2(x)\rangle = \int_{-1}^{1} dx \frac{1}{\sqrt{2}}(2x-1) = -\sqrt{2},$$

de donde

$$\begin{aligned}\tilde{P}_2(x) &= Q_2(x) - \langle P_1(x)|Q_2(x)\rangle P_1(x) \\ &= (2x-1) + \sqrt{2}\frac{1}{\sqrt{2}} \\ &= 2x,\end{aligned}$$

por lo que

$$||\tilde{P}_2(x)||^2 = \int_{-1}^{1} (2x)^2 = \frac{2^2 \cdot 2}{3},$$

que implica

$$||\tilde{P}_2(x)|| = 2\sqrt{\frac{2}{3}},$$

así el polinomio

$$P_2(x) = \frac{\tilde{P}_2(x)}{||\tilde{P}_2(x)||} \quad = \quad \sqrt{\frac{3}{2}}x$$

es unitario.

También podemos ver que

$$\begin{aligned}
\langle P_1(x)|Q_3(x)\rangle &= \frac{1}{\sqrt{2}}\int_{-1}^{1} dx(x-2x^2) = -\frac{4}{3\sqrt{2}}, \\
\langle P_2(x)|Q_3(x)\rangle &= \sqrt{\frac{3}{2}}\int_{-1}^{1} dxx(x-2x^2) = \sqrt{\frac{2}{3}},
\end{aligned}$$

entonces

$$\begin{aligned}
\tilde{P}_3(x) &= Q_3(x) - \langle P_1|Q_3(x)\rangle\, u_1 - \langle P_2(x)|Q_3(x)\rangle\, P_2(x) \\
&= (x-2x^2) + \frac{4}{3\sqrt{2}}\frac{1}{\sqrt{2}} - \sqrt{\frac{2}{3}}\sqrt{\frac{3}{2}}x \\
&= x - 2x^2 + \frac{2}{3} - x \\
&= \frac{2}{3}\left(1-3x^2\right),
\end{aligned}$$

por lo que

$$||\tilde{P}_3(x)||^2 \quad = \quad \left(\frac{2}{3}\right)^2 \int_{-1}^{1} \left(1-3x^2\right)^2 = 2^2\frac{2}{5}\left(\frac{2}{3}\right)^2,$$

de donde

$$||\tilde{P}_3(x)|| \quad = \quad 2\frac{2}{3}\sqrt{\frac{2}{5}},$$

así el polinomio

$$P_3(x) = \frac{\tilde{P}_3(x)}{||P_3(x)||} \quad = \quad \frac{1}{2}\sqrt{\frac{5}{2}}\left(x-3x^2\right),$$

es unitario.

Por lo tanto el conjunto de polinomios

$$\left\{P_1(x) = \frac{1}{\sqrt{2}}, \quad P_2(x) = \sqrt{\frac{3}{2}}x, \quad P_3(x) = \frac{1}{2}\sqrt{\frac{5}{2}}\left(x-3x^2\right)\right\}$$

es ortonormal.

Jørgen Pedersen Gram (1850-1916) fue un matemático y actuario danés. Realizó contribuciones en la teoría de números y en el estudio de la función zeta de Riemann. También se interesó en la probabilidad, estadística y en los métodos numéricos. Gram fue hijo de un granjero danés, por lo que se interesó en la silvicultura y en la aplicación de las matemáticas a la reforestación de los bosques. Gram murió en un accidente con una bicicleta.

Erhard Schmidt (1876-1959) fue un matemático alemán. Cotribuyó de forma relevante al análisis funcional. De hecho, es considerado uno de los fundadores de esta rama de las matemáticas. Fue alumno de H. A. Schwarz y de David Hilbert. En la Segunda Guerra Mundial apoyó a Hitler a quien admiraba y le deseaba la victoria. En la época nazi fue autoridad de la Universidad de Berlín y aplicó algunas medidas en contra de judíos.

El método de ortogonalización de Gram-Schmidt se le atribuye a Gram y Schmidt, pero en realidad Laplace y Cauchy ya lo habían empleado antes.

15.12. Ejercicios

1.- Realizar el producto escalar de los vectores

$$(-2i, 4, 5), \qquad (6, 5, i).$$

2.- En el intervalo $[-1, 1]$ y la función de peso $q(x) = 1$, realizar en producto escalar de las funciones

$$x, \qquad -x + 3x^2.$$

3.- Realizar el producto escalar de las matrices

$$A = \begin{pmatrix} -i & 1 \\ 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 1 \\ 5 & 0 \end{pmatrix}.$$

4.- Obtener un vector unitario del vector

$$(3, -1+2i, 4).$$

5.- De la matriz

$$A = \begin{pmatrix} 1 & -3i \\ 2 & 1 \end{pmatrix}$$

obtener una matriz de norma uno.

6.- En el intervalo $[-1, 1]$ y con la función de peso $q(x) = 1$, obtener una función unitaria de la función

$$f(x) = x^2 - 2x$$

.

7.- Determinar si el siguiente conjunto de vectores es ortonormal

$$\frac{1}{\sqrt{6}}(i, -i, 2), \quad \frac{1}{\sqrt{3}}(-i, i, 1), \quad \frac{1}{\sqrt{2}}(1, 1, 0).$$

8.- Determinar si las funciones

$$f(x) = x, \quad g(x) = x^2$$

son ortogonales con el producto escalar

$$\int_0^1 f(x)g(x).$$

9.- Determinar si las funciones de la forma

$$f_n(x) = \sqrt{\frac{2}{L}} \sin nx$$

son ortogonales con el producto escalar

$$\int_0^L f(x)g(x)dx, \quad n \in \mathbb{N}$$

10.- Determinar si las matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

son ortogonales.

11.- De los vectores

$$(2,0,0), \quad (1,3,0), \qquad (1,2,1)$$

obtener un conjunto ortnormal de vectores.

12.- Con el producto escalar

$$\langle f(x)|g(x)\rangle = \int_{-\infty}^{\infty} e^{-x^2} f^*(x)g(x)dx,$$

del conjunto de polinomios

$$\left\{Q_0 = 2, \quad Q_1(x) = x - 1, \quad Q_2(x) = 2x^2 - x\right\}$$

obtener un conjunto ortonormal de polinomios.

Capítulo 16

Valores y vectores propios

En este capítulo veremos las definiciones de vector y valor propio. Daremos algunos ejemplos y mostraremos que usando los vectores y valores propios la representación matricial de una transformación lineal se puede simplificar.

16.1. Definición

Sean M una matriz de $n \times n$, v un vector de dimensión n y λ un escalar. Se dice que $\boxed{v \text{ es vector propio de } M \text{ y } \lambda \text{ es valor propio de } V}$ si se cumple la igualdad

$$\boxed{Mv = \lambda v.}$$

Note que esta igualdad implica

$$Mv - \lambda v = (M - I\lambda)\, v = 0. \tag{16.1}$$

Esta ecuación siempre tiene como solución a $v = 0$, con λ arbitrario. Si la matriz

$$M - I\lambda$$

es invertible, la única solución de la ecuación (16.1) es $v = 0$. Por lo tanto, si queremos soluciones no triviales de dicha ecuación debemos pedir que existan valores λ tal que la matriz $M - I\lambda$ no sea invertible, que implica

$$\boxed{\det (M - I\lambda) = 0.} \tag{16.2}$$

16.2. Ejemplo

Antes de ver los resultados importantes sobre valores y vectores propios, veremos un ejemplo. Con este propósito consideremos la matriz

$$M = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & -1 \\ 0 & 0 & 4 \end{pmatrix},$$

los valores y vectores propios de esta matriz satisfacen

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

de donde se obtiene la matriz

$$M - \lambda I = \begin{pmatrix} 2-\lambda & 1 & 0 \\ 0 & 5-\lambda & -1 \\ 0 & 0 & 4-\lambda \end{pmatrix}.$$

Entonces, para tener soluciones no triviales se debe cumplir

$$\begin{aligned} \det(M - \lambda I) &= \det \begin{pmatrix} 2-\lambda & 1 & 0 \\ 0 & 5-\lambda & -1 \\ 0 & 0 & 4-\lambda \end{pmatrix} \\ &= (2-\lambda)(5-\lambda)(4-\lambda) = 0, \end{aligned}$$

por lo tanto, tenemos los valores propios

$$\lambda_1 = 2, \quad \lambda_2 = 5, \quad \lambda_3 = 4.$$

Para obtener los vectores propios se deben resolver las ecuaciones

$$\begin{aligned} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \\ \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 5 \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \\ \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 4 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

que implican

$$v_1 = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = x_2 \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad v_3 = x_3 \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix},$$

con x_1, x_2, x_3 son arbitrarios.

Note que con los valores propios se puede formar la matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

mientras que con los vectores v_1, v_2, v_3 se puede formar la matriz

$$U = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{pmatrix},$$

que tiene la inversa

$$U^{-1} = \begin{pmatrix} 1 & -\frac{1}{3} & -\frac{1}{6} \\ 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Se puede observar que con estas matrices se cumple la igualdad

$$D = U^{-1} M U,$$

es decir

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{3} & -\frac{1}{6} \\ 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Así, la matriz M es similar a la matriz diagonal D.

16.3. Vectores propios e independencia lineal

Sean $\{v_1, v_2, \cdots, v_m\}$ un conjunto de vectores propios no nulos con valores propios $\{\lambda_1, \lambda_2, \cdots, \lambda_m\}$. Si todos los valores propios son diferentes, es decir, $\lambda_i \neq \lambda_j$ con $i \neq j$, entonces el conjunto de vectores propios $\{v_1, v_2, \cdots v_m\}$ es linealmente independiente.

Veamos el caso $m = 1$, para el cual tenemos

$$\{v_1\}, \quad \{\lambda_1\},$$

en donde

$$Mv_1 = \lambda_1 v_1.$$

Como el conjunto

$$\{v_1\}$$

consta de un solo vector, claramente es linealmente independiente.

Para $m = 2$, tenemos

$$\{v_1, v_2\}, \quad \{\lambda_1, \lambda_2\},$$

en donde

$$Mv_1 = \lambda_1 v_1, \quad Mv_2 = \lambda_2 v_2, \qquad \lambda_1 \neq \lambda_2$$

Supongamos que

$$\alpha_1 v_1 + \alpha_2 v_2 = 0, \tag{16.3}$$

entonces

$$\begin{aligned} M0 &= M(\alpha_1 v_1 + \alpha_2 v_2) \\ &= \alpha_1 M v_1 + \alpha_2 M v_2 \\ &= \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 = 0, \end{aligned}$$

es decir

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 = 0. \tag{16.4}$$

Ahora, multiplicando la ecuación (16.3) por λ_2 tenemos

$$\alpha_1 \lambda_2 v_1 + \alpha_2 \lambda_2 v_2 = 0. \tag{16.5}$$

Restando la ecuaciones (16.4) y (16.5) tenemos

$$\alpha_1(\lambda_1 - \lambda_2) v_1 = 0,$$

como $\lambda_1 \neq \lambda_2$ tenemos que

$$\alpha_1 = 0.$$

Sustituyendo este valor en (16.3) encontramos

$$\alpha_2 = 0.$$

Por lo tanto, si $\lambda_1 \neq \lambda_2$ cualquier combinación lineal de la forma (16.3) implica que $\alpha_1 = \alpha_2 = 0$. Así, el conjunto de vectores propios

$$\{v_1, v_2\}$$

es linealmente independiente.

Para el caso general haremos la prueba por inducción, la base de la inducción $m = 1$ ya la hemos demostrado.

Ahora por hipótesis de inducción supondremos que todo conjunto de

$$\{v_1, v_2, \cdots, v_m\}$$

vectores propios no nulos con valores propios diferentes

$$(\lambda_i \neq \lambda_j, \qquad i, j = 1, \cdots, m)$$

es linealmente independiente.
En el paso inductivo tenemos que probar que todo conjunto de

$$\{v_1, v_2, \cdots, v_m, v_{m+1}\} \tag{16.6}$$

vectores propios no nulos con valores propios diferentes

$$\lambda_i \neq \lambda_j, \qquad i, j = 1, \cdots, m, m+1$$

es linealmente independiente.

Supongamos que tenemos una combinación lineal nula de los vectores (16.6)

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m + \alpha_{m+1} v_{m+1} = 0, \tag{16.7}$$

entonces

$$\begin{aligned} M0 &= M\left(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m + \alpha_{m+1} v_{m+1}\right) \\ &= \alpha_1 M v_1 + \alpha_2 M v_2 + \cdots + \alpha_m M v_m + \alpha_{m+1} M v_{m+1} \\ &= \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_m \lambda_m v_m + \alpha_{m+1} \lambda_{m+1} v_{m+1} = 0, \end{aligned}$$

es decir

$$\alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 + \cdots + \alpha_m\lambda_m v_m + \alpha_{m+1}\lambda_{m+1}v_{m+1} = 0. \tag{16.8}$$

Ahora, multiplicando la ecuación (16.7) por λ_{m+1} tenemos

$$\alpha_1\lambda_{m+1}v_1 + \alpha_2\lambda_{m+1}v_2 + \cdots + \alpha_m\lambda_{m+1}v_m + \alpha_{m+1}\lambda_{m+1}v_{m+1} = 0. \tag{16.9}$$

Restando la ecuaciones (16.8) y (16.9) tenemos

$$\alpha_1\left(\lambda_1 - \lambda_{m+1}\right)v_1 + \alpha_2\left(\lambda_2 - \lambda_{m+1}\right)v_2 + \cdots + \alpha_m\left(\lambda_m - \lambda_{m+1}\right)v_m = 0. \tag{16.10}$$

Note que esta última igualdad es una combinación lineal del conjunto de vectores propios no nulos $\{v_1, v_2, \cdots, v_m\}$ con valores propios diferentes

$$(\lambda_i \neq \lambda_j, \qquad i, j = 1, \cdots, m).$$

Por hipótesis de inducción, este conjunto es linealmente independiente, entonces como (16.10) es una combinación lineal nula temos que

$$\alpha_1\left(\lambda_1 - \lambda_{m+1}\right) = \alpha_2\left(\lambda_2 - \lambda_{m+1}\right) = \cdots = \alpha_m\left(\lambda_m - \lambda_{m+1}\right) = 0.$$

Además, como $\lambda_i \neq \lambda_j$ tenemos que

$$\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0.$$

sustituyendo estos valore en (16.7) encontramos

$$\alpha_{m+1} = 0.$$

Por lo tanto, si $\lambda_i \neq \lambda_j, i, j = 1, \cdots m, m+1$ cualquier combinación lineal de la forma (16.7) implica que $\alpha_1 = \alpha_2 = \cdots = \alpha_m = \alpha_{m+1} = 0$. Así, el conjunto de vectores propios

$$\{v_1, v_2, \cdots, v_m, v_{m+1}\}$$

es linealmente independiente. Esto termina la demostración.

16.4. Matriz diagonal

Sean $\{v_1, v_2, \cdots, v_n\}$ vectores propios de M, entonces M es equivalente a un matriz diagonal. Para probar esta afirmación, primero observemos que con estos vectores propios

$$v_1 = \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}, v_2 = \begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{pmatrix}, \cdots, v_n = \begin{pmatrix} a_{n1} \\ a_{n2} \\ \vdots \\ a_{nn} \end{pmatrix} \tag{16.11}$$

podemos construir los renglones de la matriz

$$a = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

la cual tiene la traspuesta

$$a^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}. \qquad (16.12)$$

Note que las columnas de esta última matriz están dadas por los vectores propios de la matriz M.

Se puede observar que si los vectores (16.11) son linealmente independientes, entonces los renglones de la matriz a son linealmente independientes y por lo tanto el determinante de a es diferente de cero. Es decir, si los vectores (16.11) son linealmente independientes, la matriz a es invertible.

Ahora, sabemos que los vectores propios satisfacen la ecuación

$$Mv_i = \lambda_i v_i,$$

por lo que usando las componentes de estos vectores (16.11) se tiene

$$M_{jk} a_{ik} = \lambda_i a_{ij}.$$

Además, usando la matriz (16.12), esta última expresión se puede escribir como

$$M_{jk} a^T_{ki} = \lambda_i a^T_{ji},$$

entonces,

$$\left(Ma^T\right)_{ji} = \lambda_i a^T_{ji}.$$

Ahora, como la matriz a es invertible tenemos

$$\left(a^T\right)^{-1}_{lj} \left(Ma^T\right)_{ji} = \lambda_i \left(a^T\right)^{-1}_{lj} a^T_{ji},$$

por lo que

$$\left(\left(a^T\right)^{-1} M a^T\right)_{li} = \lambda_i \delta_{li}.$$

Así,

$$\left(a^T\right)^{-1} M a^T = D, \tag{16.13}$$

en donde D es la matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Se puede observar que el resultado (16.13) se puede escribir como

$$M = a^T D \left(a^T\right)^{-1}, \tag{16.14}$$

es decir,

$$D = U^{-1} M U, \qquad U = a^T. \tag{16.15}$$

Como sabemos, la mayoría de operaciones con matrices son más fáciles de realizar con matrices diagonales. Por esta razón, antes de realizar operaciones con una matriz es mejor obtener sus valores propios y obtener su matriz similar diagonal.

La noción de valor propio surgió en uno de los trabajos de Euler que trata sobre métodos para resolver sistemas de ecuaciones diferenciales ordinarias lineales, en 1743. Asimismo, Lagrange usó esta idea para resolver un problema de mecánica celeste en 1774. Por las mismas fechas, D'Alembert ocupó esa idea para resolver el sistema de ecuaciones diferenciales de tres osciladores. Mientras que Cauchy usó la noción de valor propio en problemas de geometría en los cursos que impartía desde 1815.

16.5. Operaciones con Python

Si tenemos la matriz

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

para obtener sus valores y vectores se puede usar el siguiente código de Python

```python
import numpy as np
from numpy.linalg import eig
b = np.array([[1,-1,2,1],[-1,3,1,-2],[1,3,-1,1],[1,-1,0,2]
             ])
v,w=eig(b)
print('valores-propios:', v)
print('vectores-propios', w)
```

Listing 16.1: Valores y vectores propios de una matriz

con el cual se obtienen los siguientes valores y vectores propios

```
valores-propios:[4.52782655-,2.4290382,2.54377682,0.35743483]
vectores-propios␣[[0.14439693,-0.5528967,-0.5781175,-0.82056069]
[-0.83689882,-0.22159463,-0.50232403,0.10163719]
[-0.35784446,0.79975351,-0.6277125,0.03373165]
[0.38819742,0.07480226,-0.13938343,0.56143762]]
```

16.6. Ejercicios

1.- Para la matriz

$$M = \begin{pmatrix} 1 & 2 & 0 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{pmatrix},$$

obtener su forma diagonal y la matriz cambio de base.

2.- Obtener los valores y vectores propios de la matriz

$$M = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

Capítulo 17

Matrices similares

En este capítulo definiremos el concepto de matrices similares, mostraremos que dicho concepto induce una relación de equivalencia entre matrices y veremos algunas de las implicaciones de esta relación.

17.1. Relación de equivalencia

Sea G un conjunto y R un subconjunto de $G \times G$. Si $g_1, g_2 \in G$ y $(g_1, g_2) \in R$ se denota

$$g_1 \sim g_2$$

y se dice que g_1 está relacionado con g_2.

Se dice que R es una Relación de Equivalencia si se satisface las condiciones

$$\forall g \in G, \quad g \sim g, \tag{17.1}$$
$$\forall g_1, g_2 \in G, \quad \text{si} \quad g_1 \sim g_2, \quad \text{entonces} \quad g_2 \sim g_1, \tag{17.2}$$
$$\forall g_1, g_2, g_3 \in G, \quad \text{si} \quad g_1 \sim g_2 \quad \text{y} \quad g_2 \sim g_3, \quad \text{entonces} \quad g_1 \sim g_3. \tag{17.3}$$

Ahora, para cada $g \in G$ definamos los conjuntos

$$\boxed{\chi_g = \{g' \in G | g \sim g'\}.}$$

Los conjuntos χ_g son importantes porque cuando la relación R es de equivalencia el conjunto G se puede separar en conjuntos disjuntos. En otras palabras, cuando la relación R es de equivalencia y se cumple

$$\boxed{\chi_{g_1} \cap \chi_{g_2} \neq \emptyset, \qquad g_1, g_2 \in G}$$

si satisface

$$g_1 \sim g_2$$

entonces

$$\boxed{\chi_{g_1} = \chi_{g_2}.}$$

Esta afirmación es verdadera. Para iniciar, mostraremos que se cumple

$$\chi_{g_1} \subset \chi_{g_2}.$$

Supongamos que

$$g \in \chi_{g_1}$$

que implica

$$g \sim g_1.$$

Ahora, como

$$g_1 \sim g_2,$$

y R es una relación de equivalencia obtenemos

$$g \sim g_2.$$

Este último resultado nos indica que

$$g \in \chi_{g_2}$$

y por lo tanto $\chi_{g_1} \subset \chi_{g_2}$.

Ahora probaremos que se cumple

$$\chi_{g_2} \subset \chi_{g_1}.$$

Supongamos que

$$\tilde{g} \in \chi_{g_2}$$

entonces

$$\tilde{g} \sim g_2$$

y como $g_2 \sim g_1$ entonces se cumple que

$$\tilde{g} \sim g_1,$$

que implica

$$\tilde{g} \in \chi_{g_1}$$

Por lo tanto

$$\chi_{g_2} \subset \chi_{g_1}.$$

Así, hemos demostrado que se cumple

$$\chi_{g_1} \subset \chi_{g_2} \quad \text{y} \quad \chi_{g_2} \subset \chi_{g_1}$$

de donde

$$\chi_{g_1} = \chi_{g_2},$$

que es lo queríamos demostrar. Nótese que este resultado también se puede escribir como

$$\boxed{\chi_{g_1} \neq \chi_{g_2}, \qquad \text{entonces} \qquad \chi_{g_1} \cap \chi_{g_2} = \emptyset.} \tag{17.4}$$

También se puede probar que se cumple que

$$\cup \chi_{g\in G} = G, \tag{17.5}$$
$$\chi_g \neq \emptyset. \tag{17.6}$$

Cuando una familia de subconjuntos de un conjunto G satisface las propiedades (17.4)-(17.6) se dice que forma una Partición de G.

La ventaja de tener una partición en un conjunto es que en lugar de trabajar con los elementos se puede trabajar con las particiones.

17.2. Equivalencia entre matrices

Ahora veremos una relación de equivalencia entre matrices. Supongamos que M y A son dos matrices de $n \times n$. Definiremos la relación

$$\boxed{M \sim A}$$

si existe una matriz invertibe U de $n \times n$ tal que se cumple

$$\boxed{M = U^{-1}AU.} \tag{17.7}$$

Por ejemplo, consideremos las matrices

$$A = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}, \qquad B = \begin{pmatrix} 10 & 2 \\ 4 & 0 \end{pmatrix}.$$

Estas matrices están relacionadas, pues con la matriz

$$U = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

y su inversa

$$U^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

se cumple

$$\begin{pmatrix} 10 & 2 \\ 4 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Cuando se satisface la ecuación (17.7) diremos que las matrices M y A son similares y a R la llamaremos relación de similitud.

La relación de similitud es de equivalencia. Para mostrar esta afirmación, notemos que para cualquier matriz de $n \times n$, M, se cumple

$$M = I^{-1}MI$$

Donde I es la matriz identidad. Por lo tanto, cualquier matriz de $n \times n$ siempre es similar a si misma y se cumple la propiedad (17.1).

Para probar la propiedad (17.2), supongamos que M es similar a A, es decir existe una matriz invertible U tal que se cumple la ecuación (17.7). Nótese que la ecuación (17.7) implica

$$UMU^{-1} = UU^{-1}AUU^{-1} = A,$$

por lo tanto, definiendo

$$U' = U^{-1}$$

se tiene

$$A = U'^{-1} M U'.$$

Así, podemos afirmar que si M es similar a A, entonces A es similar a M. De esta forma hemos probado que se cumple la propiedad (17.1).

Para probar la propiedad (17.3) supongamos que M_1, M_2, M_3 son matrices de $n \times n$ y que M_1 es similar a M_2 y M_2 es similar a M_3. Entonces existen matrices invertibles U_1 y U_2 tales que

$$\begin{aligned} M_1 &= U_1^{-1} M_2 U_1, && (17.8) \\ M_2 &= U_2^{-1} M_3 U_2. && (17.9) \end{aligned}$$

Usando la ecuación (17.9) en la ecuación (17.8) y ocupando las propiedades del producto de las matrices obtenemos

$$\begin{aligned} M_1 &= U_1^{-1} M_2 U_1 \\ &= U_1^{-1} \left(U_2^{-1} M_3 U_2 \right) U_1 \\ &= U_1^{-1} U_2^{-1} M_3 U_2 U_1 \\ &= \left(U_2 U_1 \right)^{-1} M_3 \left(U_2 U_1 \right). \end{aligned}$$

Ahora, definiendo la matriz invertible

$$U_3 = U_1 U_2,$$

tenemos

$$M_1 = U_3^{-1} M_3 U_3,$$

que indica que la matriz M_1 es similar a M_3. Así, hemos probado que si M_1 es similar a M_2 y M_2 es similar a M_3, entonces M_1 es similar a M_3. Por lo tanto, la relación de similitud cumple la propiedad (17.3).

Como la relación de similitud es de equivalencia, podemos concluir que los conjuntos

$$\boxed{\chi_M = \{A \in M_{n\times n} | \exists U \in M_{n\times n} \quad \text{invertible} \qquad M = U^{-1} A U\}}$$

forman una partición del conjunto de matrices de $n \times n$.

17.3. Propiedades

Las matrices que pertenecen a la misma partición tienen propiedades semejantes. Por ejemplo, si A es similar a M entonces

$$\boxed{\begin{aligned} \det M =& \det A, \\ Tr(M) =& Tr(A), \end{aligned}}$$

Efectivamente, usando las propiedades del determinante y la traza, tenemos

$$\begin{aligned} \det M &= \det(U^{-1}AU) \\ &= (\det U^{-1})(\det A)(\det U) \\ &= \left(\frac{\det U}{\det U}\right)\det A \\ &= \det A, \\ Tr(M) &= Tr\left(U^{-1}AU\right) \\ &= Tr\left(UU^{-1}A\right) \\ &= Tr(A), \end{aligned}$$

que es lo que se quería demostrar. Así, dos matrices similares tienen el mismo determinante y la misma traza.

17.4. Polinomios y funciones

Si las matrices M y A son similares se cumple

$$\boxed{M^n = U^{-1}A^nU,} \tag{17.10}$$

esta afirmación es correcta pues

$$\begin{aligned} M^n &= \left(U^{-1}AU\right)^n \\ &= \underbrace{\left(U^{-1}AU\right)\left(U^{-1}AU\right)\left(U^{-1}AU\right)\cdots\left(U^{-1}AU\right)}_{n} \\ &= U^{-1}\underbrace{AA\cdots AA}_{n}U \\ &= U^{-1}A^nU. \end{aligned}$$

Ahora, supongamos que x es una variable real, entonces podemos formar un polinomio de la forma

$$\boxed{P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,}$$

donde a_i $(i = 0, 1, 2, \ldots, n)$ son constantes reales o complejas. De la misma forma se puede construir la matriz polinomio

$$\boxed{P(M) = a_0 I + a_1 M + a_2 M^2 + \cdots + a_n M^n.}$$

Si M es similar a A, entonces

$$\boxed{P(M) = U^{-1} P(A) U.}$$

En efecto, usando la identidad (17.10) se encuentra

$$\begin{aligned} P(M) &= P(U^{-1}AU) \\ &= a_0 I + a_1 U^{-1}AU + a_2 (U^{-1}AU)^2 + \cdots + a_n (U^{-1}AU)^n \\ &= a_0 U^{-1}U + a_1 U^{-1}AU + a_2 U^{-1}A^2 U + \cdots + a_n U^{-1}A^n U \\ &= U^{-1}\left(a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n\right) U \\ &= U^{-1}P(A)U, \end{aligned}$$

que es lo que se quería demostrar. De este resultado se puede deducir que cuando la matrices M y A son similares si

$$P(M) = 0,$$

entonces

$$P(A) = 0.$$

Lo que quiere decir que las matrices similares son raíces de los mismos polinomios.

De forma general, si x es una variable real y se define la función

$$\boxed{f(x) = \sum_{n\geq 0} a_n x^n, \qquad a_n \in \mathbb{R},}$$

también se puede definir la matriz

$$\boxed{f(M) = \sum_{n\geq 0} a_n M^n,}$$

donde

$$M^0 = I.$$

Se puede mostrar que si la matriz M es similar a la matriz A se cumple

$$\boxed{f(M) = U^{-1} f(A) U.}$$

La idea de matriz similar se puede encontrar en los trabajos de diferentes matemáticos como Cauchy. Sin embargo, el que notó la importancia del concepto y lo definió de manera formal fue George Frobenius en uno de sus trabajos en 1878.

F. George Frobenius (1849-1917) fue un matemático alemán que dio fundamentos sólidos a diferentes áreas de las matemáticas. Nació en el seno de una familia protestante. Realizó sus estudios en la Universidad de Berlín en donde fue alumno de Kronecker, Kummer y Weierstrass. Su tesis doctoral la realizó sobre soluciones a ecuaciones diferenciales bajo la asesoría de Weierstrass. Fue profesor de la Universidad de Berlín y de Eidgenössische Polytechnikum de Zürich. En estas universidades realizó trabajos relevantes en álgebra, funciones analíticas, geometría y teoría de grupos. En álgebra lineal, completó la prueba del Teorema de Caley-Hamilton y sus trabajos fueron importantes para dar solidez al concepto de transformación lineal. Como profesor fue director de tesis doctoral de matemáticos importantes como Issai Schur, con quien estudió la teoría de representaciones de grupos. Frobenius solía defender con pasión su enfoque de las matemáticas, por lo que tuvo varios conflictos con otros matemáticos, entre ellos Klein y Sophus Lie.

17.5. Ejercicios

1.- Con la matriz

$$U = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

y su inversa

$$U^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

a) mostrar que las siguientes matrices son equivalentes

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} -1 & -1 \\ -1 & 5 \end{pmatrix},$$

es decir, mostrar que

$$B = U^{-1}AU.$$

b) Obtener el determinante y la traza de las matrices A y B.

c) Considerando el polinomio

$$P(x) = 2 + 4x + 5x^2 - x^3$$

obtener $P(A)$ y $P(B)$.

Capítulo 18

Exponencial de una matriz

En este capítulo estudiaremos la exponencial de una matriz y daremos varios ejemplos.

18.1. Definición

Supongamos que x es una variable real y se define la función

$$\boxed{f(x) = \sum_{n\geq 0} a_n x^n,}$$

entonces, si M es una matriz de $n \times n$ se puede definir la matriz

$$\boxed{f(M) = \sum_{n\geq 0} a_n M^n,}$$

supondremos que

$$\boxed{M^0 = I,}$$

con I la matriz identidad de $n \times n$. En particular, para la función exponencial

$$e^x = \sum_{n\geq 0} \frac{1}{n!} x^n,$$

se puede definir la exponencial de una matriz

$$\boxed{e^M = \sum_{n\geq 0} \frac{1}{n!} M^n.}$$

En diferentes problemas de aplicaciones surgen expresiones de la forma

$$\boxed{e^{\alpha M} = \sum_{n\geq 0} \frac{\alpha^n}{n!} M^n}$$

en donde α es un escalar real o complejo.

18.2. Ejemplo 1

Antes de continuar recordemos las series de Taylor de las funciones coseno y seno

$$\boxed{\begin{aligned} \cos z &= \sum_{n\geq 0} \frac{(-)^n}{(2n)!} z^{2n}, \\ \sin z &= \sum_{n\geq 0} \frac{(-)^n}{(2n+1)!} z^{2n+1}. \end{aligned}}$$

Ahora consideremos la matriz

$$\boxed{M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},} \tag{18.1}$$

en este caso podemos plantear la exponencial

$$\boxed{e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\alpha},} \tag{18.2}$$

se puede observar que la serie de la exponencial se puede separar en potencias pares e impares, es decir

$$\boxed{\begin{aligned} e^{M\alpha} &= \sum_{n\geq 0} \frac{1}{n!} M^n \alpha^n \\ &= \sum_{n\geq 0} \frac{1}{(2n)!} M^{2n} \alpha^{2n} + \sum_{n\geq 0} \frac{1}{(2n+1)!} M^{2n+1} \alpha^{2n-1}. \end{aligned}}$$

Además se puede observar que

$$\begin{aligned} M^2 &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= (-)I, \end{aligned}$$

esto es

$$M^2 = (-)I.$$

Este resultado implica las igualdades

$$\boxed{\begin{aligned} M^{2n} &= (-)^n I, \\ M^{2n} M &= (-)^n I M = (-1)^n M. \end{aligned}}$$

Por lo tanto,

$$
\begin{aligned}
e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\alpha} &= \sum_{n\geq 0} \frac{1}{n!}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{n}\alpha^{n} \\
&= \sum_{n\geq 0} \frac{1}{(2n)!}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{2n}\alpha^{2n} + \sum_{n\geq 0} \frac{1}{(2n+1)!}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{2n+1}\alpha^{2n+1} \\
&= \sum_{n\geq 0} \frac{1}{(2n)!}(-)^{n} I\alpha^{2n} + \sum_{n\geq 0} \frac{1}{(2n+1)!}(-)^{n} M\alpha^{2n+1} \\
&= I\sum_{n\geq 0} \frac{1}{(2n)!}(-)^{n}\alpha^{2n} + M\sum_{n\geq 0} \frac{1}{(2n+1)!}(-)^{n}\alpha^{2n+1} \\
&= I\cos\alpha + M\sin\alpha \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\cos\alpha + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\sin\alpha \\
&= \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha t & \cos\alpha \end{pmatrix},
\end{aligned}
$$

en conclusión

$$
\boxed{e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\alpha} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.}
$$

Se puede observar que la matriz

$$
\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}
$$

representa las rotaciones en dos dimensiones. Por lo tanto, la exponencial (18.2) representa rotaciones en dos dimensiones y se dice que la matriz (18.1) es generadora de las rotaciones en dos dimensiones.

18.3. Matrices de rotación

Ahora estudiaremos las matrices

$$\boxed{M_1 = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad M_2 = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad M_3 = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.}$$

y veremos qué forma tienen las matrices

$$\boxed{\Lambda(\vec{\alpha}) = e^{-i\vec{\alpha}\cdot\vec{M}}.} \tag{18.3}$$

Primero observemos que si $n \geq 1$, se tiene

$$M_1^{2n} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = T_1, \quad M_1^{2n+1} = M_1, \tag{18.4}$$

$$M_2^{2n} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = T_2, \quad M_2^{2n+1} = M_2, \tag{18.5}$$

$$M_3^{2n} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = T_3, \quad M_3^{2n+1} = M_3, \tag{18.6}$$

es decir

$$\boxed{M_j^{2n} = T_j, \quad M_j^{2n+1} = M_j, \qquad j = 1, 2, 3.}$$

Considerando estos resultados, junto con las series de $\cos\alpha$ y $\sin\alpha$, se tiene

$$\begin{aligned}
e^{-i\alpha M_j} &= \sum_{n\geq 0}\frac{1}{n!}\left(-i\alpha M_j\right)^n \\
&= \sum_{n\geq 0}\frac{1}{(2n)!}\left(-i\alpha\right)^{2n}\left(M_j\right)^{2n} + \sum_{n\geq 0}\frac{1}{(2n+1)!}\left(-i\alpha\right)^{2n+1}\left(M_j\right)^{2n+1} \\
&= I + T_j\sum_{n\geq 1}\frac{1}{(2n)!}\left(-i\alpha\right)^{2n} + M_j\sum_{n\geq 0}\frac{1}{(2n+1)!}\left(-i\alpha\right)^{2n+1} \\
&= I - T_j + T_j + \sum_{n\geq 1}\frac{(-1)^n}{(2n)!}\alpha^{2n} - iM_j\sum_{n\geq 0}\frac{(-1)^n}{(2n+1)!}\alpha^{2n+1} \\
&= I - T_j + T_j\sum_{n\geq 0}\frac{(-)^n\alpha^{2n}}{(2n)!} - iM_j\sin\alpha \\
&= I - T_j + T_j\cos\alpha - iM_j\sin\alpha.
\end{aligned}$$

De donde, tomando en cuenta las matrices (18.4)-(18.6), tenemos

$$\boxed{\begin{aligned}
e^{-i\alpha M_1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, \\
e^{-i\alpha M_2} &= \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix}, \\
e^{-i\alpha M_3} &= \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}}$$

Así, $e^{-i\alpha M_1}$ representa una rotación sobre el eje x, $e^{-i\alpha M_2}$ representa una rotación sobre el eje y, mientras que $e^{-i\alpha M_3}$ representa una rotación sobre el eje z.

18.4. Matrices de Pauli

Las matrices de Pauli tienen aplicaciones en diversas áreas. Estas matrices se definen como

$$\boxed{\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.}$$

En este caso podemos ver que se cumple

$$\boxed{\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I,}$$

por lo que,

$$\boxed{\begin{aligned} \sigma_1^{2n} &= \sigma_2^{2n} = \sigma_3^{2n} = I, \\ \sigma_1^{2n+1} &= \sigma_1, \\ \sigma_2^{2n+1} &= \sigma_2, \\ \sigma_3^{2n+1} &= \sigma_3. \end{aligned}}$$

Así,

$$\begin{aligned} e^{i\alpha\sigma_1} &= \sum_{n\geq 0} \frac{(i\alpha)^n}{n!}\sigma_1^n \\ &= \sum_{n\geq 0} \frac{(i)^{2n}(\alpha)^{2n}}{(2n)!}\sigma_1^{2n} + \sum_{n\geq 0} \frac{(i)^{2n+1}(\alpha)^{2n+1}}{(2n+1)!}\sigma_1^{2n+1} \\ &= \sum_{n\geq 0} \frac{(-)^n(\alpha)^{2n}}{(2n)!} I + i \sum_{n\geq 0} \frac{(-)^n(\alpha)^{2n+1}}{(2n+1)!}\sigma_1 \\ &= \sum_{n\geq 0} \frac{(-)^n(\alpha)^{2n}}{(2n)!} I + i\sigma_1 \sum_{n\geq 0} \frac{(-)^n(\alpha)^{2n+1}}{(2n+1)!} \\ &= \cos\alpha I + i\sigma_1 \sin\alpha, \end{aligned}$$

es decir

$$\boxed{e^{i\alpha\sigma_1} = \cos\alpha I + i\sigma_1 \sin\alpha.}$$

De la misma forma, se obtiene

$$\boxed{\begin{aligned} e^{i\alpha\sigma_2} =& I\cos\alpha + i\sigma_2 \sin\alpha, \\ e^{i\alpha\sigma_3} =& I\cos\alpha + i\sigma_3 \sin\alpha. \end{aligned}}$$

Ahora, si $\hat{n} = (n_1, n_2, n_3)$ es un vector unitario, podemos definir la matriz

$$\boxed{\begin{aligned} \hat{n}\cdot\vec{\sigma} =& n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3 \\ =& \begin{pmatrix} n_3 & n_1 - in_2 \\ n_1 + in_2 & -n_3 \end{pmatrix}, \end{aligned}}$$

la cual cumple

$$\boxed{(\hat{n}\cdot\vec{\sigma})^2 = \begin{pmatrix} n_1^2 + n_2^2 + n_3^2 & 0 \\ 0 & n_1^2 + n_2^2 + n_3^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I,}$$

de donde

$$\boxed{(\hat{n}\cdot\vec{\sigma})^{2n} = I, \qquad (\hat{n}\cdot\vec{\sigma})^{2n+1} = \hat{n}\cdot\vec{\sigma}.}$$

Por lo tanto, si a es un escalar, tenemos que

$$\begin{aligned} e^{ia\hat{n}\cdot\vec{\sigma}} &= \sum_{n\geq 0} \frac{(ia)^n}{n!} (\hat{n}\cdot\vec{\sigma})^n \\ &= \sum_{n\geq 0} \frac{(i)^{2n} a^{2n}}{(2n)!} (\hat{n}\cdot\vec{\sigma})^{2n} + \sum_{n\geq 0} \frac{(i)^{2n+1} a^{2n+1}}{(2n+1)!} (\hat{n}\cdot\vec{\sigma})^{2n+1} \\ &= \sum_{n\geq 0} \frac{(-)^n a^{2n}}{(2n)!} I + i \sum_{n\geq 0} \frac{(-)^n a^{2n+1}}{(2n+1)!} \hat{n}\cdot\vec{\sigma} \\ &= I \sum_{n\geq 0} \frac{(-)^n a^{2n}}{(2n)!} + i\hat{n}\cdot\vec{\sigma} \sum_{n\geq 0} \frac{(-)^n a^{2n+1}}{(2n+1)!} \\ &= I\cos a + i\hat{n}\cdot\vec{\sigma}\sin a, \end{aligned}$$

es decir,

$$\boxed{e^{ia\hat{n}\cdot\vec{\sigma}} = I\cos a + i\hat{n}\cdot\vec{\sigma}\sin a.}$$

Wolfgang Pauli (1900-1958) fue un físico austríaco que realizó trabajos fundamentales para el desarrollo de la mecánica cuántica. A los 18 años terminó su licenciatura y después realizó el doctorado bajo la dirección de Arnold Sommerfeld en la Universidad de Múnich. Posteriomente realizó una estancia con Max Born de quien aprendió Álgebra Lineal y la empezó a usar en la mecánica cuántica. Las matrices de Pauli las obtuvo al estudiar propiedades del electrón que se descubrieron experimentalmente. En 1945 le fue otorgado el Premio Nobel por sus trabajos en mecánica cuántica. Cuando Hitler anexó a Austria a Alemania, Pauli decidió irse a Suiza. En Suiza, Pauli recibió información sobre el proyecto de la bomba atómica nazi y decidió alertar a la comunidad científica del peligro que existía con esa arma. Posteriormente, viajó a Estados Unidos y se convirtió en ciudadano norteamericano. Cuando terminó la segunda guerra mundial regresó a Suiza.

18.5. Exponencial de una matriz diagonal

Si D es un matriz de $n \times n$ diagonal, sus componentes se puedes escribir como

$$\boxed{D_{ij} = \lambda_i \delta_{ij}, \qquad i, j = 1, 2, \cdots, n.}$$

donde λ_i son las componentes de la diagonal de la matriz D.

Así, considerando las propiedades de la delta de Kronecker, encontramos que las componentes de D^2 son

$$\begin{aligned} D^2_{ij} &= (DD)_{ij} \\ &= D_{ik}D_{kj} \\ &= \lambda_i \delta_{ik} \lambda_k \delta_{kj} \\ &= \lambda_i \lambda_k \delta_{ik} \delta_{kj} \\ &= \lambda_i^2 \delta_{ij}, \end{aligned}$$

es decir,

$$\boxed{D^2_{ij} = \lambda_i^2 \delta_{ij}.}$$

En general, se obtiene

$$\begin{aligned} D^n_{ij} &= (DD\cdots D)_{ij} \\ &= D_{ik_1} D_{k_1k_2} D_{k_2k_3} \cdots D_{k_{n-1}j} \\ &= (\lambda_i \delta_{ik_1})(\lambda_{k_1}\delta_{k_1k_2})(\lambda_{k_2}\delta_{k_2k_3})\cdots(\lambda_{k_{n-1}}\delta_{k_{n-1}j}) \\ &= \lambda_i \lambda_{k_1}\lambda_{k_2}\cdots\lambda_{k_{n-1}}\delta_{ik_1}\delta_{k_1k_2}\delta_{k_2k_3}\cdots\lambda_{k_{n-1}j} \\ &= \lambda_i^n \delta_{ij}, \end{aligned}$$

por lo que

$$\boxed{D^n_{ij} = \lambda_i^n \delta_{ij}.}$$

Por lo tanto, si tenemos la matriz

$$f(D) = \sum_{n\geq 0} a_n D^n,$$

sus componentes son

$$\begin{aligned} (f(D))_{ij} &= \sum_{n\geq 0} a_n \left(D^n\right)_{ij} \\ &= \sum_{n\geq 0} a_n \lambda_i^n \delta_{ij} \\ &= \left(\sum_{n\geq 0} a_n \lambda_i^n\right)\delta_{ij} \\ &= f(\lambda_i)\delta_{ij}, \end{aligned}$$

de donde

$$\boxed{(f(D))_{ij} = f(\lambda_i)\delta_{ij}.}$$

Por ejemplo, si D es una matriz diagonal de 2×2, es decir

$$\boxed{D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},} \tag{18.7}$$

tenemos

$$\boxed{D^2 = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix},}$$

y para cualquier potencia se tiene

$$\boxed{D^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix},}$$

por lo tanto,

$$\boxed{f(D) = \sum \begin{pmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2) \end{pmatrix}.}$$

En particular, la exponencial de una matriz diagonal de 2×2 es

$$\boxed{e^{\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}} = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}.}$$

Como podemos ver, no es complicado obtener la exponencial de una matriz diagonal.

18.6. Exponencial de matrices similares

Por lo general es complicado calcular de forma explícita la exponencial de una matriz, pero si la matriz es diagonalizable, el problema no es tan complicado.

Supongamos que M es diagonalizable, entonces esta matriz es similar a la matriz diagonal

$$\boxed{D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.}$$

Donde los escalares λ_i son los valores propios de M. Así, existe una matriz U tal que se cumple

$$M = UDU^{-1},$$

de hecho la matriz U es la matriz de cambio de base para pasar a la base de los vectores propios de M.

Además, podemos observar que se cumple

$$\begin{aligned} M^n &= \underbrace{(UDU^{-1})(UDU^{-1})(UDU^{-1})\cdots(UDU^{-1})}_{n} \\ &= U\underbrace{DD\cdots D}_{n}U^{-1} \\ &= UD^nU^{-1}, \end{aligned}$$

esto es

$$M^n = UD^nU^{-1}.$$

Por lo que

$$\begin{aligned} e^M &= \sum_{n\geq 0}\frac{1}{n!}M^n \\ &= \sum_{n\geq 0}\frac{1}{n!}UD^nU^{-1} \\ &= U\left(\sum_{n\geq 0}\frac{1}{n!}D^n\right)U^{-1} \\ &= Ue^MU^{-1}, \end{aligned}$$

es decir

$$e^M = Ue^DU^{-1},$$

que se puede escribir como

$$\boxed{e^M = U\begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix}U^{-1}.}$$

18.6.1. Ejemplo

Por ejemplo, para la matriz

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

es complicado calcular directamente su exponencial. Sin embargo, podemos ver que los vectores propios de esta matriz son

$$\lambda_1 = 1, \qquad \lambda_2 = 2.$$

Por lo que la matriz M es similar a la matriz diagonal

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

Además usando los vectores propios de M, podemos construir las matriz de cambio de base

$$U = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$$

cuya inversa es

$$U^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix},$$

se puede probar que

$$D = U^{-1}MU,$$

por lo que

$$M = UDU^{-1},$$

es decir

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

Entonces

$$\begin{aligned} e^{\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}} &= \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} e^{\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}} \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & e^2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 2e - e^2 & 2e - 2e^2 \\ -e + e^2 & -e + 2e^2 \end{pmatrix}. \end{aligned}$$

18.7. Inversa de exponencial de una matriz

Supongamos que A es una matriz de $n \times n$ entradas, entonces podemos definir las dos exponenciales

$$\boxed{e^{A}, \qquad e^{-A},}$$

esta matrices son inversas. Para probar esta afirmación notemos que

$$\begin{aligned}
& e^{A}e^{-A} = \left(\sum_{n\geq 0}\frac{A^n}{n!}\right)\left(\sum_{m\geq 0}\frac{(-)^m A^m}{m!}\right) \\
= & \left(I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^n}{n!} + \cdots\right) \\
& \times\left(I - A + \frac{A^2}{2!} - \frac{A^3}{3!} + \cdots + (-)^n\frac{A^n}{n!} + \cdots\right) \\
= & \left(I + (-A + A) + \left(-\frac{A^2}{2!} + A^2 - \frac{A^2}{2!}\right)\right. \\
+ & \left(-\frac{A^3}{3!} + \frac{A^3}{2!} - \frac{A^3}{2!} + \frac{A^3}{3!}\right) + \cdots + \\
+ & A^n\left(\frac{(-)^n}{n!} + \frac{(-)^{n-1}}{(n-1)!} + \frac{(-)^{n-2}}{(n-2)!2!} + \cdots + \frac{(-)^{n-k}}{(n-k)!k!} + \cdots + \frac{1}{n!}\right) + \\
& \left. + \cdots\right),
\end{aligned}$$

además observe que

$$\begin{aligned}
\frac{(1-1)^n}{n!} &= \sum_{k=0}^{n}\frac{(-1)^{n-k}}{(n-k)!k!} \\
&= \left(\frac{(-)^n}{n!} + \frac{(-)^{n-1}}{(n-1)!} + \frac{(-)^{n-2}}{(n-2)!2!} + \cdots + \frac{(-)^{n-k}}{(n-k)!k!} + \cdots + \frac{1}{n!}\right).
\end{aligned}$$

Por lo tanto,

$$\boxed{e^{A}e^{-A} = I,}$$

así, la matriz inversa de e^{A} es e^{-A}.

18.8. Operaciones con Python

Si tenemos la matriz

$$A = \begin{pmatrix} -5 & 2 & 0 & 8 \\ 2 & -6 & 3 & 9 \\ -2 & -1 & 5 & -2 \\ 0 & 1 & 3 & -1 \end{pmatrix}$$

para obtener su exponencial podemos ocupar el siguiente código de Python

```python
import math
import numpy as np
from scipy.linalg import expm
A = np.array([[-5, 2,0,8], [2,-6,3,9], [-2, -1,5,-2],[0,
    1,3,-1]])
expm(A)
```

Listing 18.1: Exponencial de una matriz

con el cual se consigue la siguiente matriz

```
array([[-3.99479428,-4.87126039,3.33780779,-24.77763456],
[-3.05198432,-4.75947525,-3.86016899,-26.31391653],
[4.91414347,1.87132571,-29.50944907,1.17754784],
[-1.28927097,-2.7143525,-5.98103362,-16.11975232]])
```

18.9. Ejercicios

1.- Obtener la exponencial de la matriz

$$M = \begin{pmatrix} 6 & 1 \\ -2 & 3 \end{pmatrix}.$$

Capítulo 19

Sistema de ecuaciones diferenciales lineales de primer orden con coeficientes constantes

Como una aplicación de los temas que hemos visto en otros capítulos, veremos soluciones de sistema de ecuaciones diferenciales lineales de primer orden con coeficientes constantes.

19.1. Sistema de ecuaciones diferenciales

Para ver como surge una expresión de esta forma notemos que la ecuación diferencial ordinaria

$$\boxed{\frac{dx}{dt} = \alpha x,}$$

donde α es una constante, tiene como solución la función

$$\boxed{x(t) = e^{\alpha t} x_0,}$$

en donde x_0 es una constante que expresa una condición inicial.

Ahora consideremos el sistema lineal de ecuaciones diferenciales ordinarias

$$\boxed{\begin{aligned} \frac{dx}{dt} &= ax + by, \qquad (19.1)\\ \frac{dy}{dt} &= cx + dy, \qquad (19.2)\end{aligned}}$$

en donde a, b, c, d son constantes reales. Podemos observar que usando las definiciones

$$\boxed{X = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},}$$

el sistema de ecuaciones diferenciales ordinarias (19.1)-(19.2) se puede expresar de una forma más sugerente

$$\boxed{\frac{dX}{dt} = MX.}$$

Por lo que podemos proponer la solución

$$\boxed{X(t) = e^{Mt}X_0, \qquad X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}}$$

con X_0 es una condiciones inicial. De forma explícita, la solución tiene la forma

$$\boxed{\begin{pmatrix} x \\ y \end{pmatrix} = e^{Mt}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix},}$$

en donde

$$\boxed{e^{Mt} = e^{\begin{pmatrix} a & b \\ c & d \end{pmatrix}t}.}$$

Así, para resolver un sistema lineal de ecuaciones diferenciales ordinarias es necesario saber calcular la exponencial de una matriz.

19.1.1. Ejemplo 1

Por ejemplo, consideremos el sistema de ecuaciones lineales

$$\begin{aligned} \dot{x} &= \beta y, && (19.3) \\ \dot{y} &= -\beta x, && (19.4) \end{aligned}$$

que se puede escribir como

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Por lo tanto, la solución es

$$\begin{pmatrix} x \\ y \end{pmatrix} = e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \beta t} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Antes de continuar, observemos que la serie de la exponencial se puede escribir como

$$\begin{aligned} e^{M\beta t} &= \sum_{n\geq 0} \frac{1}{n!} M^n (\beta t)^n \\ &= \sum_{n\geq 0} \frac{1}{(2n)!} M^{2n} (\beta t)^{2n} + \sum_{n\geq 0} \frac{1}{(2n+1)!} M^{2n+1} (\beta t)^{2n-1}, \end{aligned}$$

supondremos que $M^0 = I$, con I la matriz identidad de 2×2. También recordemos las series

$$\begin{aligned} \cos z &= \sum_{n\geq 0} \frac{(-)^n}{(2n)!} z^{2n}, \\ \sin z &= \sum_{n\geq 0} \frac{(-)^n}{(2n+1)!} z^{2n+1}. \end{aligned}$$

Ahora definamos la matriz

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

de donde encontramos que se cumple

$$\begin{aligned} M^2 &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= (-)I, \end{aligned}$$

es decir,

$$M^2 = (-)I.$$

Este resultado implica la expresión

$$M^{2n} = (-)^n I.$$

También se encuentra

$$\begin{aligned} M^{2n+1} &= M^{2n}M \\ &= (-)^n IM \\ &= (-1)^n M, \end{aligned}$$

esto es

$$M^{2n+1} = (-1)^n M.$$

Por lo tanto,

$$\begin{aligned} e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\beta t} &= \sum_{n\geq 0} \frac{1}{n!} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^n (\beta t)^n \\ &= \sum_{n\geq 0} \frac{1}{(2n)!} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{2n} (\beta t)^{2n} + \sum_{n\geq 0} \frac{1}{(2n+1)!} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{2n+1} (\beta t)^{2n+1} \\ &= \sum_{n\geq 0} \frac{1}{(2n)!}(-)^n I(\beta t)^{2n} + \sum_{n\geq 0} \frac{1}{(2n+1)!}(-)^n M(\beta t)^{2n+1} \\ &= I\sum_{n\geq 0} \frac{1}{(2n)!}(-)^n (\beta t)^{2n} + M\sum_{n\geq 0} \frac{1}{(2n+1)!}(-)^n (\beta t)^{2n+1} \\ &= I\cos\beta t + M\sin\beta t \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\cos\beta t + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\sin\beta t \\ &= \begin{pmatrix} \cos\beta t & \sin\beta t \\ -\sin\beta t & \cos\beta t \end{pmatrix}, \end{aligned}$$

es decir

$$e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\beta t} = \begin{pmatrix} \cos\beta t & \sin\beta t \\ -\sin\beta t & \cos\beta t \end{pmatrix}.$$

Por lo tanto la solución al sistema de ecuaciones diferenciales ordinarias (19.3)-(19.4) es

$$
\begin{aligned}
\begin{pmatrix} x \\ y \end{pmatrix} &= e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\beta t} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \cos\beta t & \sin\beta t \\ -\sin\beta t & \cos\beta t \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\
&= \begin{pmatrix} x_0\cos\beta t + y_0 \sin\beta t \\ -x_0 \sin\beta t + y_0 \cos\beta t \end{pmatrix},
\end{aligned}
$$

es decir

$$
\begin{aligned}
x &= x_0\cos\beta t + y_0\sin\beta t, \\
y &= -x_0\sin\beta t + y_0\cos\beta t.
\end{aligned}
$$

19.2. Ejemplo 2

Consideremos el sistema de ecuaciones lineales

$$
\begin{aligned}
\dot{x} &= -2\beta y, && (19.5) \\
\dot{y} &= \beta x + 3\beta y, && (19.6)
\end{aligned}
$$

que se puede escribir como

$$
\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \beta \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.
$$

Por lo tanto, la solución es

$$
\begin{pmatrix} x \\ y \end{pmatrix} = e^{\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}\beta t} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.
$$

Los valores propios de la matriz

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

son

$$
\lambda_1 = 2, \qquad \lambda_2 = 1
$$

mientras que los vectores propios están dados por

$$
v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}
$$

Entonces, la matriz de cambio de base y su inversa son

$$\begin{aligned} U &= \begin{pmatrix} 1 & 1 \\ -1 & -\frac{1}{2} \end{pmatrix}, \\ U^{-1} &= \begin{pmatrix} -1 & -2 \\ 2 & 2 \end{pmatrix}. \end{aligned}$$

Por lo que

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = U^{-1}MU,$$

de donde

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

Así,

$$\begin{aligned} e^{\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}\beta t} &= Ue^{\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\beta t} U^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} e^{2\beta t} & 0 \\ 0 & e^{\beta t} \end{pmatrix} \begin{pmatrix} -1 & -2 \\ 2 & 2 \end{pmatrix} \\ &= \begin{pmatrix} -e^{2\beta t} + 2e^{\beta t} & -2e^{2\beta t} + 2e^{\beta t} \\ e^{2\beta t} - e^{\beta t} & 2e^{2\beta t} - e^{\beta t} \end{pmatrix}. \end{aligned}$$

De donde, las soluciones del sistema de ecuaciones (19.5)-(19.6) son

$$\begin{aligned} x &= \left(2e^{\beta t} - e^{2\beta t}\right) x_0 + 2\left(e^{2\beta t} - e^{\beta t}\right) y_0, \\ y &= \left(e^{2\beta t} - e^{\beta t}\right) x_0 + \left(2e^{2\beta t} - e^{\beta t}\right) y_0. \end{aligned}$$

19.3. Ejemplo 3

Ahora veamos como resolver el sistema de ecuaciones

$$\begin{aligned} \dot{x} &= \beta x + \beta y, && (19.7) \\ \dot{y} &= 4\beta x + \beta y. && (19.8) \end{aligned}$$

Primero definamos

$$X = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad M = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix},$$

por lo que el sistema de ecuaciones (19.7)-(19.8) toma la forma

$$\frac{dX}{dt} = \beta M X. \tag{19.9}$$

Los valores propios de M son

$$\lambda_1 = 3, \qquad \lambda_2 = -1,$$

cuyos vectores propios son

$$v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Además, la matriz de cambio de base y su inversa son

$$\begin{aligned} U &= \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}, \\ U^{-1} &= \frac{1}{4}\begin{pmatrix} 2 & 1 \\ 2 & -1 \end{pmatrix} \end{aligned}$$

y se cumple

$$D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} = U^{-1} M U.$$

Ahora, definiendo

$$\bar{X} = U^{-1} X = \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix},$$

el sistema de ecuaciones (19.9) toma la forma

$$\frac{d(U^{-1}X)}{dt} = \beta (U^{-1} M U)(U^{-1} X),$$

es decir

$$\frac{d\bar{X}}{dt} = \beta D \bar{X}.$$

Note que en este caso las condiciones iniciales están dadas por

$$\bar{X}_0 = \begin{pmatrix} \bar{x}_0 \\ \bar{y}_0 \end{pmatrix} = U^{-1} X_0.$$

Este sistema de ecuaciones tiene la forma

$$\frac{d}{dt}\left(\begin{array}{c}\bar{x}\\ \bar{y}\end{array}\right)=\beta\left(\begin{array}{cc}3 & 0\\ 0 & -1\end{array}\right)\left(\begin{array}{c}\bar{x}\\ \bar{y}\end{array}\right),$$

lo que implica

$$\begin{aligned}\frac{d\bar{x}}{dt} &= 3\beta\bar{x},\\ \frac{d\bar{y}}{dt} &= -\beta\bar{y}.\end{aligned}$$

Cuyas soluciones son

$$\begin{aligned}\bar{x} &= e^{3\beta t}\bar{x}_0,\\ \bar{y} &= e^{-\beta t}\bar{y}_0.\end{aligned}$$

Entonces

$$\begin{aligned}X &= U\bar{X}=U\left(\begin{array}{c}\bar{x}\\ \bar{y}\end{array}\right)=U\left(\begin{array}{c}e^{3\beta t}\bar{x}_0\\ e^{-\beta t}\bar{y}_0\end{array}\right)\\ &= U\left(\begin{array}{cc}e^{3\beta t} & 0\\ 0 & e^{-\beta t}\end{array}\right)\left(\begin{array}{c}\bar{x}_0\\ \bar{y}_0\end{array}\right)\\ &= U\left(\begin{array}{cc}e^{3\beta t} & 0\\ 0 & e^{-\beta t}\end{array}\right)U^{-1}\left(\begin{array}{c}x_0\\ y_0\end{array}\right),\end{aligned}$$

es decir,

$$X = \frac{1}{4}\left(\begin{array}{cc}1 & 1\\ 2 & -2\end{array}\right)\left(\begin{array}{cc}e^{3\beta t} & 0\\ 0 & e^{-\beta t}\end{array}\right)\left(\begin{array}{cc}2 & 1\\ 2 & -1\end{array}\right)\left(\begin{array}{c}x_0\\ y_0\end{array}\right).$$

Así las soluciones del sistema de ecuaciones (19.7)-(19.8) son

$$\begin{aligned}x &= \frac{1}{2}\left(e^{3\beta t}+e^{-\beta t}\right)x_0+\frac{1}{4}\left(e^{3\beta t}-e^{-\beta t}\right)y_0,\\ y &= \frac{1}{2}\left(e^{3\beta t}+e^{-\beta t}\right)y_0+\left(e^{3\beta t}-e^{-\beta t}\right)x_0.\end{aligned}$$

19.4. Sistema de osciladores acoplados

Antes de iniciar esta sección recordemos que la ecuación diferencial

$$\frac{d^2z}{dt^2}=-\omega^2 z$$

tiene como solución

$$z = z_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t.$$

Ahora, un ejemplo de sistema de dos osciladores acoplados está descrito por el sistema de ecuaciones diferenciales

$$\begin{aligned} m\frac{d^2x}{dt^2} &= -k_1 x + k_2(y - x), \\ m\frac{d^2y}{dt^2} &= -k_1 y - k_2(y - x). \end{aligned}$$

Este sistema de ecuaciones se puede escribir como

$$m\frac{d^2x}{dt^2} = -(k_1 + k_2)x + k_2 y, \tag{19.10}$$

$$m\frac{d^2y}{dt^2} = k_2 x - (k_1 + k_2)y. \tag{19.11}$$

Así, definiendo

$$X = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad M = \begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_1 + k_2 \end{pmatrix},$$

el sistema de ecuaciones (19.10)-(19.11) se puede escribir como

$$m\frac{d^2X}{dt^2} = -MX \tag{19.12}$$

Además, se puede probar que los valores propios de M son

$$\lambda_1 = k_1, \qquad \lambda_2 = k_1 + 2k_2,$$

cuyos vectores propios normalizados son

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Con estos vectores propios, la matriz de cambio de base y su inversa toman la forma

$$U = U^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

y se cumple

$$D = \begin{pmatrix} k & 0 \\ 0 & k + 2k_2 \end{pmatrix} = U^{-1} M U.$$

Ahora, usando la definición

$$\bar{X} = U^{-1}X = \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}, \tag{19.13}$$

el sistema de ecuaciones (19.12) toma la forma

$$m\frac{d^2(U^{-1}X)}{dt^2} = -(U^{-1}MU)(U^{-1}X),$$

es decir

$$m\frac{d^2\bar{X}}{dt^2} = -D\bar{X}.$$

Por lo tanto, tenemos la ecuaciones diferenciales

$$\begin{aligned} m\frac{d^2\bar{x}}{dt^2} &= -k_1\bar{x}, \\ m\frac{d^2\bar{y}}{dt^2} &= -(k_1+2k_2)\bar{y}, \end{aligned}$$

cuyas soluciones son

$$\bar{x} = \bar{x}_0\cos\omega_1 t + \frac{\bar{v}_{0x}}{\omega_1}\sin\omega_1 t, \tag{19.14}$$

$$\bar{y} = \bar{y}_0\cos\omega_2 t + \frac{\bar{v}_{0y}}{\omega_2}\sin\omega_2 t, \tag{19.15}$$

con

$$\omega_1^2 = \frac{k_1}{m}, \qquad \omega_2^2 = \frac{k_1+2k_2}{m}.$$

Observe que las soluciones (19.14)-(19.15) se pueden escribir como

$$\bar{X} = \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} \cos\omega_1 t & 0 \\ 0 & \cos\omega_2 t \end{pmatrix}\begin{pmatrix} \bar{x}_0 \\ \bar{y}_0 \end{pmatrix} + \begin{pmatrix} \frac{\sin\omega_1 t}{\omega_1} & 0 \\ 0 & \frac{\sin\omega_2 t}{\omega_2} \end{pmatrix}\begin{pmatrix} \bar{v}_{0x} \\ \bar{v}_{0y} \end{pmatrix}$$

y considerando la identidad (19.13), se obtiene

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= U\begin{pmatrix} \cos\omega_1 t & 0 \\ 0 & \cos\omega_2 t \end{pmatrix}U^{-1}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\ &\quad + U\begin{pmatrix} \frac{\sin\omega_1 t}{\omega_1} & 0 \\ 0 & \frac{\sin\omega_2 t}{\omega_2} \end{pmatrix}U^{-1}\begin{pmatrix} v_{0x} \\ v_{0y} \end{pmatrix}, \end{aligned}$$

por lo tanto,

$$\begin{aligned}
x &= \frac{1}{2}\left(\cos\omega_1 t + \cos\omega_2 t\right) x_0 + \frac{1}{2}\left(\cos\omega_1 t - \cos\omega_2 t\right) y_0 \\
&\quad + \frac{1}{2}\left(\frac{\sin\omega_1 t}{\omega_1} + \frac{\sin\omega_2 t}{\omega_2}\right) v_{0x} + \frac{1}{2}\left(\frac{\sin\omega_1 t}{\omega_1} - \frac{\sin\omega_2 t}{\omega_2}\right) v_{0y}, \\
y &= \frac{1}{2}\left(\cos\omega_1 t + \cos\omega_2 t\right) y_0 + \frac{1}{2}\left(\cos\omega_1 t - \cos\omega_2 t\right) x_0 \\
&\quad + \frac{1}{2}\left(\frac{\sin\omega_1 t}{\omega_1} + \frac{\sin\omega_2 t}{\omega_2}\right) v_{0y} + \frac{1}{2}\left(\frac{\sin\omega_1 t}{\omega_1} - \frac{\sin\omega_2 t}{\omega_2}\right) v_{0x}.
\end{aligned}$$

El desarrollo del álgebra lineal y de los sistemas de ecuaciones diferenciales están íntimamente correlacionados. Por ejemplo, Euler (1707-1783) muestra que una ecuación ordinaria de orden n de coeficientes constantes se puede reducir a un sistema de n ecuaciones diferenciales de primer orden y este sistema de n ecuaciones ordinarias se puede relacionar con un problema de valores propios. Asimismo, D' Alembert (1717-1783) mostró que el problema de las ecuaciones de movimiento para osciladores acoplados se puede reducir a un problema de valores propios. Vale la pena mencionar que Lagrange (1736-1813) ocupó métodos que hoy en día se reconocen como del álgebra lineal para estudiar problemas de mecánica celeste. Actualmente sabemos que diferentes problemas de ecuaciones diferenciales, ordinarias y parciales, en realidad son problemas del álgebra lineal en el espacio vectorial de funciones.

Jean le Rond D' Alembert (1717-1783) fue un físico-matemático y filosofo francés que realizó aportaciones fundamentales a la mecánica, al álgebra y a las ecuaciones diferenciales. Siendo muy joven escribió un libro sobre cálculo integral y realizó una demostración del teorema fundamental del álgebra. Además, se interesó en dar solución a problemas de mecánica celeste, dinámica de fluidos, del movimiento de una cuerda vibrante y de diferentes ecuaciones diferenciales. Como profesor, fue asesor doctoral de Laplace. En su papel de intelectual se interesó en que el conocimiento saliera de la academia y se divulgara en la sociedad. Junto con Denis Diderot (1713-1784) emprendió el proyecto llamado la *Enciclopedia*, una obra escrita en la cual se intentó colocar todo el conocimiento que se tenía hasta esos días. La convicción de D' Alembert y Diderot era que el conocimiento se divulgara en la sociedad para el desarrollo humano. En ese proyecto también participaron personajes como Voltaire, Montesquieu, Jacques Rousseau y Adam Smith. La *Enciclopedia* fue un gran éxito editorial y causó revuelo en distintos ámbitos sociales de su época. Debido a que D' Alembert era un ateo declarado, la Iglesia y la Monarquía francesa presionaron para que D' Alembert saliera del proyecto. D' Alembert fue controvertido y polémico desde su nacimiento hasta el último día de su vida. En efecto, fue hijo natural de dos miembros de la élite francesa de la época, pero nunca fue reconocido por sus padres y poco tiempo después de nacer su madre lo abandonó en la puerta de la iglesia de Saint-Jean-le Rond, de esa iglesia tomó su nombre. El niño D' Alembert fue adoptado por una mujer humilde que tomó el papel de su madre. Cabe señalar que hasta el final de sus días sostuvo su condición de ateo por lo que fue enterrado en una tumba común sin nombre.

19.5. Ejercicios

1.- Resolver el sistema de ecuaciones diferenciales ordinarias

$$\begin{aligned} \dot{x} &= 2x - y, \\ \dot{y} &= -2x + 3y. \end{aligned}$$

Capítulo 20

Relación entre el determinante y la traza de una matriz

En este capítulo mostraremos que el determinante de una matriz se puede relacionar con su traza y daremos algunos ejemplos de esta relación.

20.1. Relación

Supongamos que A es una matriz de $n \times n$ diagonalizable. Entonces existe una matriz U invertible de tal forma que se cumple

$$\boxed{A = U \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} U^{-1},}$$

donde $\lambda_i, i = 1, 2, \cdots, n$ son los valores propios de A. Se puede mostrar que si $f(x)$ es una función de la forma

$$\boxed{f(x) = \sum_{n \geq 0} a_n x^n,}$$

entonces

$$\boxed{f(A) = U \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix} U^{-1},}$$

por lo que

$$\boxed{Tr(f(A)) = \sum_{i=1}^{n} f(\lambda_i).}$$

Además, como A es diagonizable se cumple

$$\det A = \Pi_{i=1}^{n} \lambda_i,$$

de esta expresión se encuentra

$$\begin{aligned} \det A &= \Pi_{i=1}^{n} \lambda_i \\ &= e^{\ln\left(\Pi_{i=1}^{n} \lambda_i\right)} \\ &= e^{\sum_{i=1}^{n} \ln \lambda_i} \\ &= e^{Tr(\ln A)}, \end{aligned}$$

es decir,

$$\boxed{\det A = e^{Tr(\ln A)}.} \tag{20.1}$$

Por lo que el determinante de la matriz A está relacionado con su traza.

20.2. Determinante de la exponencial de una matriz

Supongamos que la matriz A es de la forma

$$\boxed{A = e^{M}}$$

donde M es una matriz de $n \times n$, entonces

$$\begin{aligned}
\det A &= e^{Tr(\ln A)} \\
&= e^{Tr(\ln e^{M})} \\
&= e^{Tr(M)},
\end{aligned}$$

por lo que

$$\boxed{\det\left(e^{M}\right) = e^{Tr(M)}.}$$

Por ejemplo, para la matriz

$$e^{\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}}$$

tenemos que

$$\det\left(e^{\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}}\right) = e^{3}.$$

Note que si M es una matriz antisimétrica de $n \times n$ se cumple $Tr(M) = 0$, entonces

$$\det\left(e^{M}\right) = 1.$$

20.3. Ejercicios

1.- Obtener el determinante de la matriz

$$M = e^{\begin{pmatrix} -3 & 1 & 2 \\ 2 & 1 & 0 \\ 1 & 1 & -2 \end{pmatrix}}.$$

2.- Si la matriz M se puede diagonalizar y $Tr(M) = 0$, determinar el valor de

$$\det\left(e^{M}\right).$$

Capítulo 21

Operador adjunto

En este capítulo daremos el concepto operador adjunto y veremos algunos ejemplos. Este concepto nos ayudará a obtener los valores propios de diferentes transformaciones lineales.

21.1. Definición

Dado un operador A lineal definiremos el operador adjunto, $A^{\dagger}$, como el operador que satisface

$$\boxed{\langle Av|u\rangle = \left\langle v|A^{\dagger}u\right\rangle,}$$

con u y v dos vectores arbitrarios.

21.1.1. Matrices

Por ejemplo, para $\mathbb{C}^n$ se tiene

$$\begin{aligned}
\langle Av|u\rangle &= (Av)^{*T}u \\
&= v^{*T}A^{*T}u \\
&= \left\langle v|A^{\dagger}u\right\rangle \\
&= v^{*T}A^{\dagger}u,
\end{aligned}$$

de donde, para una matriz cuadrada con entradas complejas, la matriz adjunta es

$$\boxed{A^{\dagger} = A^{*T}.} \tag{21.1}$$

21.1.2. Derivada

Para el espacio vectorial de las funciones, el adjunto de un operador depende fuertemente del dominio, las condiciones de borde que se satisfacen y del producto escalar. En el espacio de las funciones suaves e integrables $\psi(x)$ definidas en el intervalo $[a, b]$ y que cumplen $\psi(a) = \psi(b) = 0$. se puede definir el operador

$$\boxed{A = \alpha \frac{\partial}{\partial x},}$$

con α un número complejo. Veamos cuál es el operador adjunto de este operador. Para ello, tomaremos el producto escalar

$$\langle \psi_1 | \psi_2 \rangle = \int_a^b \psi_1^*(x) \psi_2(x) dx. \tag{21.2}$$

En este caso, se puede notar que

$$\begin{aligned}
\langle A\psi_1 | \psi_2 \rangle &= \int_a^b dx (A\psi_1(x))^* \psi_2(x) \\
&= \int_a^b dx \left(\alpha \frac{\partial}{\partial x} \psi_1(x) \right)^* \psi_2(x) \\
&= \alpha^* \int_a^b dx \frac{\partial \psi_1^*(x)}{\partial x} \psi_2(x) \\
&= \alpha^* \int_a^b dx \left(\frac{\partial \left(\psi_1^*(x)\psi_2(x)\right)}{\partial x} - \psi_1^*(x) \frac{\partial \psi_2(x)}{\partial x} \right) \\
&= \alpha^* \left(\psi_1^*(x)\psi_2(x)\right) \Big|_a^b + \int_a^b dx \psi_1^*(x) \left(-\alpha^* \frac{\partial \psi_2(x)}{\partial x} \right) \\
&= \int_a^b dx \psi_1^*(x) \left(-\alpha^* \frac{\partial}{\partial x} \right) \psi_2(x),
\end{aligned}$$

esto es

$$\langle A\psi_1 | \psi_2 \rangle = \int_a^b dx \psi_1^*(x) \left(-\alpha^* \frac{\partial}{\partial x} \right) \psi_2(x).$$

Ahora

$$\left\langle \psi_1 | A^\dagger \psi_2 \right\rangle = \int_a^b dx \psi_1^*(x) A^\dagger \psi_2(x),$$

que implica la igualdad

$$\boxed{\left(\alpha\frac{\partial}{\partial x}\right)^{\dagger} = -\alpha^{*}\frac{\partial}{\partial x}.} \tag{21.3}$$

Note que este resultado depende fuertemente de que se cumpla $\psi(a) = \psi(b) = 0$.

21.1.3. Derivada con peso

Se puede observar que el adjunto del operador

$$A = \alpha\frac{\partial}{\partial x}$$

no está bien definido con un producto escalar general

$$\langle f|g\rangle = \int_a^b dx q(x) f^{*}(x) g(x), \qquad q(x) > 0$$

donde se satisface

$$q(x) > 0 \qquad \forall a \in [a, b]. \tag{21.4}$$

Para este caso es más conveniente considerar el operador

$$\boxed{\tilde{A} = \frac{\alpha}{q(x)}\frac{\partial}{\partial x},}$$

quien sí tiene bien definido su adjunto. En efecto,

$$\begin{aligned}
\langle \tilde{A}\psi_1|\psi_2\rangle &= \int_a^b dx q(x) (\tilde{A}\psi_1(x))^{*}\psi_2(x) \\
&= \int_a^b dx q(x) \left(\frac{\alpha}{q(x)}\frac{\partial}{\partial x}\psi_1(x)\right)^{*}\psi_2(x) \\
&= \alpha^{*}\int_a^b dx \frac{\partial\psi_1^{*}(x)}{\partial x}\psi_2(x) \\
&= \int_a^b dx \psi_1^{*}(x)\left(-\alpha^{*}\frac{\partial}{\partial x}\right)\psi_2(x) \\
&= \int_a^b dx q(x)\psi_1^{*}(x)\left(\frac{-\alpha^{*}}{q(x)}\frac{\partial}{\partial x}\right)\psi_2(x)
\end{aligned}$$

es decir,

$$\left\langle \tilde{A}\psi_1|\psi_2 \right\rangle = \int_a^b dxq(x)\psi_1^*(x)\left(\frac{-\alpha^*}{q(x)}\frac{\partial}{\partial x}\right)\psi_2(x)$$

Además, se debe cumplir

$$\left\langle \psi_1|\tilde{A}^\dagger\psi_2 \right\rangle = \int_a^b dxq(x)\psi_1^*(x)\tilde{A}^\dagger\psi_2(x),$$

por lo cual se obtiene

$$\boxed{\tilde{A}^\dagger = -\frac{\alpha^*}{q(x)}\frac{\partial}{\partial x}.}$$

21.1.4. Propiedades del operador adjunto

Hay dos propiedades importantes de los operadores adjuntos. La primera propiedad está relacionada con las suma. Si A, B son dos operadores lineales, se tiene

$$\langle (A+B)\, v|u\rangle = \left\langle v|\, (A+B)^\dagger\, u \right\rangle,$$

pero

$$\begin{aligned}
\langle (A+B)\, v|u\rangle &= \langle (Av+Bv)\,|u\rangle \\
&= \langle Av|u\rangle + \langle Bv|u\rangle \\
&= \langle v|A^\dagger u\rangle + \langle v|B^\dagger u\rangle \\
&= \langle v|\left(A^\dagger + B^\dagger\right) u\rangle .
\end{aligned}$$

Por lo tanto,

$$\left\langle v|\, (A+B)^\dagger\, |u \right\rangle = \langle v|\left(A^\dagger + B^\dagger\right) u\rangle ,$$

este resultado es válido para cualquier par de vectores v y u, entonces

$$\boxed{(A+B)^\dagger = A^\dagger + B^\dagger.} \tag{21.5}$$

La segunda propiedad está relacionada con el producto de dos operadores:

$$\begin{aligned}
\langle (AB)\, v|u\rangle &= \langle A\,(Bv)\,|u\rangle \\
&= \langle Bv|A^\dagger u\rangle \\
&= \langle v|B^\dagger A^\dagger u\rangle \\
&= \left\langle v|\, (AB)^\dagger\, u \right\rangle ,
\end{aligned}$$

que implica

$$\boxed{(AB)^\dagger = B^\dagger A^\dagger.} \tag{21.6}$$

21.2. Ejercicios

1.- Obtener el operador adjunto de la matriz

$$M = \begin{pmatrix} 3 & -i & 1 \\ 4+2i & -1 & i \\ 1 & 1+2i & 1 \end{pmatrix}.$$

2.- Obtener el operador adjunto de la matriz

$$M = \begin{pmatrix} 1 & 1+2i & 3 \\ 1-2i & 2 & -i \\ 3 & i & 4 \end{pmatrix}.$$

3.- Obtener el operador adjunto de la matriz

$$M = \begin{pmatrix} 3 & 4-2i & 1 \\ -i & -1 & 1-2i \\ 1 & -i & 1-2i \end{pmatrix}.$$

Capítulo 22

Operadores hermíticos

En este capítulo veremos la definición de operador hermítico y algunas de sus propiedades. Los operadores hermíticos tienen aplicaciones en diversas áreas como la mecánica cuántica.

22.1. Definición

Una clase importante de operadores son los autoadjuntos, que satisfacen

$$\boxed{A^\dagger = A,}$$

a estos operadores también se les llama hermíticos o hermitianos.

Ahora recordemos que en el espacio de la matrices de entradas complejas de $n \times n$ el operador adjunto de una matriz Λ es

$$\Lambda^\dagger = \Lambda^{*T}$$

Por lo que si Λ es una matriz hermítica se cumple

$$\boxed{(\Lambda)^{*T} = \Lambda,} \tag{22.1}$$

Un ejemplo de matriz hermítica es la matriz identidad

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

además las matrices

$$
\begin{aligned}
M_1 &= i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \\
M_2 &= i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \\
M_3 &= i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

son hermíticas.

Note que si la matriz Λ solo tiene coeficientes reales, entonces la condición (22.1) toma la forma

$$\boxed{(\Lambda)^T = \Lambda,} \tag{22.2}$$

así las matrices simétricas de coeficientes reales son hermíticas. Por ejemplo las matrices

$$
\begin{aligned}
A &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
B &= \begin{pmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 5 \end{pmatrix}, \\
C &= \begin{pmatrix} 9 & 2 & 0 \\ 2 & 6 & -1 \\ 0 & -1 & 5 \end{pmatrix}
\end{aligned}
$$

son simétricas y por lo tanto hermíticas.

22.2. Valores propios de operadores hermíticos

Los operadores Hermíticos tienen valores propios reales. En efecto, si v es un vector propio de un operador hermítico A con valor propio λ, es decir, si se satisface

$$Av = \lambda v,$$

se tiene

$$\begin{aligned}\langle Av|v\rangle &= \langle \lambda v|v\rangle \\ &= \lambda^* \langle v|v\rangle ,\end{aligned}$$

pero también tenemos

$$\begin{aligned}\langle Av|v\rangle &= \langle v|A^\dagger v\rangle \\ &= \langle v|Av\rangle \\ &= \langle v|\lambda v\rangle \\ &= \lambda \langle v|v\rangle ,\end{aligned}$$

que implica la igualdad

$$\lambda^* \langle v|v\rangle = \lambda \langle v|v\rangle$$

que nos conduce al resultado

$$\boxed{\lambda^* = \lambda.}$$

Así, los valores propios de los operadores hermíticos son reales.

De la ecuación (21.5) se puede ver que la suma de dos operadores hermíticos es otro operador hermítico. Además, de la expresión (21.6) es claro que si A y B son operadores hermíticos y **conmutan**, es decir $AB = BA$, entonces

$$(AB)^\dagger = B^\dagger A^\dagger = BA = AB.$$

así

$$\boxed{(AB)^\dagger = AB.}$$

Por lo tanto, el producto de dos operadores hermíticos que **conmutan** es hermítico.

22.3. Ortogonalidad de vectores propios de matrices hermíticas

Sea M una matriz hermítica de $n\times n$, note que en este caso todos los valores propios de M son reales. Ahora, probaremos que si v_a y v_b son vectores propios de M con valores propios λ_a y λ_b diferentes, entonces v_a y v_b son ortogonales, es decir

$$\langle v_a | v_b \rangle = 0.$$

Para probar esta afirmación, su pongamos que se cumplen la ecuaciones

$$\begin{aligned} M v_1 &= \lambda_1 v_1, \\ M v_2 &= \lambda_2 v_2. \end{aligned}$$

Considerando estas dos ecuaciones y que M es una matriz es hermítica, se tiene

$$\begin{aligned} \lambda_1 \langle v_1 | v_2 \rangle &= \langle \lambda_1 v_1 | v_2 \rangle \\ &= \langle M v_1 | v_2 \rangle \\ &= \langle v_1 | M^\dagger v_2 \rangle \\ &= \langle v_1 | M v_2 \rangle \\ &= \langle v_1 | \lambda_2 v_2 \rangle \\ &= \lambda_2 \langle v_1 | v_2 \rangle, \end{aligned}$$

esto es

$$\lambda_1 \langle v_1 | v_2 \rangle = \lambda_2 \langle v_1 | v_2 \rangle,$$

de donde

$$(\lambda_1 - \lambda_2) \langle v_1 | v_2 \rangle = 0.$$

Ahora, si

$$\lambda_a - \lambda_b \neq 0,$$

se llega a

$$\langle v_1 | v_2 \rangle = 0.$$

que es lo que queríamos demostrar.

Observe que si todos los valores propios de M son diferentes, entonces todos vectores propios de M son ortogonales entre si. En otras palabras, si todos los n valores propios de M son diferentes, entonces con los vectores propios de M se puede formar una base ortonormal. Note que en este caso la matriz de cambio de base formada con los vectores propios de M tendrá renglones ortogonales entre si. Posteriormente a este tipo de matrices les llamaremos unitarias u ortogonales.

22.4. Ejemplo

Consideremos la matriz hermítica

$$M = \begin{pmatrix} 1 & 1-i \\ 1+i & -1 \end{pmatrix},$$

los valores propios de esta matriz son

$$\lambda_1 = \sqrt{3}, \qquad \lambda_2 = -\sqrt{3}$$

y los vectores propios normalizados son

$$v_1 = \frac{1}{\sqrt{3-\sqrt{3}}} \begin{pmatrix} 1 \\ \frac{\sqrt{3}-1}{2}(1+i) \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{3+\sqrt{3}}} \begin{pmatrix} 1 \\ -\frac{\sqrt{3}+1}{2}(1+i) \end{pmatrix},$$

los cuales son ortogonales. Además la matriz de cambio de base es

$$U = \begin{pmatrix} \frac{1}{\sqrt{3-\sqrt{3}}} & \frac{1}{\sqrt{3+\sqrt{3}}} \\ \frac{\sqrt{3}-1}{2\sqrt{3-\sqrt{3}}}(1+i) & -\frac{\sqrt{3}+1}{2\sqrt{3+\sqrt{3}}}(1+i) \end{pmatrix}$$

la cual satisface

$$U^\dagger = U^{*T} = \begin{pmatrix} \frac{1}{\sqrt{3-\sqrt{3}}} & \frac{\sqrt{3}-1}{2\sqrt{3-\sqrt{3}}}(1-i) \\ \frac{1}{\sqrt{3+\sqrt{3}}} & -\frac{\sqrt{3}+1}{2\sqrt{3+\sqrt{3}}}(1-i) \end{pmatrix} = U^{-1}.$$

22.5. Ejemplo con matriz simétrica

Consideremos la matriz

$$M = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$$

la cual es simétrica. En este caso los valores propios son

$$\lambda_1 = 6, \qquad \lambda_2 = 1,$$

mientras que los vectores propios normalizados son

$$v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

los cuales son ortogonormales. Además la matriz de cambio de base es

$$U = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix},$$

la cual cumple

$$U^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix} = U^{-1}.$$

Charle Hermite (1822-1901) fue un matemático francés que realizó diferentes aportaciones a las matemáticas. En particular, demostró que cualquier matriz simétrica tiene valores propios reales y que el número e es trascendente. Debido a su afición por los trabajos de Lagrange, de adolescente le llamaban el pequeño Lagrange. A los 20 años inició sus estudios en la École Polytechnique. Sin embargo, se le detectó un defecto en el pie derecho y por esa absurda razón las autoridades escolares no le permitieron seguir sus estudios. Posteriormente, continuó sus estudios en una Ecoles d' Applications. A Hermite no le gustaban los exámenes, él prefería leer el trabajo de matemáticos clásicos. Con el tiempo logró obtener resultados importantes, así que se convirtió en un matemático de renombre. En 1869 regresó a la École Polytechnique como profesor y tiempo después fue nombrado miembro de la Academia de Paris.

22.5.1. Parámetros libres

Sea Λ una matriz compleja de $n \times n$ entradas. Si separamos esta matriz en su parte real e imaginaria se tiene

$$\Lambda = A + iB,$$

con A y B dos matrices de $n \times n$ con entradas reales. Usando esta descomposición, la igualdad (22.1) toma la forma

$$(A + iB)^{*T} = A + iB,$$

es decir

$$A^T - iB^T = A + iB,$$

de donde

$$A^T = A, \qquad B^T = -B,$$

Por lo tanto, A debe ser una matriz simétrica y B debe ser una matriz antisimétrica.

Ahora, si una matriz real de $n \times n$ es simétrica, entonces los coeficientes que están sobre la diagonal son iguales a los coeficientes que están de bajo de la diagonal. Por lo que el primer renglón tiene n coeficientes reales libres, el segundo renglón tiene $n-1$ coeficientes reales libres, el renglón k tiene $n-(k-1)$ coeficientes reales libres. Así, el total de coeficientes libres de una matriz simétrica es

$$\boxed{\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.}$$

Además, si una matriz de $n \times n$ es antisimétrica, sus elementos de la diagonal son nulos y los coeficientes que están sobre la diagonal son inversos aditivos a los coeficientes que están de bajo de la diagonal. En consecuencia, el primer renglón tiene $n-1$ coeficientes reales libres, el segundo renglón tiene $n-2$ coeficientes reales libres, el renglón k tiene $n-k$ coeficientes reales libres. Entonces, el total de coeficientes reales libres de una matriz antisimétrica es

$$\boxed{\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}.}$$

Considerando estos resultados, podemos ver que la matriz A tiene

$$\frac{n(n+1)}{2}$$

parámetros reales libres, mientras que la matriz B tiene

$$\frac{n(n-1)}{2}$$

parámetros reales libres. Por lo cual, considerando que

$$n^2 = \frac{n(n+1)}{2} + \frac{n(n-1)}{2},$$

llegamos a la conclusión que una matriz hermítica Λ de $n \times n$ con entradas complejas tiene

$$\boxed{n^2}$$

parámetros reales libres.

22.6. Matrices de Pauli

Supongamos que tenemos la siguiente matriz de 2×2 con entradas complejas

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

de donde

$$A^{*T} = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}.$$

Si A es una matriz hermítica, la igualdad $A = A^{*T}$ implica

$$\begin{aligned} a &= a^*, \\ d &= d^*, \\ c &= b^*. \end{aligned}$$

Por lo tanto, a y b deben ser número reales y se debe cumplir $c = b^*$. Así, para obtener todas las matrices hermíticas de 2×2 tenemos cuatros parámetros libres. Es decir, el espacio de las matrices hermíticas de 2×2 es de dimensión 4.

Supongamos que

$$\beta_1, \quad \beta_2, \quad \beta_3, \quad \beta_4$$

son número reales, entonces la matriz

$$\boxed{A = \begin{pmatrix} \beta_1 & \beta_3 - i\beta_4 \\ \beta_3 + i\beta_4 & \beta_2 \end{pmatrix},}$$

es hermítica. Se puede observar que toda matriz hermítica se puede escribir como

$$\boxed{A = \beta_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \beta_3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \beta_4 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.}$$

Por lo tanto, cualquier matriz hermítica de 2×2 se puede generar con las matrices

$$\boxed{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.}$$

Como estas 4 matrices son linealmente independientes, son base del espacio de las matrices hermíticas de 2×2.

Ahora, supongamos que

$$\alpha_0, \quad \alpha_1, \quad \alpha_2, \quad \alpha_4$$

son número reales. Entonces, otra forma de parametrizar a las matrices hermíticas de 2×2 es

$$\boxed{A = \begin{pmatrix} \alpha_0 + \alpha_3 & \alpha_1 - i\alpha_2 \\ \alpha_1 + i\alpha_2 & \alpha_0 - \alpha_3 \end{pmatrix}.}$$

Observe que en este caso tenemos

$$\boxed{A = \alpha_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \alpha_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.}$$

Por lo tanto, cualquier matriz hermítica de 2×2 se puede generar con las matrices

$$\boxed{\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.}$$

A las matrices $\sigma_1, \sigma_2, \sigma_3$ se les llama de Pauli, las cuales son linealmente independientes.

Así, cualquier matriz hermítica de 2×2 se puede escribir como combinación lineal de la matriz identidad y de las matrices de Pauli de la siguiente forma

$$\boxed{A = \alpha_0\sigma_0 + \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3.}$$

Por lo que la matriz identidad junto con las matrices de Pauli forman una base para las matrices hermíticas de 2×2.

22.7. Matrices de Gell-Mann

Supongamos que tenemos la siguiente matriz de 3×3 con entradas complejas

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

de donde

$$A^{*T} = \begin{pmatrix} a^*_{11} & a^*_{21} & a^*_{31} \\ a^*_{12} & a^*_{22} & a^*_{32} \\ a^*_{13} & a^*_{23} & a^*_{33} \end{pmatrix}.$$

Si A es una matriz hermítica, la igualdad $A = A^{*T}$ implica

$$a_{11} = a^*_{11}, \quad (22.3)$$
$$a_{22} = a^*_{22}, \quad (22.4)$$
$$a_{33} = a^*_{33}, \quad (22.5)$$
$$a_{21} = a^*_{12}, \quad (22.6)$$
$$a_{31} = a^*_{13}, \quad (22.7)$$
$$a_{23} = a^*_{32} \quad (22.8)$$

Por lo tanto los números a_{11}, a_{22}, a_{33} deben ser reales. Además, de las ecuaciones (22.6)-(22.8), tenemos seis parámetros reales. Así, las matrices hermíticas de 3×3 tienen 9 parámetros libres y por lo tanto forman un espacio de dimesión nueve.

Ahora, supongamos que

$$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9$$

son número reales, entonces la matriz de 3×3

$$\boxed{A = \begin{pmatrix} \beta_3 & \beta_1 - i\beta_2 & \beta_4 - i\beta_5 \\ \beta_1 + i\beta_2 & \beta_8 & \beta_6 - i\beta_7 \\ \beta_4 + i\beta_5 & \beta_6 + i\beta_7 & \beta_9 \end{pmatrix},}$$

es hermítica. Por lo tanto, cualquier matriz hermítica de 3×3 se puede escribir como

$$\boxed{\begin{aligned} A = &\beta_1 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \beta_2 \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \beta_3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &+ \beta_4 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \beta_5 \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} + \beta_6 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ &+ \beta_7 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} + \beta_8 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \beta_9 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}}$$

Así, las matrices hermíticas

$$\boxed{\begin{matrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{matrix}}$$

generan cualquier matriz hermítica de 3×3.

Existen otras formas de parametrizar a las matrices hermíticas de 3×3. Por ejemplo, supongamos que

$$\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8$$

son número reales, entonces la matriz

$$\boxed{A = \begin{pmatrix} \alpha_0 + \alpha_3 + \frac{1}{\sqrt{3}}\alpha_8 & \alpha_1 - i\alpha_2 & \alpha_4 - i\alpha_5 \\ \alpha_1 + i\alpha_2 & \alpha_0 - \alpha_3 + \frac{1}{\sqrt{3}}\alpha_8 & \alpha_6 - i\alpha_7 \\ \alpha_4 + i\alpha_5 & \alpha_6 + i\alpha_7 & \alpha_0 - \frac{2}{\sqrt{3}}\alpha_8 \end{pmatrix},}$$

es hermítica. De donde,

$$\boxed{\begin{aligned} A = {} & \alpha_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \alpha_1 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ & + \alpha_3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \alpha_5 \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \\ & + \alpha_6 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \alpha_7 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} + \alpha_8 \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned}}$$

Por lo cual, las matrices

$$\boxed{\begin{aligned} & \lambda_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ & \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \\ & \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned}}$$

matrices generan las matrices hermíticas de 3×3.

A las matrices

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8$$

se les llama matrices de Gell-Mann y surgieron en el estudio de las partículas llamadas quarks.

Así, las matrices hermíticas de 3×3 se generan con la matriz identidad y las matrices de Gell-Mann de la forma

$$A = \sum_{k=0}^{8} \alpha_k \lambda_k.$$

Murray Gell-Mann (1929-2019) fue un físico norteamericano que realizó contribuciones fundamentales en el estudio de partículas elementales y en sistemas complejos. Para modelar a las partículas que se encuentran dentro del núcleo atómico propuso el empleo de las ahora llamadas matrices de Gell-Mann. Los resultados experimentales de los aceleradores de partículas mostraron que su modelo es correcto. Debido a sus modelos teóricos, en 1969 le otorgaron el premio Nobel de Física. También se interesó en otras disciplinas como la literatura, historia, arqueología, lingüística, etc. Curiosamente se basó en una novela de James Joyce para nombrar como quarks a las componentes del interior del núcleo atómico. Aprendió diferentes lenguas, entre ellos el Maya. Fue director del Instituto de Santa Fe de Estados Unidos en donde colaboró con expertos de diferentes disciplinas.

22.8. Operadores con valores propios positivos

Sea B de la forma

$$\boxed{B = A^\dagger A,}$$

con A un matriz de $n \times n$. Note que

$$\begin{aligned} B^{\dagger} &= \left(A^{\dagger}A\right)^{\dagger} \\ &= A^{\dagger}\left(A^{\dagger}\right)^{\dagger} \\ &= A^{\dagger}A \\ &= B, \end{aligned}$$

es decir

$$B^{\dagger} = B,$$

este resultado implica que B es una matriz hermítica.

Además se puede probar que si B es diagonalizable, entonces sus valores propios son positivos. En efecto, si v es un vector no nulo y

$$Bv = \lambda v$$

con λ un escalar, entonces

$$\langle v|Bv\rangle = \langle v|\lambda v\rangle = \lambda\langle v|v\rangle, \tag{22.9}$$

además

$$\langle v|Bv\rangle = \langle v|A^{\dagger}Av\rangle = \langle Av|Av\rangle \geq 0. \tag{22.10}$$

Usando las ecuaciones (22.9)-(22.10), encontramos que

$$\lambda\langle v|v\rangle = \langle Av|Av\rangle \geq 0,$$

así

$$\lambda = \frac{\langle Av|Av\rangle}{\langle v|v\rangle} \geq 0,$$

esto es

$$\boxed{\lambda \geq 0,}$$

que es lo que se quería demostrar.

22.9. Ejercicios

1.- Obtener los valores propios de la matriz hermítica

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

2.- Sea el espacio de las funciones bien portadas definidas en el intervalo $[a, b,]$ que cumplen $\psi(a) = \psi(b) = 0$. En este espacio, determinar bajo qué condiciones el operador

$$A = \alpha \frac{\partial}{\partial x}, \quad \alpha \in \mathbb{C}$$

es hermítico.

Capítulo 23

Matrices antihermíticas

En este capítulo estudiaremos las matrices antihermíticas y sus propiedades. En particular, mostraremos que toda matriz de $n \times n$ con entradas complejas se puede escribir como las suma de una matriz hemítica con un matriz antihermítica.

23.1. Definición

Λ es una matriz antihermítica si satisface

$$\boxed{\Lambda^\dagger = -\Lambda} \tag{23.1}$$

23.2. Ejemplo, las matrices de Dirac

Las matrices de Dirac se usan en mecánica cuántica relativista. Estas matrices están dadas por

$$\boxed{\begin{array}{c} \gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \\ \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \\ \gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \\ \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \end{array}}$$

Observe que la matriz γ^0 es hermítica y las matrices $\gamma^i, (i = 1, 2, 3)$ son antihermíticas.

23.3. Espectro de operadores antihermíticos

Los operadores antihermíticos tienen valores propios imaginarios. En efecto, si v es un vector propio de un operador antihermítico A con valor propio λ, es decir

$$Av = \lambda v,$$

se tiene

$$\begin{aligned} \langle Av|v\rangle &= \langle \lambda v|v\rangle \\ &= \lambda^* \langle v|v\rangle , \end{aligned}$$

adicionalmente se encuentra

$$\begin{aligned}
\langle Av|v\rangle &= \langle v|A^{\dagger}v\rangle \\
&= \langle v|(-Av)\rangle \\
&= -\langle v|\lambda v\rangle \\
&= -\lambda \langle v|v\rangle,
\end{aligned}$$

de donde

$$\lambda^{*} \langle v|v\rangle = -\lambda \langle v|v\rangle,$$

que nos lleva al resultado

$$\boxed{\lambda^{*} = -\lambda.}$$

Así, los valores propios de los operadores antihermíticos son imaginarios.

También se puede observar que si Λ es una matriz antihermítica y α es un número real, entonces la matriz

$$\boxed{A = e^{i\alpha\Lambda}}$$

es hermítica. En efecto, tenemos que

$$\begin{aligned}
A^{\dagger} &= \left(e^{i\alpha\Lambda}\right)^{\dagger} \\
&= e^{-i\alpha\Lambda^{\dagger}} \\
&= e^{i\alpha\Lambda} \\
&= A,
\end{aligned}$$

es decir,

$$A^{\dagger} = A.$$

En particular, si Λ es una matriz de entradas reales y antisimétrica, es decir, si

$$\Lambda^{T} = -\Lambda,$$

entonces la matriz

$$A = e^{i\alpha\Lambda}$$

es hermítica.

23.3.1. Parámetros libres

Si separamos a una matriz compleja Λ en su parte real e imaginaria se tiene

$$\Lambda = A + iB,$$

con A y B dos matrices de $n \times n$ con entradas reales. Usando esta descomposición, la igualdad (23.1) toma la forma

$$(A + iB)^{*T} = -(A + iB),$$

es decir,

$$A^T - iB^T = -A - iB,$$

de donde

$$A^T = -A, \qquad B^T = B.$$

Por lo tanto, A debe ser una matriz antisimétrica y B debe ser una matriz simétrica.

Ahora, como A es una matriz antisimétrica real de $n \times n$ por lo que tiene

$$\frac{n(n-1)}{2}$$

parámetros reales libres, mientras que B es una matriz simétrica real de $n \times n$ entonces tiene

$$\frac{n(n+1)}{2}$$

parámetros reales libres. Además, considerando que

$$n^2 = \frac{n(n+1)}{2} + \frac{n(n-1)}{2},$$

una matriz antihermítica Λ con entradas complejas tiene

$$\boxed{n^2.}$$

parámetros reales libres.

23.3.2. Descomposición en operadores hermíticos y anti-hermíticos

Con cualquier matriz Λ se pueden construir las matrices

$$\boxed{\begin{aligned} \Lambda_+ &= \frac{\Lambda + \Lambda^\dagger}{2}, \\ \Lambda_- &= \frac{\Lambda - \Lambda^\dagger}{2}, \end{aligned}}$$

estas matrices satisfacen

$$\begin{aligned} \Lambda_+^\dagger &= \left(\frac{\Lambda + \Lambda^\dagger}{2}\right)^\dagger \\ &= \frac{\Lambda^\dagger + \Lambda}{2} \\ &= \Lambda_+^\dagger, \\ \Lambda_-^\dagger &= \left(\frac{\Lambda - \Lambda^\dagger}{2}\right)^\dagger \\ &= \frac{\Lambda^\dagger - \Lambda}{2} \\ &= -\Lambda_-. \end{aligned}$$

Por lo tanto, Λ_+ es una matriz hermítica y Λ_- es una matriz antihermítica, además se cumple

$$\Lambda = \frac{\Lambda + \Lambda^\dagger}{2} + \frac{\Lambda - \Lambda^\dagger}{2},$$

es decir,

$$\boxed{\Lambda = \Lambda_+ + \Lambda_-,}$$

así cualquier matriz Λ compleja de $n \times n$ se puede descomponer en la suma de una matriz hermítica y una matriz antihermítica.

Note que si Λ solo tienen componentes reales, entonces

$$\boxed{\begin{aligned} \Lambda_+ &= \frac{\Lambda + \Lambda^T}{2}, \\ \Lambda_- &= \frac{\Lambda - \Lambda^T}{2}, \end{aligned}}$$

donde Λ_+ es una matriz simétrica y Λ_- es una matriz antisimétrica. Además, se sigue cumpliendo la igualdad

$$\boxed{\Lambda = \Lambda_+ + \Lambda_-}$$

por lo que cualquier matriz Λ de $n \times n$ de entradas reales se puede descomponer como la suma de una matriz simétrica y una matriz antisimétrica.

23.4. Ejercicios

1.- Obtener la parte simétrica y antisimétrica de la matriz

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

2.- Obtener la parte hermítica y antihermítica de la matriz

$$\begin{pmatrix} -i & 2 & 3i \\ \sqrt{2} & i & -2 \\ 2i & 4 & 5 \end{pmatrix}.$$

Capítulo 24

Matrices ortogonales

Como una aplicación de los temas vistos anteriormente ahora estudiaremos las matrices ortogonales, las cuales se relacionan con las matrices de rotaciones.

24.1. Definición

Una matriz Λ de entradas reales de $n \times n$ es ortogonal si cumple

$$\boxed{\Lambda^T \Lambda = I} \tag{24.1}$$

donde I es la matriz identidad.

Por ejemplo, la matriz de rotación sobre el eje x y su transpuesta

$$\begin{aligned}
\Lambda_x(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \\
\Lambda_x^T(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},
\end{aligned}$$

satisfacen

$$\begin{aligned}
\Lambda_x^T(\theta)\,\Lambda_x(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos^2\theta + \sin^2\theta & \sin\theta\cos\theta - \sin\theta\cos\theta \\ 0 & \sin\theta\cos\theta - \sin\theta\cos\theta & \cos^2\theta + \sin^2\theta \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},
\end{aligned}$$

es decir

$$\Lambda_x^T(\theta)\,\Lambda_x(\theta) = I.$$

Por lo tanto la matriz $\Lambda_x(\theta)$ es ortogonal.

De la misma forma se puede mostrar que las matrices de rotación sobre los ejes x y y

$$\begin{aligned}
\Lambda_y(\psi) &= \begin{pmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{pmatrix}, \\
\Lambda_z(\phi) &= \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

también son ortogonales.

24.2. Grupo ortogonal $O(n)$ y grupo $SO(n)$

Antes de continuar recordemos lo que es un grupo. Sea G un conjunto con una operación

$$\cdot : G \times G \to G.$$

El par $(G,\cdot)$ es un grupo si cumple

1) Axioma de cerradura:

$$\boxed{g_1 \in G, g_2 \in G \qquad \Longrightarrow \qquad g_1 \cdot g_2 \in G}$$

2) Axioma de asociatividad:

$$\boxed{g_1 \in G, g_2 \in G, g_3 \in G, \qquad \Longrightarrow \qquad g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3.}$$

3) Axioma del neutro:

$$\boxed{\exists e \in G, \qquad g_1 \in G \Longrightarrow \qquad g_1 \cdot e = e \cdot g_1 = g_1.}$$

4) Axioma del inverso:

$$\boxed{\forall g_1 \in G, \exists g_1^{-1} \in G, \quad g_1 \cdot g_1^{-1} = g_1^{-1} \cdot g_1 = e.}$$

Definamos a $O(n)$ como el conjunto de matrices Λ con entradas reales de $n \times n$ que cumplen la ecuación de ortogonalidad (24.1). Probaremos que $O(n)$ es un grupo.

Supongamos que Λ_1 y Λ_2 son matrices ortogonales. Entonces cumplen

$$\boxed{\Lambda_1^T \Lambda_1 = I, \qquad \Lambda_2^T \Lambda_2 = I,}$$

de donde

$$\begin{aligned} (\Lambda_1 \Lambda_2)^T (\Lambda_1 \Lambda_2) &= \Lambda_2^T \Lambda_1^T \Lambda_1 \Lambda_2 \\ &= \Lambda_2^T \Lambda_2 = I, \end{aligned}$$

esto es

$$(\Lambda_1 \Lambda_2)^T (\Lambda_1 \Lambda_2) = I.$$

Este resultado implica $\Lambda_1 \Lambda_2 \in O(n)$, es decir

$$\boxed{\Lambda_1, \Lambda_2 \in O(n) \quad \Longrightarrow \quad \Lambda_1 \Lambda_2 \in O(n).}$$

Por lo tanto, se cumple el axioma la cerradura.

El producto de matrices es asociativo, en particular el producto de las matrices ortogonales. Además, la identidad I satisface la ecuación (24.1), es decir, $I \in O(n)$. Así, se cumplen el axioma de la asociatividad y el del elemento neutro.

Ahora, si Λ cumple la ecuación (24.1), entonces

$$\Lambda^T = \Lambda^{-1},$$

por lo que

$$\begin{aligned}\left(\Lambda^{-1}\right)^T\Lambda^{-1} &= \left(\Lambda^T\right)^T\Lambda^{-1}\\ &= \Lambda\Lambda^{-1} = I,\end{aligned}$$

esto es

$$\boxed{\left(\Lambda^{-1}\right)^T\Lambda^{-1} = I.}$$

Es decir, si la matriz Λ es ortogonal, entonces Λ^{-1} también es ortogonal.

Así, las matrices Λ de $n \times n$ entradas reales que cumplen la ecuación (24.1) forman un grupo. A este grupo ortogonal se llama grupo ortogonal $O(n)$.

24.3. Determinantes de una matriz ortogonal

Ahora, recordemos que para cualquier matriz A se cumple

$$\boxed{\det A = \det A^T.}$$

Entonces, de la ecuación (24.1) se tiene

$$\begin{aligned}\det\left(\Lambda^T\Lambda\right) &= \det\Lambda^T\det\Lambda\\ &= \det\Lambda\det\Lambda = \det I = 1,\end{aligned}$$

de donde

$$\boxed{(\det\Lambda)^2 = 1,}$$

es decir

$$\boxed{\det\Lambda = \pm 1.}$$

Las matrices Λ que cumplen

$$\boxed{\det\Lambda = 1.}$$

sí forman un grupo, este es el grupo $SO(n)$.

El subconjunto de matrices Λ que cumplen $\det\Lambda = -1$ no forman un grupo. Por ejemplo, la identidad no está en ese subconjunto. Algunas matrices de este tipo son

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$
$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Esta matrices cumplen la ecuación (24.1), pero no son de rotación ni de $SO(3)$.

24.4. Parámetros

Observe que la matriz de $n \times n$

$$\boxed{\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \cdots & \Lambda_{1n} \\ \Lambda_{21} & \Lambda_{22} & \cdots & \Lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{n1} & \Lambda_{n2} & \cdots & \Lambda_{nn} \end{pmatrix}}$$

se pueden formar con los vectores columna

$$\boxed{C_1 = \begin{pmatrix} \Lambda_{11} \\ \Lambda_{21} \\ \vdots \\ \Lambda_{n1} \end{pmatrix}, C_2 = \begin{pmatrix} \Lambda_{12} \\ \Lambda_{22} \\ \vdots \\ \Lambda_{n2} \end{pmatrix}, \cdots, C_n = \begin{pmatrix} \Lambda_{1n} \\ \Lambda_{2n} \\ \vdots \\ \Lambda_{nn} \end{pmatrix}.}$$

Claramente, estos vectores representan los renglones de la matriz traspuesta, Λ^T. Por lo tanto, la condición (24.1) se puede escribir como

$$\boxed{\begin{aligned}\Lambda^T\Lambda &= \begin{pmatrix} \Lambda_{11} & \Lambda_{21} & \cdots & \Lambda_{n1} \\ \Lambda_{12} & \Lambda_{22} & \cdots & \Lambda_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{1n} & \Lambda_{2n} & \cdots & \Lambda_{nn} \end{pmatrix} \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \cdots & \Lambda_{1n} \\ \Lambda_{21} & \Lambda_{22} & \cdots & \Lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{n1} & \Lambda_{n2} & \cdots & \Lambda_{nn} \end{pmatrix} \\ &= \begin{pmatrix} C_1\cdot C_1 & C_1\cdot C_2 & \cdots & C_1\cdot C_n \\ C_2\cdot C_1 & C_2\cdot C_2 & \cdots & C_n\cdot C_2 \\ \vdots & \vdots & \ddots & \vdots \\ C_n\cdot C_1 & C_n\cdot C_2 & \cdots & C_n\cdot C_n \end{pmatrix} = I. \end{aligned}} \tag{24.2}$$

Otra forma de expresar esta igualdad es

$$\boxed{C_i \cdot C_j = \delta_{ij},}$$

es decir, si una matriz satisface la condición (24.1), sus columnas son ortonormales.

Ahora, una matriz Λ de $n \times n$ con componentes reales tiene n^2 parámetros libres. Pero si satisface la ecuación (24.1), entonces no todos los parámetros son libres. De la igualdad (24.2) se puede ver que $\Lambda^T\Lambda$ es una matriz simétrica, por lo que (24.2) sólo tiene

$$\frac{n(n+1)}{2}$$

ecuaciones independientes. Así, los parámetros libres de una matriz ortogonal son

$$n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

Así el grupo de matrices Λ de $n \times n$ de entradas reales ortogonales $O(n)$ tiene

$$\boxed{\frac{n(n-1)}{2}} \tag{24.3}$$

parámetros libres.

24.5. Matriz antisimétrica y matrices ortogonales

Sea A una matriz antisimétrica, es decir que satisface

$$\boxed{A^T = -A}$$

y α un número real, entonces la matriz

$$\boxed{\Lambda = e^{\alpha A}} \tag{24.4}$$

es ortogonal. En efecto, note que

$$\begin{aligned}
\Lambda^T &= \left(e^{\alpha A}\right)^T \\
&= e^{\alpha A^T} \\
&= e^{-\alpha A} \\
&= \Lambda^{-1},
\end{aligned}$$

esto es

$$\Lambda^T = \Lambda^{-1},$$

por lo tanto

$$\left(e^{\alpha A}\right)^T e^{\alpha A} = I.$$

Así, la matriz (24.4) es ortogonal. Además, si A es un matriz de $n \times n$ anti simétrica, tiene

$$\frac{n(n-1)}{2}$$

parámetros libres, que coincide con el número de grados de libertad de las matrices ortogonales (24.3).

También se puede observar que considerando la identidad

$$\boxed{\det\left(e^M\right) = e^{Tr(M)}}$$

y que toda matriz antisimétrica tiene traza nula

$$\boxed{Tr(A) = 0,}$$

entonces

$$\begin{aligned} \det\left(e^{A}\right) &= e^{Tr(A)} \\ &= e^{0} = 1, \end{aligned}$$

de donde

$$\boxed{e^{A} \in SO(n).}$$

Se dice que las matrices antisimétricas general al grupo $SO(n)$.

24.5.1. Matrices de $SO(2)$

La única matriz antisimétrica con entradas reales de 2×2 es

$$A = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}$$

con la cual se pueden construir las siguientes matrices ortogonales de 2×2

$$\Lambda = e^{A} = e^{\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix},$$

así

$$SO(2) = \left\{ \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \right\}$$

24.5.2. Matrices ortogonales de $SO(3)$

Ahora, cualquier matriz antisimétricas de 3×3 con entradas reales se puede escribir como

$$A = \begin{pmatrix} 0 & -\alpha_3 & \alpha_2 \\ \alpha_3 & 0 & -\alpha_1 \\ -\alpha_2 & \alpha_1 & 0 \end{pmatrix},$$

por lo que

$$A = \alpha_1 m_1 + \alpha_2 m_2 + \alpha_3 m_3$$

con

$$m_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad m_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad m_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Así

$$SO(3) = \left\{ e^{\alpha_1 m_1 + \alpha_2 m_2 + \alpha_3 m_3} \right\},$$

en particular, las matrices

$$\begin{aligned} e^{\alpha_1 m_1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_1 \end{pmatrix}, \\ e^{\alpha_2 m_2} &= \begin{pmatrix} \cos\alpha_2 & 0 & \sin\alpha_2 \\ 0 & 1 & 0 \\ -\sin\alpha_2 & 0 & \cos\alpha_2 \end{pmatrix}, \\ e^{\alpha_3 m_3} &= \begin{pmatrix} \cos\alpha_3 & -\sin\alpha_3 & 0 \\ \sin\alpha_3 & \cos\alpha_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

pertenecen a $SO(3)$.

24.6. Distancia

Sea $\vec{x} = (x_1, \cdots, x_n)$ un vector en $\mathbb{R}^n$ y definamos la forma cuadrática

$$\boxed{l^2 = x_1^2 + x_2^2 + \cdots + x_n^2,}$$

la cual representa la distancia de $\vec{x}$ al origen. Observe que si definimos la matriz columna

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

y la matriz renglón

$$X^T = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix},$$

la distancia se puede escribir como

$$\boxed{l^2 = X^T X = X^T I X.}$$

Ahora, si Λ es una matriz de $n \times n$ y se hace la transformación de coordenadas

$$\boxed{X' = \Lambda X,}$$

se tiene la distancia

$$\begin{aligned} l'^2 &= X'^T I X' \\ &= (\Lambda X)^T \Lambda X \\ &= X^T \left(\Lambda^T I \Lambda\right) X, \end{aligned}$$

es decir

$$l'^2 = X^T \left(\Lambda^T I \Lambda\right) X,$$

Por lo tanto, si Λ es tal que deja la distancia invariante,

$$\boxed{l^2 = l'^2,}$$

se debe cumplir que

$$X^T I X = X^T \left(\Lambda^T I \Lambda\right) X,$$

de donde

$$\boxed{\Lambda^T I \Lambda = I.}$$

Así, la distancia es invariante bajo transformaciones ortogonales.

24.7. Ejercicios

1.- Determinar si la siguiente matriz es ortogonal

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

2.- Determinar si la siguiente matriz es ortogonal

$$M = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{5}} \end{pmatrix}.$$

Capítulo 25

Matrices unitarias

Una generalización de las matrices ortogonales son las llamadas matrices unitarias, que son muy útiles para entender la mecánica cuántica. En este capítulo daremos la definición de matrices unitarias y estudiaremos algunas de sus aplicaciones.

25.1. Definición

Una matriz de $n \times n$ con entradas complejas U es unitaria si cumple

$$\boxed{U^\dagger U = I, \qquad U^\dagger = U^{*T}} \tag{25.1}$$

donde I es la matriz identidad. Note que esta igualdad implica

$$\boxed{U^\dagger = U^{-1}.}$$

Por ejemplo, la matriz

$$\frac{1}{\sqrt{2}}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$$

es unitaria.

25.2. Grupos unitarios $U(n)$ y $SU(n)$

Definiremos a $U(n)$ como el conjunto de las matrices U de $n \times n$ que cumplen la ecuación de unitariedad (25.1), probaremos $U(n)$ que con la operación del producto de las matrices forma un grupo. Antes de continuar, recordemos que un grupo cumple con el axioma de la cerradura, el axioma de la asociatividad, el axioma del neutro y el axioma del inverso.

Ahora, supongamos que U_1 y U_2 son matrices ortogonales, entonces cumplen

$$\boxed{U_1^\dagger U_1 = I, \qquad U_2^\dagger U_2 = I,}$$

de donde

$$\begin{aligned} (U_1 U_2)^\dagger (U_1 U_2) &= U_2^\dagger U_1^\dagger U_1 U_2 \\ &= U_2^\dagger I U_2 = I, \end{aligned}$$

es decir

$$\boxed{(U_1 U_2)^\dagger (U_1 U_2) == I.}$$

Esto implica que $U_1 U_2 \in U(n)$ y por lo tanto se cumple el axioma la cerradura.

El producto de matrices es asociativo. En particular, el producto de las matrices unitarias. Además, la identidad I satisface la ecuación (25.1), es decir, $I \in U(n)$. Así, se cumplen el axioma de la asociatividad y el del elemento neutro.

Ahora, si U cumple la ecuación (25.1), entonces

$$U^\dagger = U^{-1},$$

por lo que

$$\begin{aligned} \left(U^{-1}\right)^\dagger U^{-1} &= \left(U^\dagger\right)^\dagger U^{-1} \\ &= U U^{-1} = I, \end{aligned}$$

esto es

$$\boxed{\left(U^{-1}\right)^\dagger U^{-1} = I.}$$

Es decir, si la matriz U es unitaria, entonces U^{-1} también es unitaria.

Así, las matrices U de $n \times n$ con entradas complejas que cumplen la ecuación (25.1) forman un grupo. A este grupo se llama grupo unitario $U(n)$.

Ahora, recordemos que para cualquier matriz A se cumple

$$\det A = \det A^T.$$

Entonces, las matrices unitarias satisfacen

$$\begin{aligned}
\det\left(U^\dagger U\right) &= \left(\det U^\dagger\right)(\det U)\\
&= \left(\det U^{*T}\right)(\det U)\\
&= (\det U^*)(\det U)\\
&= (\det U)^*(\det U)\\
&= |\det U|^2\\
&= \det I = 1,
\end{aligned}$$

que implica

$$\boxed{|\det U|^2 = 1,}$$

de donde

$$\boxed{\det U = e^{i\alpha},}$$

donde α es un parámero real.

Las matrices U que cumplen

$$\det U = 1$$

forman un grupo, este es el grupo $SU(n)$.

25.3. Parámetros libres

Observe que la matriz de $n \times n$

$$\boxed{U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n}\\ u_{21} & u_{22} & \cdots & u_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ u_{n1} & u_{n2} & \cdots & u_{nn}\end{pmatrix}}$$

se pueden formar con los vectores columna

$$\boxed{V_1 = \begin{pmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{n1} \end{pmatrix}, V_2 = \begin{pmatrix} u_{12} \\ u_{22} \\ \vdots \\ u_{n2} \end{pmatrix}, \cdots, V_n = \begin{pmatrix} u_{1n} \\ u_{2n} \\ \vdots \\ u_{nn} \end{pmatrix}.}$$

Claramente, estos vectores representan los renglones de la matriz traspuesta, U^T. Por lo tanto, la condición (25.1) se puede escribir como

$$\boxed{\begin{aligned} U^\dagger U =& U^{T*}U \\ =& \begin{pmatrix} u_{11}^* & u_{21}^* & \cdots & u_{n1}^* \\ u_{12}^* & u_{22}^* & \cdots & u_{n2}^* \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n}^* & u_{2n}^* & \cdots & u_{nn}^* \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix} \\ =& \begin{pmatrix} V_1^* \cdot V_1 & V_1^* \cdot V_2 & \cdots & V_1^* \cdot V_n \\ V_2^* \cdot V_1 & V_2^* \cdot V_2 & \cdots & V_n^* \cdot V_2 \\ \vdots & \vdots & \ddots & \vdots \\ V_n^* \cdot V_1 & V_n^* \cdot V_2 & \cdots & V_n^* \cdot V_n \end{pmatrix} = I, \end{aligned}} \tag{25.2}$$

otra forma de expresar esta igualdad es

$$\boxed{V_i^* \cdot V_j = \delta_{ij}.}$$

Note que la ecuación (25.2) tiene n ecuaciones reales en la diagonal. Además, podemos ver que las ecuaciones arriba de la diagonal de (25.2) son las mismas que están abajo de la diagonal. Así, fuera de la diagonal hay

$$2\frac{n(n-1)}{2} = n(n-1)$$

ecuaciones reales independientes. Por lo tanto, considerando que una matriz de $n \times n$ con entradas complejas tiene $2n^2$ parámetros, los parámetros libres de una matriz unitaria de $n \times n$ son

$$2n^2 - n - n(n-1) = n^2.$$

Así el grupo unitario $U(n)$ tiene

$$\boxed{n^2}$$

parámetros reales libres. La matrices que tienen determinante uno se llaman $SU(n)$, estas matrices cumplen una ecuación extra, por lo que el grupo $SU(n)$ tiene

$$\boxed{n^2 - 1}$$

parámetros reales libres.

25.4. Matrices hermíticas y matrices unitarias

B es una matriz es hermítica ($B = B^\dagger$) y α un número real, entonces la matriz

$$\boxed{U = e^{i\alpha B}}$$

es unitaria. En efecto

$$\begin{aligned} U^\dagger &= \left(e^{i\alpha B}\right)^\dagger \\ &= e^{-i\alpha B^\dagger} \\ &= e^{-i\alpha B} \\ &= U^{-1} \end{aligned}$$

es decir

$$U^\dagger = U^{-1}.$$

25.5. $U(1)$

Sea los números complejos de la forma

$$\boxed{U = e^{i\alpha}} \tag{25.3}$$

con α un número real. Observe que

$$\begin{aligned} U^\dagger &= \left(e^{i\alpha}\right)^* \\ &= e^{-i\alpha} \\ &= U^{-1}. \end{aligned}$$

Por lo tanto, el conjunto de números complejos de la forma (25.3) forman la matriz de 1×1 unitarias. A este grupo se le llama $U(1)$ y es fundamental para el estudio de la elecrodinámica. También podemos notar que este conjunto representa un círculo en el plano complejo, como se muestra en la figura (25.1).

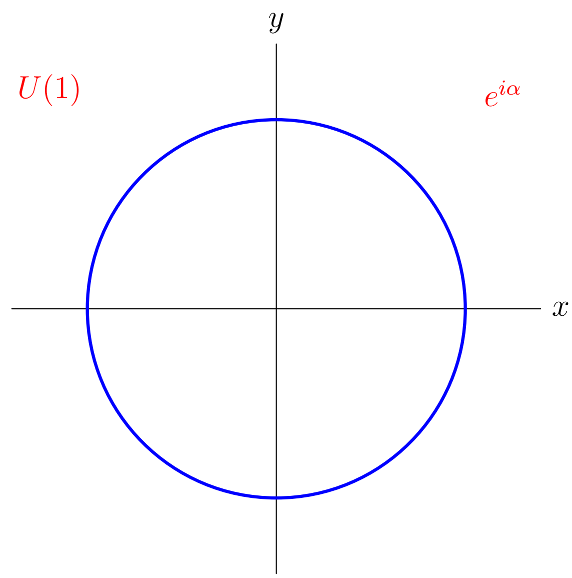

Figura 25.1: Círculo unitario

25.6. Matrices de $U(2)$ y $SU(2)$

Las matrices unitarias de 2×2 tienen 4 parámetros libres, además recordemos que si $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ son números reales, cualquier matriz A hermítica de 2×2 se puede escribir como

$$A = \alpha_0\sigma_0 + \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3,$$

con σ_0 la matriz identidad y

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

las matrices de Pauli.

Por lo tanto, las matrices unitarias de 2×2 tienen la forma

$$\boxed{U = e^{i(\alpha_0\sigma_0+\alpha_1\sigma_1+\alpha_2\sigma_2+\alpha_3\sigma_3)}.}$$

Ahora, considerando la identidad (20.1)

$$\det M = e^{Tr(\ln M)},$$

se encuentra

$$\begin{aligned}
\det U &= e^{Tr(\ln U)} \\
&= e^{Tr\left(\ln\left(e^{i(\alpha_0\sigma_0+\alpha_1\sigma_1+\alpha_2\sigma_2+\alpha_3\sigma_3)}\right)\right)} \\
&= e^{iTr(\alpha_0\sigma_0+\alpha_1\sigma_1+\alpha_2\sigma_2+\alpha_3\sigma_3)} \\
&= e^{i(\alpha_0 Tr(\sigma_0)+\alpha_1 Tr(\sigma_1)+\alpha_2 Tr(\sigma_2)+\alpha_3 Tr(\sigma_3))} \\
&= e^{i2\alpha_0}.
\end{aligned}$$

Por lo tanto, las matrices unitarias de 2×2 con determinante uno se pueden escribir en términos de las matrices de Pauli de la forma

$$\boxed{U = e^{i(\alpha_1\sigma_1+\alpha_2\sigma_2+\alpha_3\sigma_3)}.}$$

A estas matrices se les llama $SU(2)$ y se usan para estudiar diferentes propiedades del electrón.

25.7. $SU(3)$

Las matrices unitarias de 3×3 tienen 9 parámetros libres. Además las matrices hermíticas de 3×3 se pueden generan con la matriz identidad y las matrices de Gell-Mann de la forma

$$A = \sum_{k=0}^{8} \alpha_k \lambda_k$$

con

$$\boxed{\begin{array}{ll} \lambda_0 = I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & \lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \\ \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \\ \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} & \end{array}}$$

las matrices de Gell-Mann.

Por lo tanto las matrices unitarias de 3×3 se pueden escribir de la forma

$$\boxed{U = e^{i \sum_{k=0}^{8} \alpha_k \lambda_k}.}$$

Ahora, considerando que el determinante se puede expresar como (20.1)

$$\det M = e^{Tr(\ln M)},$$

se encuentra

$$\begin{aligned} \det U &= e^{Tr(\ln U)} \\ &= e^{Tr\left(\ln\left(e^{i \sum_{k=0}^{8} \alpha_k \lambda_k}\right)\right)} \\ &= e^{Tr\left(i \sum_{k=0}^{8} \alpha_k \lambda_k\right)} \\ &= e^{i \sum_{k=0}^{8} \alpha_k Tr(\lambda_k)} \\ &= e^{i 3 \alpha_0}. \end{aligned}$$

Por lo tanto, las matrices unitarias de 3×3 con determinante uno se pueden escribir en términos de las matrices de Gell-Mann de la forma

$$\boxed{U = e^{i \sum_{k=1}^{8} \alpha_k \lambda_k}.}$$

A estas matrices se les llama $SU(3)$ y se usan para estudiar diferentes propiedades de la cromodinámica cuántica.

25.8. Matrices antihermíticas y matrices unitarias

Si B es una matriz es hermítica $(B = -B^{\dagger})$ y α un número real, entonces la matriz

$$U = e^{-\alpha B}$$

es unitaria. En efecto,

$$\begin{aligned} U^{\dagger} &= \left(e^{-\alpha B}\right)^{\dagger} \\ &= e^{-\alpha B^{\dagger}} \\ &= e^{\alpha B} \\ &= U^{-1} \end{aligned}$$

es decir,

$$U^{\dagger} = U^{-1}.$$

25.9. Matrices unitarias y mecánica cuántica

La mecánica cuántica se basa en el estudio de los estados de partículas, un estado se puede como un vector que describe las propiedades de las partículas. Por ejemplo, si un sistema tiene los estados

$$\psi_1 = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \qquad \psi_2 = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

La probabilidad de que la partícula pase del estado ψ_1 al estado ψ_2 es

$$\langle \psi_1 | \psi_2 \rangle = \psi_1^{T*} \psi_2. \tag{25.4}$$

Esta igualdad es invariante ante transformaciones unitarias. En efecto, si

$$\psi_1' = U\psi_1, \qquad \psi_2' = U\psi_2,$$

entonces

$$\begin{aligned}\langle\psi_1'|\psi_2'\rangle &= (U\psi_1)^{*T} U\psi_2\\ &= \psi_1^{*T}U^{*T}U\psi_2.\end{aligned}$$

Por lo tanto, la igualdad

$$\langle\psi_1'|\psi_2'\rangle = \langle\psi_1|\psi_2\rangle$$

implica

$$U^{*T}U = 1,$$

es decir, las matrices unitarias dejan invariante la cantidad (25.4).

25.10. Ejercicios

1.- Determinar si la siguiente matriz es unitaria

$$M = \begin{pmatrix} \frac{i}{\sqrt{6}} & -\frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}}\\ -\frac{i}{\sqrt{6}} & \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}}\\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \end{pmatrix}.$$

2.- Determinar si la siguiente matriz es unitaria

$$M = \begin{pmatrix} \frac{i}{\sqrt{6}} & \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}}\\ -\frac{i}{\sqrt{6}} & \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}}\\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \end{pmatrix}.$$

Capítulo 26

Integrales Gaussianas

Como una aplicación de la matrices simétricas estudiaremos algunas integrales Gaussianas. Estas integrales son importantes en diferentes áreas como la mecánica cuántica, probabillidad, finanzas, machine learning, etc.

26.1. Integrales Gaussianas

Para iniciar, recordemos que en una dimesión se tiene el resultado

$$\boxed{I_1 = \int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}, \quad \alpha > 0.} \tag{26.1}$$

Para estudiar el caso de dos dimensiones primero veamos un ejemplo. Consideremos la integral

$$I_2 = \int_{\mathbb{R}^2} e^{-\left(3x^2 - 2xy + 3y^2\right)} dxdy, \tag{26.2}$$

observe que definiendo el vector

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

y la matriz

$$M = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

la integral (26.2) se puede escribir como

$$I_2 = \int_{\mathbb{R}^2} e^{-X^T M X} dxdy. \tag{26.3}$$

Ahora, los valores propios de M son

$$\lambda_1 = 2, \qquad \lambda_2 = 4$$

mientras que los vectores propios normalizados son

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

los cuales son ortogonales. Además la matriz de cambio de base es

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

la cual es ortogonal, pues cumple

$$U^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = U^{-1}.$$

Además se satisface la igualdad

$$M = UDU^T$$

con

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}. \tag{26.4}$$

Entonces, definiendo

$$X' = \begin{pmatrix} x' \\ y' \end{pmatrix} = U^T X$$

y considerando que

$$\det U = 1,$$

la integral (26.3) toma la forma

$$\begin{aligned} I_2 &= \int_{\mathbb{R}^2} e^{-X^T M X} dxdy \\ &= \int_{\mathbb{R}^2} e^{-X^T U D U^T X} dxdy \\ &= \int_{\mathbb{R}^2} e^{-\left(U^T X\right)^T D \left(U^T X\right)} dxdy \\ &= \int_{\mathbb{R}^2} e^{-X'^T D X'} dx'dy'. \end{aligned}$$

Por lo tanto, considerando la matriz D (26.4) y la integral Gaussiana (26.1) se llega a

$$\begin{aligned} I_2 &= \int_{\mathbb{R}^2} e^{-(2x'^2+4y'^2)} dx'dy' \\ &= \int_{\mathbb{R}} e^{-2x'^2} dx' \int_{\mathbb{R}} e^{-4y'^2} dy' \\ &= \sqrt{\frac{\pi}{2}}\sqrt{\frac{\pi}{4}} = \sqrt{\frac{\pi^2}{8}}, \end{aligned}$$

es decir

$$\int_{\mathbb{R}^2} e^{-\left(x^2-2xy+3y^2\right)} dxdy = \sqrt{\frac{\pi^2}{8}}.$$

26.1.1. Caso general $2d$

Ahora consideremos el caso general

$$I_{2d} = \int_{\mathbb{R}^2} e^{-\left(ax^2+bxy+cy^2\right)} dxdy, \tag{26.5}$$

donde a, b, c son números reales. Note que definiendo

$$\boxed{\begin{aligned} X &= \begin{pmatrix} x \\ y \end{pmatrix}, \\ M &= \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \end{aligned}} \tag{26.6}$$

la integral (26.5) se puede escribir como

$$I_{2d} = \int_{\mathbb{R}^2} e^{-X^T MX} dxdy. \tag{26.7}$$

Supongamos que M es una matriz diagonalizable y que sus valores propios son

$$\lambda_1, \qquad \lambda_2,$$

también supongamos que los vectores propios normalizados de M son

$$v_1, \quad v_2.$$

Como la matriz (26.6) es simétrica, los vectores propios son ortonormales y la matriz de cambio de base

$$U$$

es ortogonal y satisface

$$M = UDU^T$$

con

$$D \;=\; \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{26.8}$$

Por lo que, definiendo

$$X' = \begin{pmatrix} x' \\ y' \end{pmatrix} = U^T X$$

y considerando que

$$\det U = 1,$$

la integral (26.5) se puede escribir como

$$\begin{aligned}
I_{2d} &= \int_{\mathbb{R}^2} dxdye^{-X^T MX} \\
&= \int_{\mathbb{R}^2} e^{-X^T UDU^T X} dxdy \\
&= \int_{\mathbb{R}^2} e^{-\left(U^T X\right)^T D\left(U^T X\right)} dxdy \\
&= \int_{\mathbb{R}^2} e^{-X'^T DX'} dx'dy'.
\end{aligned}$$

Por lo tanto, considerando la matriz D (26.8) se llega a

$$\begin{aligned}
I_{2d} &= \int_{\mathbb{R}^2} e^{-(\lambda_1 x'^2 + \lambda_2 y'^2)} dx'dy' \\
&= \int_{\mathbb{R}} e^{-\lambda_1 x'^2} dx' \int_{\mathbb{R}} e^{-\lambda_2 y'^2} dy'.
\end{aligned}$$

Si alguno de los valores propios de matriz (26.6) es negativo, esta integral diverge. Pero si los dos valores propios son positivos, usando la integral Gaussiana (26.1), se encuentra

$$
\begin{aligned}
I_{2d} &= \int_{\mathbb{R}^2} e^{-(\lambda_1 x'^2+\lambda_2 y'^2)} dx'dy' \\
&= \sqrt{\frac{\pi}{\lambda_1}}\sqrt{\frac{\pi}{\lambda_2}} = \sqrt{\frac{\pi^2}{\lambda_1\lambda_2}}.
\end{aligned}
$$

Ahora como

$$\det M = \lambda_1\lambda_2,$$

tenemos que

$$
\begin{aligned}
I_{2d} &= \int_{\mathbb{R}^2} e^{-(\lambda_1 x'^2+\lambda_2 y'^2)} dx'dy' \\
&= \sqrt{\frac{\pi^2}{\det M}}.
\end{aligned}
$$

Por lo tanto, si los valores propios de la matriz (26.6) son positivos, entonces

$$\int_{\mathbb{R}^2} e^{-\left(ax^2+bxy+cy^2\right)} dxdy = \sqrt{\frac{\pi^2}{\left(ac-\frac{b^2}{4}\right)}},$$

es decir

$$\boxed{\begin{aligned}
&I_{2d} = \int_{\mathbb{R}^2} e^{-\left(ax^2+bxy+cy^2\right)} dxdy = \int_{\mathbb{R}^2} e^{-X^T M X} dxdy = \sqrt{\frac{\pi^2}{\det M}}, \\
&X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad M = \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}.
\end{aligned}}$$

Usando el mismo procedimiento, se puede probar que si M es una matriz simétrica de $n\times n$ con entradas reales y valores propios positivos, se cumple

$$\boxed{\int_{\mathbb{R}^n} e^{-X^T M X} \Pi_{i=1}^n dx_i = \sqrt{\frac{\pi^n}{\det M}}.} \tag{26.9}$$

26.2. Integrales Gaussianas 2

Otra integral importante es

$$I = \int_{-\infty}^{\infty} e^{-\alpha x^2 + jx} dx, \qquad \alpha > 0.$$

Para obtener esta integral observemos que

$$\begin{aligned} -\alpha x^2 + jx &= -\alpha\left(x^2 - \frac{jx}{\alpha}\right) \\ &= -\alpha\left(x^2 - \frac{jx}{\alpha} + \frac{j^2}{4\alpha^2} - \frac{j^2}{4\alpha^2}\right) \\ &= -\alpha\left(x - \frac{j}{2\alpha}\right)^2 + \frac{j^2}{4\alpha}, \end{aligned}$$

es decir

$$-\alpha x^2 + jx = -\alpha\left(x - \frac{j}{2\alpha}\right)^2 + \frac{j^2}{4\alpha}.$$

Usando este resultado obtenemos

$$I = \int_{-\infty}^{\infty} e^{-\alpha x^2 + Jx} dx = \int_{-\infty}^{\infty} e^{-\alpha\left(x - \frac{j}{2\alpha}\right)^2 + \frac{j^2}{4\alpha}} dx,$$

por lo que considerando el cambio de variable

$$x' = x - \frac{j}{2\alpha}$$

y la integral Gaussiana (26.1), llegamos a

$$\begin{aligned} I &= \int_{-\infty}^{\infty} e^{-\alpha x^2 + jx} dx = e^{\frac{j^2}{4\alpha}} \int_{-\infty}^{\infty} e^{-\alpha x'^2} dx' \\ &= \sqrt{\frac{\pi}{\alpha}} e^{\frac{j^2}{4\alpha}}, \end{aligned}$$

en otras palabras

$$\boxed{\int_{-\infty}^{\infty} e^{-\alpha x^2 + jx} dx = \sqrt{\frac{\pi}{\alpha}} e^{\frac{j^2}{4\alpha}}.}$$

Ahora veamos el caso de n dimensiones

$$\int_{\mathbb{R}^n} e^{-X^T M X + J \cdot X} \Pi_{i=1}^n dx_i,$$

donde J es un vector constante de n dimensiones.

Primero notemos que debido a que M es una matriz simétrica se cumple

$$\begin{aligned} M^T &= M, \\ M^{-1} &= (M^T)^{-1} = (M^{-1})^T, \end{aligned}$$

por lo que

$$\begin{aligned} &\left(X - \frac{M^{-1}J}{2}\right)^T M \left(X - \frac{M^{-1}J}{2}\right) = \\ = &\ X^T M X - \frac{1}{2}\left(J^T (M^{-1})^T M X + X^T M M^{-1} J\right) + \frac{1}{4} J^T (M^{-1})^T M M^{-1} X \\ = &\ X^T M X - \frac{1}{2}\left(J^T X + X^T J\right) + \frac{1}{4} J^T M^{-1} J. \end{aligned}$$

Además como

$$J \cdot X = J^T X = X J^T,$$

se llega a

$$\left(X - \frac{M^{-1}J}{2}\right)^T M \left(X - \frac{M^{-1}J}{2}\right) = X^T M X - J \cdot X + \frac{1}{4} J^T M^{-1} J,$$

de donde

$$X^T M X - J \cdot X = \left(X - \frac{M^{-1}J}{2}\right)^T M \left(X - \frac{M^{-1}J}{2}\right) - \frac{1}{4} J^T M^{-1} J.$$

Por lo que

$$\begin{aligned} \int_{\mathbb{R}^n} e^{-X^T M X + J \cdot X} \Pi_{i=1}^n dx_i &= \int_{\mathbb{R}^n} e^{-\left(X - \frac{M^{-1}J}{2}\right)^T M \left(X - \frac{M^{-1}J}{2}\right) + \frac{1}{4} J^T M^{-1} J} \Pi_{i=1}^n dx_i. \\ &= e^{\frac{1}{4} J^T M^{-1} J} \int_{\mathbb{R}^n} e^{-\left(X - \frac{M^{-1}J}{2}\right)^T M \left(X - \frac{M^{-1}J}{2}\right)} \Pi_{i=1}^n dx_i, \end{aligned}$$

entonces con el cambio de variable

$$X' = X - \frac{M^{-1}J}{2}$$

se encuentra

$$\int_{\mathbb{R}^n} e^{-X^T M X + J \cdot X} \Pi_{i=1}^n dx_i = e^{\frac{1}{4} J^T M^{-1} J} \int_{\mathbb{R}^n} e^{-X'^T M X'} \Pi_{i=1}^n dx_i$$

y usando la integral (26.9) se obtiene

$$\boxed{\int_{\mathbb{R}^n} e^{-X^T M X + J \cdot X} \Pi_{i=1}^n dx_i = \sqrt{\frac{\pi^n}{\det M}} e^{\frac{1}{4} J^T M^{-1} J}.}$$

26.3. Ejercicios

1.- Obtener la integral

$$\int_{\mathbb{R}^2} e^{-3x^2-2xy-2y^2+3x+8y}dxdy.$$

2.- Obtener la integral

$$\int_{\mathbb{R}^3} e^{-3x^2+2xy-4xz-2yz-5y^2-2z^2+5x+3y-z}dxdydz.$$

Capítulo 27

Transformada discreta de Fourier y la matriz de Fourier

Como una aplicación de las matrices unitarias, en este capítulo estudiaremos la Transformada discreta de Fourier, en la cual aparece de manera natural una matriz unitaria.

La Transformada discreta de Fourier es muy útil para diversas aplicaciones, en particular para analizar señales eléctricas.

27.1. Definición

Supongamos que en el espacio vectorial $\mathbb{C}^N$, tenemos el vector

$$x = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{pmatrix}, \tag{27.1}$$

se define la transformada Discreta de Fourier del vector x como

$$\boxed{X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn}, \qquad k = 0, 1, \cdots, N-1.} \tag{27.2}$$

Note que para cada valor de k tenemos un valor de X_k. Por ejemplo, si $k=0$ se tiene

$$X_0 = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n = \frac{x_0 + x_1 + \cdots + x_{N-1}}{\sqrt{N}},$$

mientras que para $k=1$ se llega a

$$\begin{aligned} X_1 &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} n} \\ &= \frac{1}{\sqrt{N}} \left(x_0 + x_1 e^{-\frac{2\pi i}{N}} + x_2 e^{-\frac{4\pi i}{N}} + \cdots + x_{N-1} e^{-\frac{2\pi i (N-1)}{N}} \right), \end{aligned}$$

para $k=2$ se obtiene

$$\begin{aligned} X_2 &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-\frac{4\pi i}{N} n} \\ &= \frac{1}{\sqrt{N}} \left(x_0 + x_1 e^{-\frac{4\pi i}{N}} + x_2 e^{-\frac{8\pi i}{N}} + \cdots + x_{N-1} e^{-\frac{4\pi i (N-1)}{N}} \right), \end{aligned}$$

además para $k = N-1$ se encuentra

$$\begin{aligned} X_{N-1} &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-\frac{2(N-1)\pi i}{N} n} \\ &= \frac{1}{\sqrt{N}} \left(x_0 + x_1 e^{-\frac{2(N-1)\pi i}{N}} + x_2 e^{-\frac{4(N-1)\pi i}{N}} + \cdots + x_{N-1} e^{-\frac{2\pi i (N-1)^2}{N}} \right). \end{aligned}$$

Por lo tanto, mediante la transformada discreta de Fourier pasamos del vector x definido en (27.1) al vector

$$X = \begin{pmatrix} X_0 \\ X_2 \\ \vdots \\ X_{N-1} \end{pmatrix}. \qquad (27.3)$$

Así, la transformada discreta de Fourier define una transformación entre vectores del mismo espacio.

27.2. Matriz de Fourier

Se puede observar que definiendo

$$\boxed{\omega = e^{-\frac{2\pi i}{N}},} \tag{27.4}$$

la transformada discreta de Fourier (27.2) se puede escribir como

$$\boxed{X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n \omega^{kn},} \tag{27.5}$$

en particular se encuentra

$$\boxed{\begin{aligned}
X_0 =& \frac{1}{\sqrt{N}} \left(x_0 + x_1 + \cdots + x_{N-1}\right), \\
X_1 =& \frac{1}{\sqrt{N}} \left(x_0 + x_1\omega + x_2\omega^2 + \cdots + x_{N-1}\omega^{N-1}\right), \\
X_2 =& \frac{1}{\sqrt{N}} \left(x_0 + x_1\omega^2 + x_2\omega^4 + \cdots + x_{N-1}\omega^{2(N-1)}\right), \\
&\vdots \\
X_{N-1} =& \frac{1}{\sqrt{N}} \left(x_0 + x_1\omega^{(N-1)} + x_2\omega^{2(N-1)} + \cdots + x_{N-1}\omega^{(N-1)^2}\right).
\end{aligned}}$$

Por lo que, definiendo la matriz de Fourier

$$\boxed{\mathbb{W} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix}} \tag{27.6}$$

tenemos

$$\boxed{\begin{pmatrix} X_0 \\ X_2 \\ \vdots \\ X_{N-1} \end{pmatrix} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{pmatrix}.}$$

Así, con la matriz de Fourier (27.6) la transformada discreta de Fourier (38.1) se puede escribir como

$$X = \mathbb{W}x.$$

Por ejemplo, para el caso $N = 2$ tenemos

$$\boxed{\mathbb{W}_2 = H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},}$$

a la matriz H se le llama matriz de Hadamard y es utilizada como una compuerta lógica en diferentes algoritmos de la computación cuántica [15].

Cuando $N = 3$ se obtiene

$$\mathbb{W}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \frac{-1+i\sqrt{3}}{2} & \frac{-1-i\sqrt{3}}{2} \\ 1 & \frac{-1-i\sqrt{3}}{2} & \frac{-1+i\sqrt{3}}{2} \end{pmatrix}$$

y si $N = 4$ se llega a

$$\mathbb{W}_4 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}.$$

27.3. La inversa de la matriz de Fourier

Se puede observar que las componentes de la matriz de Fourier $\mathbb{W}$ (27.6) se pueden escribir como

$$\mathbb{W}_{ij} = \frac{1}{\sqrt{N}} \omega^{ij}, \qquad i, j = 0, 1, 2, 3, \cdots, N-1.$$

Note que en esta matriz los índices empiezan en cero y teminan en N, por lo que tenemos una matriz de $N \times N$.

Ahora, considerando la definición de ω dada en (27.4), las componentes de la matriz adjunta de la matriz de Fourier son

$$(\mathbb{W}^\dagger)_{ij} = \frac{1}{\sqrt{N}} \left(\omega^{ji}\right)^* = \frac{1}{\sqrt{N}}\omega^{-ji},$$

es decir

$$(\mathbb{W}^\dagger)_{ij} = \frac{1}{\sqrt{N}}\omega^{-ji}.$$

De donde

$$\begin{aligned}
(\mathbb{W}\mathbb{W}^\dagger)_{ij} &= \mathbb{W}_{ik}\mathbb{W}^\dagger_{kj} \\
&= \frac{1}{\sqrt{N}}\omega^{ik}\frac{1}{\sqrt{N}}\omega^{-jk} \\
&= \frac{1}{N}\omega^{ik-jk} \\
&= \frac{1}{N}\omega^{k(i-j)},
\end{aligned}$$

esto es

$$(\mathbb{W}\mathbb{W}^\dagger)_{ij} = \frac{1}{N}\omega^{k(i-j)}.$$

Recordemos que, de acuerdo a la convención de la suma de Einstein, en esta expresión tenemos una suma sobre k. Así, por razones de claridad, escribiremos

$$(\mathbb{W}\mathbb{W}^\dagger)_{ij} = \frac{1}{N}\sum_{k=0}^{N-1}\omega^{k(i-j)}. \qquad (27.7)$$

Para el caso en que

$$i = j$$

tenemos

$$\begin{aligned}
(\mathbb{W}\mathbb{W}^\dagger)_{ii} &= \frac{1}{N}\sum_{k=0}^{N-1}\omega^{k0} \\
&= \frac{1}{N}\sum_{k=0}^{N-1}1 \\
&= \frac{N}{N} = 1,
\end{aligned}$$

entonces

$$(\mathbb{W}\mathbb{W}^{\dagger})_{ii} \quad = \quad 1.$$

Mientras que en el caso

$$i \neq j$$

podemos definir

$$z = \omega^{i-j} \tag{27.8}$$

y la suma (27.7) se puede escribir como

$$(\mathbb{W}\mathbb{W}^{\dagger})_{ij} \quad = \quad \frac{1}{N}\sum_{k=0}^{N-1} z^k.$$

Antes de continuar recordemos la suma geométrica

$$1 + q + q^2 + \cdots + q^N = \frac{q^{N+1} - 1}{q - 1}, \qquad q \neq 1,$$

que implica la expresión

$$(\mathbb{W}\mathbb{W}^{\dagger})_{ij} \quad = \quad \frac{1}{N}\sum_{k=0}^{N-1} z^k = \frac{1}{N}\frac{z^N - 1}{z - 1},$$

es decir

$$(\mathbb{W}\mathbb{W}^{\dagger})_{ij} \quad = \quad \frac{1}{N}\frac{z^N - 1}{z - 1}.$$

Ahora, considerando la definición de z dada en (27.8) y la definición ω dada en (27.4) se encuentra

$$\begin{aligned} z^N &= \left(\omega^{i-j}\right)^N \\ &= \left(e^{\frac{2\pi}{N}}\right)^{(i-j)N} \\ &= e^{\frac{2\pi(i-j)N}{N}} \\ &= e^{2\pi(i-j)} = 1, \end{aligned}$$

es decir,

$$z^N = 1.$$

Por lo tanto, si $i \neq j$, se encuentra

$$(\mathbb{W}\mathbb{W}^\dagger)_{ij} \quad = \quad 0.$$

Estos resultados nos indican que se cumple la igualdad

$$(\mathbb{W}\mathbb{W}^\dagger)_{ii} \quad = \quad \delta_{ij},$$

en otras palabras

$$\mathbb{W}\mathbb{W}^\dagger = I.$$

Por lo tanto, se cumple la igualdad

$$\boxed{\mathbb{W}^\dagger = \mathbb{W}^{-1}}$$

y la matriz de Fourier $\mathbb{W}$ es una matriz unitaria.

En particular se encuentra

$$\begin{aligned}
\mathbb{W}_2^{-1} &= \mathbb{W}_2^\dagger = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \\
\mathbb{W}_3^{-1} &= \mathbb{W}_3^\dagger = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \frac{-1-i\sqrt{3}}{2} & \frac{-1+i\sqrt{3}}{2} \\ 1 & \frac{-1+i\sqrt{3}}{2} & \frac{-1-i\sqrt{3}}{2} \end{pmatrix}, \\
\mathbb{W}_4^{-1} &= \mathbb{W}_4^\dagger = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}.
\end{aligned}$$

Además, con matriz

$$\mathbb{W}^\dagger = \frac{1}{\sqrt{N}}\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \cdots & \omega^{-(N-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \cdots & \omega^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{-(N-1)} & \omega^{-2(N-1)} & \cdots & \omega^{-(N-1)(N-1)} \end{pmatrix} \tag{27.9}$$

se encuentra la transformada discreta Inversa de Fourier, la cual está dada por

$$\boxed{x_n = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi i}{N}kn}.}$$

27.4. Definición alterna

La transformada discreta de Fourier tiene aplicaciones en diferentes áreas y en algunos casos se toman la siguiente definición

$$\boxed{X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N}kn}, \qquad k = 0, 1, \cdots, N-1.} \tag{27.10}$$

cuya inversa es

$$\boxed{x_n = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi i}{N}kn}.}$$

La definición (27.10) es la más usada en diferentes ingenierías y por esa razón es la que se calcula en varios algortimos computacionales. Así, si se emplea algún programa computacional es importante saber que definición se usa.

Para grandes cantidades de datos los algoritmos deben realizar una cantidad enorme de operaciones, lo que los vuelve lentos. En estos casos es mejor emplear otros algortimos, como la llamada Transformada Rápida de Fourier (FFT, por su siglas inglés), que es un algoritmo más rápido para obtener la Transformada discreta de Fourier.

27.5. Operaciones con Python

Para obtener la Transformada Rápida de Fourier del vector

$$(1, 3, 4, 7, 3, 5, -1)$$

se usa el código de Python

```
from scipy.fft import fft
import numpy as np
x = np.array([1.0, 3.0, 4.0, 7.0, 3.0, 5.0, -1.0])
y = fft(x)
y
```

que da como resultado

```
array([22.-0.j,-8.76539748-3.88793297j,
-1.31886366-1.20626946j,2.58426114-6.41707809j,
2.58426114+6.41707809j,-1.31886366+1.20626946j,
-8.76539748+3.88793297j])
```

en este caso se emplea la definción (27.10).

Para obtener obtener la transformada discreta de Fourier de la definición (27.2) debemos multiplicar el resultado obtenido por $1/\sqrt{7}$.

Capítulo 28

Polinomio característico

Para obtener los valores propios de una matriz se deben obtener las raíces de un polinomio. En este capítulo veremos algunas propiedades de dicho polinomio.

28.1. Definición y propiedades

Para obtener los valores propios debemos resolver la ecuación

$$\boxed{\det (A - \lambda I) = 0.}$$

En particular, en el caso de una matriz de 2×2

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

tenemos

$$\det (A - \lambda I) = \lambda^2 - \lambda (A_{11} + A_{22}) + (A_{11}A_{22} - A_{12}A_{21}),$$

es decir

$$\boxed{\det (A - \lambda I) = \lambda^2 - \lambda Tr(A) + \det A.}$$

Para el caso de una matriz de 3×3

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix},$$

se encuentra

$$\begin{aligned}
\det(A-\lambda I) &= (-)\lambda^3 + \lambda^2 (A_{11}+A_{22}+A_{33}) \\
&\quad -\lambda\Big(A_{22}A_{33} + A_{33}A_{11} + A_{22}A_{11} \\
&\quad -A_{13}A_{31}A_{23}A_{32} - A_{12}A_{21}\Big) \\
&\quad +A_{11}(A_{22}A_{33} - A_{23}A_{32}) \\
&\quad -A_{12}(A_{21}A_{32} - A_{23}A_{31}) \\
&\quad +A_{13}(A_{21}A_{32} - A_{22}A_{31})
\end{aligned}$$

es decir

$$\begin{aligned}
\det(A-\lambda I) &= (-)\lambda^3 + \lambda^2 Tr(A) \\
&\quad -\lambda\left(A_{22}A_{33} + A_{33}A_{11} + A_{22}A_{11} - A_{13}A_{31}A_{23}A_{32} - A_{12}A_{21}\right) \\
&\quad +\det A.
\end{aligned}$$

En general, para una matriz de $n\times n$ tenemos

$$\boxed{\det(A-\lambda I) = (-)^n\lambda^n + (-)^{n-1}\lambda^{n-1}Tr(A) + \cdots + \det A.}$$

Para probar este resultado observemos que

$$(A-\lambda I)_{ij} = A_{ij} - \delta_{ij}\lambda,$$

por lo que, usando la definición del determinante de una matriz (4.12), obtenemos

$$\det(A-\lambda I) = \epsilon_{i_1 i_2\cdots i_n}(A_{1i_1} - \delta_{1i_1}\lambda)(A_{2i_2} - \delta_{2i_2}\lambda)\cdots(A_{ni_n} - \delta_{ni_n}\lambda).$$

Desarrollando este producto en potencias de λ encontramos

$$\begin{aligned}
&\det(A-\lambda I) = (-)^n\lambda^n\epsilon_{i_1i_2\cdots i_n}\delta_{1i_1}\delta_{2i_2}\cdots\delta_{ni_n} + (-)^{n-1}\lambda^{n-1}\epsilon_{i_1i_2\cdots i_n} \\
&\left(\delta_{2i_2}\delta_{3i_3}\cdots\delta_{ni_n}A_{1i_1} + \delta_{1i_1}\delta_{3i_3}\cdots\delta_{ni_n}A_{2i_2} + \cdots + \delta_{1i_1}\delta_{2i_2}\cdots\delta_{(n-1)i_{n-1}}A_{ni_n}\right) \\
&+\cdots+\epsilon_{i_1i_2\cdots i_n}A_{1i_1}A_{2i_2}\cdots A_{ni_n} \\
&= (-)^n\lambda^n\epsilon_{12\cdots n} \\
&+(-)^{n-1}\lambda^{n-1}\Big(\epsilon_{i_123\cdots n}A_{1i_1} + \epsilon_{1i_23\cdots n}A_{2i_2} + \cdots + \epsilon_{12\cdots(n-1)i_n}A_{ni_n}\Big) \\
&+\cdots+\det A \\
&= (-)^n\lambda^n + (-)^{n-1}\lambda^{n-1}(A_{11} + A_{22} + \cdots A_{nn}) + \cdots + \det A \\
&= (-)^n\lambda^n + (-)^{n-1}\lambda^{n-1}Tr(A) + \cdots + \det A,
\end{aligned}$$

es decir,

$$\det (A - \lambda I) = (-)^n \lambda^n + (-)^{n-1} \lambda^{n-1} Tr(A) + \cdots + \det A.$$

que es lo queríamos demostrar

Así, tenemos el polinomio

$$\boxed{P(\lambda) = \det (A - \lambda I) = a_0 + a_1 \lambda + \cdots + a_n \lambda^{n-1} + a_n \lambda^n,} \tag{28.1}$$

con

$$\boxed{a_0 = \det A, \cdots, a_{n-1} = (-)^{n-1} Tr(A), a_n = (-)^n.} \tag{28.2}$$

Un resultado interesate sobre el polinomio característico de una matriz es que dos matrices similares tienen el mismo polinomio característico, para probar esta afirmación supongamos que

$$A' = UAU^{-1}$$

entonces

$$\begin{aligned} P'(\lambda) &= \det (A' - \lambda I) \\ &= \det \left(UAU^{-1} - \lambda I \right) \\ &= \det \left(UAU^{-1} - \lambda UIU^{-1} \right) \\ &= \det \left(U (A - \lambda I) U^{-1} \right) \\ &= \det U \det (A - \lambda I) \det U^{-1} \\ &= \frac{\det U}{\det U} \det (A - \lambda I) \\ &= \det (A - \lambda I) \\ &= P(\lambda), \end{aligned}$$

esto es,

$$P'(\lambda) = P(\lambda),$$

que es lo que queríamos demostrar.

28.2. Teorema de Cayley-Hamilton

Otro resultado interesante es que la matriz A es raíz de su polinomio característico, es decir, se satisface

$$\boxed{P(A) = a_0 + a_1 A + \cdots + a_{n-1}A^{n-1} + a_n A^n = 0,}$$

con los coeficientes dados por (28.2).

Para probar este resultado recordemos que dada una matriz B de $n \times n$ su matriz adjunta de está dada por (5.2)

$$(adj B)_{ij} = \frac{1}{(n-1)!}\epsilon_{ii_2i_3\cdots i_n}\epsilon_{jj_2j_3\cdots j_n}B_{i_2j_2}B_{i_3j_3}\cdots B_{i_nj_n},$$

también recordemos que esta matriz satisface

$$B(adj B) = (\det B)I. \tag{28.3}$$

En particular, para la matriz $A - \lambda I$ tenemos

$$\begin{aligned}&(adj(A-\lambda I))_{ij}\\ &= \frac{1}{(n-1)!}\epsilon_{ii_2i_3\cdots i_n}\epsilon_{jj_2j_3\cdots j_n}(A-\lambda I)_{i_2j_2}(A-\lambda I)_{i_3j_3}\cdots(A-\lambda I)_{i_nj_n},\end{aligned}$$

se puede notar que la potencia máxima potencia de λ de esta expresión es $n-1$, por lo que

$$adj(A-\lambda I) = C_0 + \lambda C_1 + \lambda^2 C_2 + \cdots + \lambda^{n-1}C_{n-1}, \tag{28.4}$$

con

$$C_0, C_1, \cdots C_{n-1}$$

un conjunto de matrices.

Ahora, de la igualdad (28.3) se obtiene

$$(A-\lambda I)(adj(A-\lambda I)) = \det(A-\lambda I).$$

Usando las igualdades (28.4) y (28.1) en esta última ecuación, tenemos

$$(A-\lambda I)\left(C_0 + \lambda C_1 + \cdots + \lambda^{n-1}C_{n-1}\right) = \left(a_0 + a_1\lambda + \cdots + a_n\lambda^n\right)I,$$

por lo que

$$\begin{aligned}&AC_0 + \lambda\left(AC_1 - C_0\right) + \lambda^2\left(AC_2 - C_1\right) + \cdots + \lambda^{n-1}\left(AC_{n-1} - C_{n-2}\right)\\ &-\lambda^n C_{n-1} = (a_0 + a_1\lambda + \cdots + a_n\lambda^n)I.\end{aligned}$$

De donde, igualando potencias de λ encontramos

$$\begin{aligned}a_0 I &= AC_0,\\ a_1 I &= AC_1 - C_0,\\ a_2 I &= AC_2 - C_1,\\ &\vdots\\ a_{n-1} I &= AC_{n-1} - C_{n-2},\\ a_n I &= -C_{n-1},\end{aligned}$$

de donde

$$\begin{aligned}a_0 I &= AC_0,\\ a_1 A &= A^2C_1 - AC_0,\\ a_2 A^2 &= A^3C_2 - A^2C_1,\\ &\vdots\\ a_{n-1} A^{n-1} &= A^nC_{n-1} - A^{n-1}C_{n-2},\\ a_n A^n &= -A^nC_{n-1}.\end{aligned}$$

Así,

$$\boxed{\begin{aligned}P(A) =& a_0 + a_1A + \cdots + a_{n-1}A^{n-1} + a_nA^n\\ =& AC_0 + A^2C_1 - AC_0 + A^3C_2 - A^2C_1 + \cdots +\\ & + A^nC_{n-1} - A^{n-1}C_{n-2} - A^nC_{n-1} = 0,\end{aligned}}$$

que es lo que se quería demostrar.

Independientemente de la belleza del teorema de Caley-Hamilton, es muy últil para algunos métodos numéricos. Por ejemplo, el método Krylov para resolver sistemas lineales se basa en este teorema.

El teorema de Caley-Hamilton fue demostrado por Caley para el caso de tres dimensiones y Hamilton realizó una demostración para tres dimensiones con vectores y para cuatro dimensiones con cuaterniones. La demostración general fue dada por Georg Frobenius en 1878. Debido al respeto que le inspiraban Caley y Hamilton, Frobenius decidió dar los nombres de estos dos matemáticos al teorema.

William R. Hamilton (1805-1865) fue un físico-matemático irlandés que trabajó en mecánica clásica, astronomía, geometría, óptica, teoría de gráficas y álgebra. En particular, ayudó a reformular la mecánica de Newton en lo que hoy se conoce como mecánica hamiltoniana. En temas geométricos fundó las bases de lo que hoy se llama geometría simpléctica y en temas algebraicos se le atribuyen los fundamentos de los llamados cuaterniones. Su padre fue un abogado y activista político que apoyó la independencia de Irlanda de Inglaterra. Por problemas económicos de su padre, siendo muy pequeño fue separado de su familia y a los 12 años quedó huerfano de padre y madre. Después de la muerte de sus padres, quedó bajo el cuidado de uno de sus tíos. Siendo un adolescente, se refugió en las obras clásicas de Newton, Lagrange y Laplace, notablemente encontró algunos errores en la obra de Laplace. Realizó sus estudios en el Trinity College de Dublín. Apesar de ser un gran matemático, era timido y sufría de depresiones, incluso estuvo apunto de cometer suicidio. Al final de su vida, la salud de Hamilton se deterioró debido a su alcoholismo y a sus malos hábitos alimenticios. Después de Newton, es considerado el más grande genio que han dado las islas Británicas.

28.3. Ejercicios

1.- Para la matriz

$$A = \begin{pmatrix} 5 & 2 \\ -8 & -3 \end{pmatrix}$$

obtener su polinomio característico

$$P(\lambda)$$

y mostrar que se cumple

$$P(A) = 0.$$

2.- Para la matriz

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

obtener su polinomio característico

$$P(\lambda)$$

y mostrar que se cumple

$$P(A) = 0.$$

Capítulo 29

Espacios invariantes y bloques de Jordan

Ahora estudiaremos matrices que tienen valores propios repetidos. Para lograr este objetivo, definiremos el concepto de espacio invariante.

29.1. Espacio invariantes

Sean V un espacio vectorial de dimensión n sobre el campo K y

$$T : V \to V$$

una transformación lineal. Si U es un subespacio vectorial de V tal que

$$\boxed{T(U) \subset U,}$$

se dice que U es un espacio invariante bajo T.

Por ejemplo, debido a que T es una transformación lineal se cumple $T(0) = 0$, entonces el conjunto

$$\boxed{U = \{0\}}$$

es un espacio invariante.

También podemos observar

$$Ker(T) \subset V \qquad \text{y} \qquad T\left(Ker(T)\right) = \{0\} \subset Ker(T).$$

Por lo tanto, el conjunto $Ker(T)$ es un subespacio invariante bajo T.

Otro ejemplo trivial de espacio invariante está dado por V.

Ahora, supongamos que $v \in V, v \neq 0$ y es vector propio de T con valor propio $\lambda \in K$, es decir

$$\boxed{T(v) = \lambda v.}$$

Entonces, el espacio vectorial

$$\boxed{\langle v \rangle = \{w \in V | \exists \alpha \in K, w = \alpha v\},}$$

es un espacio invariante. En efecto, sea $w_1 \in T(\langle v \rangle)$ entonces existe $w_2 \in \langle v \rangle$ tal que $T(w_2) = w_1$. Como $w_2 \in \langle v \rangle$, entonces existe $\alpha \in K$ tal que $w_2 = \alpha v$. Por lo tanto,

$$\begin{aligned} w_1 &= T(w_2) \\ &= T(\alpha v) \\ &= \alpha T(v) \\ &= \alpha \lambda v \in \langle v \rangle, \end{aligned}$$

es decir,

$$w_1 \in \langle v \rangle,$$

que implica

$$T(\langle v \rangle) \subset \langle v \rangle.$$

Así, el espacio generado por un vector propio de T es un espacio invariante.

Además, se puede mostrar que si los elementos del conjunto de vectores

$$\boxed{\{v_1, v_2, \cdots, v_m\}, \qquad m \leq n,} \tag{29.1}$$

son vectores propios de T, entonces el conjunto generado por estos vectores definido por

$$\boxed{\langle v_1, v_2, \cdots, v_m \rangle = \{w \in V | \exists \alpha_i \in K, w = \alpha_i v_i\}}$$

es un espacio vectorial invariante bajo T. En efecto, recordemos que si los vectores (29.1) son vectores propios de T, entonces existen $\lambda_i \in K(i = 1, 2, \cdots, m)$, tal que

$$\begin{aligned} T(v_1) &= \lambda_1 v_1, \\ T(v_2) &= \lambda_2 v_2, \\ &\vdots \\ T(v_m) = \lambda_m v_m, \end{aligned} \tag{29.2}$$

ahora supongamos que

$$w_1 \in T\left(\langle v_1, v_2, \cdots, v_m \rangle\right),$$

entonces existe $w_2 \in \langle v_1, v_2, \cdots, v_m \rangle$ tal que

$$T(w_2) = w_1.$$

Además, debido a que w_2 es parte del espacio vectorial generado por el conjunto (29.1), entonces existen m escalares $\alpha_i \in K$ tal que

$$w_2 = \alpha_i v_i,$$

de donde

$$\begin{aligned} w_1 &= T(w_2) \\ &= T(\alpha_i v_i) \\ &= \alpha_i T(v_i) \\ &= \alpha_i \lambda_i v_i \in \langle v_1, v_2, \cdots, v_m \rangle, \end{aligned}$$

es decir,

$$w_1 \in \langle v_1, v_2, \cdots, v_m \rangle,$$

que implica

$$T\left(\langle v_1, v_2, \cdots, v_m \rangle\right) \subset \langle v_1, v_2, \cdots, v_m \rangle.$$

Esto quiere decir que el espacio generado por los vector propios de T es un espacio invariante.

Ahora, supongamos que U_1 y U_2 son dos subespacios vectoriales de V y cumplen

$$\boxed{U_1 \cup U_2 = V, \qquad U_1 \cap U_2 = \emptyset,}$$

también supongamos que estos subespacios son invariantes bajo T. Entonces si

$$\{u_1, u_2, \cdots, u_m\}, \qquad m < n,$$

es base de U_1 y

$$\{w_1, w_2, \cdots, w_k\}, \qquad m + k = n,$$

es base de U_2, el conjunto

$$\{u_1, u_2, \cdots, u_m, w_1, w_2, \cdots, w_k\},$$

es base de V. En este caso se tiene

$$\boxed{\begin{aligned}
T(u_1) =& \alpha_{11}u_1 + \alpha_{12}u_2 + \cdots + \alpha_{1m}u_m + 0w_1 + 0w_2 + \cdots + 0w_k,\\
T(u_2) =& \alpha_{21}u_1 + \alpha_{22}u_2 + \cdots + \alpha_{2m}u_m + 0w_1 + 0w_2 + \cdots + 0w_k,\\
&\vdots\\
T(u_m) =& \alpha_{m1}u_1 + \alpha_{m2}u_2 + \cdots + \alpha_{mm}u_m + 0w_1 + 0w_2 + \cdots + 0w_k,\\
T(w_1) =& 0u_1 + 0u_2 + \cdots + 0u_m + \beta_{11}w_1 + \beta_{12}w_2 + \cdots + \beta_{1k}w_k,\\
T(w_2) =& 0u_1 + 0u_2 + \cdots + 0u_m + \beta_{21}w_1 + \beta_{22}w_2 + \cdots + \beta_{2k}w_k,\\
&\vdots\\
T(w_k) =& 0u_1 + 0u_2 + \cdots + 0u_m + \beta_{21}w_1 + \beta_{k2}w_2 + \cdots + \beta_{kk}w_k,
\end{aligned}}$$

que se puede escribir como

$$\boxed{\begin{pmatrix} T(u_1)\\ T(u_2)\\ \vdots\\ T(u_m)\\ T(w_1)\\ T(w_2)\\ \vdots\\ T(w_k) \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} & 0 & \cdots & 0\\ \alpha_{21} & \cdots & \alpha_{2m} & 0 & \cdots & 0\\ \vdots & & \vdots & \vdots & & \vdots\\ \alpha_{m1} & \cdots & \alpha_{mm} & 0 & \cdots & 0\\ 0 & \cdots & 0 & \beta_{11} & \cdots & \beta_{1k}\\ 0 & \cdots & 0 & \beta_{21} & \cdots & \beta_{2k}\\ \vdots & \cdots & \vdots & \vdots & & \vdots\\ 0 & \cdots & 0 & \beta_{k1} & \cdots & \beta_{kk} \end{pmatrix} \begin{pmatrix} u_1\\ u_2\\ \vdots\\ u_m\\ w_1\\ w_2\\ \vdots\\ w_k \end{pmatrix}.}$$

Por lo tanto, usando base de estos espacios invariantes U_1 y U_2 la matriz asociada a la transformación lineal T se simplifica y toma la forma

$$\boxed{M = \begin{pmatrix} X_1 & 0\\ 0 & X_2 \end{pmatrix}}$$

con

$$\boxed{X_{1ij} = \alpha_{ij}^T, \qquad X_{2ij} = \beta_{ij}^T.}$$

En general, si los subespacios vectoriales de V

$$\boxed{U_1, U_2, \cdots U_r, \qquad r \leq n}$$

son invariantes bajo T y satisfacen

$$\boxed{V = U_1 \cup U_2 \cup \cdots \cup U_r, \qquad U_i \cap U_j = \emptyset, \quad i \neq j \qquad i, j = 1, 2, \cdots r,}$$

entonces, usando las bases de estos subespacios invariantes, la matriz asociada a la transformación lineal T tiene la forma

$$\boxed{M = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_r \end{pmatrix},}$$

donde X_i $(i = 1, 2, \cdots, r)$ es la matriz asociada a la transformación T restringida al espacio invariante U_i. Se puede observar que X_i es una matriz de $d_i \times d_i$, donde d_i es la dimensión del subespacio U_i.

Así la forma de la matriz asociada a la transformación lineal T se simplifica usando las bases de los subespacios invariantes de V bajo T.

Nótese que el máximo número de espacios invariantes diferentes que puede tener V es n. Cuando V tiene n vectores invariantes, cada matriz X_i tiene dimensión uno y es un elemento de K. También note que cuando la matriz asociada a la transformación T es diagonalizable, cada vector propio está asociado a un espacio invariante de dimensión uno y la matriz asociada a la transformación T se reduce a una matriz diagonal de la forma

$$M = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Así, si V tiene dimensión n y tiene n subespacios vectoriales invariantes bajo T, la matriz asociada a la trasformación T es diagonalizable.

29.2. Valores propios repetidos

Sabemos que el polinomio característico de la matriz asociada a la transformación T es de la forma

$$P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).$$

Si todos los valores propios son diferentes, la matriz es diagonalizable y los subespacios invariantes bajo T están dados por los espacios generado por cada vector propio.

Supongamos que el valor propio λ_1 se repite k veces y los demás no se repiten, entonces tenemos los valores propios

$$\boxed{\lambda_1, \lambda_2, \cdots \lambda_{n-k}}$$

y los vectores propios

$$\boxed{v_1, v_2, \cdots v_{n-k}.}$$

Además, en este caso el polinomio característico de la matriz asociada a la transformación T es de la forma

$$\boxed{P(\lambda) = (\lambda - \lambda_1)^k (\lambda - \lambda_2)\cdots(\lambda - \lambda_{n-k}).}$$

Observe que si restringimos la transformación T al espacio generado por los vectores propios

$$v_2, \cdots, v_{n-k} \tag{29.3}$$

el polinomio característico de la matriz asociada a la transformación T restringida a este subespacio es

$$\boxed{P_{\text{restringido}}(\lambda) = (\lambda - \lambda_2)\cdots(\lambda - \lambda_{n-k}),}$$

mientras que la matriz asociada asociada a la transformación T restringida a este subespacio es

$$\boxed{M_{\text{restringido}} = \begin{pmatrix} \lambda_2 & 0 & \cdots & 0 \\ 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-r} \end{pmatrix}.}$$

En este caso cada, uno de los vectores (29.3) nos dan un subespacio invariante bajo T.

Por lo tanto, es posible que el valor propio λ_1 esté asociado a un espacio invariante de dimensión r. Si existen r vectores propios de T linealmente independientes y asociados al valor propio λ_1, esos vectores propios formarían el subespacio vectorial invariante de dimensión r.

También es posible que existan $l < r$ vectores propios de T asociados al valor propio λ_1 o que v_1 sea el único vector propio de T asociado al valor propio λ_1. En todo caso debemos buscar los $r-1$ vectores linealmente independientes restantes que junto con v_1 formen el subespacio invariante de dimensión r.

Ahora, tomemos la redefinición

$$\boxed{w_0 = v_1, \qquad \lambda = \lambda_1}$$

que implica

$$\boxed{T(w_0) = \lambda w_0.}$$

Supongamos que existe un vector w_1 tal que

$$\boxed{(T - \lambda I)(w_1) = w_0,} \tag{29.4}$$

donde I es la función identidad. Note que otra forma de escribir la ecuación (29.4) es

$$\boxed{T(w_1) = \lambda w_1 + w_0.} \tag{29.5}$$

Si existe w_1 que satisface la ecuación (29.4) entonces los vectores

$$\boxed{\{w_1, w_0\}} \tag{29.6}$$

son linealmente independientes. Para probar esta afirmación supongamos que existen $\alpha_0, \alpha_1 \in K$ tal que

$$\alpha_1 w_1 + \alpha_0 w_0 = 0. \tag{29.7}$$

Aplicando la función $T - \lambda I$ a esta última ecuación, se obtiene

$$\alpha_1 (T - \lambda I)(w_1) + \alpha_0 (T - \lambda I)(w_0) = 0,$$

es decir,

$$\alpha_1 w_0 = 0.$$

Por lo tanto,

$$\alpha_1 = 0,$$

sustituyendo este resultado en la ecuación (29.7) se llegaba a

$$\alpha_0 w_0 = 0,$$

que implica

$$\alpha_0 = 0,$$

así

$$\alpha_1 = \alpha_0 = 0.$$

Por lo tanto, el conjunto de vectores (29.6) es linealmente independiente.

Además, el espacio vectorial generado por conjunto de vectores (29.6) es invariante bajo T. En efecto, supongamos que

$$\boxed{v \in \langle \{w_1, w_0\} \rangle\,,}$$

entonces existen $\beta_0, \beta_1 \in K$ tal que

$$v = \beta_1 w_1 + \beta_0 w_0,$$

de donde, recurriendo a la ecuación (29.5), se llega a

$$\begin{aligned} T(v) &= T\left(\beta_1 w_1 + \beta_0 w_0\right) \\ &= \beta_1 T(w_1) + \beta_0 T(w_0) \\ &= \beta_1\left(\lambda w_1 + w_0\right) + \beta_0 \lambda w_0 \\ &= \beta_1 \lambda w_1 + \left(\beta_1 + \beta_0 \lambda\right) w_0, \end{aligned}$$

esto es,

$$T(v) = \beta_1 \lambda w_1 + \left(\beta_1 + \beta_0 \lambda\right) w_0,$$

que implica

$$T(v) \in \langle \{w_1, w_0\} \rangle ,$$

por lo que

$$T\left(\langle \{w_1, w_0\} \rangle\right) \subset \langle \{w_1, w_0\} \rangle .$$

Así, el conjunto

$$\boxed{\langle \{w_1, w_0\} \rangle .}$$

es invariante bajo T.

En general, si existen $r-1$ vectores

$$\boxed{\{w_0, w_1, w_2, \cdots, w_{r-2}, w_{r-1}\}} \tag{29.8}$$

tal que

$$\boxed{\begin{aligned} (T-\lambda I)\left(w_k\right) =& w_{k-1}, \qquad k = 1, 2, \cdots, r-1, & (29.9) \\ (T-\lambda I)\left(w_0\right) =& 0, & (29.10) \end{aligned}}$$

entonces el conjunto de vectores (29.8) es linealmente independientes y el espacio vectorial generado por dichos vectores es invariante bajo T.

Antes de probar estas dos afirmaciones, notemos que

$$(T-\lambda I)^2(w_k) = (T-\lambda I)\left(w_{k-1}\right) = w_{k-2},$$

es decir
$$(T-\lambda I)^2(w_k)=w_{k-2},$$
en general se cumple
$$(T-\lambda I)^l(w_k)=w_{k-l},$$
que implica
$$\begin{aligned}(T-\lambda I)^k(w_k) &= w_0, && (29.11)\\ (T-\lambda I)^l(w_k) &= 0, \qquad l>k. && (29.12)\end{aligned}$$
Ahora, supongamos que existen $\alpha_0, \alpha_1, \cdots, \alpha_{r-1} \in K$ tal que
$$\alpha_0 w_0+\alpha_1 w_1+\alpha_2 w_2+\cdots+\alpha_{r-2}w_{r-2}+\alpha_{r-1}w_{r-1}=0. \qquad (29.13)$$
Entonces aplicando la función $(T-\lambda I)^{r-1}$ a esta combinación lineal y usando las igualdades (29.11)-(29.12), se llega a
$$\begin{aligned}&(T-\lambda I)^{r-1}\left(\alpha_0 w_0+\alpha_1 w_1+\alpha_2 w_2+\cdots+\alpha_{r-2}w_{r-2}+\alpha_{r-1}w_{r-1}\right)=\\ &=\alpha_0(T-\lambda I)^{r-1}(w_0)+\alpha_1(T-\lambda I)^{r-1}(w_1)+\alpha_2(T-\lambda I)^{r-1}(w_2)+\cdots\\ &+\alpha_{r-2}(T-\lambda I)^{r-1}(w_{r-2})+\alpha_{r-1}(T-\lambda I)^{r-1}(w_{r-1})\\ &=\alpha_{r-1}w_0=0\end{aligned}$$
es decir,
$$\alpha_{r-1}w_0=0,$$
por lo que
$$\alpha_{r-1}=0.$$
Usando este resultado en la ecuación (29.13) se llega a
$$\alpha_0 w_0+\alpha_1 w_1+\alpha_2 w_2+\cdots+\alpha_{r-2}w_{r-2}=0. \qquad (29.14)$$
De la misma forma, aplicando la función $(T-\lambda I)^{r-2}$ a la combinación lineal (29.14) se llega a
$$\alpha_{r-2}=0,$$
que implica
$$\alpha_0 w_0+\alpha_1 w_1+\alpha_2 w_2+\cdots+\alpha_{r-3}w_{r-3}=0.$$
Realizando varias veces el mismo procedimiento, encontramos que
$$\alpha_0=\alpha_1=\cdots=\alpha_{r-2}=\alpha_{r-1}=0,$$

por lo tanto, el conjunto de vectores (29.8) es linealmente independiente.

Además, si

$$v \in \langle w_0, w_1, w_2, \cdots, w_{r-2}, w_{r-1} \rangle,$$

entonces existen $\beta_0, \beta_1, \cdots, \beta_{r-1} \in K$ tal que

$$v = \beta_0 w_0 + \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_{r-2} w_{r-2} + \beta_{r-1} w_{r-1},$$

por lo que, notando que la ecuación (29.9) se puede escribir como

$$T(w_k) = \lambda w_k + w_{k-1},$$

se obtiene

$$\begin{aligned} T(v) &= T\left(\beta_0 w_0 + \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_{r-2} w_{r-2} + \beta_{r-1} w_{r-1}\right) \\ &= \beta_0 T(w_0) + \beta_1 T(w_1) + \beta_2 T(w_2) + \cdots + \beta_{r-2} T(w_{r-2}) + \beta_{r-1} T(w_{r-1}) \\ &= \beta_0 \lambda w_0 + \beta_1 \left(\lambda w_1 + w_0\right) + \beta_2 \left(\lambda w_2 + w_1\right) + \cdots \\ &\quad + \beta_{r-2} \left(\lambda w_{r-2} + w_{r-3}\right) + \beta_{r-1} \left(\lambda w_{r-1} + w_{r-2}\right) \\ &= \left(\lambda \beta_0 + \beta_1\right) w_0 + \left(\lambda \beta_1 + \beta_2\right) w_2 + \cdots \\ &\quad + \left(\lambda \beta_{r-2} + \beta_{r-1}\right) w_{r-2} + \beta_{r-1} \lambda w_{r-1}, \end{aligned}$$

así podemos concluir que

$$T(v) \in \langle w_0, w_1, w_2, \cdots, w_{r-2}, w_{r-1} \rangle,$$

que implica

$$T\left(\langle w_0, w_1, w_2, \cdots, w_{r-2}, w_{r-1} \rangle\right) \subset \langle w_0, w_1, w_2, \cdots, w_{r-2}, w_{r-1} \rangle.$$

Entonces el espacio vectorial generado por los vectores (29.8) es invariante bajo la transformación T.

Por lo cual, el conjunto de vectores

$$\boxed{w_0 = v_1, w_1, w_2, \cdots, w_{r-1}, v_2, v_3, \cdots, v_{r-n}}$$

son n vectores linealmente independientes y por lo tanto forman una base de V. Además, los subespacios vectoriales

$$\boxed{< w_0 = v_1, w_2, \cdots, w_{r-1} >, < v_2 >, < v_3 >, \cdots, < v_{r-n} >}$$

son invariantes bajo T.

Note que para $r = 2$ tenemos la base del espacio invariante $\{w_0, w_1\}$, que cumple

$$\boxed{\begin{aligned} T(w_0) =& \lambda w_0, \\ T(w_1) =& \lambda w_1 + w_0, \end{aligned}}$$

esta relación se puede escribir como

$$\boxed{\begin{pmatrix} T(w_0) \\ T(w_1) \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \end{pmatrix},}$$

así la matriz asociada asociada restringida esta base es

$$\boxed{J_2 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}}$$

a esta matriz se le llama bloque de Jordan de orden 2.

Si $r = 3$, la base del espacio invariante $\{w_0, w_1, w_2\}$ cumple

$$\boxed{\begin{aligned} T(w_0) =& \lambda w_0, \\ T(w_1) =& \lambda w_1 + w_0, \\ T(w_2) =& \lambda w_2 + w_1. \end{aligned}}$$

Esta relación se puede escribir como

$$\boxed{\begin{pmatrix} T(w_0) \\ T(w_1) \\ T(w_2) \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \end{pmatrix}.}$$

Así, con esta base se tiene la matriz asociada

$$\boxed{J_3 = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},}$$

que es bloque de Jordan de orden 3.

Para $r = 4$ tenemos que la base del espacio invariante es $\{w_0, w_1, w_2, w_3\}$ y cumple

$$\boxed{\begin{aligned} T(w_0) =& \lambda w_0, \\ T(w_1) =& \lambda w_1 + w_0, \\ T(w_2) =& \lambda w_2 + w_1, \\ T(w_3) =& \lambda w_3 + w_2, \end{aligned}}$$

que se puede escribir como

$$\boxed{\begin{pmatrix} T(w_0) \\ T(w_1) \\ T(w_2) \\ T(w_3) \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix},}$$

entonces tenemos la matriz asociada

$$\boxed{J_4 = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix},}$$

que es el bloque de Jordan de orden 4.

Para el caso general la base del espacio invariante es $\{w_0, w_1, \cdots, w_{r-1}\}$ y cumple las relaciones

$$\boxed{\begin{aligned} T(w_0) =& \lambda w_0, \\ T(w_1) =& \lambda w_1 + w_0, \\ T(w_2) =& \lambda w_2 + w_1, \\ & \vdots \\ T(w_{r-1}) =& \lambda w_{r-1} + w_{r-2} \end{aligned}}$$

que implican la matriz asociada

$$\boxed{J_r = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}}$$

que es bloque de Jordan de orden r.

Por lo tanto, usando la bases de los espacios invariantes, la matriz asociada a la transformación T es

$$\boxed{M = \begin{pmatrix} J_r & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-r} \end{pmatrix}}.$$

En general, el polinomio característico de una matriz puede tomar la forma

$$\boxed{P(\lambda) = (\lambda - \lambda_1)^{i_1} (\lambda - \lambda_2)^{i_2} \cdots (\lambda - \lambda_k)^{i_k}}$$

con

$$\boxed{i_1 + i_2 + \cdots + i_k = n.}$$

Por lo que, usando el mismo algoritmo que se emplea para el caso de un valor propio repetido, se puede mostrar que con los espacios invariantes para cada valor propio repetido la matriz asociada a la transformación T se puede escribir como

$$\boxed{M = \begin{pmatrix} J_{i_1} & 0 & \cdots & 0 \\ 0 & J_{i_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{i_k} \end{pmatrix}}.$$

en donde J_{i_j} es el bloque de Jordan de orden i_j.

29.3. Ejemplos

Como un ejemplo consideremos la matriz

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

de donde

$$A - \lambda I = \begin{pmatrix} 1-\lambda & 2 & -4 \\ 0 & 3-\lambda & 0 \\ 0 & 1 & 3-\lambda \end{pmatrix}$$

cuyo determinantes es

$$\det (A - \lambda I) = (3-\lambda)^2 (1-\lambda)$$

que implica los valores característicos

$$\lambda_1 = 1, \qquad \lambda_2 = 3.$$

Como podemos ver, el valor propio $\lambda_2 = 3$ tiene multiplicidad 2.

Para obtener los valores característicos debemos resolver los sistemas de ecuaciones

$$\begin{aligned} (A - \lambda_1 I) v_1 &= \begin{pmatrix} 0 & 2 & -4 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \\ (A - \lambda_2 I) v_2 &= \begin{pmatrix} -2 & 2 & -4 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \end{aligned}$$

de donde obtenemos los vectores propios

$$\begin{aligned} v_1 &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \\ v_2 &= \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Para obtener el tercer vector linealmente dependiente debemos buscar un vector v_3 que cumpla

$$(A - \lambda_2 I)\, v_3 \;=\; v_2,$$

es decir debemos resolver el sistema de ecuaciones

$$\begin{pmatrix} -2 & 2 & -4 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix},$$

que implica

$$v_3 \;=\; \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Con los vectores propios v_1, v_2, v_3 podemos formar las columnas de la matriz de cambio de base

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

cuya inversa es

$$U^{-1} = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

con estas matrices se puede comprobar que se cumple

$$\begin{aligned} U^{-1}AU \;&=\; \begin{pmatrix} 1 & -2 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & -4 \\ 0 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \\ &=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}. \end{aligned}$$

Así, la forma canónica de A es

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

29.4. Operaciones con Python

Si tenemos la matriz

$$A = \begin{pmatrix} 7 & 1 & -3 & 2 & 2 \\ -6 & 2 & 4 & -2 & -2 \\ 0 & 1 & 3 & 1 & -1 \\ -2 & -1 & 6 & 0 & -3 \\ -4 & 0 & 3 & -1 & -1 \end{pmatrix}$$

para obtener la forma canónica de Jordan podemos ocupar el siguiente código de Python

```
from sympy import Matrix
A = Matrix([[7,1,-3,2,1],
            [-6,2,4,-2,-2],
            [0,1,3,1,-1],
            [-2,-1,6,0,-3],
            [-4,0,3,-1,1]])
J,P=A.jordan_form()
P
```

con el cual se consigue la siguiente matriz

$$\begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2-2\sqrt{3} & 0 \\ 0 & 0 & 0 & 0 & 2-2\sqrt{3} \end{pmatrix}.$$

Marie Camille Jordan (1838-1922) fue un ingeniero y matemático francés que realizó diferentes aportaciones a las matemáticas. Su primer trabajo fue como ingeniero de minas, pero poco a poco se inclinó más por las matemáticas y realizó un doctorado en esta disciplina bajo la dirección de Victor Puiseux y Joseph Alfred Serret. En 1876 se convirtió en profesor de matemáticas de la École Polytechnique. Trabajó en diversos temas como grupos finitos, geometría, topología, variable compleja y álgebra. Fue de los primeros matemáticos que entendieron la importancia de los trabajos de Galois. También fue de los primeros matemáticos en emplear conceptos y herramientas del álgebra para estudiar problemas de geometría. En particular, el concepto de grupo. En este aspecto a Jordan se le atribuye el concepto de homotopía, que es esencial en la topología algebraica. Sus trabajos influyeron a otros matemáticos importantes como Sophus Lie y Felix Klein. En 1912 se retiró de la academia, murió en depresión debido a que tres de sus seis hijos murieron asesinados en la primera guerra Mundial.

29.5. Ejercicios

1.- Obtener la forma canónica de Jordan de la matriz

$$A = \begin{pmatrix} 2 & 4 & -6 & 0 \\ 4 & 1 & -3 & -4 \\ 1 & 0 & 4 & -1 \\ 0 & 4 & -6 & 2 \end{pmatrix}.$$

Capítulo 30

Espacio dual

En este capítulo estudiaremos el Espacio Dual de un espacio vectorial. Nos enfocaremos en los espacios vectoriales de dimensión finita. El material que veremos en este capítulo es importantes para entender temas posteriores.

30.1. Definición

Sea V un espacio vectorial sobre el campo K. Denotaremos el conjunto de las transformaciones lineales de V en K como

$$\boxed{V^* = \{T : V \to K | T \quad \text{lineal}\} \,,}$$

a este conjunto se le llama espacio dual de V. Lo primero que podemos obsevar es que, dado que V^* es el conjunto de todas las funciones lineales de V en K, el espacio V^* es un espacio vectorial sobre K.

30.2. Ejemplos

30.2.1. Caso $\mathbb{C}^n$

Como primer ejemplo, veamos un elemento del espacio dual de $\mathbb{C}^n$. Supongamos que v_0 es un vector fijo que pertenece a $\mathbb{C}^n$ y que N es una matriz de entradas complejas de $n \times n$, entonces la función

$$T : \mathbb{C}^n \to \mathbb{C}$$

definida como

$$T(v) = v_0^{*T} N v \tag{30.1}$$

cumple

$$\begin{aligned}
T(\lambda_1 v_1 + \lambda_2 v_2) &= v_0^{*T} N \left(\lambda_1 v_1 + \lambda_2 v_2\right) \\
&= v_0^{*T} N \lambda_1 v_1 + v_0^{*T} N \lambda_2 v_2 \\
&= \lambda_1 v_0^{*T} N v_1 + \lambda_2 v_0^{*T} N v_2 \\
&= \lambda_1 T(v_1) + \lambda_2 T(v_2).
\end{aligned}$$

Así, la función T definida en (30.1) es lineal y por lo tanto pertenece al espacio dual $(\mathbb{C}^n)^*$.

30.2.2. Caso $(\mathcal{M}_{n\times n}(\mathbb{C}))^*$

Ahora veamos un elemento de $(\mathcal{M}_{n\times n}(\mathbb{C}))^*$. Recoremos que $\mathcal{M}_{n\times n}(\mathbb{C})$ es el espacio vectorial de las matrices de entradas complejas de $n \times n$. Para este caso, supongamos que A_0 es una matriz fija que pertenece a $\mathcal{M}_{n\times n}(\mathbb{C})$, entonces se puede definir la función

$$T : \mathcal{M}_{nn}(\mathbb{C}) \to \mathbb{C},$$

tal que si $A \in \mathcal{M}_{nn}(\mathbb{C})$ se cumple

$$T(A) = Tr\left(A_0 A\right). \tag{30.2}$$

Considerando las propiedades de la traza, se puede ver que si $A, B \in \mathcal{M}_{n\times n}(\mathbb{C})$ y $\lambda, \gamma \in \mathbb{C}$, se cumple

$$\begin{aligned}
T\left(\lambda A + \gamma B\right) &= Tr\left(A_0 \cdot (\lambda A + \gamma B)\right) \\
&= Tr\left(\lambda A_0 \cdot A + \gamma A_0 \cdot B\right) \\
&= \lambda Tr\left(A_0 \cdot A\right) + \gamma Tr\left(A_0 \cdot B\right) \\
&= \lambda T\left(A\right) + \gamma T\left(B\right),
\end{aligned}$$

es decir

$$T\left(\lambda A + \gamma B\right) = \lambda T\left(A\right) + \gamma T\left(B\right).$$

Por lo tanto, la función T definida en (30.2) es lineal, lo cual nos indica que pertenece al espacio dual $(\mathcal{M}_{nn}(\mathbb{C}))^*$.

30.3. Producto escalar y espacio dual

Si un espacio vectorial tiene producto escalar es posible obtener elementos de su espacio dual con la ayuda de dicho producto. En efecto, supongamos que V es un espacio vectorial sobre $\mathbb{C}$ con producto escalar, entonces si v_0 es un vector fijo en V se puede definir la función

$$T_{v_0} : V \to \mathbb{C},$$

tal que

$$\boxed{T_{v_0}(v) = \langle v_0 | v \rangle .} \tag{30.3}$$

Usando las propiedades del producto escalar se puede demostrar que T_{v_0} es una función lineal. Esta última afirmación es vedadera, pues si $v_1, v_2 \in V$ y $\lambda_1, \lambda_2 \in \mathbb{C}$, se cumple

$$\begin{aligned} T_{v_0}(\lambda_1 v_1 + \lambda_2 v_2) &= \langle v_0 | \lambda_1 v_1 + \lambda_2 v_2 \rangle \\ &= \langle v_0 | \lambda_1 v_1 \rangle + \langle v_0 | \lambda_2 v_2 \rangle \\ &= \lambda_1 \langle v_0 | v_1 \rangle + \lambda_2 \langle v_0 | v_2 \rangle \\ &= \lambda_1 T_{v_0}(v_1) + \lambda_2 T_{v_0}(v_2), \end{aligned}$$

de donde

$$T_{v_0}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T_{v_0}(v_1) + \lambda_2 T_{v_0}(v_2).$$

En consecuencia la función T_{v_0} definida en (30.3) es lineal y pertenece al espacio dual V^*.

30.3.1. Ejemplo $\mathbb{R}^3$

Por ejemplo, en el espacio vectorial $\mathbb{R}^3$, con el vector constante fijo

$$v_0 = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

se puede definir la función T_{v_0} tal que si

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3,$$

entonces

$$T_{v_0}(v) = \begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

Esta función es lineal y por lo tanto pertenece al espacio dual $(\mathbb{R}^3)^*$.

30.3.2. Ejemplo $\mathcal{M}_{nn}(\mathbb{C})$

De la misma forma, en el espacio de las matrices de entradas complejas de $n \times n$, $\mathcal{M}_{nn}(\mathbb{C})$, si N es una matriz fija que pertenece a $\mathcal{M}_{nn}(\mathbb{C})$ se puede definir la función T_N tal que para cualquier matriz M en $\mathcal{M}_{nn}(\mathbb{C})$ se cumple

$$T_N(M) = Tr\left(N^{*T} M\right).$$

La cual es lineal y entonces pertenece al espacio dual $(\mathcal{M}_{nn}(\mathbb{C}))^*$.

30.3.3. Ejemplo en el espacio de las funciones

Además, en el espacio vectorial de las funciones integrables en el intervalo $[a, b]$, con la función fija

$$g(x) = \operatorname{sen} x,$$

se puede definir la función T_g tal que para cualquier función intengrable $f(x)$ en el intervalo $[a, b]$ se cumpla

$$T_g(f(x)) = \int_a^b (\sin x)\, f(x).$$

Esta función es lineal y por lo tanto pertenece al espacio dual del espacio vectorial de las funciones integrales en el intervalo $[a, b]$.

30.4. Base del espacio dual V^*

Ahora para el caso de un espacio vectorial V de dimensión finita veremos como construir una base para su espacio dual V^*. Primero recordemos que, dada una base de V, una función lineal sobre V está definida de forma única por el modo en que actúa en los vectores de dicha base.

Suponagmos que V tiene dimesión n y los vectores

$$v_1, v_2, \cdots, v_n \tag{30.4}$$

forman una base de V, entonces si $f \in V^*$ existen n elementos en K

$$\alpha_1, \cdots, \alpha_n$$

tales que

$$f(v_1) = \alpha_1, f(v_2) = \alpha_2, \cdots, f(v_n) = \alpha_n, \tag{30.5}$$

estas últimas igualdades definen completamente a f. En efecto, si $v \in V$, existen n números

$$\lambda_1, \lambda_2, \cdots, \lambda_n \in K$$

tales que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n, \tag{30.6}$$

por lo que, debido a que f es lineal, se cumple

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) \\ &= \lambda_1 f(v_1) + \lambda_2 f(v_2) + \cdots + \lambda_n f(v_n) \\ &= \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_n \alpha_n, \end{aligned}$$

es decir,

$$f(v) = \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_n \alpha_n.$$

Este resultado nos ayudará a encontrar una base para el espacio dual V^*.

De nuevo supongamos que los vectores (30.4) forman una base de V, entonces definiremos las funciones lineales

$$\{f_1, f_2, \cdots, f_n\}, \tag{30.7}$$

tales que

$$\boxed{f_i(v_j) = \delta_{ij}, \qquad i, j = 1, 2, \cdots, n.}$$

Probaremos que las funciones (30.7) forman una base del espacio dual V^*.

Antes realizar la prueba, notemos que debido a que los vectores (30.4) forman una base de V, cualquier vector $v \in V$ se puede escribir de la forma (30.6), por lo que se cumplen

$$\begin{aligned} f_i(v) &= f_i(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) \\ &= \lambda_1 f_i(v_1) + \lambda_2 f_i(v_2) + \cdots + \lambda_i f_i(v_i) + \cdots + \lambda_n f_i(v_n) \\ &= \lambda_i, \end{aligned}$$

esto es,

$$f_i(v) = \lambda_i. \tag{30.8}$$

Ahora probaremos que las funciones (30.7) generan al espacio dual V^*, es decir, probaremos que si $f \in V^*$, entonces f es una combinación lineal de las funciones (30.7). Primero, notemos que si $f \in V^*$ y el conjunto de vectores (30.4) es una base de V, entonces podemos definir la siguiente combinación lineal de funciones

$$f' = f(v_1)f_1 + f(v_2)f_2 + \cdots + f(v_n)f_n.$$

Se puede observar que si $v \in V$, se satisface

$$\begin{aligned} f'(v) &= \left(f(v_1)f_1 + f(v_2)f_2 + \cdots + f(v_n)f_n\right)(v) \\ &= f(v_1)f_1(v) + f(v_2)f_2(v) + \cdots + f(v_n)f_n(v), \end{aligned}$$

además, considerando que se cumple la igualdad (30.8) y que f es una función lineal, se encuentra

$$\begin{aligned} f'(v) &= f(v_1)\lambda_1 + f(v_2)\lambda_2 + \cdots f(v_n)\lambda_n \\ &= \lambda_1 f(v_1) + \lambda_2 f(v_2) + \cdots \lambda_n f(v_n) \\ &= f\left(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n\right) \\ &= f(v). \end{aligned}$$

En consecuencia, para cualquier vector $v \in V$, se cumple

$$f(v) = f'(v),$$

lo que prueba que cualquier función $f \in V^*$ se puede escribir como una combinación lineal de las funciones (30.7) de la forma

$$f = f(v_1)f_1 + f(v_2)f_2 + \cdots + f(v_n)f_n.$$

Así, las funciones (30.7) generan al espacio dual V^*.

Para probar que el conjunto de funciones (30.7) es linealmente independiente supongamos que existen los valores

$$\gamma_1, \gamma_2, \cdots, \gamma_n \in K$$

tal que

$$\gamma_1 f_1 + \gamma_2 f_2 + \cdots + \gamma_n f_n = 0.$$

En particular, para v_1 se encuentra

$$\begin{aligned} 0 &= (\gamma_1 f_1 + \gamma_2 f_2 + \cdots + \gamma_n f_n)(v_1) \\ &= \gamma_1 f_1(v_1) + \gamma_2 f_2(v_2) + \cdots + \gamma_n f_n(v_1), \end{aligned}$$

entonces, considerado la igualdad (30.8), se llega a

$$\begin{aligned} 0 &= \gamma_1 f_1(v_1) + \gamma_2 f_2(v_2) + \cdots + \gamma_n f_n(v_1) \\ &= \gamma_1, \end{aligned}$$

por lo que

$$\gamma_1 = 0.$$

Con el mismo procedimiento se puede probar que

$$\gamma_2 = \gamma_3 = \cdots = \gamma_n = 0,$$

por lo tanto, el conjunto de funciones (30.7) es lineamente independiente.

En consecuencia, el conjunto de funciones (30.7) es una base del espacio dual V^*.

Se puede observar que este resultado también prueba que si la dimensión de V es n, entonces la dimensión del espacio dual V^* también es n, es decir

$$\boxed{\dim(V^*) = n.}$$

30.5. Cambio de base en el espacio dual

Ya demostramos que si V tiene dimensión n y que el conjunto de vectores

$$\{v_1, v_2, \cdots, v_n\} \tag{30.9}$$

es base de V. Entonces, si las funciones

$$f_1, f_2, \cdots, f_n \tag{30.10}$$

satisfacen

$$f_i(v_j) = \delta_{ij}, \qquad i, j = 1, 2, \cdots, n$$

forman una base del espacio dual V^*. De la misma forma si el conjunto de vectores

$$\{v'_1, v'_2, \cdots, v'_n\} \tag{30.11}$$

es base de V, entonces, si las funciones

$$f'_1, f'_2, \cdots, f'_n \tag{30.12}$$

satisfacen

$$f'_i\left(v'_j\right) = \delta_{ij},$$

tenemos una base del espacio dual V^*.

Debido a que los vectores (30.9) y (30.11) son bases de V, existe una matriz $U \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que

$$v'_i = U_{ij} v_j.$$

De la misma forma, debido a que las funciones (30.10) y (30.12) son bases del espacio dual V^*, existe una matriz $U' \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que

$$f'_i = U'_{ij} f_j.$$

Se puede probar que se cumple

$$U' = \left(U^{-1}\right)^T.$$

En efecto, podemos ver que

$$\begin{aligned}
\delta_{ij} &= f_i'(v_j') \\
&= f_i'(U_{jk}v_k) \\
&= U_{jk}f_i'(v_k) \\
&= U_{jk}U_{il}'f_l(v_k) \\
&= U_{jk}U_{il}'\delta_{lk} \\
&= U_{jk}U_{ik}' \\
&= U_{jk}U_{ki}'^{T} \\
&= \left(UU'^{T}\right)_{ji}
\end{aligned}$$

esto es,

$$\delta_{ij} = \left(UU'^{T}\right)_{ji}$$

de donde

$$UU'^{T} = I,$$

es decir,

$$U = \left(U'^{-1}\right)^{T},$$

que implica

$$U' = \left(U^{-1}\right)^{T}.$$

Por lo tanto, bajo un cambio de base se tiene

$$\boxed{\begin{aligned} &v_i' = U_{ij}v_j, \\ &f_i' = \left(\left(U^{-1}\right)^{T}\right)_{ij} f_j = \left(U^{-1}\right)_{ji} f_j. \end{aligned}}$$

30.6. Ejemplos

30.6.1. Ejemplo en $\mathbb{R}^3$

Para $\mathbb{R}^3$ tenemos la base

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

entonces la base del espacio dual de $\mathbb{R}^3$ son las funciones lineales

$$f_1, \quad f_2, \quad f_3,$$

que satisfacen

$$\begin{aligned}
f_1(e_1) &= 1, \quad f_1(e_2) = 0 \quad f_1(e_3) = 0, \\
f_2(e_1) &= 0, \quad f_2(e_2) = 1, \quad f_2(e_3) = 0, \\
f_3(e_1) &= 0, \quad f_3(e_2) = 0, \quad f_3(e_3) = 1.
\end{aligned}$$

Ahora, consideremos un vector arbitrario de $\mathbb{R}^3$

$$v = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

entonces, debido a que f_1 es una función lineal, se cumple

$$\begin{aligned}
f_1 \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= f_1 \left(a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\
&= a f_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b f_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c f_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
&= a,
\end{aligned}$$

es decir,

$$f_1 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = a.$$

De la misma forma, se puede comprobar que se cumple

$$f_2 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = b, \qquad f_3 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = c.$$

Además, debido a que f_1, f_2, f_3 es una base de $(\mathbb{R}^3)^*$, entonces si $f \in (\mathbb{R}^3)^*$ existen $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tal que

$$f = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$$

de donde

$$\begin{aligned} f\begin{pmatrix} a \\ b \\ c \end{pmatrix} &= (\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3)\begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &= \lambda_1 f_1 \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \lambda_2 f_2 \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \lambda_3 f_3 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &= \lambda_1 a + \lambda_2 b + \lambda_3 c. \end{aligned}$$

Así, todas las funciones del espacio dual $(\mathbb{R}^3)^*$, son de la forma

$$f\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \lambda_1 a + \lambda_2 b + \lambda_3 c.$$

30.6.2. Ejemplo en el espacio de las matrices $\mathcal{M}_{2\times 2}(\mathbb{R})$

Para $\mathcal{M}_{2\times 2}(\mathbb{R})$ tenemos la base

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

entonces, una base del espacio dual de $\mathcal{M}_{2\times 2}(\mathbb{R})$ son las funciones lineales

$$f_1, \quad f_2, \quad f_3, \quad f_4,$$

que satisfacen

$$\begin{aligned} f_1(e_1) &= 1, \quad f_1(e_2) = 0 \quad f_1(e_3) = 0, \quad f_1(e_4) = 0, \\ f_2(e_1) &= 0, \quad f_2(e_2) = 1, \quad f_2(e_3) = 0, \quad f_2(e_4) = 0, \\ f_3(e_1) &= 0, \quad f_3(e_2) = 0, \quad f_3(e_3) = 1, \quad f_3(e_4) = 0, \\ f_4(e_1) &= 0, \quad f_4(e_2) = 0, \quad f_4(e_3) = 0, \quad f_4(e_4) = 1. \end{aligned}$$

Así, para cualquier matriz de $\mathcal{M}_{2\times 2}(\mathbb{R})$

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

debido a que f_1 es una función lineal, se cumple

$$\begin{aligned} f_1\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= f_1\left(a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ a & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= af_1\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + bf_1\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + cf_1\begin{pmatrix} 0 & 0 \\ a & 0 \end{pmatrix} + df_1\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= a, \end{aligned}$$

es decir

$$f_1\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a.$$

De la misma forma, se puede comprobar que se cumple

$$f_2\begin{pmatrix} a & b \\ c & d \end{pmatrix} = b, \quad f_3\begin{pmatrix} a & b \\ c & d \end{pmatrix} = c, \quad f_4\begin{pmatrix} a & b \\ c & d \end{pmatrix} = d.$$

Además, debido a que f_1, f_2, f_3, f_4 forman una base de $(\mathcal{M}_{2\times 2}(\mathbb{R}))^*$, entonces si $f \in (\mathcal{M}_{2\times 2}(\mathbb{R}))^*$, existen $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$ tal que

$$f = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 + \lambda_4 f_4$$

de donde

$$\begin{aligned} f\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= (\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 + \lambda_4 f_4)\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \lambda_1 f_1\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \lambda_2 f_2\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \lambda_3 f_3\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \lambda_4 f_4\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d. \end{aligned}$$

Así, toda las funciones del espacio dual $(\mathcal{M}_{2\times 2}(\mathbb{R}))^*$ son de la forma

$$f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d.$$

30.7. Ejercicios

1.- Para el espacio dual $(\mathcal{M}_{3\times 3}(\mathbb{R}))^*$

a) Obtener una base.

b) Obtener la forma general de una función de este espacio dual.

Capítulo 31

Funciones bilineales

En este capítulo estudiaremos el concepto de función bilineal, el cual extiende el concepto de función lineal y tiene una amplia gama de aplicaciones.

Los temas de este capítulo son importantes para entender conceptos y aplicaciones avanzadas del álgebra lineal.

31.1. Definición

Sean V, U y W espacios vectoriales en el campo K, se dice que la función

$$T : V \times U \to W$$

es bilineal si cumple las condiciones

$$\boxed{\begin{array}{ll} \forall v_1, v_2 \in V, u \in U, & T(v_1 + v_2, u) = T(v_1, u) + T(v_2, u), \\ \forall v \in V, u_1, u_2 \in U, & T(v, u_1 + u_2) = T(v, u_1) + T(v, u_2), \\ \forall v \in V, u \in U, \lambda \in K, & T(\lambda v, u) = \lambda T(v, u), \\ \forall v \in V, u \in U, \lambda \in K, & T(v, \lambda u) = \lambda T(v, u). \end{array}}$$

Note que esta definición nos dice que la función T es bilineal si es lineal en cada una de sus entradas.

De hecho se puede probar que una función es bilineall si y solo si cumplen las dos condiciones

$$\boxed{\begin{array}{ll} \forall\lambda_1,\lambda_2\in K,\forall v_1,v_2\in V,u\in U, & T\left(\lambda_1 v_1+\lambda_2 v_2,u\right)=\lambda_1 T(v_1,u)+\lambda_2 T(v_2,u), \\ \forall\lambda_1,\lambda_2\in K,\forall v\in V,u_1,u_2\in U, & T(v,\lambda_1 u_1+\lambda_2 u_2)=\lambda_1 T(v,u_1)+\lambda_2 T(v,u_2). \end{array}}$$

Ver ejercicio 1.

31.2. Ejemplos

31.2.1. Ejemplo en $\mathbb{C}\times\mathbb{C}$

Para construir el primer ejemplo, supongamos λ es una constante compleja, entonces definamos la función

$$T:\mathbb{C}\times\mathbb{C}\to\mathbb{C}$$

tal que si $z,w\in\mathbb{C}$ se cumple

$$T(z,w)=\lambda zw. \tag{31.1}$$

Esta función es bilineal, pues si $\gamma_1,\gamma_2\in\mathbb{C}$, para cualquier $z_1,z_2\in\mathbb{C}$ se tiene

$$\begin{aligned} T\left(\gamma_1 z_1+\gamma_2 z_2,w\right) &= \lambda\left(\gamma_1 z_1+\gamma_2 z_2\right)w \\ &= \lambda\gamma_1 z_1 w+\lambda\gamma_2 z_2 w \\ &= \gamma_1\left(\lambda z_1 w\right)+\gamma_2\left(\lambda z_2 w\right) \\ &= \gamma_1 T(z_1,w)+\gamma_2 T(z_2,w). \end{aligned}$$

Adicionalmente, para cualquier $w_1,w_2\in\mathbb{C}$ se cumple

$$\begin{aligned} T\left(z,\gamma_1 w_1+\gamma_2 w_2\right) &= \lambda z\left(\gamma_1 w_1+\gamma_2 w_2\right) \\ &= \lambda\gamma_1 z w_1+\lambda\gamma_2 z w_2 \\ &= \gamma_1 z\left(\lambda w_1\right)+\gamma_2\left(\lambda z w_2\right) \\ &= \gamma_1 T(z,w_1)+\gamma_2 T(z,w_2). \end{aligned}$$

Por lo tanto, la función definida en (31.1) es una función bilineal.

31.2.2. Ejemplo en $\mathbb{R}^m \times \mathbb{R}^n$

Ahora veamos un ejemplo en $\mathbb{R}^m \times \mathbb{R}^n$, supongamos M es una matriz de entradas reales de $m \times n$,

$$M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix},$$

entonces definamos la función

$$T : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$$

tal que si

$$u = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{R}^m, \qquad v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$$

entonces

$$\begin{aligned} T(u,v) &= u^T M v \\ &= \begin{pmatrix} a_1 & a_2 & \cdots & a_m \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \end{aligned} \tag{31.2}$$

Esta función es bilineal, pues si $\gamma_1, \gamma_2 \in \mathbb{R}$, para cualquier $u, u' \in \mathbb{R}^m$ se tiene

$$\begin{aligned} T(\gamma_1 u + \gamma_2 u', v) &= (\gamma_1 u + \gamma_2 u')^T M v \\ &= \left(\gamma_1 u^T + \gamma_2 u'^T\right) M v \\ &= \gamma_1 u^T M v + \gamma_2 u'^T M v \\ &= \gamma_1 T(u,v) + \gamma_2 T(u',v). \end{aligned}$$

Adicionalmente, para cualquier $v, v', \in \mathbb{R}^n$ se cumple

$$\begin{aligned} T(u, \gamma_1 v + \gamma_2 v') &= u^T M (\gamma_1 v + \gamma_2 v') \\ &= u^T M \gamma_1 v + u^T M \gamma_2 v' \\ &= \gamma_1 u^T M v + \gamma_2 u^T M v' \\ &= \gamma_1 T(u,v) + \gamma_2 T(u,v'). \end{aligned}$$

Por lo tanto, la función definida en (31.2) es una función bilineal.

31.2.3. Ejemplo con matrices

Como otro ejemplo, consideremos el espacio de las matrices reales de $n \times n$, $\mathcal{M}_{n\times n}(\mathbb{R})$. En este espacio supogamos que N es una matriz fija de $n \times n$. Entonces, podemos definir función

$$T : \mathcal{M}_{n\times n}(\mathbb{R}) \times \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathcal{M}_{n\times n}(\mathbb{R})$$

tal que si $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ se cumple

$$T(A, B) = ANB. \tag{31.3}$$

Note que, si $\gamma_1, \gamma_2 \in \mathbb{R}$ y $A_1, A_2 \in \mathcal{M}_{n\times n}(\mathbb{R})$, se encuentra

$$\begin{aligned} T(\gamma_1 A_1 + \gamma_2 A_2, B) &= (\gamma_1 A_1 + \gamma_2 A_2) NB \\ &= \gamma_1 A_1 NB + \gamma_2 A_2 NB \\ &= \gamma_1 T(A_1 B) + \gamma_2 T(A_2 B). \end{aligned}$$

Además, si $B_1, B_2 \in \mathcal{M}_{n\times n}(\mathbb{R})$, se obtiene

$$\begin{aligned} T(A, \gamma_1 B_1 + \gamma_2 B_2) &= AN(\gamma_1 B_1 + \gamma_2 B_2) \\ &= AN\gamma_1 B_1 + AN\gamma_2 B_2 \\ &= \gamma_1 (ANB_1) + \gamma_2 (ANB_2) \\ &= \gamma_1 T(A, B_1) + \gamma_2 T(A, B_2). \end{aligned}$$

Así, la función definida en (31.3) es una función bilineal.

31.2.4. Ejemplo con funciones

Veamos un ejemplo en el espacio de las funciones de una variable real derivables, denotaremos a este espacio como A. Supongamos que λ es un constante real, entonces definamos la función

$$T : A \times A \to A$$

con la regla

$$T(f, g)(x) = \lambda f(x) \frac{\partial}{\partial x} g(x), \qquad f, g \in A. \tag{31.4}$$

Supongamos que $\gamma_1, \gamma_2 \in \mathbb{R}$ y que $f_1, f_2, g \in A$, entonces se cumple

$$
\begin{aligned}
T\left(\gamma_1 f_1 + \gamma_2 f_2, g\right) &= \lambda\left(\gamma_1 f_1(x) + \gamma_2 f_2(x)\right) \frac{\partial}{\partial x} g(x) \\
&= \lambda \gamma_1 f_1(x) \frac{\partial}{\partial x} g(x) + \lambda \gamma_2 f_2(x) \frac{\partial}{\partial x} g(x) \\
&= \gamma_1 \left(\lambda f_1(x) \frac{\partial}{\partial x} g(x)\right) + \gamma_2 \left(\lambda f_2(x) \frac{\partial}{\partial x} g(x)\right) \\
&= \gamma_1 T(f_1, g) + \gamma_2 T(f_2, g).
\end{aligned}
$$

Además, si $g_1, g_2, f \in A$, entonces se cumple

$$
\begin{aligned}
T\left(f, \gamma_1 g_1 + \gamma_2 g_2\right) &= \lambda f(x) \frac{\partial}{\partial x}\left(\gamma_1 g_1(x) + \gamma_2 g_2(x)\right) \\
&= \lambda f(x) \frac{\partial}{\partial x} \gamma_1 g_1(x) + \lambda f(x) \frac{\partial}{\partial x} \gamma_1 g_2(x) \\
&= \lambda \gamma_1 f(x) \frac{\partial}{\partial x} g_1(x) + \lambda \gamma_2 f(x) \frac{\partial}{\partial x} g_2(x) \\
&= \gamma_1 \left(\lambda f(x) \frac{\partial}{\partial x} g_1(x)\right) + \gamma_2 \left(\lambda f(x) \frac{\partial}{\partial x} g_1(x)\right) \\
&= \gamma_1 T(f, g_1) + \gamma_2 T(f, g_2).
\end{aligned}
$$

Por lo tanto la función definida en (31.4) es una función bilineal.

31.3. Funciones bilineales actuando en una base

Recordemos que si V_1 y V_2 son espacios vectoriales sobre K y la función

$$F : V_1 \to V_2$$

es lineal, entonce si F está definida en una base del dominio V_1, automáticamente está definida en todo el dominio V_1. También recordemos que si dos funciones lineales coinciden en una base del dominio, esa funciones son las mismas. Probaremos que tenemos un resultado similar con las funciones bilineales.

Supongamos que U, V, W son espacios vectoriales de dimensión finita sobre K con

$$
\begin{aligned}
\dim(U) &= n, \\
\dim(V) &= m.
\end{aligned}
$$

También supongamos que la función

$$T : U \times V \to W$$

es bilineal. Probaremos que si T está definida es una base de U y en una base de V, automáticamente está definida en todo el dominio $U \times V$.

Para iniciar la prueba notemos que debido a que T es una función sobre $U \times V$, si

$$u_1, u_2, \cdots, u_m$$

es una base de U y

$$v_1, v_2, \cdots, v_n$$

es una base de V, entonces existen nm vectores en W tales que

$$T(u_i, v_j) = w_{ij}, \quad i = 1, 2, \cdots, m, j = 1, 2, \cdots, n, \quad w_{ij} \in W. \tag{31.5}$$

Ahora supongamos que $u \in U$ y $v \in V$, entonces existen m elementos en K

$$\lambda_1, \lambda_2, \cdots, \lambda_m,$$

tales que

$$u = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_m u_m,$$

y existen n elementos en K

$$\gamma_1, \gamma_2, \cdots, \gamma_n$$

de tal forma que

$$v = \gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_n v_n.$$

De donde,

$$T(u, v) = T(\lambda_i u_i, \gamma_j v_j) = \lambda_i \gamma_j T(u_i, v_j) = \lambda_i \gamma_j w_{ij},$$

esto es

$$T(u, v) = \lambda_i \gamma_j w_{ij}, \tag{31.6}$$

Por lo tanto, si una función bilineal se define sobre una base U y una base V, automáticamente está definida sobre todo el dominio $U \times V$.

Ahora, supongamos que

$$G : U \times V \to W$$

es otra función bilineal que satisface (31.5), es decir cumple

$$G(u_i, v_j) = w_{ij}.$$

Por lo cual, considerando la igualdad (31.6) y que G es una función bilineal, se encuentra

$$\begin{aligned} G(u, v) &= G(\lambda_i u_i, \gamma_j v_j) \\ &= \lambda_i \gamma_j G(u_i, v_j) \\ &= \lambda_i \gamma_j w_{ij} \\ &= T(u, v), \end{aligned}$$

de donde

$$G(u, v) = T(u, v).$$

Como este resultado se cumple para cualquier $u \in U$ y cualquier $v \in V$ entonces

$$G = T.$$

Por lo tanto, si dos funciones bilineales coinciden en una base de U y una base de V, coinciden en todo el dominio $U \times V$, es decir son iguales

31.3.1. Ejemplo

Por ejemplo, sea $V = \mathbb{R}^4$ con la base

$$(1, 0, 0, 0), \quad (0, 1, 0, 0), \quad (0, 0, 1, 0) \quad (0, 0, 0, 1)$$

y U el espacio vectorial de las matrices de entradas reales de 2×2 antisimétricas con la base

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

Entonces definiremos la función bilineal

$$T : V \times U \to \mathcal{M}_{2\times 2}(\mathbb{R})$$

tal que

$$\begin{aligned}
T\left((1,0,0,0), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\
T\left((0,1,0,0), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\
T\left((0,0,1,0), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\
T\left((0,0,0,1), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}$$

Por lo cual, la forma general de esta función bilineal es

$$\begin{aligned}
& T\left((a,b,c,d), \begin{pmatrix} 0 & -e \\ e & 0 \end{pmatrix}\right) \\
& = T\left(a(1,0,0,0) + b(0,1,0,0) + c(0,0,1,0) + d(0,0,0,1), e\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) \\
& = aeT\left((1,0,0,0), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) + beT\left((0,1,0,0), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) + \\
& + ceT\left((0,0,1,0), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) + deT\left((0,0,0,1), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) \\
& = ae\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + be\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + ce\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + de\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\
= & \begin{pmatrix} ae & be \\ ce & de \end{pmatrix},
\end{aligned}$$

entonces la bilineal función bilineal es

$$T\left((a,b,c,d), \begin{pmatrix} 0 & -e \\ e & 0 \end{pmatrix}\right) = \begin{pmatrix} ae & be \\ ce & de \end{pmatrix}.$$

31.4. Ejercicios

1.- Sean V, U y W espacios vectoriales en el campo K, y

$$T : V \times U \to W$$

T una función. Probar que T es una función bilineal si y solo si satisface las condiciones

$$\forall \lambda_1, \lambda_2 \in K, \forall v_1, v_2 \in V, u \in U, \quad T(\lambda_1 v_1 + \lambda_2 v_2, u) = \lambda_1 T(v_1, u) + \lambda_2 T(v_2, u),$$
$$\forall \lambda_1, \lambda_2 \in K, \forall v \in V, u_1, u_2 \in U, \quad T(v, \lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 T(v, u_1) + \lambda_2 T(v, u_2).$$

2.- Sea $V = \mathbb{R}^2$ con la base

$$(1,0), \quad (0,1)$$

y U el espacio vectorial de las matrices de entradas reales de 2×2 simétricas con la base

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Encontrar la función bilineal

$$T : V \times U \to \mathbb{R}^6$$

tal que

$$\begin{aligned}
T\left((1,0), \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) &= (1,0,0,0,0), \\
T\left((0,1), \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) &= (0,1,0,0,0,0), \\
T\left((1,0), \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) &= (0,0,0,1,0), \\
T\left((0,1), \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) &= (0,0,0,0,0,1), \\
T\left((1,0), \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) &= (0,0,1,0,0,0), \\
T\left((0,1), \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) &= (0,0,0,1,0,0).
\end{aligned}$$

Capítulo 32

Formas bilineales

En este capítulo veremos el concepto de forma bilineal. En particular, veremos que este concepto está relacionado con el concepto de espacio dual, el cual se define en el espacio de las funciones lineales.

32.1. Definición

Sean V y U espacios vectoriales sobre el campo K, si la función

$$T : V \times U \to K$$

es una función bilineal, se dice que T que es una forma bilineal. El conjunto de funciones bilineales de V y U en K se denota como

$$L^2\left(V, U, K\right).$$

Si $U = V$ se denota

$$L^2\left(V, V, K\right) = Bil\left(V\right).$$

32.1.1. Ejemplo en $\mathbb{C}^n \times \mathbb{C}^m$

Por ejemplo, si M_{nm} es una matriz de coeficientes complejos, la función

$$T : \mathbb{C}^n \times \mathbb{C}^m \to \mathbb{C}$$

definida como

$$T(v_1, v_2) = v_1^T M v_2$$

es una función bilineal y por lo tanto una forma bilineal

32.1.2. Ejemplo con matrices

Como otro ejemplo, consideremos el espacio de las matrices reales de $n \times n$, $\mathcal{M}_{n\times n}(\mathbb{R})$. En este espacio supogamos que N es una matriz fija de $n \times n$. Entonces, podemos definir función

$$T : \mathcal{M}_{n\times n}(\mathbb{R}) \times \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$$

tal que si $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$ se cumple

$$T(A,B) = Tr\left(ANB\right). \tag{32.1}$$

Note que, si $\gamma_1, \gamma_2 \in \mathbb{R}$ y $A_1, A_2 \in \mathcal{M}_{n\times n}(\mathbb{R})$, se encuentra

$$\begin{aligned} T\left(\gamma_1 A_1 + \gamma_2 A_2, B\right) &= Tr\left(\left(\gamma_1 A_1 + \gamma_2 A_2\right) NB\right) \\ &= Tr\left(\gamma_1 A_1 NB + \gamma_2 A_2 NB\right) \\ &= \gamma_1 Tr\left(A_1 B\right) + \gamma_2 Tr\left(A_2 B\right). \\ &= \gamma_1 T\left(A_1 B\right) + \gamma_2 T\left(A_2 B\right). \end{aligned}$$

Además, si $B_1, B_2 \in \mathcal{M}_{n\times n}(\mathbb{R})$, se obtiene

$$\begin{aligned} T\left(A, \gamma_1 B_1 + \gamma_2 B_2\right) &= Tr\left(AN\left(\gamma_1 B_1 + \gamma_2 B_2\right)\right) \\ &= Tr\left(AN\gamma_1 B_1 + AN\gamma_2 B_2\right) \\ &= \gamma_1 Tr\left(ANB_1\right) + \gamma_2 Tr\left(ANB_2\right) \\ &= \gamma_1 T\left(A, B_1\right) + \gamma_2 T\left(A, B_2\right). \end{aligned}$$

Así, la función definida en (32.1) es una función bilineal y por lo tanto una forma bilineal.

32.1.3. Ejemplo con funciones

Veamos un ejemplo en el espacio de las funciones de una variable real derivables e integrables en el intervalo $[a, b]$. Denotaremos a este espacio como A. Supongamos que λ es un constante real, entonces definamos la función

$$T : A \times A \to \mathbb{R}$$

con la regla

$$T(f,g) = \int_a^b \lambda f(x) \frac{\partial}{\partial x} g(x) dx, \qquad f, g \in A. \tag{32.2}$$

Supongamos que $\gamma_1, \gamma_2 \in \mathbb{R}$ y que $f_1, f_2, g \in A$, entonces se cumple

$$
\begin{aligned}
T\left(\gamma_1 f_1 + \gamma_2 f_2, g\right) = \int_a^b \lambda\left(\gamma_1 f_1(x) + \gamma_2 f_2(x)\right) dx \frac{\partial}{\partial x} g(x) & \\
= \int_a^b \left(\lambda \gamma_1 f_1(x) \frac{\partial}{\partial x} g(x) + \lambda \gamma_2 f_2(x) \frac{\partial}{\partial x} g(x)\right) dx & \\
= \gamma_1 \int_a^b \left(\lambda f_1(x) \frac{\partial}{\partial x} g(x)\right) dx + \gamma_2 \int_a^b \left(\lambda f_2(x) \frac{\partial}{\partial x} g(x)\right) dx & \\
= \gamma_1 T(f_1, g) + \gamma_2 T(f_2, g). &
\end{aligned}
$$

Además, si $g_1, g_2, f \in A$, entonces se cumple

$$
\begin{aligned}
T\left(f, \gamma_1 g_1 + \gamma_2 g_2\right) = \int_a^b \lambda f(x) \frac{\partial}{\partial x}\left(\gamma_1 g_1(x) + \gamma_2 g_2(x)\right) dx & \\
= \int_a^b \left(\lambda f(x) \frac{\partial}{\partial x} \gamma_1 g_1(x) + \lambda f(x) \frac{\partial}{\partial x} \gamma_1 g_2(x)\right) dx & \\
= \gamma_1 \int_a^b \left(\lambda f(x) \frac{\partial}{\partial x} g_1(x)\right) dx + \gamma_2 \int_a^b \left(\lambda f(x) \frac{\partial}{\partial x} g_1(x)\right) dx & \\
= \gamma_1 T(f, g_1) + \gamma_2 T(f, g_2). &
\end{aligned}
$$

Entonces la función definida en (32.2) es una función bilineal y por lo tanto una forma bilineal.

32.2. Matriz asociada

Cuando los espacios vectoriales tienen dimensión finita de forma natural las formas bilineales tienen asociada una matriz.

En efecto, supongamos que U, V son espacios vectoriales sobre el campo K, con

$$
\begin{aligned}
\dim(U) &= m, \\
\dim(V) &= n.
\end{aligned}
$$

También supongamos que

$$u_1, u_2, \cdots, u_m$$

es una base de U y

$$v_1, v_2, \cdots, v_n$$

es una base de V. Entonces si

$$T : U \times V \to K$$

es una función bilineal, existen nm vectores en K tales que

$$\boxed{T(u_i, v_j) = M_{ij}, \quad i = 1, 2, \cdots, m, j = 1, 2, \cdots, n, \quad M_{ij} \in K.} \tag{32.3}$$

Así, podemos definir la matriz

$$\boxed{M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix}.}$$

Esta matriz define completamente a la forma bilineal. De hecho, si $u \in U$ y $v \in V$, entonces existen m elementos en K

$$\lambda_1, \lambda_2, \cdots, \lambda_m,$$

tal que

$$u = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_m u_m$$

y existen n elementos en K

$$\gamma_1, \gamma_2, \cdots, \gamma_n,$$

tal que

$$v = \gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_n v_n.$$

Por lo tanto,

$$\begin{aligned} T(u, v) &= T(\lambda_i u_i, \gamma_j v_j) \\ &= \lambda_i \gamma_j T(u_i, v_j) \\ &= \lambda_i \gamma_j M_{ij} \\ &= \lambda_i M_{ij} \gamma_j \\ &= \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}. \end{aligned}$$

esto es

$$T(u,v) = \begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_m\end{pmatrix} \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix} \begin{pmatrix}\gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n\end{pmatrix}.$$

32.3. Ejemplos

32.3.1. Ejemplo 1

Por ejemplo, sea $K = \mathbb{R}$ y $U = \mathbb{R}^2, V = \mathbb{R}^3$, con las bases

$$\left\{ u_1 = \begin{pmatrix}1\\0\end{pmatrix}, u_2 = \begin{pmatrix}0\\1\end{pmatrix} \right\}, \left\{ v_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, v_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, v_3 = \begin{pmatrix}0\\0\\1\end{pmatrix} \right\}.$$

Entonces podemos definir la forma bilineal tal que

$$\begin{aligned}
F(u_1, v_1) &= F\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}\right) = a, \\
F(u_1, v_2) &= F\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right) = b, \\
F(u_1, v_2) &= F\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right) = c, \\
F(u_2, v_1) &= F\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}\right) = d, \\
F(u_2, v_2) &= F\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right) = e, \\
F(u_2, v_3) &= F\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right) = f.
\end{aligned}$$

De donde

$$M = \begin{pmatrix} F(u_1,v_1) & F(u_1,v_2) & F(u_1,v_3) \\ F(u_2,v_1) & F(u_2,v_2) & F(u_2,v_3) \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

De donde, para cualquier dos vectores de $\mathbb{R}^2$ y $\mathbb{R}^3$ se tiene

$$F\left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

32.3.2. Ejemplo 2

Supongamos que U un espacio vectorial de dimensión m, V un espacio vectorial de dimensión n. Entonces, si M es una matriz que pertenece a $\mathcal{M}_{m\times m}(K)$,, dada una base de U, $\{u_i\}_{i=1}^m$, y una base de V, $\{v_j\}_{j=1}^n$, se puede definir la forma bilineal F tal que

$$F(u_i, v_j) = M_{ij}.$$

Por ejemplo, con la matriz

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & q \end{pmatrix} \in \mathcal{M}_{3\times 3}(\mathbb{R}),$$

se puede definir la función bilineal

$$F : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$$

tal que

$$F\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & q \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

32.4. Espacio dual y formas bilineales

Ahora veremos que todas la formas bilineales se pueden expresar en términos de espacios duales.

Para iniciar, supongamos que U y V son espacios vectoriales sobre K y U^*, V^* son sus espacios duales. Entonces, si $f \in U^*$ y $g \in V^*$ podemos definir la función

$$T : U \times V \to K$$

tal que si $u \in U$ y $v \in V$ se cumple

$$\boxed{T(u,v) = f(u)g(v)} \tag{32.4}$$

Probaremos que T es una función bilineal y por lo tanto una forma bilineal. Para realizar la prueba supongamos que $\lambda, \gamma \in K, u_1, u_2 \in U$ y $v \in V$, por lo que

$$\begin{aligned} T\left(\lambda u_1 + \gamma u_2, v\right) &= f\left(\lambda u_1 + \gamma u_2\right) g(v) \\ &= \left(\lambda f(u_1) + \gamma f(u_2)\right) g(v) \\ &= \lambda f(u_1)g(v) + \gamma f(u_2)g(v) \\ &= \lambda T(u_1, v) + \gamma T(u_2, v). \end{aligned}$$

Además, si $u \in U$ y $v_1, v_2 \in V$ se encuentra

$$\begin{aligned} T\left(u, \lambda v_1 + \gamma v_2\right) &= f\left(u\right) g\left(\lambda v_1 + \gamma v_2\right) \\ &= f(u)\left(\lambda g(v_1) + \gamma g(v_2)\right) \\ &= \lambda f(u)g(v_1) + \gamma f(u)g(v_2) \\ &= \lambda T(u, v_1) + \gamma T(u, v_2). \end{aligned}$$

Así, la función (32.4) es una forma bilineal.

32.5. Base de las Formas Bilineales

Con las bases del espacio dual de U y V se puede obtener una base de las formas bilineales de $U \times V$.

Para probar esta afirmación recordemos que si U es un espacio vectorial de dimensión n sobre K y

$$\{u_1, u_2, \cdots u_m\}$$

es una de sus bases, entonces el conjunto de funciones

$$\{f_1, f_2, \cdots, f_m\} \tag{32.5}$$

que satisfacen

$$f_i(u_j) = \delta_{ij}, \qquad i, j = 1, 2, \cdots, m$$

son una base del espacio dual U^*. De la misma forma, si V es un espacio vectorial sobre K de dimensión n y

$$\{v_1, v_2, \cdots v_n$$

es una de sus bases, entonces el conjunto de funciones

$$\{g_1, g_2, \cdots, g_n\} \tag{32.6}$$

que satisfacen

$$g_k(v_l) = \delta_{kl}, \qquad k, l = 1, 2, \cdots, n$$

son una base del espacio dual V^*.

Con las funciones (32.5) y (32.6), se pueden construir las funciones de $U \times V$ en K,

$$f_1g_1, f_1g_2, \cdots, f_1g_n, f_2g_1, f_2g_2, \cdots, f_2g_n, \cdots, f_mg_1, f_mg_2, \cdots, f_mg_n \tag{32.7}$$

las cuales son formas bilineales. Las funciones (32.7) son una base del espacio de las formas bilineales de $U \times V$ en K,

$$Bil(U, V, K) = \{T : U \times V \to K | T \quad \text{Bilineal}\}. \tag{32.8}$$

Primero probaremos que las formas bilineales (32.7) general al espacio (32.8). Así, supongamos que la función

$$T : U \times V \to K$$

es una forma bilineal, entonces existen nm elementos de K tal que

$$T(u_i, v_k) = \alpha_{ik},$$

además, podemos definir la función bilineal

$$T' = \alpha_{ik} f_i g_k.$$

Ahora, si $u \in U$, entonces existen m elementos de K

$$\lambda_1, \lambda_2, \cdots, \lambda_m$$

tales que

$$u = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_m u_m,$$

de forma análoga si $v \in V$, entonces existen n elementos de K

$$\gamma_1, \gamma_2, \cdots, \gamma_n$$

de tal forma que

$$v = \gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_n v_n.$$

Entonces,

$$\begin{aligned} T(u,v) &= T(\lambda_i u_i, \gamma_k v_k) \\ &= \lambda_i \gamma_k T(u_i, v_k) \\ &= \lambda_i \gamma_k \alpha_{ik}, \end{aligned}$$

además,

$$\begin{aligned} T'(u,v) &= \alpha_{jl} f_j(u) g_l(v) \\ &= \alpha_{jl} f_j(\lambda_i u_i) g_l(\gamma_k v_k) \\ &= \alpha_{jl} \lambda_i f_j(u_i) \gamma_k g_l(v_k) \\ &= \lambda_i \gamma_k \alpha_{jl} f_j(u_i) g_l(v_k) \\ &= \lambda_i \gamma_k \alpha_{jl} \delta_{ji} \delta_{lk} \\ &= \lambda_i \gamma_k \alpha_{ik}. \end{aligned}$$

Por lo tanto, para cualquier $(u,v) \in U \times V$ se cumple

$$T(u,v) = T'(u,v),$$

en consecuencia cualquier forma bilineal se puede escribir de la forma

$$T = \alpha_{ik} f_i g_k$$

que es una combinación lineal de las formas bilineales (32.7) Por lo tanto, el espacio de las formas bilineales (32.8) es generado por las funciones (32.7).

Ahora probaremos que el conjunto de las formas bilineales (32.7) son linealmente independientes. Para realizar esta prueba, supongamos que existen nm elementos $\beta_{ik} \in K$ tales que

$$\beta_{ik} f_i g_k = 0,$$

entonces

$$\begin{aligned} 0 &= (\beta_{ik} f_i g_k)(u_1, v_1) \\ &= \beta_{ik} f_i(u_1) g_k(v_1) \\ &= \beta_{ik} \delta_{i1} \delta_{k1} \\ &= \beta_{11}, \end{aligned}$$

es decir,

$$\beta_{11} = 0.$$

Con el mismo procedimiento, se puede mostrar que para cualquiera i, k se obtiene

$$\beta_{ik} = 0.$$

Por lo tanto, el conjunto de funciones (32.7) es una base del espacio de las formas bilineales

Note que si U tiene dimensión m y V tiene dimensión n, entonces el espacio de las formas bilineales tiene dimensión nm, es decir

$$\boxed{dim\,(Bil(U, V, K)) = nm.}$$

En particular, si $U = V$ se obtiene

$$\boxed{dim\,(Bil(V, V, K)) = n^2.}$$

32.5.1. Ejemplo

Por ejemplo, para $\mathbb{R}^2$ tenemos la base

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

por lo que el espacio dual tiene como base las funciones lineales

$$f_1, \quad f_2,$$

que satisfacen

$$\begin{aligned} f_1(e_1) &= 1, \quad f_1(e_2) = 0, \\ f_2(e_1) &= 0, \quad f_2(e_2) = 1. \end{aligned}$$

Mientras que para el espacio $\mathcal{M}_{2\times 2}(\mathbb{R})$, tenemos la base

$$m_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad m_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad m_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad m_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

entonces una base del espacio dual de $\mathcal{M}_{2\times 2}(\mathbb{R})$ son las funciones lineales

$$g_1, \quad g_2, \quad g_3, \quad g_4,$$

que satisfacen

$$\begin{array}{rcl} g_1(m_1) & = & 1, \quad g_1(m_2) = 0 \quad g_1(m_3) = 0, \quad g_1(m_4) = 0, \\ g_2(m_1) & = & 0, \quad g_2(m_2) = 1, \quad g_2(m_3) = 0, \quad g_2(m_4) = 0, \\ g_3(m_1) & = & 0, \quad g_3(m_2) = 0, \quad g_3(m_3) = 1, \quad g_3(m_4) = 0, \\ g_4(m_1) & = & 0, \quad g_4(m_2) = 0, \quad g_4(m_3) = 0, \quad g_4(m_4) = 1. \end{array}$$

Así, una base de las formas bilineales de $\mathbb{R}^2 \times \mathcal{M}_{2\times 2}(\mathbb{R})$ son la funciones

$$f_1g_1, f_1g_2, f_1g_3, f_1g_4, f_2g_1, f_2g_2, f_2g_3, f_2g_4,$$

y una cualquier forma bilineal de $\mathbb{R}^3 \times \mathcal{M}_{2\times 2}(\mathbb{R})$ se puede escribir como

$$T = \alpha_{11}f_1g_1 + \alpha_{12}f_1g_2 + \alpha_{13}f_1g_3 + \alpha_{13}f_1g_4 + \alpha_{21}f_2g_1 + \alpha_{22}f_2g_2 + \alpha_{23}f_2g_3 + \alpha_{24}f_2g_4,$$

con $\alpha_{ij} \in \mathbb{R}$. De donde

$$\begin{array}{rcl} T\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) & = & \alpha_{11}xa + \alpha_{12}xb + \alpha_{13}xc + \alpha_{13}xd + \\ & & +\alpha_{21}ya + \alpha_{22}yb + \alpha_{23}yc + \alpha_{24}yd. \end{array}$$

Se puede observar que con la matriz y el vector

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \end{pmatrix}, \qquad \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$$

se tiene

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

32.6. Ejercicios

1.- Para el espacio $\mathbb{R}^3 \times \mathcal{M}_{2\times 2}(\mathbb{R})$

a) Obtener una base de sus formas bilineales.

b) Obtener la expresión generales sus formas bilineales.

c) Obtener la expresión matricial sus formas bilineales

Capítulo 33

Formas cuadráticas

En este capítulo estudiaremos las formas cuadráticas y algunas de sus aplicaciones en geometría y física.

En particular, estudiamos la forma cuadrática que fundamenta la relatividad especial y que implica las transformaciones de Lorentz.

33.1. Definición

Supongamos que V es un espacio vectorial sobre el campo K y que la función

$$T : V \times V \to K$$

es una forma bilineal. Entonces podemos definir la función de V en K

$$\boxed{q : V \to K}$$

de tal forma que si $v \in V$ se tiene la regla

$$\boxed{q(v) = T(v, v).}$$

A esta función se le llama forma cuadrática.

33.2. Matriz asociada

Si q es una forma cuádrica, entonces es una forma bilineal y debe tener asociada una matriz.

Para encontrar la matriz asociada, supongamos que V tiene dimensión n y que los vectores

$$v_1, v_2, \cdots, v_n$$

forman una base de sus bases. Por lo que si $v \in V$, entonces existen n elementos de K

$$\lambda_1, \lambda_2, \cdots, \lambda_n,$$

tales que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

y por lo tanto tenemos

$$\begin{aligned} q(v) &= T(v, v) \\ &= T(\lambda_i v_i, \lambda_j v_j) \\ &= \lambda_i T(v_i, v_j) \lambda_j \\ &= \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n \end{pmatrix} \eta \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \end{aligned}$$

con

$$\boxed{\eta_{ij} = T(v_i, v_j), \qquad i, j = 1, 2, \cdots, n.}$$

Se puede observar que definiendo

$$X = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix},$$

se tiene

$$\boxed{q(v) = q(X) = X^T \eta X.} \tag{33.1}$$

33.3. Transformaciones invariantes

En la geometría y en varias áreas de la física son importantes las transformaciones lineales

$$\boxed{X' = \Lambda X, \qquad \Lambda \in \mathcal{M}_{n\times n}(\mathbb{R})} \tag{33.2}$$

que dejan invariante a la forma cuádratica (33.1), es decir, las transformaciones lineales que cumplen

$$\boxed{q(X') = q(X).} \tag{33.3}$$

Note que considerando la expresión (33.1), la igualdad (33.3) se puede escribir como

$$X'^T \eta X' = X^T \eta X. \tag{33.4}$$

Además, se puede observar que sustituyendo la ecuación (33.2) en la igualdad (33.4) se obtiene

$$\begin{aligned} X'^T \eta X' &= \left(\Lambda X^T\right) \eta \Lambda X \\ &= X^T \left(\Lambda^T \eta \Lambda\right) X \\ &= X^T \eta X, \end{aligned}$$

de donde

$$X^T \left(\Lambda^T \eta \Lambda\right) X = X^T \eta X.$$

Esta última igualdad se cumple si

$$\Lambda^T \eta \Lambda = \eta \tag{33.5}$$

que implica

$$\det\left(\Lambda^T \eta \Lambda\right) = \det \eta.$$

Así, tomando en cuenta que $\det \Lambda^T = \det \Lambda$, se obtiene

$$\left(\det \Lambda^T\right) (\det \eta) (\det \Lambda) = (\det \Lambda)^2 (\det \eta) = \det \eta. \tag{33.6}$$

Por lo que si η es una matriz invertible, es decir si

$$\det \eta \neq 0,$$

entonces la ecuación (33.6) implica

$$(\det \Lambda)^2 = 1,$$

esto es

$$\det \Lambda = \pm 1.$$

Por lo tanto, cuando η es invertible, las matrices que dejan invariante la forma cuádratica (33.1) son invertibles.

Si η es invertible, podemos definir el conjunto de matrices invertibles

$$\boxed{G_\eta = \left\{ \Lambda \in \mathcal{M}_{n\times n}(\mathbb{R}) | \Lambda^T \eta \Lambda = \eta, \quad \det \eta \neq 0 \right\}.}$$

Con el producto de matrices, el conjunto G_η es un grupo.

Antes de realizar la prueba recordemos que para que G_η sea un grupo debe cumplir las siguientes propiedades:

1) Axioma de cerradura:

$$\boxed{\Lambda_1, \Lambda_2 \in G_\eta \qquad \Longrightarrow \qquad \Lambda_1\Lambda_2 \in G_\eta.}$$

2) Axioma de asociatividad:

$$\boxed{\Lambda_1, \Lambda_2, \Lambda_3 \in G_\eta, \qquad \Longrightarrow \qquad \Lambda_1 (\Lambda_2 \cdot \lambda_3) = (\Lambda_1\Lambda_2) \Lambda_3.}$$

3) Axioma del neutro:

$$\boxed{\exists I \in G_\eta, \forall \Lambda \in G_\eta \Longrightarrow \qquad \Lambda \cdot I = I \cdot \Lambda = \Lambda.}$$

4) Axioma del inverso:

$$\boxed{\forall\Lambda \in G_\eta, \exists\Lambda^{-1} \in G_\eta, \quad \Lambda\Lambda^{-1} = \Lambda^{-1}\Lambda = I.}$$

Se puede observar que $I \in G_\eta$, por lo que se cumple el axioma del neutro. Además, debido a que los elementos de G_η son matrices, se cumple el axioma de la asociatividad. Por lo tanto, para mostrar que G_η es un grupo solo debemos mostrar el axioma de la cerradura y el axioma del inverso.

Primero mostraremos el axioma de la cerradura, para ello supongamos que $\Lambda_1, \Lambda_2 \in G_\eta$, entonces se cumple

$$\Lambda_1^T \eta \Lambda_1 = \eta, \qquad \Lambda_2^T \eta \Lambda_2 = \eta,$$

por lo que

$$\begin{aligned}
(\Lambda_1\Lambda_2)^T \eta (\Lambda_1\Lambda_2) &= \Lambda_2^T \left(\Lambda_1^T \eta \Lambda_1\right) \Lambda_2 \\
&= \Lambda_2^T \eta \Lambda_2 \\
&= \eta,
\end{aligned}$$

es decir,

$$(\Lambda_1\Lambda_2)^T \eta (\Lambda_1\Lambda_2) = \eta.$$

De donde, si $\Lambda_1, \Lambda_2 \in G_\eta$, se cumple $\Lambda_1\Lambda_2 \in G_\eta$, lo que prueba que se cumple el axioma de la cerradura.

Ahora probaremos el axioma del inverso. Supongamos que $\Lambda \in G_\eta$, entonces sabemos que existe la matriz inversa Λ^{-1} y se cumple la ecuación (33.5). Por lo tanto, de la ecuación (33.5), se obtiene

$$\begin{aligned}
\left(\Lambda^{-1}\right)^T \eta &= \left(\Lambda^{-1}\right)^T \left(\Lambda^T \eta \Lambda\right) \\
&= \left(\Lambda^{-1}\right)^T \Lambda^T \eta \Lambda \\
&= \left(\Lambda\Lambda^{-1}\right)^T \eta \Lambda \\
&= \eta\Lambda,
\end{aligned}$$

es decir,

$$\left(\Lambda^{-1}\right)^T \eta = \eta\Lambda,$$

de donde

$$\left(\Lambda^{-1}\right)^T \eta\Lambda^{-1} = \eta\Lambda\Lambda^{-1} = \eta,$$

esto es,

$$\left(\Lambda^{-1}\right)^T \eta\Lambda^{-1} = \eta.$$

Esta igualdad nos indica que si $\Lambda^{-1} \in G_\eta$. Por lo tanto, si $\Lambda \in G_\eta$, entoces $\Lambda^{-1} \in G_\eta$, en otras palabras, se cumple el exioma del inverso.

Por lo tanto, el conjunto G_η es un grupo, es decir, si $\det \eta \neq 0$ el conjunto de las transformaciones lineales que dejan invariante la forma cuadrática (33.1) es un grupo. Estos tipos de grupos son fundamentales para diferentes leyes de la física.

33.4. Distancia

Como primer ejemplo, supongamos que $V = \mathbb{R}^n$ y

$$\eta = I.$$

Entonces, si

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

la forma cuadrática es

$$\boxed{q(X) = X^T I X = X^T X = x_i \delta_{ij} x_j = x_1^2 + x_2^2 + \cdots + x_n^2 = r^2.}$$

Note que con la matriz identidad, la forma cuadrática nos da la distancia euclidiana de un vector de $\mathbb{R}^n$ al origen. Por esta razón, se dice que la matriz identidad es la métrica del espacio euclidiano.

En particular, para $\mathbb{R}^3$ se tiene la forma cuadrática

$$\boxed{r^2 = x_1^2 + x_2^2 + x_3^2.} \tag{33.7}$$

Esta cantidad es muy importante en la Física. Por ejemplo, la fuerza gravitacional entre dos cuerpos masivos depende de las masas y del inverso de r^2. De la misma forma, la fuerza eléctrica entre dos partículas cargadas depende de las cargas y del inverso de r^2. En este caso, el grupo que deja invariante a la forma cuadrática r^2 es el grupo de rotaciones y se ha mostrado experimentalmente que la fuerza gravitacional y la fuerza electrostática son invariantes bajo rotaciones. Adicionalmente, a nivel cosmológico, de forma observacional se ha mostrado que aspecto del universo no cambia según el ángulo con el que se enfocan los telescopios. Por lo tanto, cualquier teoría cosmológica que intente describir al universo debe ser invariante bajo rotaciones.

Así, la forma cuadrática (33.7) y su grupo de transformaciones que la dejan invariantes son importantes para describir la naturaleza.

33.5. Pseudodistancia

Ahora veamos un ejemplo en la relativida especial de Einstein. En esta teoría, si c es la velocidad de la luz, t es el tiempo y x, y, z son las coordenadas del espacio tridimensional, un cuadrivector se define como

$$X = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}. \tag{33.8}$$

En este espacio-tiempo se define la llamada métrica de Minkowski como la matriz

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

De donde, se puede construir la forma cuadrática

$$
\begin{aligned}
q(X) &= s^2 = X^T \eta X \\
&= \begin{pmatrix} ct & x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \\
&= c^2t^2 - x^2 - y^2 - z^2,
\end{aligned}
$$

es decir

$$
s^2 = c^2t^2 - x^2 - y^2 - z^2. \tag{33.9}
$$

A esta forma cuadrática se le llama pseudodistancia y fue propuesta por A. Einstein para poder explicar que la velocidad de luz es la misma para cualquier observador que se mueve con velocidad constante con respecto a la fuente de luz. Así, la forma cuadrática es la piedra angular de la relatividad de A. Einstein.

Al grupo de transformaciones lineales que dejan invariantes la forma cuadrática (33.9) se le llama transformaciones de Lorentz. Un estudio completo del grupo de Lorentz rebasa los propósitos de este libro, pero veremos algunos miembros de este grupo y algunas de sus implicaciones. Por simplicidad, consideremos la matriz

$$
\Lambda = \begin{pmatrix} a_1 & a_2 & 0 & 0 \\ a_3 & a_4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
$$

entonces, la ecuación

$$
\Lambda^T \eta \Lambda = \eta
$$

se puede escribir explícitamente como

$$
\begin{aligned}
\Lambda^T \eta \Lambda &= \begin{pmatrix} a_1 & a_3 & 0 & 0 \\ a_2 & a_4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & 0 & 0 \\ a_3 & a_4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} a_1 & a_3 & 0 & 0 \\ a_2 & a_4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & 0 & 0 \\ -a_3 & -a_4 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} a_1^2 - a_3^2 & a_1 a_2 - a_3 a_4 & 0 & 0 \\ a_1 a_2 - a_3 a_4 & a_2^2 - a_4^2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= \eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},
\end{aligned}
$$

que implica

$$a_1^2 - a_3^2 = 1, \tag{33.10}$$

$$a_2^2 - a_4^2 = -1, \tag{33.11}$$

$$a_1 a_2 - a_3 a_4 = 0. \tag{33.12}$$

Ahora, definiendo

$$
\begin{aligned}
\beta &= \frac{v}{c}, \\
\gamma &= \frac{1}{\sqrt{1-\beta^2}},
\end{aligned}
$$

se obtienen las soluciones

$$
\begin{aligned}
a_1 &= a_4 = \gamma, \\
a_2 &= a_3 = -\beta\gamma.
\end{aligned}
$$

Por lo que una transformación de Lorentz está dada por

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

de donde

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix},$$

que implica

$$\begin{aligned} ct' &= \gamma\left(ct - \beta x\right), \\ x' &= \gamma\left(x - \beta ct\right), \\ y' &= y, \\ z' &= z. \end{aligned}$$

Uno de los postulados que impuso Albert Einstein para construir su teoría especial de la relatividad es que las leyes de la física deben ser invariantes bajo las transformaciones de simetría de la forma cuadrática (33.9). Otra forma de expresar este postulado es que las leyes de la física deben ser invariantes bajo las transformaciones de Lorentz.

Se puede observar que la ecuación

$$ct' = \gamma\left(ct - \beta x\right)$$

nos indica que el tiempo no es un cantidad absoluta, pues depende de β que a su vez depende de la velocidad del observador v. Esta ley de transformación ha sido comprobada experimentalmente de diversas formas. El hecho de que el tiempo no sea una variable absoluta, si no que dependa del observador, fue un paso revolucionario en la Física y en todas la ciencias.

Al estudiar las simetrías de la ecuación de onda, H. Lorentz descrubrió las transformaciones que llevan su nombre. Sin embargo, fue Albert Einstein quien propuso cambiar todas las leyes de la física para que fueran compatibles con dichas transformaciones.

Hendrik A. Lorentz (1853-1928) fue un físico y matemático holandés. En su infancia recibió una eduación en la cual podía elegir los temas y el ritmo para abordarlos. Se graduó de física y matemáticas en la universidad de Leyden. Posteriormente, trabajó como profesor e inició su doctorado en la misma universidad. Como investigador realizó trabajos que dieron claridad sobre temas de óptica y diversos fenómenos electromagnéticos, también escribió un tratado de análisis matemático. Uno de sus trabajo más relevantes fue sobre el estudio de las simetrías de las ecuaciones de Lorentz, en donde encontró las transformaciones que llevan su nombre. Este trabajo en un inicio fue mal recibido, pero diferentes trabajos experimentales le dieron la razón y permitieron a Einstein contruir la relatividad especial. Además, junto con su alumno P. Zeeman realizó diversos estudios sobre los efectos de un campo magnético intenso en una fuente de luz de sodio, en donde descubrieron el llamado efecto Zemman. Por dicho trabajo ambos recibieron el premio Nobel de física en 1902.

33.6. Espacio de anti de Sitter

Otra forma cuadrática es la que define al llamado espacio de anti de Sitter.

Este espacio se define en $\mathbb{R}^{2+d}$. Por lo que podemos tomar

$$X = \begin{pmatrix} x_0 \\ x_{0'} \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

y la matriz asociada es η

$$\begin{aligned} \eta_{00} &= 1, \quad \eta_{0'0} = \eta_{00'} = 0, \quad \eta_{0i} = \eta_{i0} = 0, \qquad i = 1, 2, \cdots, d, \\ \eta_{0'0'} &= 1, \quad \eta_{0'i} = \eta_{i0'} = 0, \\ \eta_{ji} &= \delta_{ij}, \quad j = 1, 2, \cdots, d, \end{aligned}$$

esto es

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 \end{pmatrix}.$$

De donde se obtiene la forma cuadrática

$$\begin{aligned} q(X) &= X^T \eta X \\ &= \begin{pmatrix} x_0 & x_{0'} & x_1 & x_2 & \cdots & x_d \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_{0'} \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} \\ &= x_0^2 + x_{0'}^2 - \left(x_1^2 + x_2^2 + \cdots + x_d^2\right), \end{aligned}$$

esto es

$$q(X) = x_0^2 + x_{0'}^2 - \left(x_1^2 + x_2^2 + \cdots + x_d^2\right). \tag{33.13}$$

Debido a las dos componentes $x_0, x_{0'}$ y a los dos primeros renglones de la matriz η, se dice que es un espacio de dos tiempos y d coordenadas espaciales. El grupo de transformaciones que deja invariante a la forma cuadrática (33.13) se denota como

$$O(2, d).$$

El caso

$$q(X) = -R^2,$$

es decir,

$$x_0^2 + x_{0'}^2 - \left(x_1^2 + x_2^2 + \cdots + x_d^2\right) = -R^2$$

define un espacio hiperbólico. Este espacio es uno de los más estudiados en Física debido a que tiene propiedades que permitener ralacionar diferentes teorías.

Capítulo 34

Funciones multilineales y tensores

En este capítulo, extenderemos el concepto de función bilineal a función multilineal. De la misma forma, extenderemos el concepto forma bilineal a forma multilineal.

A las componentes de una forma multilineal se le suelen llamar tensores y tienen una aplia gamma de aplicaciones en diferentes ingenierías y en la Física.

34.1. Definición

Sean $V_1, V_2, \cdots, V_N$ espacios vectoriales en el campo K, se dice que la función

$$T : V_1 \times V_2 \times \cdots \times V_N \to W$$

es multilineal si $\lambda_1, \lambda_2 \in K$ y

$$v_1 \in V_2, v_2 \in V_2, \cdots, v_i, v_i' \in V_i, \cdots, v_N \in V_N$$

se cumple

$$\begin{aligned}
&T(v_1, v_2, \cdots, \lambda_1 v_i + \lambda_2 v_i', \cdots, v_N) \\
&= \lambda_1 T(v_1, v_2, \cdots, v_i, \cdots v_N) + \lambda_2 T(v_1, v_2, \cdots, v_i', \cdots v_N)
\end{aligned}$$

En otras palabras, la función T es multilineal si es lineal en cada una de sus entradas.

34.2. Funciones multilineales actuando en una base

Ahora probaremos que si una funciones multilineales está definida en una base del dominio $V_1 \times V_2 \times \cdots \times V_N$ automáticamente está definida en todo el dominio. También recordemos que si dos funciones lineales coinciden en una base del dominio, esa funciones multilineales son las mismas. Nos concentraremos en espacios vectoriales de dimensión finita.

Supongamos que $V_1, V_2, \cdots, V_N$ son espacios vectoriales de dimensión finita sobre K con

$$\begin{aligned} \dim(V_1) &= n_1, \\ \dim(V_2) &= n_2. \\ &\vdots \\ \dim(V_i) &= n_i, \\ &\vdots \\ \dim(V_N) &= n_N. \end{aligned}$$

También supongamos que la función

$$T : V_1 \times V_2 \times \cdots \times V_N \to W$$

es multilineal. Probaremos que si T está definida en una base de cada uno de los espacios vectoriales V_i, automáticamente está definida en todo el dominio

$$V_1 \times V_2 \times \cdots \times V_N. \tag{34.1}$$

Para iniciar la prueba, notemos que si los vectores

$$v_{1i}, v_{2i}, \cdots, v_{n_i i}$$

son base del espacio vectorial V_i, entonces existen $n_1 n_2 \cdots n_N$ vectores en W tales que

$$T(v_{j_1 1}, v_{j_2 2}, \cdots, v_{j_i i}, \cdots, v_{j_N N}) = w_{j_1 j_2 \cdots j_N}, \quad j_k = 1, 2, \cdots, n_k. \tag{34.2}$$

Ahora supongamos que $v_i \in V_i$, entonces existen n_i elementos en K

$$\lambda_{1i}, \lambda_{2i}, \cdots, \lambda_{n_i i},$$

tales que

$$v_i = \lambda_{1i}v_{1i} + \lambda_{2i}v_{2i} + \cdots + \lambda_{n_i i}v_{n_i i},$$

esto es

$$v_i = \lambda_{j_i i}v_{j_i i}, \quad j_i = 1, 2, \cdots, n_i$$

De donde,

$$\begin{aligned}
&T(v_1, v_2, \cdots, v_i, \cdots, v_N) = T(\lambda_{j_1 1}v_{j_1 1}, \lambda_{j_2 2}v_{j_2 2}, \cdots, \lambda_{j_i i}v_{j_i i}, \cdots, \lambda_{j_N N}v_{j_N N})\\
&= \lambda_{j_1 1}\lambda_{j_2 2}\cdots\lambda_{j_i i}\cdots\lambda_{j_N N}T(v_{j_1 1}, v_{j_2 2}, \cdots, v_{j_i i}, \cdots, v_{j_N N})\\
&= \lambda_{j_1 1}\lambda_{j_2 2}\cdots\lambda_{j_N N}w_{j_1 j_2 \cdots j_N},
\end{aligned}$$

esto es

$$T(v_1, v_2, \cdots, v_N) = \lambda_{j_1 1}\lambda_{j_2 2}\cdots\lambda_{j_N N}w_{j_1 j_2 \cdots j_N}. \tag{34.3}$$

Por lo tanto, si una función multilineal se define sobre cada una de las bases de los espacios V_i, automáticamente está definida sobre todo el dominio (34.1)

Ahora, supongamos que

$$G : V_1 \times V_2 \times \cdots \times V_N \to W$$

es otra función multiilineal que satisface (34.2), es decir, que cumple

$$G(v_{1j_1}, v_{2j_2}, \cdots, v_{ij_i}, \cdots, v_{Nj_N}) = w_{j_1 j_2 \cdots j_N}, \quad j_k = 1, 2, \cdots, n_k.$$

Por lo cual, considerando que G es una función multilineal y la igualdad (34.3), se encuentra

$$\begin{aligned}
&G(v_1, v_2, \cdots, v_i, \cdots, v_N) = G(\lambda_{j_1 1}v_{j_1 1}, \lambda_{j_2 2}v_{j_2 2}, \cdots, \lambda_{j_i i}v_{j_i i}, \cdots, \lambda_{j_N N}v_{j_N N})\\
&= \lambda_{j_1 1}\lambda_{j_2 2}\cdots\lambda_{j_i i}\cdots\lambda_{j_N N}G(v_{j_1 1}, v_{j_2 2}, \cdots, v_{j_i i}, \cdots, v_{j_N N})\\
&= \lambda_{j_1 1}\lambda_{j_2 2}\cdots\lambda_{j_N N}w_{j_1 j_2 \cdots j_N}\\
&= T(v_1, v_2, \cdots, v_N)
\end{aligned}$$

esto es

$$G(v_1, v_2, \cdots, v_N) = T(v_1, v_2, \cdots, v_N)$$

de donde

$$G = T.$$

Por lo tanto, si dos funciones multilineales coinciden en una base del dominio

$$V_1 \times V_2 \times \cdots \times V_N,$$

coinciden en todo ese espacio y por lo tanto son iguales.

34.3. Formas multilineales

Sean $V_1, V_2, \cdots, V_N$ espacios vectoriales en el campo K, si la función

$$\boxed{T : V_1 \times V_2 \times \cdots \times V_N \to K}$$

es multilineal se dice que T que es una forma multilineal.

Con los espacios duales V_i^* de los espacios vectoriales V_i se pueden construir formas multlineales.

En efecto, supongamos que V_i son espacios vectoriales sobre K y V_i^* son sus espacios duales. Entonces, si $f_i \in V_i^*$ podemos definir la función

$$T : V_1 \times V_2 \times \cdots \times V_N \to K$$

tal que si $v_i \in V_i$ se cumple

$$\boxed{T(v_1, v_2, \cdots, v_i, \cdots, v_n) = f_1(v_1) f_2(v_2) \cdots f_i(v_i) \cdots f_N(v_N)} \tag{34.4}$$

Probaremos que T es una función multilineal y por lo tanto una forma multilineal. Para realizar la prueba supongamos que $\lambda, \gamma \in K$ y $v_i, v_i' \in V_i$ de donde

$$\begin{aligned}
&T(v_1, v_2, \cdots, \lambda v_i + \gamma v_i', \cdots, v_n) = \\
&= f_1(v_1) f_2(v_2) \cdots f_{i-1}(v_{i-1}) f_i(\lambda v_i + \gamma v_i') f_{i+1}(v_{i+1}) \cdots f_N(v_N) \\
&= f_1(v_1) f_2(v_2) \cdots f_{i-1}(v_{i-1}) \left(\lambda f_i(v_i) + \gamma f(v_i')\right) f_{i+1}(v_{i+1}) \cdots f_N(v_N) \\
&= \lambda f_1(v_1) f_2(v_2) \cdots f_{i-1}(v_{i-1}) f_i(v_i) f_{i+1}(v_{i+1}) \cdots f_N(v_N) + \\
&+ \gamma f_1(v_1) f_2(v_2) \cdots f_{i-1}(v_{i-1}) f(v_i') f_{i+1}(v_{i+1}) \cdots f_N(v_N) \\
&= \lambda T(v_1, v_2, \cdots, v_i, \cdots, v_N) + \gamma T(v_1, v_2, \cdots, v_i', \cdots, v_N)
\end{aligned}$$

esto es

$$\begin{aligned}
&T(v_1, v_2, \cdots, \lambda v_i + \gamma v_i', \cdots, v_N) = \\
&= \lambda T(v_1, v_2, \cdots, v_i, \cdots, v_N) + \gamma T(v_1, v_2, \cdots, v_i', \cdots, v_n).
\end{aligned}$$

Así, la función (34.4) es una forma multilineal.

De hecho, cualquier forma multilineal se puede expresar en términos de los espacios duales V_i.

34.4. Base de las formas bilineales

Definamos el conjunto de formas bilineales

$$\boxed{FMul = \{T : V_1 \times V_2 \times \cdots \times V_N \to K | T \quad \text{Multilineal}\}.} \tag{34.5}$$

Mostraremos que este espacio es generado por los espacios duales V_i^*.

Para probar esta afirmación, recordemos que si V_i es un espacio vectorial de dimensión n_i sobre K y

$$\{v_{1i}, v_{2i}, \cdots v_{n_i i}\}$$

es una de sus bases, entonces el conjunto de funciones lineales

$$\{f_{1i}, f_{2i}, \cdots, f_{n_i i}\} \tag{34.6}$$

que satisfacen

$$f_{ki}(v_{li}) = \delta_{kl}, \qquad k, l = 1, 2, \cdots, n_i$$

forman una base del espacio dual V_i^*.

Por lo que podemos construir el conjunto de $n_1 n_2 \cdots n_N$ formas multilineales

$$f_{i_1 1} f_{i_2 2} f_{i_3 3} \cdots f_{i_N N} \qquad i_k = 1, 2, \cdots, n_k \tag{34.7}$$

que tienen como dominio $V_1 \times V_2 \times \cdots \times V_N$.

Primero probaremos que las formas mutilineales (34.7) general el espacio (34.5). Así, supongamos que la función

$$T : V_1 \times V_2 \times \cdots \times V_N \to K$$

es una forma multilineal, entonces existen $n_1 n_2 \cdots n_N$ elementos $\alpha_{i_1 i_2 \cdots i_N}$ en K tal que

$$T(v_{i_1 1}, v_{i_2 2}, \cdots, v_{i_N N}) = \alpha_{i_1 i_2 \cdots i_N}.$$

De manera equivalente, podemos definir la función multilineal

$$T' = \alpha_{i_1 i_2 \cdots i_N} f_{i_1 1} f_{i_2 2} \cdots f_{i_N N}$$

Ahora supongamos que $v_i \in V_i$, entonces existen n_i elementos en K

$$\lambda_{1i}, \lambda_{2i}, \cdots, \lambda_{n_i i},$$

tales que

$$v_i = \lambda_{1i} v_{1i} + \lambda_{2i} v_{2i} + \cdots + \lambda_{n_i i} v_{n_i i},$$

esto es,

$$v_i = \lambda_{j_i i} v_{j_i i}, \quad j_I = 1, 2, \cdots, n_i$$

De donde,

$$\begin{aligned}
&T(v_1, v_2, \cdots, v_i, \cdots, v_N) = T(\lambda_{j_1 1} v_{j_1 1}, \lambda_{j_2 2} v_{j_2 2}, \cdots, \lambda_{j_i i} v_{j_i i}, \cdots, \lambda_{j_N N} v_{j_N N}) \\
&= \lambda_{j_1 1} \lambda_{j_2 2} \cdots \lambda_{j_i i} \cdots \lambda_{j_N N} T(v_{j_1 1}, v_{j_2 2}, \cdots, v_{j_i i}, \cdots, v_{j_N N}) \\
&= \lambda_{j_1 1} \lambda_{j_2 2} \cdots \lambda_{j_N N} \alpha_{j_1 j_2 \cdots j_N},
\end{aligned}$$

es decir,

$$T(v_1, v_2, \cdots, v_N) = \lambda_{j_1 1} \lambda_{j_2 2} \cdots \lambda_{j_N N} \alpha_{j_1 j_2 \cdots j_N}.$$

De forma analoga, como T' es multilineal, tenemos

$$\begin{aligned}
&T'(v_1, v_2, \cdots, v_N) = \left(\alpha_{i_1 i_2 \cdots i_N} f_{i_1 1} f_{i_2 2} \cdots f_{i_N N}\right)(v_1, v_2, \cdots, v_N) \\
&= \alpha_{i_1 i_2 \cdots i_N} f_{i_1 1}(v_1) f_{i_2 2}(v_2) \cdots f_{i_N N}(v_N) \\
&= \alpha_{i_1 i_2 \cdots i_N} f_{i_1 1}(\lambda_{j_1 1} v_{j_1 1}) f_{i_2 2}(\lambda_{j_2 2} v_{j_2 2}) \cdots f_{i_N N}(\lambda_{j_N N} v_{j_N N}) \\
&= \lambda_{j_1 1} \lambda_{j_2 2} \cdots \lambda_{j_i i} \cdots \lambda_{j_N N} \alpha_{i_1 i_2 \cdots i_N} f_{i_1 1}(v_{j_1 1}) f_{i_2 2}(v_{j_2 2}) \cdots f_{i_N N}(v_{j_N N}) \\
&= \lambda_{j_1 1} \lambda_{j_2 2} \cdots \lambda_{j_i i} \cdots \lambda_{j_N N} \alpha_{i_1 i_2 \cdots i_N} \delta_{i_1 j_1} \delta_{i_2 j_2} \cdots \delta_{i_N j_N} \\
&= \lambda_{j_1 1} \lambda_{j_2 2} \cdots \lambda_{j_N N} \alpha_{j_1 j_2 \cdots j_N},
\end{aligned}$$

de donde

$$T'(v_1, v_2, \cdots, v_N) = \lambda_{j_1 1} \lambda_{j_2 2} \cdots \lambda_{j_N N} \alpha_{j_1 j_2 \cdots j_N} = T(v_1, v_2, \cdots, v_N)$$

Así, cualquier forma multilineal se puede escribir como la siguiente combinación lineal

$$\boxed{T = \alpha_{i_1 i_2 \cdots i_N} f_{i_1 1} f_{i_2 2} \cdots f_{i_N N}.}$$

Ahora probaremos que el conjunto de las formas multilineales (34.7) son linealmente independientes. Para realizar esta prueba, supongamos que existen nm elementos $\beta_{i_1 i_2 \cdots i_N} \in K$ tales que

$$\beta_{i_1 i_2 \cdots i_N} f_{i_1 1} f_{i_2 2} \cdots f_{i_N N} = 0,$$

entonces

$$\begin{aligned}
0 &= (\beta_{i_1 i_2 \cdots i_N} f_{i_1 1} f_{i_2 2} \cdots f_{i_N N})(v_{11}, v_{12}, \cdots, v_{1N}) \\
&= \beta_{i_1 i_2 \cdots i_N} f_{i_1 1}(v_{11}) f_{i_2 2}(v_{12}) \cdots f_{i_N N}(v_{1N}) \\
&= \beta_{i_1 i_2 \cdots i_N} \delta_{i_1 1} \delta_{i_2 2} \cdots \delta_{1N} \\
&= \beta_{11 \cdots 1}
\end{aligned}$$

es decir,

$$\beta_{11 \cdots 1} = 0.$$

Con el mismo procedimiento, se puede mostrar que para cualesquiera $i_1, i_2, \cdots, i_N$ se obtiene

$$\beta_{i_1 i_2 \cdots i_N} = 0.$$

Entonces, el conjunto de funciones (34.7) es linealmente independiente y por lo tanto es una base del espacio de las formas multilineales. Este resultado implica que el espacio de las formas multilineales (34.5) tiene dimensión

$$\boxed{\dim (FMul) = n_1 n_2 \cdots n_N.}$$

En particular, si todos los espacios vectoriales son los mismos, $V = V_i$, tenemos el espacio

$$FMul(V) = \{T : V \times V \times \cdots \times \to K | T \quad \text{Multilineal}\}.$$

que cuya dimensión es

$$\boxed{dim (FMul(V)) = n^N.}$$

34.5. Ejercicios

1.- Probar que si $\lambda \in \mathbb{C}$, la función

$$T : \mathbb{C} \times \mathbb{C} \times \mathbb{C} \to \mathbb{C}$$

tal que si $z_1, z_2, z_3 \in \mathbb{C}$ se cumple

$$T(z_1, z_2, z_3) = \lambda z_1 z_2 z_3$$

es multilineal.

2.- Para el espacio de las matrices $\mathcal{M}_{n\times n}(\mathbb{R})$ probar que si $\lambda \in \mathbb{R}$, la función

$$T : \mathcal{M}_{n\times n}(\mathbb{R}) \times \mathcal{M}_{n\times n}(\mathbb{R}) \times \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathcal{M}_{n\times n}(\mathbb{R})$$

tal que si $A_1, A_2, A_3 \in \mathcal{M}_{n\times n}(\mathbb{R})$ se cumple

$$T(A_1, A_2, A_3) = \lambda A_1 A_2 A_3$$

es una función multilineal.

3.- Para el espacio de las matrices $\mathcal{M}_{n\times n}(\mathbb{R})$ probar que si $\lambda \in \mathbb{R}$, la función

$$T : \mathcal{M}_{n\times n}(\mathbb{R}) \times \mathcal{M}_{n\times n}(\mathbb{R}) \times \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$$

tal que si $A_1, A_2, A_3 \in \mathcal{M}_{n\times n}(\mathbb{R})$ se cumple

$$T(A_1, A_2, A_3) = \lambda Tr(A_1 A_2 A_3)$$

es una forma multilineal.

3.- Para $\mathbb{R}^3$ la función producto vectorial

$$T : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$

se define tal que si

$$\begin{aligned} \hat{i} = (1,0,0), \hat{j} &= (0,1,0), \hat{k} = (0,0,1) \\ V = (v_1, v_2, v_3), &\qquad W = (w_1, w_2, w_3), \end{aligned}$$

se cumple

$$\begin{aligned} T(V, W) &= V \times W = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\ &= (v_2 w_3 - v_3 w_2)\hat{i} + (v_3 w_1 - v_1 w_3)\hat{j} + (v_1 w_2 - v_2 w_1)\hat{k}. \end{aligned}$$

Probar que T es una función bilineal.

4.- Para $\mathbb{R}^3$ la función triple producto vectorial

$$T : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$

es definida como

$$T(V_1, V_2, V_3) = V_1 \times (V_2 \times V_3)\,,$$

con

$$V_1, V_2, V_3 \in \mathbb{R}^3,$$

probar que T es una función multilineal.

5.- Para $\mathbb{R}^3$ la función triple producto escalar

$$T : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$

es definida como

$$T(V_1, V_2, V_3) = V_1 \cdot (V_2 \times V_3)\,,$$

con

$$V_1, V_2, V_3 \in \mathbb{R}^3,$$

probar que T es una forma multilineal.

Capítulo 35

Funciones alternantes, producto exterior y derivada exterior

En este capítulo estudiaremos las funciones multilineales alternantes con las cuales definiremos el producto exterior. También definiremos la derivada exterior, la cual nos permite relacionar el álgebra lineal con varios conceptos importantes del cálculo diferencial de varias variables como el rotacional y la divergencia.

35.1. Funciones bilineales alternantes

Sea V y W espacios vectoriales sobre el campo K, supongamos que la función

$$T : V \times V \to W$$

es bilineal, se dice que T es una función alternante si $\forall v \in V$ se cumple

$$\boxed{T(v, v) = 0.}$$

Una propiedad de estas funciones es que son antisimétricas. Esto es si T es una función alternante, entonces se cumple

$$\boxed{T(v, u) = -T(u, v) \qquad \forall u, v \in V.}$$

Para probar esta afirmación, notemos que debido a que T es una función bilineal alternante se encuentra

$$\begin{aligned} 0 &= T(u+v, u+v) \\ &= T(u, u+v) + T(v, u+v) \\ &= T(u,u) + T(u,v) + T(v,u) + T(v,v) \\ &= 0 + T(u,v) + T(v,u) + 0 \\ &= T(u,v) + T(v,u), \end{aligned}$$

esto es ,

$$T(u,v) + T(v,u) = 0,$$

de donde

$$T(u,v) = -T(v,u).$$

Que es lo que queríamos demostrar.

35.2. Ejercicios

1.- En la Mecánica Clásica es importante el espacio fase

$$x, \quad p = m\dot{x}.$$

y las funciones en dicho espacio $F(x,p)$. Ahora supongamos que tenemos dos funciones del espacio fase

$$\begin{aligned} F_1 &: \mathbb{R}^2 \to \mathbb{R}, \\ F_2 &: \mathbb{R}^2 \to \mathbb{R}, \end{aligned}$$

entonces se definen los paréntesis o corchetes de Poisson

$$\boxed{T(F_1, F_2) = \{F_1, F_2\} = \frac{\partial F_1}{\partial x}\frac{\partial F_2}{\partial p} - \frac{\partial F_2}{\partial x}\frac{\partial F_1}{\partial p}.}$$

Probar que los paréntesis de Poisson son bilineales y alternantes.

2.- En Mecánica Cuántica es importante el llamado conmutador entre operadores lineales. Para definir el conmutador entre matrices, supongamos que

$$A_1, A_2 \in \mathcal{M}_{n\times n}(\mathbb{C})$$

entonces definiremos el conmutador como

$$\boxed{T\left(A_1, A_2\right) = [A_1, A_2] = A_1A_2 - A_2A_1.}$$

Probar que el conmutador es bilineal y alternante.

35.3. Función alternante actuando en una base

Sea V un espacio vectorial en el campo K y

$$T : V \times V \to W$$

una función alternante. Debido a que T es una función bilineal queda completamente determinada si se define en una base de V. Por lo que si V tiene dimensión n y

$$\{v_1, v_2, \cdots, v_n\}$$

es una de sus bases, entonces T está completamente determinada en todo el dominio V si se definen los vectores w_{ij} en W tales que

$$T(v_i, v_j) = w_{ij}, \quad i, j = 1, 2, \cdots, n.$$

Pero debido a que se debe cumplir

$$\boxed{T(v_i, v_i) = 0, \qquad T(v_i, v_j) = -T(v_j, v_i),}$$

entonces T está completamente definida si se definen los vectores w_{ij} en W tales que

$$\boxed{T(v_i, v_j) = w_{ij}, \qquad i < j.} \tag{35.1}$$

Ahora, si $u, v \in V$, entonces existen n elementos en K

$$\lambda_1, \lambda_2, \cdots, \lambda_n,$$

tal que

$$u = \lambda_{11}v_1 + \lambda_{12}v_2 + \cdots + \lambda_{1n}v_n,$$

y existen n elementos en K

$$\lambda_{21}, \lambda_{22}, \cdots, \lambda_{2n},$$

tal que

$$v = \lambda_{21}v_1 + \lambda_{22}v_2 + \cdots + \lambda_{2n}v_n.$$

Por lo tanto,

$$\begin{aligned} T(u,v) &= T(\lambda_{1j_1}v_{j_1}, \lambda_{2j_2}v_{j_2}) \\ &= \lambda_{1j_1}\lambda_{2j_2}T(v_{j_1}, v_{j_2}), \quad j_1, j_2 = 1, 2, \cdots, n, \end{aligned}$$

es decir,

$$T(u,v) = \lambda_{1j_1}\lambda_{2j_2}T(v_{j_1}, v_{j_2}).$$

De forma explícita se tiene

$$\begin{aligned} T(u,v) &= \lambda_{11}\lambda_{22}T(v_1,v_2) + \lambda_{11}\lambda_{23}T(v_1,v_3) + \cdots + \lambda_{11}\lambda_{2n}T(v_1,v_n) + \\ &\quad +\lambda_{12}\lambda_{21}T(v_2,v_1) + \lambda_{12}\lambda_{23}T(v_2,v_3) + \cdots + \lambda_{12}\lambda_{2n}T(v_2,v_n) + \\ &\quad \vdots \\ &\quad +\lambda_{1n}\lambda_{21}T(v_n,v_1) + \lambda_{1n}\lambda_{22}T(v_n,v_2) + \cdots + \lambda_{1n}\lambda_{2n-1}T(v_n,v_{n-1}), \end{aligned}$$

que se puede reescribir como

$$\begin{aligned} T(u,v) &= \lambda_{11}\lambda_{22}T(v_1,v_2) + \lambda_{12}\lambda_{21}T(v_2,v_1) + \cdots + \lambda_{11}\lambda_{2n}T(v_1,v_n) + \lambda_{1n}\lambda_{21}T(v_n,v_1) + \\ &\quad +\lambda_{12}\lambda_{23}T(v_2,v_3) + \lambda_{13}\lambda_{22}T(v_3,v_2) + \cdots + \lambda_{12}\lambda_{2n}T(v_2,v_n) + \lambda_{1n}\lambda_{22}T(v_n,v_2) + \\ &\quad +\lambda_{13}\lambda_{24}T(v_3,v_4) + \lambda_{14}\lambda_{23}T(v_4,v_3) + \cdots + \lambda_{13}\lambda_{2n}T(v_3,v_n) + \lambda_{1n}\lambda_{23}T(v_n,v_3) + \\ &\quad \vdots \\ &\quad +\lambda_{1n-1}\lambda_{2n}T(v_{n-1},v_n) + \lambda_{1n}\lambda_{2n-1}T(v_n,v_{n-1}). \end{aligned}$$

Entonces, como T es una función alternante se tiene

$$\begin{aligned} T(u,v) &= (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\,T(v_1,v_2) + \cdots + (\lambda_{11}\lambda_{2n} - \lambda_{1n}\lambda_{21})\,T(v_1,v_n) + \\ &\quad + (\lambda_{12}\lambda_{23} - \lambda_{13}\lambda_{22})\,T(v_2,v_3) + \cdots + (\lambda_{12}\lambda_{2n} - \lambda_{1n}\lambda_{22})\,T(v_2,v_n) + \\ &\quad + (\lambda_{13}\lambda_{24} - \lambda_{14}\lambda_{23})\,T(v_3,v_4) + \cdots + (\lambda_{13}\lambda_{2n} - \lambda_{1n}\lambda_{23})\,T(v_3,v_n) + \\ &\quad \vdots \\ &\quad + (\lambda_{1n-1}\lambda_{2n} - \lambda_{1n}\lambda_{2n-1})\,T(v_{n-1},v_n). \end{aligned}$$

Así, con la convención de Einstein se tiene

$$\boxed{T(u,v) = (\lambda_{1j_1}\lambda_{2j_2} - \lambda_{1j_2}\lambda_{2j_1})\, T(v_{j_1}, v_{j_2}), \qquad j_1 < j_2.} \tag{35.2}$$

Además, recordemos que el tensor de Levi-Civita

$$\boxed{\epsilon_{ij}}$$

satisface

$$\begin{aligned} \epsilon_{12} &= 0, \\ \epsilon_{i_1 i_2} &= -\epsilon_{i_1 i_2}, \quad i_1, i_2 = 1, 2. \end{aligned}$$

Por lo que la función bilineal alternante (35.2) se puede expresar como

$$\boxed{T(u,v) = \epsilon_{i_1 i_2}\lambda_{i_1 j_1}\lambda_{i_2 j_2} T(v_{j_1}, v_{j_2}),} \tag{35.3}$$

con

$$\boxed{i_1, i_2 = 1, 2, \quad j_1, j_2 = 1, 2, \cdots, n, \quad j_1 < j_2.}$$

35.4. Dimensión de funciones bilineales alternantes

Supongamos que V tiene dimensión n y

$$\{v_1, v_2, \cdots, v_n\} \tag{35.4}$$

es una de sus bases. Entonces considerando el resultado (35.1) cualquier función bilineal artenante se puede escribir como combinación de

$$\boxed{T(v_i, v_j), \qquad i < j, \quad i, j = 1, 2, \cdots, n.}$$

De forma explícita tenemos

$$\boxed{\begin{array}{llll} T(v_1,v_2), & T(v_1,v_3), & \cdots, & T(v_1,v_n), \\ T(v_2,v_3), & T(v_2,v_4), & \cdots, & T(v_2,v_n), \\ T(v_3,v_4), & T(v_3,v_5), & \cdots, & T(v_3,v_n), \\ & & & \vdots \\ & T(v_{n-2},v_{n-1}), & & T(v_{n-2},v_n), \\ & & & T(v_{n-1},v_n), \end{array}}$$

los cuales, en total son

$$\boxed{(n-1)+(n-2)+\cdots+2+1=\frac{n(n-1)}{2}}$$

elementos linealmente independientes.

Así, cualquier función bilineal alternante se puede escribir como combinación lineal de

$$\boxed{\frac{n(n-1)}{2}}$$

términos linealmente independientes. Entonces, si la dimensión de V es n, la dimensión del espacio de las funciones alternantes en ese espacio es $n(n-1)/2$.

35.4.1. Ejemplos

35.4.2. Caso de dimensión 2

Supongamos que V tiene dimensión 2 y que

$$\{v_1,v_2\}.$$

es una de sus bases. Entonces, si $u,v\in V$ existen

$$\lambda_{11},\lambda_{12},\lambda_{21},\lambda_{22}\in K,$$

tal que

$$\begin{aligned} u &= \lambda_{11}v_1 + \lambda_{12}v_2, \\ v &= \lambda_{21}v_1 + \lambda_{22}v_2. \end{aligned}$$

Por lo que, considerando la expresión (35.2), se puede observar que, en este caso, todas las funciones alternantes se pueden escribir como

$$\boxed{T(u,v) = (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\,T(v_1,v_2).} \tag{35.5}$$

Además, definiendo la matriz

$$\boxed{M = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix},}$$

se obtiene

$$\boxed{\det M = \lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21}}$$

de donde

$$\boxed{T(u,v) = (\det M)\,T\left(v_1,v_2\right).}$$

35.4.3. Caso de dimensión 3

Supongamos que V tiene dimensión 3 y que

$$\{v_1, v_2, v_3\}$$

es una de sus bases. Entonces, si $u, v \in V$ existen

$$\lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{21}, \lambda_{22}, \lambda_{23} \in K$$

tal que

$$\begin{aligned} u &= \lambda_{11}v_1 + \lambda_{12}v_2 + \lambda_{13}v_3, \\ v &= \lambda_{21}v_1 + \lambda_{22}v_2 + \lambda_{23}v_3, \end{aligned}$$

Así, usando la expresión (35.2), todas las funciones alternantes se pueden escribir como

$$\boxed{\begin{aligned} T(u,v) = {} & (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\,T(v_1,v_2) + (\lambda_{11}\lambda_{23} - \lambda_{23}\lambda_{21})\,T(v_1,v_3) + \\ & + (\lambda_{12}\lambda_{23} - \lambda_{13}\lambda_{22})\,T(v_2,v_3). \end{aligned}}$$

35.5. Producto exterior

Si T es una función alternante, se define el producto exterior o cuña como

$$\boxed{T(u,v) = u \wedge v.}$$

Con esta defición se tiene

$$\boxed{\begin{aligned} &v \wedge v = 0, \qquad \forall v \in V, \\ &u \wedge v = -v \wedge u, \qquad \forall v, u \in V. \end{aligned}}$$

Por lo que si V tiene dimensión n y

$$\{v_1, v_2, \cdots, v_n\}$$

es una de sus bases, entonces considerando el resultado (35.1) el producto exterior entre los elementos de la base se escribe como

$$\boxed{v_i \wedge v_j, \quad i < j, \quad i, j = 1, 2, \cdots, n.} \tag{35.6}$$

De forma explícita tenemos

$$\boxed{\begin{array}{cccc} v_1 \wedge v_2, & v_1 \wedge v_3, & \cdots, & v_1 \wedge v_n, \\ v_2 \wedge v_3, & v_2 \wedge v_4, & \cdots, & v_2 \wedge v_n, \\ v_3 \wedge v_4, & v_3 \wedge v_5, & \cdots, & v_3 \wedge v_n, \\ & & & \vdots \\ & & v_{n-2} \wedge v_{n-1}, & v_{n-2} \wedge v_n, \\ & & & v_{n-1} \wedge v_n, \end{array}}$$

los cuales en total son

$$\boxed{\frac{n(n-1)}{2}}$$

productos exteriores linealmente independientes.

El producto exterior entre cualesquiera dos elementos de V se puede escribir en términos de los productos exteriores (35.6). En efecto, si $u, v \in V$, entonces existen n elementos en K

$$\lambda_{11}, \lambda_{12}, \cdots, \lambda_{1n},$$

tales que

$$u = \lambda_{11} v_1 + \lambda_{12} v_2 + \cdots + \lambda_{1n} v_n,$$

y existen n elementos en K

$$\lambda_{21}, \lambda_{22}, \cdots, \lambda_{2n},$$

tales que

$$v = \lambda_{21} v_1 + \lambda_{22} v_2 + \cdots + \lambda_{2n} v_n.$$

Por lo que, considerando el resultado (35.2), el producto exterior entre u y v se escribe como

$$\boxed{u \wedge v = \left(\lambda_{1j_1}\lambda_{2i_2} - \lambda_{1i_2}\lambda_{2j_1}\right)\left(v_{j_1} \wedge v_{j_2}\right), \quad i_1 < i_2.}$$

o usando el tensor de Levi-Civita

$$\boxed{u \wedge v = \epsilon_{i_1 i_2} \lambda_{i_1 j_1} \lambda_{i_2 j_2} (v_{j_1} \wedge v_{j_2}),}$$

con

$$\boxed{i_1, i_2 = 1, 2, \quad j_1, j_2 = 1, 2, \cdots, n, \quad j_1 < j_2.} \tag{35.7}$$

35.6. Ejemplos

35.6.1. Caso de dimensión 2

Supongamos que V tiene dimensión 2 y que

$$\{v_1, v_2\}$$

es una de sus bases. Entonces, si $u, v \in V$ existen $\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22} \in K$ tales que

$$\begin{aligned} u &= \lambda_{11} v_1 + \lambda_{12} v_2, \\ v &= \lambda_{12} v_1 + \lambda_{22} v_2. \end{aligned}$$

Así, en este caso el producto exterior entre u y v tienen la forma

$$\boxed{u \wedge v = (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\,(v_1 \wedge v_2).}$$

Además, definiendo la matriz

$$\boxed{M = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix},}$$

se obtiene

$$\boxed{\det M = \lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21}}$$

de donde

$$\boxed{u \wedge v = (\det M)\,(v_1 \wedge v_2)\,.}$$

35.6.2. Caso de dimensión 3

Ahora supongamos que V tiene dimensión 3 y que

$$\{v_1, v_2, v_3\}$$

es una de sus bases. Entonces, si $u, v \in V$ existen

$$\lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{21}, \lambda_{22}, \lambda_{23} \in K$$

tales que

$$\begin{aligned} u &= \lambda_{11} v_1 + \lambda_{12} v_2 + \lambda_{13} v_3, \\ v &= \lambda_{21} v_1 + \lambda_{22} v_2 + \lambda_{23} v_3, \end{aligned}$$

Así, usando la expresión (35.2), el producto exterior se pueden escribir como

$$\boxed{\begin{aligned} u \wedge v = {} & (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\,(v_1 \wedge v_2) + (\lambda_{11}\lambda_{23} - \lambda_{23}\lambda_{21})\,(v_1 \wedge v_3) + \\ & + (\lambda_{12}\lambda_{23} - \lambda_{13}\lambda_{22})\,(v_2 \wedge v_3). \end{aligned}}$$

35.7. Funciones multilineales alternantes

Sea V un espacio vectorial sobre el campo K y tiene dimensión n, entonces podemos construir el espacio vectorial

$$V^m = \underbrace{V \times V \cdots \times V}_{m}.$$

Por lo que si W un espacio vectorial sobre el campo K y la función

$$\boxed{T : V^m \to W, \qquad\qquad m \leq n} \tag{35.8}$$

es función multilineal, se dice que T es alternante si para todo $u \in V, u_i \in V$

$$\boxed{(u_1, u_2, \cdots, u_{i-1}, u, u_{i+1}, \cdots, u_{j-1}, u, u_{j+1}, \cdots, u_m) \in V^m,}$$

con

$$i \neq j$$

se cumple

$$\boxed{T(u_1, u_2, \cdots, u_{i-1}, u, u_{i+1}, \cdots, u_{j-1}, u, u_{j+1}, \cdots, u_n) = 0.}$$

Además, se dice que una función multilineal T es totalmente antisimétrica si al intercambiar dos vectores del dominio V^m la función cambia en un signo. En otras palabras, T es totalmente antisimétrica si para todo $i \neq j$ se cumple

$$\boxed{\begin{array}{l} T(u_1, u_2, \cdots, u_{i-1}, u_i, u_{i+1}, \cdots, u_{j-1}, u_j, u_{j+1}, \cdots, u_n) = \\ -T(u_1, u_2, \cdots, u_{i-1}, u_j, u_{i+1}, \cdots, u_{j-1}, u_i, u_{j+1}, \cdots, u_n). \end{array}}$$

Si T es una función multilineal alternante, entonces es totalmente antisimétrica. Para probar esta afirmación, primero notemos que

$$T(u_1, u_2, \cdots, u_{i-1}, u+v, u_{i+1}, \cdots, u_{j-1}, u+v, u_{j+1}, \cdots, u_n) = 0,$$

de donde

$$
\begin{aligned}
0 &= T(u_1,u_2,\cdots,u_{i-1},u+v,u_{i+1},\cdots,u_{j-1},u+v,u_{j+1},\cdots,u_n) = \\
&= T(u_1,u_2,\cdots,u_{i-1},u,u_{i+1},\cdots,u_{j-1},u+v,u_{j+1},\cdots,u_n) + \\
&+ T(u_1,u_2,\cdots,u_{i-1},v,u_{i+1},\cdots,u_{j-1},u+v,u_{j+1},\cdots,u_n) = \\
&= T(u_1,u_2,\cdots,u_{i-1},u,u_{i+1},\cdots,u_{j-1},u,u_{j+1},\cdots,u_n) + \\
&+ T(u_1,u_2,\cdots,u_{i-1},u,u_{i+1},\cdots,u_{j-1},v,u_{j+1},\cdots,u_n) + \\
&+ T(u_1,u_2,\cdots,u_{i-1},v,u_{i+1},\cdots,u_{j-1},u,u_{j+1},\cdots,u_n) + \\
&+ T(u_1,v_2,\cdots,u_{i-1},v,u_{i+1},\cdots,u_{j-1},v,u_{j+1},\cdots,u_n) = \\
&= T(u_1,u_2,\cdots,u_{i-1},u,u_{i+1},\cdots,u_{j-1},v,u_{j+1},\cdots,u_n) + \\
&+ T(u_1,u_2,\cdots,u_{i-1},v,u_{i+1},\cdots,u_{j-1},u,u_{j+1},\cdots,u_n),
\end{aligned}
$$

así

$$
\begin{aligned}
0 &= T(u_1,u_2,\cdots,u_{i-1},u,u_{i+1},\cdots,u_{j-1},v,u_{j+1},\cdots,u_n) + \\
&\quad +T(u_1,u_2,\cdots,u_{i-1},v,u_{i+1},\cdots,u_{j-1},u,u_{j+1},\cdots,u_n).
\end{aligned}
$$

que implica

$$
\begin{aligned}
&T(u_1,u_2,\cdots,u_{i-1},u,u_{i+1},\cdots,u_{j-1},v,u_{j+1},\cdots,u_n) = \\
&-T(u_1,u_2,\cdots,u_{i-1},v,u_{i+1},\cdots,u_{j-1},u,u_{j+1},\cdots,u_n).
\end{aligned}
$$

Esto muestra que si T es una función alternate entonces es totalmente antisimétrica.

Ahora, como T es una función multilineal está completamente definida si se define en una base de V. Por lo que si V tiene dimensión n y

$$\{v_1, v_2, \cdots, v_n\}$$

es una de sus bases, entonces los vectores

$$T(v_{i_1}, v_{i_2}, \cdots, v_{i_m}), \quad i_1, i_2, \cdots, i_m = 1, 2, \cdots, n$$

definen complentamente a la función multilineal T.

Pero como T es una función alternante, es diferente de cero en el arreglo de vectores

$$(v_{i_1}, v_{i_2}, \cdots, v_{i_m})$$

solo si se cumple

$$v_{i_k} \neq v_{i_l}, \qquad k,l = 1,2,\cdots,m.$$

Así, de la base de n vectores debemos tomar combinaciones de m elementos sin repetición, el total de estas combinaciones sin repetición son

$$\boxed{D = \frac{n!}{(n-m)!m!}.}$$

Esto nos indica que el espacio de las funciones alternantes (35.8) tiene dimensión D.

Así, si T es una función alternante, los únicos vectores diferentes de cero son

$$\boxed{T(v_{i_1}, v_{i_2}, \cdots, v_{i_m}), \quad i_1 < i_2 < i_3 < \cdots < i_m.}$$

Por lo que si

$$u_1, u_2, \cdots, u_m \in V$$

existen

$$\boxed{\frac{n!}{(n-m)!m!}}$$

constantes $\alpha_{i_1 i_2 \cdots i_m} \in K$ tales que

$$\boxed{T(u_1, u_2, \cdots, u_m) = \alpha_{i_1 i_2 \cdots i_m} T(v_{i_1}, v_{i_2}, \cdots, v_{i_m}), \quad i_1 < i_2 < i_3 < \cdots < i_m.}$$

Ahora, usando la notación del producto exterior, si

$$u_1, u_2, \cdots, u_m \in V$$

se define

$$\boxed{T(u_1, u_2, \cdots, u_m) = u_1 \wedge u_2 \wedge \cdots \wedge u_m.}$$

En particular, para los elementos de la base de V tenemos

$$\boxed{T(v_{i_1}, v_{i_2}, \cdots, v_{i_m}) = v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_m}.}$$

de donde

$$\boxed{u_1 \wedge u_2 \wedge \cdots \wedge u_m = \alpha_{i_1 i_2 \cdots i_m} v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_m}, \quad i_1 < i_2 < i_3 < \cdots < i_m.}$$

35.8. Caso de dimensión 3

Si V tiene dimensión 2, las únicas funciones multilineales alternantes que podemos tener son las bilineales, las cuales ya estudiamos.

Ahora, si la dimensión de V es 3, las únicas funciones multilineales alternantes son de la forma

$$\begin{aligned} T_2 &: V \times V \to W, \\ T_3 &: V \times V \times V \to W. \end{aligned}$$

Podemos obsevar que T_2 es una función bilineal alternante, ya estudiada anteriormente. Por lo que nos concentraremos en las funciones multilineales alternantes de la forma T_3.

Para iniciar supongamos que

$$\{v_1, \quad v_2, \quad v_3\}$$

es una de las bases de V, entonces, si $u, v, w \in V$ existen

$$\lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{21}, \lambda_{22}, \lambda_{23}, \lambda_{31}, \lambda_{32}, \lambda_{33} \in K$$

tal que

$$\boxed{\begin{aligned} u &= \lambda_{11} v_1 + \lambda_{12} v_2 + \lambda_{13} v_3, \\ v &= \lambda_{21} v_1 + \lambda_{22} v_2 + \lambda_{23} v_3, \\ w &= \lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3. \end{aligned}}$$

Además, T_3 se puede expresar en términos de

$$\boxed{v_1 \wedge v_2 \wedge v_3,}$$

en este caso tenemos

$$\begin{aligned}
u \wedge v \wedge w &= (\lambda_{11} v_1 + \lambda_{12} v_2 + \lambda_{13} v_3) \wedge (\lambda_{21} v_1 + \lambda_{22} v_2 + \lambda_{23} v_3) \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) \\
&= \lambda_{11} v_1 \wedge (\lambda_{21} v_1 + \lambda_{22} v_2 + \lambda_{23} v_3) \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{12} v_2 \wedge (\lambda_{21} v_1 + \lambda_{22} v_2 + \lambda_{23} v_3) \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{13} v_3 \wedge (\lambda_{21} v_1 + \lambda_{22} v_2 + \lambda_{23} v_3) \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) \\
&= \lambda_{11} v_1 \wedge (\lambda_{22} v_2 + \lambda_{23} v_3) \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{12} v_2 \wedge (\lambda_{21} v_1 + \lambda_{23} v_3) \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{13} v_3 \wedge (\lambda_{21} v_1 + \lambda_{22} v_2) \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) \\
&= \lambda_{11} v_1 \wedge \lambda_{22} v_2 \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{11} v_1 \wedge \lambda_{23} v_3 \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{12} v_2 \wedge \lambda_{21} v_1 \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{12} v_2 \wedge \lambda_{23} v_3 \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{13} v_3 \wedge \lambda_{21} v_1 \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) + \\
&\quad + \lambda_{13} v_3 \wedge \lambda_{22} v_2 \wedge (\lambda_{31} v_1 + \lambda_{32} v_2 + \lambda_{33} v_3) \\
&= \lambda_{11} \lambda_{22} \lambda_{33} (v_1 \wedge v_2 \wedge v_3) + \\
&\quad + \lambda_{11} \lambda_{23} \lambda_{32} (v_1 \wedge v_3 \wedge v_2) + \\
&\quad + \lambda_{12} \lambda_{21} \lambda_{33} (v_2 \wedge v_1 \wedge v_3) + \\
&\quad + \lambda_{12} \lambda_{23} \lambda_{31} (v_2 \wedge v_3 \wedge v_1) + \\
&\quad + \lambda_{13} \lambda_{21} \lambda_{32} (v_3 \wedge v_1 \wedge v_2) + \\
&\quad + \lambda_{13} \lambda_{22} \lambda_{31} (v_3 \wedge v_2 \wedge v_1),
\end{aligned}$$

esto es,

$$\begin{aligned}
u \wedge v \wedge w &= \lambda_{11} \lambda_{22} \lambda_{33} (v_1 \wedge v_2 \wedge v_3) + \lambda_{11} \lambda_{23} \lambda_{32} (v_1 \wedge v_3 \wedge v_2) + \\
&\quad + \lambda_{12} \lambda_{21} \lambda_{33} (v_2 \wedge v_1 \wedge v_3) + \lambda_{12} \lambda_{23} \lambda_{31} (v_2 \wedge v_3 \wedge v_1) + \\
&\quad + \lambda_{13} \lambda_{21} \lambda_{32} (v_3 \wedge v_1 \wedge v_2) + \lambda_{13} \lambda_{22} \lambda_{31} (v_3 \wedge v_2 \wedge v_1).
\end{aligned}$$

Ahora, considerando que el producto exterior es completamente antisimétrico tenemos

$$\begin{aligned}
u \wedge v \wedge w &= \lambda_{11} \lambda_{22} \lambda_{33} (v_1 \wedge v_2 \wedge v_3) - \lambda_{11} \lambda_{23} \lambda_{32} (v_1 \wedge v_2 \wedge v_3) + \\
&\quad - \lambda_{12} \lambda_{21} \lambda_{33} (v_1 \wedge v_2 \wedge v_3) + \lambda_{12} \lambda_{23} \lambda_{31} (v_1 \wedge v_2 \wedge v_3) + \\
&\quad + \lambda_{13} \lambda_{21} \lambda_{32} (v_1 \wedge v_2 \wedge v_3) - \lambda_{13} \lambda_{22} \lambda_{31} (v_1 \wedge v_2 \wedge v_3),
\end{aligned}$$

es decir ,

$$u \wedge v \wedge w = \Big(\lambda_{11}\lambda_{22}\lambda_{33} - \lambda_{11}\lambda_{23}\lambda_{32} - \lambda_{12}\lambda_{21}\lambda_{33} + \lambda_{12}\lambda_{23}\lambda_{31} + +\lambda_{13}\lambda_{21}\lambda_{32} - \lambda_{13}\lambda_{22}\lambda_{31}\Big)(v_1 \wedge v_2 \wedge v_3)\,. \tag{35.9}$$

Además, reordenando términos tenemos

$$u \wedge v \wedge w = \Big(\lambda_{11}\lambda_{22}\lambda_{33} - \lambda_{11}\lambda_{32}\lambda_{23} - \lambda_{21}\lambda_{12}\lambda_{33} + \lambda_{21}\lambda_{13}\lambda_{32} +\lambda_{31}\lambda_{12}\lambda_{23} - \lambda_{31}\lambda_{13}\lambda_{22}\Big)(v_1 \wedge v_2 \wedge v_3)\,.$$

Por lo que, usando el tensor de Levi-Civita en tres dimesiones

$$\epsilon_{i_1 i_2 i_3}, \qquad i_1, i_2, i_3 = 1, 2, 3$$

se obtiene

$$u \wedge v \wedge w = \Big(\epsilon_{1 i_2 i_3}\lambda_{11}\lambda_{i_2 2}\lambda_{i_3 3} + \epsilon_{2 i_2 i_3}\lambda_{21}\lambda_{i_2 2}\lambda_{i_3 3} + \epsilon_{3 i_2 i_3}\lambda_{31}\lambda_{i_2 2}\lambda_{i_3 3}\Big)(v_1 \wedge v_2 \wedge v_3)\,.$$

de donde

$$\boxed{u \wedge v \wedge w = \epsilon_{i_1 i_2 i_3}\lambda_{i_1 1}\lambda_{i_2 2}\lambda_{i_3 3}\,(v_1 \wedge v_2 \wedge v_3)\,, \quad i_1, i_2, i_3 = 1, 2, 3.}$$

que se puede escribir como

$$\boxed{\begin{array}{l} u_1 \wedge u_2 \wedge u_3 = \epsilon_{i_1 i_2 i_3}\lambda_{i_1 j_1}\lambda_{i_2 j_2}\lambda_{i_3 j_3}\,(v_{j_1} \wedge v_{j_2} \wedge v_{j_3})\,, \\ i_1, i_2, i_3, j_1, j_2, j_3 = 1, 2, 3, \qquad j_1 < j_2 < j_3. \end{array}}$$

Adicionalmente, se puede observar que definiendo la matriz

$$\boxed{M = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix},}$$

se obtiene

$$\boxed{\det M = \lambda_{11}\left(\lambda_{22}\lambda_{33} - \lambda_{23}\lambda_{32}\right) - \lambda_{12}\left(\lambda_{21}\lambda_{33} - \lambda_{23}\lambda_{31}\right) + \lambda_{13}\left(\lambda_{21}\lambda_{32} - \lambda_{22}\lambda_{31}\right),}$$

que coincide con la expresión (35.9). Por lo tanto, se encuentra

$$\boxed{u \wedge v \wedge w = (\det M)\left(v_1 \wedge v_2 \wedge v_3\right).}$$

También se puede observar que si definimos los vectores

$$\boxed{\vec{\lambda} = (\lambda_1, \lambda_2, \lambda_3), \quad \vec{\gamma} = (\gamma_1, \gamma_2, \gamma_3), \quad \vec{\zeta} = (\zeta_1, \zeta_2, \zeta_3),}$$

usando el llamado triple producto vectorial se encuentra

$$\boxed{u \wedge v \wedge w = \vec{\lambda} \cdot \left(\vec{\gamma} \times \vec{\zeta}\right)\left(v_1 \wedge v_2 \wedge v_3\right).}$$

35.9. Caso de dimensión n

Ahora supongamos que V es un espacio vectorial sobre K con

$$\dim(V) = n.$$

En este caso, podemos construir funciones alternantes de las siguienes $n-1$ formas diferentes

$$\begin{aligned} &T_2 : V \times V \to W, \\ &T_3 : V \times V \times V \to W. \\ &\vdots \\ &T_n : V^n \to W. \end{aligned}$$

Además, supongamos que

$$v_1, v_2, \cdots, v_n$$

es una base del espacio vectorial V y que

$$T_m : V^m \to W$$

es una función alternante con

$$m \leq n.$$

Entonces, si

$$u_1, u_2, \cdots, u_m \in V$$

existen nm escalares λ_{ij} en K tales que

$$\begin{aligned}
u_1 &= \lambda_{11}v_1 + \lambda_{12}v_2 + \cdots + \lambda_{1n}v_n, \\
&\vdots \\
u_m &= \lambda_{m1}v_1 + \lambda_{m2}v_2 + \cdots + \lambda_{mn}v_n,
\end{aligned}$$

En este caso el producto exterior se puede expresar como

$$\boxed{\begin{array}{l}
u_1 \wedge u_2 \wedge \cdots \wedge u_m = \epsilon_{i_1 i_2 \cdots i_m} \lambda_{i_1 j_1} \lambda_{i_2 j_2} \cdots \lambda_{i_m j_m} v_{j_1} \wedge v_{j_2} \cdots \wedge v_{j_m}, \\
i_1, i_2, \cdots, i_m = 1, 2, \cdots, m, \qquad j_1, j_2, \cdots, j_m = 1, 2, \cdots, n, \\
j_1 < j_2 < \cdots < j_m.
\end{array}}$$

En particular, si $n = m$ tenemos

$$\begin{aligned}
u_1 &= \lambda_{11}v_1 + \lambda_{12}v_2 + \cdots + \lambda_{1n}v_n, \\
&\vdots \\
u_n &= \lambda_{n1}v_1 + \lambda_{n2}v_2 + \cdots + \lambda_{nn}v_n,
\end{aligned}$$

por lo que se puede definir la matriz de $n \times n$

$$\boxed{M = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & \cdots & \vdots \\ \lambda_{n1} & \cdots & \lambda_{nn} \end{pmatrix},}$$

y se obtiene

$$\boxed{u_1 \wedge u_2 \wedge \cdots \wedge u_n = (\det M)\, v_1 \wedge v_2 \wedge \cdots \wedge v_n.}$$

35.10. Derivada exterior

Consideremos la siguientes funciones

$$
\begin{aligned}
&F : \mathbb{R}^n \to \mathbb{R}. \\
&F_a : \mathbb{R}^n \to \mathbb{R}, \\
&F_{ab} : \mathbb{R}^n \to \mathbb{R}, \quad a, b = 1, 2, \cdots, n.
\end{aligned}
$$

A F la definiremos como una $0-$Forma y su derivada exterior la definiremos como

$$\boxed{dF = \frac{\partial F}{\partial x^a} dx^a.}$$

También definiremos una $1-$Forma como

$$\boxed{f = F_a dx^a.}$$

y su derivada exterior la definiremos como

$$\boxed{df = \frac{\partial F_a}{\partial x^b} dx^b \wedge dx^a.}$$

Además, definiremos una $2-$Forma como

$$\boxed{g = F_{ab} dx^a \wedge dx^b.}$$

y su derivada exterior la definiremos como

$$\boxed{dg = \frac{\partial F_{ab}}{\partial x^c} dx^c \wedge dx^a \wedge dx^b.}$$

Siguiendo la misma lógica se puede definir una n-forma y su derivada exterior.

35.11. Caso $n=3$

En este caso una 0-forma es una función como

$$F:\mathbb{R}^3\to\mathbb{R}.$$

y su derivada exterior es

$$\boxed{dF=\frac{\partial F}{\partial x^1}dx^1+\frac{\partial F}{\partial x^2}dx^2+\frac{\partial F}{\partial x^3}dx^3}$$

Ahora, del cálculo de varias variables sabemos que se cumple

$$\vec{\nabla}F=\left(\frac{\partial F}{\partial x^1},\frac{\partial F}{\partial x^2},\frac{\partial F}{\partial x^3}\right),$$

de donde

$$\vec{\nabla}F\cdot d\vec{x}=\frac{\partial F}{\partial x^1}dx^1+\frac{\partial F}{\partial x^2}dx^2+\frac{\partial F}{\partial x^3}dx^3,$$

lo por lo que

$$\boxed{dF=\vec{\nabla}F\cdot d\vec{x}.}$$

Ahora, si tenemos las funciones

$$\begin{aligned}F_1&:\mathbb{R}^3\to\mathbb{R},\\ F_2&:\mathbb{R}^3\to\mathbb{R},\\ F_3&:\mathbb{R}^3\to\mathbb{R},\end{aligned}$$

una $1-$Forma se puede escribir como

$$\boxed{\omega=F_1dx^1+F_2dx^2+F_3dx^3.}$$

Por lo que su derivada exterior es

$$
\begin{aligned}
d\omega &= d(F_1 dx^1) + d(F_2 dx^2) + d(F_3 dx^3) \\
&= \left(\frac{\partial F_1}{\partial x^1} dx^1 + \frac{\partial F_1}{\partial x^2} dx^2 + \frac{\partial F_1}{\partial x^3} dx^3\right) \wedge dx^1 \\
&\quad + \left(\frac{\partial F_2}{\partial x^1} dx^1 + \frac{\partial F_2}{\partial x^2} dx^2 + \frac{\partial F_2}{\partial x^3} dx^3\right) \wedge dx^2 \\
&\quad + \left(\frac{\partial F_3}{\partial x^1} dx^1 + \frac{\partial F_3}{\partial x^2} dx^2 + \frac{\partial F_3}{\partial x^3} dx^3\right) \wedge dx^3 \\
&= \frac{\partial F_1}{\partial x^1} dx^1 \wedge dx^1 + \frac{\partial F_1}{\partial x^2} dx^2 \wedge dx^1 + \frac{\partial F_1}{\partial x^3} dx^3 \wedge dx^1 \\
&\quad + \frac{\partial F_2}{\partial x^1} dx^1 \wedge dx^2 + \frac{\partial F_2}{\partial x^2} dx^2 \wedge dx^2 + \frac{\partial F_2}{\partial x^3} dx^3 \wedge dx^2 \\
&\quad + \frac{\partial F_3}{\partial x^1} dx^1 \wedge dx^3 + \frac{\partial F_3}{\partial x^2} dx^2 \wedge dx^3 + \frac{\partial F_3}{\partial x^3} dx^3 \wedge dx^3 \\
&= \left(\frac{\partial F_2}{\partial x^1} - \frac{\partial F_1}{\partial x^2}\right) dx^1 \wedge dx^2 + \left(\frac{\partial F_3}{\partial x^2} - \frac{\partial F_2}{\partial x^3}\right) dx^2 \wedge dx^3 + \left(\frac{\partial F_1}{\partial x^3} - \frac{\partial F_3}{\partial x^1}\right) dx^3 \wedge dx^1,
\end{aligned}
$$

es decir

$$
\boxed{d\omega = \left(\frac{\partial F_2}{\partial x^1} - \frac{\partial F_1}{\partial x^2}\right) dx^1 \wedge dx^2 + \left(\frac{\partial F_3}{\partial x^2} - \frac{\partial F_2}{\partial x^3}\right) dx^2 \wedge dx^3 + \left(\frac{\partial F_1}{\partial x^3} - \frac{\partial F_3}{\partial x^1}\right) dx^3 \wedge dx^1.}
$$

Ahora, del cálculo de varias variables sabemos que para la función vectorial

$$
\boxed{\vec{\omega} = (F_1, F_2, F_3)}
$$

definiendo

$$
\hat{i} = (1, 0, 0), \quad \hat{j} = (0, 1, 0), \quad \hat{k} = (0, 0, 1)
$$

el rotacional se define como

$$
\begin{aligned}
\vec{\nabla} \times \vec{\omega} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1 & F_2 & F_3 \end{vmatrix} \\
&= \left(\frac{\partial F_3}{\partial x^2} - \frac{\partial F_2}{\partial x^3}\right) \hat{i} + \left(\frac{\partial F_1}{\partial x^3} - \frac{\partial F_3}{\partial x^1}\right) \hat{j} + \left(\frac{\partial F_2}{\partial x^1} - \frac{\partial F_1}{\partial x^2}\right) \hat{k},
\end{aligned}
$$

esto es

$$\boxed{\vec{\nabla}\times\vec{\omega} = \left(\frac{\partial F_3}{\partial x^2} - \frac{\partial F_2}{\partial x^3}\right)\hat{i} + \left(\frac{\partial F_1}{\partial x^3} - \frac{\partial F_3}{\partial x^1}\right)\hat{j} + \left(\frac{\partial F_2}{\partial x^1} - \frac{\partial F_1}{\partial x^2}\right)\hat{k},}$$

de donde podemos definir

$$\boxed{\begin{aligned} \left(\vec{\nabla}\times\vec{\omega}\right)_1 &= \left(\frac{\partial F_3}{\partial x^2} - \frac{\partial F_2}{\partial x^3}\right), \\ \left(\vec{\nabla}\times\vec{\omega}\right)_2 &= \left(\frac{\partial F_1}{\partial x^3} - \frac{\partial F_3}{\partial x^1}\right), \\ \left(\vec{\nabla}\times\vec{\omega}\right)_3 &= \left(\frac{\partial F_2}{\partial x^1} - \frac{\partial F_1}{\partial x^2}\right). \end{aligned}}$$

Por lo que la derivada exterior de una $1-$Forma se puede escribir como

$$\boxed{d\omega = \left(\vec{\nabla}\times\vec{\omega}\right)_1 dx^2\wedge dx^3 + \left(\vec{\nabla}\times\vec{\omega}\right)_2 dx^3\wedge dx^1 + \left(\vec{\nabla}\times\vec{\omega}\right)_3 dx^1\wedge dx^2.}$$

Como podemos ver la derivada exterior de una $1-$Forma define el rotacional.

Ahora, para la $2-$Forma

$$\boxed{\Omega = F_1 dx^2\wedge dx^3 + F_2 dx^3\wedge dx^1 + F_3 dx^1\wedge dx^2,}$$

tenemos la derivada exterior

$$
\begin{aligned}
d\Omega &= d\left(F_1 dx^2 \wedge dx^3 + F_2 dx^3 \wedge dx^1 + F_3 dx^1 \wedge dx^2\right) \\
&= d\left(F_1 dx^2 \wedge dx^3\right) + d\left(F_2 dx^3 \wedge dx^1\right) + d\left(F_3 dx^1 \wedge dx^2\right) \\
&= \left(\frac{\partial F_1}{\partial x^1} dx^1 + \frac{\partial F_1}{\partial x^2} dx^2 + \frac{\partial F_1}{\partial x^3} dx^3\right) \wedge dx^2 \wedge dx^3 + \\
&\quad + \left(\frac{\partial F_2}{\partial x^1} dx^1 + \frac{\partial F_2}{\partial x^2} dx^2 + \frac{\partial F_2}{\partial x^3} dx^3\right) \wedge dx^3 \wedge dx^1 \\
&\quad + \left(\frac{\partial F_3}{\partial x^1} dx^1 + \frac{\partial F_3}{\partial x^2} dx^2 + \frac{\partial F_3}{\partial x^3} dx^3\right) \wedge dx^1 \wedge dx^2 \\
&= \frac{\partial F_1}{\partial x^1} dx^1 \wedge dx^2 \wedge dx^3 + \frac{\partial F_2}{\partial x^2} dx^2 \wedge dx^3 \wedge dx^1 + \frac{\partial F_3}{\partial x^3} dx^3 \wedge dx^1 \wedge dx^2 \\
&= \left(\frac{\partial F_1}{\partial x^1} + \frac{\partial F_2}{\partial x^2} + \frac{\partial F_3}{\partial x^3}\right) dx^1 \wedge dx^2 \wedge dx^3,
\end{aligned}
$$

de donde

$$\boxed{d\Omega = \left(\frac{\partial F_1}{\partial x^1} + \frac{\partial F_2}{\partial x^2} + \frac{\partial F_3}{\partial x^3}\right) dx^1 \wedge dx^2 \wedge dx^3.}$$

Ahora, si definimos

$$\boxed{\vec{\Omega} = (F_1, F_2, F_3)\,,}$$

tenemos la divergencia

$$\boxed{\vec{\nabla} \cdot \vec{\Omega} = \frac{\partial F_1}{\partial x^1} + \frac{\partial F_2}{\partial x^2} + \frac{\partial F_3}{\partial x^3}.}$$

por lo cual

$$\boxed{d\Omega = \left(\vec{\nabla} \cdot \vec{\Omega}\right) dx^1 \wedge dx^2 \wedge dx^3.}$$

Así, la derivada exterior de una $2-$Forma permite definir la divergencia.

Capítulo 36

Producto tensorial

En este capítulo estudiaremos el producto tensorial el cual nos permite entender la estructura general de las funciones multilineales.

El producto tensorial tiene aplicaciones en diversas áreas como la computación, la mecánica cuántica, la relatividad general, etc.

36.1. Producto tensorial de dos espacios vectoriales

Supongamos que V_1, V_1 y U son espacios vectoriales sobre K y la función

$$F : V_1 \times V_2 \to U$$

es bilineal. Se dice que el par (F, U) es un producto tensorial de V_1 y V_2 si para cualquier espacio vectorial W sobre K y para cualquier función bilineal

$$T : V_1 \times V_2 \to W,$$

existe una única función lineal

$$H : U \to W,$$

tal que

$$T = H \circ F$$

Se puede observar que si existe el producto tensorial (F, U), entonces cualquier función bilineal de $V_1 \times V_2$ sobre cualquier espacio vectorial W se puede escribir en términos de F.

Veamos el caso de espacios vectoriales de dimensión finita. Supongamos que V_1, V_2, W tiene dimensión finita tal que

$$\begin{aligned} \dim(V_1) &= n_1, \\ \dim(V_2) &= n_2. \end{aligned}$$

y que

$$\{v_{11}, v_{21}, \cdots, v_{n1}\}$$

es base de V_1 y

$$\{v_{12}, v_{22}, \cdots, v_{m2}\}$$

es base de V_2. También supongamos que la función

$$T : V_1 \times V_2 \to W, \tag{36.1}$$

es bilineal. Debido a que T es una función bilineal existen nm vectores de W tales que

$$T(v_{i1}, v_{j2}) = w_{ij} \in W, \quad i = 1, 2, \cdots n, j = 1, 2, \cdots, m. \tag{36.2}$$

Debido a que cualquier función bilineal queda determinada por la forma en que actúa en una base de su dominio, la expresión (36.2) determina completamente a T.

Ahora, para U tomaremos un espacio vectorial sobre el campo K de dimensión

$$nm.$$

En U tomaremos una de sus bases y la enumeraremos de la forma

$$e_{ij}, \qquad i = 1, \cdots, m, \quad j = 1, \cdots, n.$$

Así, podemos definir la función bilineal F

$$F : V_1 \times V_2 \to U, \tag{36.3}$$

tal que

$$F(v_{i1}, v_{j2}) = e_{ij}. \tag{36.4}$$

Además, considerando que cualquier función lineal está definida por la forma en que actúa en la base del dominio, existe una única función lineal

$$H : U \to W,$$

tal que

$$H(e_{ij}) = w_{ij} = T(v_i, v_j).$$

Es posible probar que se cumple la igualdad

$$T = H \circ F. \tag{36.5}$$

En efecto podemos observar que

$$(H \circ F)(v_{i1}, v_{j2}) = H(F(v_{i1}, v_{j2})) = H(e_{ij}) = w_{ij} = T(v_{i1}, v_{j1}).$$

Por lo tanto, para cualquier par de elementos (v_{i1}, v_{j2}) de la base del dominio se cumple la igualdad

$$(H \circ F)(v_{i1}, v_{j2}) = G(v_{i1}, v_{j2}),$$

lo que implica que cualquier par de vectores

$$(v_1, v_2) \in V_1 \times V_2$$

se cumple la ecuación (36.5).

En consecuencia, cualquier función bilineal de $V_1 \times V_2$ se puede escribir en términos del espacio vectorial U y la función F definida en (36.3)-(36.4). Es decir, la pareja

$$(U, F),$$

es un producto tensorial de V_1 y V_2.

La existencia del producto tesorial es muy útil porque nos dice que todas las funciones bilineales se relacionan con la función universal F y basta estudiar esta. Se suele denotar a U como

$$\boxed{U = V_1 \otimes_K V_2,}$$

también se suele usar la notación

$$F(v_{i1}, v_{j2}) = v_{i1} \otimes v_{j2}.$$

$$\boxed{F(v_{i1}, v_{j2}) = v_{i1} \otimes v_{j2}.}$$

Se puede observar debido a que F es una función bilineal, si $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$ y $\lambda, \gamma \in K$ se cumple

$$\boxed{(\lambda v_1 + \gamma v_1') \otimes v_2 = \lambda v_1 \otimes v_2 + \gamma v_1' \otimes v_2.}$$

De la misma forma, considerando que F es una función bilineal, se puede probar que se cumple

$$\boxed{\begin{aligned} & v_1 \otimes (\lambda v_2 + \gamma v_2') = \lambda v_1 \otimes v_2 + \gamma v_1 \otimes v_2', \\ & \lambda(v_1 \otimes v_2) = (\lambda v_1) \otimes v_2 = v_1 \otimes (\lambda v_2), \\ & -v_1 \otimes v_2 = v_1 \otimes (-v_2), \\ & 0 \otimes v_2 = v_1 \otimes 0 = 0. \end{aligned}}$$

36.2. Caso de $\mathbb{R}^2$

Como un ejemplo, veamos el caso en que $V_1 = V_2 = \mathbb{R}^2$. Estos dos espacios vectoriales tienen como base canónica

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Como T podemos tomar $\mathbb{R}^4$, cuya base canónica es

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, E_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Debido a que $V_1 = V_2 = \mathbb{R}^2$ y $U = \mathbb{R}^4$ la función F debe ser una función bilineal que relacione la base de $\mathbb{R}^2 \times \mathbb{R}^2$ con la base $\mathbb{R}^4$. Una forma de relacionar estas bases es con la regla

$$
\begin{aligned}
\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} &= F\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\begin{pmatrix}1\\0\end{pmatrix}\\0\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix},\\
\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} &= F\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}1\begin{pmatrix}0\\1\end{pmatrix}\\0\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\\
\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} &= F\left(\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}0\begin{pmatrix}1\\0\end{pmatrix}\\1\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} = \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\\
\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} &= F\left(\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}0\begin{pmatrix}0\\1\end{pmatrix}\\1\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.
\end{aligned}
$$

Así, para dos vectores de $\mathbb{R}^2$ tenemos

$$
\begin{aligned}
\begin{pmatrix}a\\b\end{pmatrix}\otimes\begin{pmatrix}c\\d\end{pmatrix} &= \left(a\begin{pmatrix}1\\0\end{pmatrix}+b\begin{pmatrix}0\\1\end{pmatrix}\right)\otimes\left(c\begin{pmatrix}1\\0\end{pmatrix}+d\begin{pmatrix}0\\1\end{pmatrix}\right)\\
&= ac\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}+ad\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}\\
&\quad +bc\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}+bd\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}\\
&= ac\begin{pmatrix}1\\0\\0\\0\end{pmatrix}+ad\begin{pmatrix}0\\1\\0\\0\end{pmatrix}+bc\begin{pmatrix}0\\0\\1\\0\end{pmatrix}+bd\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\\
&= \begin{pmatrix}ac\\ad\\bc\\bd\end{pmatrix}
\end{aligned}
$$

esto es

$$\boxed{\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \begin{pmatrix} c \\ d \end{pmatrix} \\ b \begin{pmatrix} c \\ d \end{pmatrix} \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bd \\ bd \end{pmatrix},}$$

De forma análoga, para un vector de $\mathbb{R}^m$

$$u = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

y un vector de $\mathbb{R}^n$

$$v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

se puede definir un producto vectorial de la forma

$$\boxed{u \otimes v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{m-1} \\ a_m \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 v \\ a_2 v \\ \vdots \\ a_{m-1} v \\ a_m v \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ \vdots \\ a_m b_{n-1} \\ a_m b_n \end{pmatrix},}$$

note que el vector final pertenece al espacio vectorial R^{nm}.

36.2.1. Producto de Kronecker

Otra opción de producto tensorial para $\mathbb{R}^2 \times \mathbb{R}^2$ es tomar como U a $M_{2\times 2}(\mathbb{R})$, que tiene dimensión 4. Si para $\mathbb{R}^2$ tomamos la base canónica

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

y para $M_{2\times 2}(\mathbb{R})$ tomamos la base

$$\begin{aligned}
e_1 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\
e_2 &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\
e_3 &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\
e_4 &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},
\end{aligned}$$

la función bilineal F debe asociar estas dos bases. En este caso podemos tomar

$$\begin{aligned}
\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= F\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) \\
&= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \\
&= \begin{pmatrix} 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \\ 0 \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \end{pmatrix} \\
&= \begin{pmatrix} 1 \begin{pmatrix} 1 & 0 \end{pmatrix} \\ 0 \begin{pmatrix} 1 & 0 \end{pmatrix} \end{pmatrix} \\
&= \begin{pmatrix} 1 \cdot 1 & 1 \cdot 0 \\ 0 \cdot 1 & 0 \cdot 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},
\end{aligned}$$

que se suele escribir como

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

además

$$\begin{aligned}
\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} &= F\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)\\
&= \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}^T\\
&= \begin{pmatrix}1\begin{pmatrix}0\\1\end{pmatrix}^T\\ 0\begin{pmatrix}0\\1\end{pmatrix}^T\end{pmatrix}\\
&= \begin{pmatrix}1\begin{pmatrix}0 & 1\end{pmatrix}\\ 0\begin{pmatrix}0 & 1\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}1\cdot 0 & 1\cdot 1\\ 0\cdot 0 & 0\cdot 1\end{pmatrix}\\
&= \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},
\end{aligned}$$

que se puede expresar como

$$\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.$$

También tenemos el caso

$$\begin{aligned}
\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} &= F\left(\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}^T\\
&= \begin{pmatrix}0\begin{pmatrix}1\\0\end{pmatrix}^T\\ 1\begin{pmatrix}1\\0\end{pmatrix}^T\end{pmatrix}\\
&= \begin{pmatrix}0\begin{pmatrix}1 & 0\end{pmatrix}\\ 1\begin{pmatrix}1 & 0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0\cdot 1 & 0\cdot 0\\ 1\cdot 1 & 1\cdot 0\end{pmatrix}\\
&= \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},
\end{aligned}$$

esto es

$$\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix}=\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}.$$

El último caso es

$$\begin{aligned}\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} &= F\left(\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)\\ &= \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}^T\\ &= \begin{pmatrix}0\begin{pmatrix}0\\1\end{pmatrix}^T\\ 1\begin{pmatrix}0\\1\end{pmatrix}^T\end{pmatrix}\\ &= \begin{pmatrix}0\begin{pmatrix}0 & 1\end{pmatrix}\\ 1\begin{pmatrix}0 & 1\end{pmatrix}\end{pmatrix}\\ &= \begin{pmatrix}0\cdot 1 & 0\cdot 0\\ 1\cdot 0 & 1\cdot 1\end{pmatrix}\\ &= \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix},\end{aligned}$$

que se suele escribir como

$$\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix}=\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}.$$

Así tenemos

$$\begin{aligned}\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} &= \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix}=\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix},\\ \begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} &= \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix}=\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix},\\ \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} &= \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix}=\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix},\\ \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} &= \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix}=\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}.\end{aligned}$$

Considerando que F debe ser una función bilineal, se puede mostrar que para cualquier par de vectores de $\mathbb{R}^2$ se tiene

$$\boxed{\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} c & d \end{pmatrix} = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}.} \tag{36.6}$$

Por ejemplo,

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} \otimes \begin{pmatrix} -1 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \begin{pmatrix} -1 & 5 \end{pmatrix} = \begin{pmatrix} -2 & 10 \\ -3 & 15 \end{pmatrix} \tag{36.7}$$

La regla (36.6) se puede generalizar para obtener un producto tensorial entre un vector de $\mathbb{C}^m$ y un vector de $\mathbb{C}^n$ de la siguiente forma

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} &= \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}^{*T} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \begin{pmatrix} b_1^* & b_2^* \cdots & b_n^* \end{pmatrix} \\ &= \begin{pmatrix} a_1 b_1^* & a_1 b_2^* & \cdots & a_1 b_n^* \\ a_2 b_1^* & a_2 b_2^* & \cdots & a_2 b_n^* \\ \vdots & \vdots & \cdots & \vdots \\ a_n b_1^* & a_n b_2^* & \cdots & a_m b_n^* \end{pmatrix}, \end{aligned}$$

es decir,

$$\boxed{\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1^* & a_1 b_2^* & \cdots & a_1 b_n^* \\ a_2 b_1^* & a_2 b_2^* & \cdots & a_2 b_n^* \\ \vdots & \vdots & \cdots & \vdots \\ a_n b_1^* & a_n b_2^* & \cdots & a_m b_n^* \end{pmatrix},}$$

a este producto tensorial se suele llamar producto de Kronecker.

36.2.2. Matrices

Como otro ejemplo, consideremos el espacio de las matrices reales de 2×2, es decir $V_1 = V_2 = M_{2\times 2}(\mathbb{R})$. En este caso podemos tomar la base

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \; e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Se puede observar que el espacio vectorial $M_{2\times 2}(\mathbb{R}) \times M_{2\times 2}(\mathbb{R})$ tiene dimensión 16. Entonces, el espacio U debe tener dimensión 16. Por ejemplo, podemos tomar como U al espacio vectorial $M_{4\times 4}(\mathbb{R})$, cuya base canónica es

$$
\begin{aligned}
M_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
M_4 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
M_7 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_8 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_9 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
M_{10} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_{11} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M_{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
M_{13} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, M_{14} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, M_{15} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\
M_{16} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\end{aligned}
$$

La función F debe relacionar un par de elementos de la base de $M_{2\times 2}(\mathbb{R})$ con un elemento de la base de $M_{4\times 4}(\mathbb{R})$. Una forma de definir a F es con la regla

$$\begin{aligned}
\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \otimes \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} &= F\left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right)\\
&= \begin{pmatrix} 1\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} & 0\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\\ 0\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} & 0\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = M_1,
\end{aligned}$$

esto es,

$$\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \otimes \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = M_1.$$

Tambiém tenemos

$$\begin{aligned}
\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \otimes \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} &= F\left(\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\right)\\
&= \begin{pmatrix} 1\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} & 0\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\\ 0\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} & 0\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = M_2,
\end{aligned}$$

es decir,

$$\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \otimes \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = M_2.$$

Adicionalmente tenemos

$$
\begin{aligned}
\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} &= F\left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\right)\\
&= \begin{pmatrix}1\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} & 0\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\\ 0\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} & 0\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = M_5,\\
\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} &= F\left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right)\\
&= \begin{pmatrix}1\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} & 0\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\\ 0\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} & 0\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = M_6,\\
\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\otimes\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} &= F\left(\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} & 1\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\\ 0\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} & 0\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = M_3,
\end{aligned}
$$

$$
\begin{aligned}
\begin{pmatrix}0&1\\0&0\end{pmatrix}\otimes\begin{pmatrix}0&1\\0&0\end{pmatrix} &= F\left(\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&1\\0&0\end{pmatrix} & 1\begin{pmatrix}0&1\\0&0\end{pmatrix}\\ 0\begin{pmatrix}0&1\\0&0\end{pmatrix} & 0\begin{pmatrix}0&1\\0&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix} = M_4,\\
\begin{pmatrix}0&1\\0&0\end{pmatrix}\otimes\begin{pmatrix}0&0\\1&0\end{pmatrix} &= F\left(\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&0\\1&0\end{pmatrix} & 1\begin{pmatrix}0&0\\1&0\end{pmatrix}\\ 0\begin{pmatrix}0&0\\1&0\end{pmatrix} & 0\begin{pmatrix}0&0\\1&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{pmatrix} = M_7,\\
\begin{pmatrix}0&1\\0&0\end{pmatrix}\otimes\begin{pmatrix}0&0\\0&1\end{pmatrix} &= F\left(\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&0\\0&1\end{pmatrix} & 1\begin{pmatrix}0&0\\0&1\end{pmatrix}\\ 0\begin{pmatrix}0&0\\0&1\end{pmatrix} & 0\begin{pmatrix}0&0\\0&1\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix} = M_8,
\end{aligned}
$$

$$
\begin{aligned}
\begin{pmatrix}0&0\\1&0\end{pmatrix}\otimes\begin{pmatrix}1&0\\0&0\end{pmatrix} &= F\left(\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}1&0\\0&0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}1&0\\0&0\end{pmatrix} & 0\begin{pmatrix}1&0\\0&0\end{pmatrix}\\ 1\begin{pmatrix}1&0\\0&0\end{pmatrix} & 0\begin{pmatrix}1&0\\0&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\1&0&0&0\\0&0&0&0\end{pmatrix}=M_9,\\
\begin{pmatrix}0&0\\1&0\end{pmatrix}\otimes\begin{pmatrix}0&1\\0&0\end{pmatrix} &= F\left(\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&1\\0&0\end{pmatrix} & 0\begin{pmatrix}0&1\\0&0\end{pmatrix}\\ 1\begin{pmatrix}0&1\\0&0\end{pmatrix} & 0\begin{pmatrix}0&1\\0&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}=M_{10},\\
\begin{pmatrix}0&0\\1&0\end{pmatrix}\otimes\begin{pmatrix}0&0\\1&0\end{pmatrix} &= F\left(\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&0\\1&0\end{pmatrix} & 0\begin{pmatrix}0&0\\1&0\end{pmatrix}\\ 1\begin{pmatrix}0&0\\1&0\end{pmatrix} & 0\begin{pmatrix}0&0\\1&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\1&0&0&0\end{pmatrix}=M_{13},
\end{aligned}
$$

$$
\begin{aligned}
\begin{pmatrix}0&0\\1&0\end{pmatrix}\otimes\begin{pmatrix}0&0\\0&1\end{pmatrix} &= F\left(\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&0\\0&1\end{pmatrix} & 0\begin{pmatrix}0&0\\0&1\end{pmatrix}\\ 1\begin{pmatrix}0&0\\0&1\end{pmatrix} & 0\begin{pmatrix}0&0\\0&1\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&1&0&0\end{pmatrix} = M_{14},\\
\begin{pmatrix}0&0\\0&1\end{pmatrix}\otimes\begin{pmatrix}1&0\\0&0\end{pmatrix} &= F\left(\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}1&0\\0&0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}1&0\\0&0\end{pmatrix} & 0\begin{pmatrix}1&0\\0&0\end{pmatrix}\\ 0\begin{pmatrix}1&0\\0&0\end{pmatrix} & 1\begin{pmatrix}1&0\\0&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0\end{pmatrix} = M_{11}, \begin{pmatrix}0&0\\0&1\end{pmatrix}\otimes\begin{pmatrix}0&1\\0&0\end{pmatrix} \quad =\\
F\left(\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}\right)&\\
&= \begin{pmatrix}0\begin{pmatrix}0&1\\0&0\end{pmatrix} & 0\begin{pmatrix}0&1\\0&0\end{pmatrix}\\ 0\begin{pmatrix}0&1\\0&0\end{pmatrix} & 1\begin{pmatrix}0&1\\0&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix} = M_{12},
\end{aligned}
$$

$$
\begin{aligned}
\begin{pmatrix}0&0\\0&1\end{pmatrix}\otimes\begin{pmatrix}0&0\\1&0\end{pmatrix} &= F\left(\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&0\\1&0\end{pmatrix} & 0\begin{pmatrix}0&0\\1&0\end{pmatrix}\\ 0\begin{pmatrix}0&0\\1&0\end{pmatrix} & 1\begin{pmatrix}0&0\\1&0\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&1&0\end{pmatrix} = M_{15},\\
\begin{pmatrix}0&0\\0&1\end{pmatrix}\otimes\begin{pmatrix}0&0\\0&1\end{pmatrix} &= F\left(\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right)\\
&= \begin{pmatrix}0\begin{pmatrix}0&0\\0&1\end{pmatrix} & 0\begin{pmatrix}0&0\\0&1\end{pmatrix}\\ 0\begin{pmatrix}0&0\\0&1\end{pmatrix} & 1\begin{pmatrix}0&0\\0&1\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix} = M_{16}.
\end{aligned}
$$

Usando que F es una función bilineal, se puede probar que en general se cumple

$$
\begin{aligned}
\begin{pmatrix}a&b\\c&d\end{pmatrix}\otimes\begin{pmatrix}e&f\\g&h\end{pmatrix} &= F\left(\begin{pmatrix}a&b\\c&d\end{pmatrix},\begin{pmatrix}e&f\\g&h\end{pmatrix}\right)\\
&= \begin{pmatrix}a\begin{pmatrix}e&f\\g&h\end{pmatrix} & b\begin{pmatrix}e&f\\g&h\end{pmatrix}\\ c\begin{pmatrix}e&f\\g&h\end{pmatrix} & d\begin{pmatrix}e&f\\g&h\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}ae&af&be&bf\\ag&ah&bg&bh\\ce&cf&de&df\\cg&ch&dg&dh\end{pmatrix},
\end{aligned}
$$

es decir

$$\boxed{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a\begin{pmatrix} e & f \\ g & h \end{pmatrix} & b\begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ c\begin{pmatrix} e & f \\ g & h \end{pmatrix} & d\begin{pmatrix} e & f \\ g & h \end{pmatrix} \end{pmatrix} = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}.}$$

Por ejemplo

$$\begin{aligned} \begin{pmatrix} -2 & 5 \\ 3 & 4 \end{pmatrix} \otimes \begin{pmatrix} -1 & 8 \\ 3 & 2 \end{pmatrix} &= \begin{pmatrix} -2\begin{pmatrix} -1 & 8 \\ 3 & 2 \end{pmatrix} & 5\begin{pmatrix} -1 & 8 \\ 3 & 2 \end{pmatrix} \\ 3\begin{pmatrix} -1 & 8 \\ 3 & 2 \end{pmatrix} & 4\begin{pmatrix} -1 & 8 \\ 3 & 2 \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} 2 & -16 & -5 & 40 \\ -6 & -4 & 15 & 10 \\ -3 & 24 & -4 & 32 \\ 9 & 6 & 12 & 8 \end{pmatrix}. \end{aligned}$$

Se puede mostrar que para cualquier dos matrices

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots \\ b_{s1} & \cdots & b_{sr} \end{pmatrix},$$

se puede definir un producto tensorial como

$$\boxed{A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{pmatrix}.}$$

36.3. Producto de polinomios con el método Maya

Como una aplicación del producto de Kronecker entre vectores, veremos el método maya para obtener el producto de dos polinomios.

Supongamos que K es un campo y que tenemos los polinomios

$$\begin{aligned} P(x) &= a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n, \\ Q(x) &= b_0 + b_1x + b_2x^2 + \cdots + b_{m-1}x^{m-1} + b_mx^m. \end{aligned}$$

con

$$a_0, a_2, \cdots, a_n, b_0, b_1, \cdots, b_m \in K.$$

Entonces, el producto de estos polinomios es

$$\begin{aligned} P(x)Q(x) &= \big(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n\big) \\ &\quad \big(b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4 + \cdots + b_mx^m\big) \\ &= a_0b_0 + (a_0b_1 + b_0a_1)x + (a_0b_2 + a_1b_1 + b_0a_2)x^2 + \cdots + \\ &\quad +(a_{n-2}b_m + a_{n-1}b_{m-1} + a_nb_{n-2})x^{n+m-2} + \\ &\quad +(a_{n-1}b_m + a_nb_{m-1})x^{n+m-1} + \\ &\quad +a_nb_mx^{n+m}, \end{aligned}$$

que se puede escribir como

$$P(x)Q(x) = c_0 + c_1x + c_2x^2 + \cdots + c_{n+m-2}x^{n+m-2} + c_{n+m-1}x^{n+m-1} + c_{n+m}x^{n+m}$$

con

$$\begin{aligned} c_{n+m} &= a_nb_m, \\ c_{n+m-1} &= a_{n-1}b_m + a_nb_{m-1}, \\ c_{n+m-2} &= a_{n-2}b_m + a_{n-}b_{m-1} + a_nb_{m-2}, \\ &\vdots \\ c_2 &= a_0b_2 + a_1b_1 + b_0a_2, \\ c_1 &= a_0b_1 + b_0a_1, \\ c_0 &= a_0b_0, \end{aligned}$$

Se puede observar que los coeficientes c_i determinan el producto de los polinomios $P(x)$ y $Q(x)$.

Los mayas tenían un método muy eficiente para realizar este producto en el cual se usa el producto de Kronecker. Primero, al polinomio

$$P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n,$$

le asociamos el vector columna

$$a = \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_2 \\ a_1 \\ a_0 \end{pmatrix},$$

mientras que al polinomio

$$Q(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_{m-1} x^{m-1} + b_m x^m$$

le asociamos el vector columna

$$b = \begin{pmatrix} b_n \\ b_{n-1} \\ \vdots \\ b_2 \\ b_1 \\ b_0 \end{pmatrix}.$$

Posteriormente hacemos el producto de Kronecker de estos vectores

$$\begin{aligned} a \otimes b &= \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} \otimes \begin{pmatrix} b_m \\ b_{m-1} \\ \vdots \\ b_1 \\ b_2 \\ b_0 \end{pmatrix} = \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} \begin{pmatrix} b_m & b_{m-1} \cdots & b_2 & b_1 & b_0 \end{pmatrix} \\ &= \begin{pmatrix} a_n b_m & a_n b_{m-1} & a_n b_{m-2} & \cdots & a_n b_2 & a_n b_1 & a_n b_0 \\ a_{n-1} b_m & a_{n-1} b_{m-1} & a_{n-1} b_{m-2} & \cdots & a_{n-1} b_2 & a_{n-1} b_1 & a_{n-1} b_0 \\ a_{n-2} b_m & a_{n-2} b_{m-1} & a_{n-2} b_{m-2} & \cdots & a_{n-2} b_2 & a_{n-1} b_1 & a_{n-2} b_0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ a_2 b_m & a_2 b_{m-1} & a_2 b_{m-2} & \cdots & a_2 b_2 & a_2 b_1 & a_2 b_0 \\ a_1 b_m & a_1 b_{m-1} & a_1 b_{m-2} & \cdots & a_1 b_2 & a_1 b_1 & a_1 b_0 \\ a_0 b_m & a_0 b_{m-1} & a_0 b_{m-2} & \cdots & a_0 b_2 & a_0 b_1 & a_0 b_0 \end{pmatrix}, \end{aligned}$$

es decir

$$a \otimes b = \begin{pmatrix} a_n b_m & a_n b_{m-1} & a_n b_{m-2} & \cdots & a_n b_2 & a_n b_1 & a_n b_0 \\ a_{n-1} b_m & a_{n-1} b_{m-1} & a_{n-1} b_{m-2} & \cdots & a_{n-1} b_2 & a_{n-1} b_1 & a_{n-1} b_0 \\ a_{n-2} b_m & a_{n-2} b_{m-1} & a_{n-2} b_{m-2} & \cdots & a_{n-2} b_2 & a_{n-1} b_1 & a_{n-2} b_0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ a_2 b_m & a_2 b_{m-1} & a_2 b_{m-2} & \cdots & a_2 b_2 & a_2 b_1 & a_2 b_0 \\ a_1 b_m & a_1 b_{m-1} & a_1 b_{m-2} & \cdots & a_1 b_2 & a_1 b_1 & a_1 b_0 \\ a_0 b_m & a_0 b_{m-1} & a_0 b_{m-2} & \cdots & a_0 b_2 & a_0 b_1 & a_0 b_0 \end{pmatrix}.$$

Ahora realizaremos una operación que llamaremos traza maya, la cual consiste en sumar los elementos de las diagonales de la matriz de acuerdo con ell color y colocaremos el resultado en un vector. En este caso tenemos

$$Tr_{maya}\left(a \otimes b\right) = \begin{pmatrix} a_n b_m \\ a_{n-1} b_m + a_n b_{m-1} \\ a_{n-2} b_m + a_{n-1} b_{m-1} + a_n b_{m-2} \\ \vdots \\ a_0 b_2 + a_1 b_1 + a_2 b_2 \\ a_0 b_1 + a_1 b_0 \\ a_0 b_0 \end{pmatrix}.$$

Se puede observar que las entradas de este vector coincide con los coeficientes del producto de los polinomios $P(x)Q(x)$.

36.3.1. Ejemplo

Consideremos los polinomios

$$\begin{aligned} P(x) &= 2 - 3x + 4x^2 + x^3, \\ Q(x) &= 1 + 2x - x^2. \end{aligned}$$

En este caso tenemos los vectores

$$a = \begin{pmatrix} 1 \\ 4 \\ -3 \\ 2 \end{pmatrix}, \qquad b = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$$

cuyo producto de Kronecker es

$$
\begin{aligned}
a \otimes b &= \begin{pmatrix} 1 \\ 4 \\ -3 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ 4 \\ -3 \\ 2 \end{pmatrix} \begin{pmatrix} -1 & 2 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -1 & 2 & 1 \\ -4 & 8 & 4 \\ 3 & -6 & -3 \\ -2 & 4 & 2 \end{pmatrix},
\end{aligned}
$$

es decir

$$
a \otimes b = \begin{pmatrix} -1 & 2 & 1 \\ -4 & 8 & 4 \\ 3 & -6 & -3 \\ -2 & 4 & 2 \end{pmatrix},
$$

de donde

$$
Tr_{maya}\left(a \otimes b\right) = \begin{pmatrix} -1 \\ -2 \\ 12 \\ -4 \\ 1 \\ 2 \end{pmatrix},
$$

que implica el resultado

$$
\begin{aligned}
P(x)Q(x) &= \left(2 - 3x + 4x^2 + x^3\right)\left(1 + 2x - x^2\right) \\
&= 2 + x - 4x^2 + 12x^3 - 2x^4 - x^5.
\end{aligned}
$$

Este método para obtener el producto de dos polinomios se le suele llamar método de la diagonal, de las columnas o de las casillas (lattice). Existen documentos que indican que este método se conocía en la India en el siglo X y después fue introducido por Fibonacci a Europa. Pero también existen códices que muestran que el algorimo fue conocido por los mayas varios siglos antes del siglo X. Notablemente, el producto de Kronecker es desarrollado hasta el

siglo XIX en las matemáticas de Europa. Actualmente este algoritmo se suele usar en criptografía.
Un desarrollo semejante al mostrado en este capítulo se encuentra en el artículo [19].

36.4. Caso general

Supongamos que $V_1, V_2, \cdots, V_N$ y U son espacios vectoriales sobre K y la función

$$F : V_1 \times V_2 \times \cdots \times V_N \to U$$

es multilineal. Se dice que el par (F, U) es un producto tensorial si para cualquier espacio vectorial W sobre K y para cualquier función multilineal

$$T : V_1 \times V_2 \times \cdots \times V_N \to W,$$

existe una única función lineal

$$H : T \to W,$$

tal que

$$T = H \circ F.$$

Note que si existe el producto tensorial (F, U), entonces cualquier función multilineal de $V_1 \times V_2 \times \cdots \times V_N$ sobre cualquier espacio vectorial W, se puede escribir en términos de U.

Veamos el caso de espacios vectoriales de dimensión finita. Supongamos que cada espacio vectorial V_i $(i = 1, 2, \cdots, N)$ tiene dimensión finita n_i y que el conjunto de vectores

$$\{v_{i1}, v_{i2}, \cdots, v_{in_i}\} \tag{36.8}$$

es una de sus bases, también supongamos que la función

$$T : V_1 \times V_2 \times \cdots \times V_n \to W,$$

es multilineal. Ahora, sabemos que cualquier función multilineal T queda determinada por la forma en que actúa en la bases del dominio, que en este caso es

$$V_1 \times V_2 \times \cdots \times V_N.$$

Debido a que W es multilineal, existen $n_1 n_2 \cdots n_N$ vectores en W, tales que

$$T(v_{1i_1}, v_{2i_2}, \cdots, v_{Ni_N}) = w_{i_1 i_2 \cdots i_N} \in W,$$

donde

$$i_l = 1, 2, \cdots, n_l, \qquad l = 1, 2, \cdots, N.$$

Para este caso, el espacio vectoria U lo tomaremos como un espacio vectorial de dimensión

$$n_1 \cdots n_N,$$

también tomaremos una base de U y la enumeraremos de la forma

$$e_{i_1 i_2 \cdots i_N}, \qquad i_l = 1, \cdots, n_l, \quad l = 1, \cdots, N.$$

Por lo que podemos definir la función multilineal

$$F : V_1 \times V_2 \times \cdots \times V_N \to U,$$

tal que

$$F(v_{1i_1}, \cdots, v_{Ni_N}) = e_{i_1 i_2 \cdots i_N}.$$

Además, debido a que cualquier función lineal está definida por la forma en que actúa en la base del dominio, existe una única función lineal

$$H : U \to W,$$

tal que

$$H(e_{i_1 i_2 \cdots i_n}) = w_{i_1 i_2 \cdots i_N} = T(v_{1i_1}, v_{2i_2}, \cdots, v_{Ni_N}).$$

Se puede probar que se cumple,

$$T = H \circ F, \tag{36.9}$$

en efecto, podemos ver que

$$\begin{aligned}
(H \circ F)(v_{1i_1}, \cdots, v_{Ni_N}) &= H(F(v_{1i_1}, \cdots, v_{Ni_N})) = h(e_{i_1 i_2 \cdots i_N}) = w_{i_1 i_2 \cdots i_N} \\
&= T(v_{1i_1}, v_{2i_2}, \cdots, v_{Ni_N})
\end{aligned}$$

Así, para las bases (36.8) se cumple la igualdad

$$(H \circ F)(v_{1i_1}, \cdots, v_{Ni_N}) = T(v_{1i_1}, \cdots, v_{Ni_N})$$

lo que implica que para cualquier vector

$$(v_1, v_2, \cdots, v_N) \in V_1 \times V_2 \times \cdots \times V_N$$

se cumple la igualdad (36.9).

La existencia del producto tensorial es muy útil porque nos dice que todas las funciones multilineales se relacionan con la función universal F y basta estudiar ésta. Se suele denotar a U como

$$\boxed{U = V_1 \otimes_K V_2 \otimes_K \cdots \otimes_K V_N,}$$

también se suele usar la notación

$$\boxed{F(v_1, v_2, \cdots, v_N) = v_1 \otimes v_2 \otimes \cdots \otimes v_N.}$$

Se puede observar que debido a que F es una función multilineal, si

$$v_1 \in V_1, v_2 \in V_2, \cdots, v_{i-1} \in V_i, v_i, v_i' \in V_i, v_{i+1} \in V_{i+1}, \cdots, v_N \in V_N$$

y

$$\lambda, \gamma \in K$$

entonces se cumple

$$\begin{aligned} & v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes (\lambda v_i + \gamma v_i') \otimes v_{i+1} \otimes \cdots \otimes v_N = \\ = \; & \lambda v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes v_i \otimes v_{i+1} \otimes \cdots \otimes v_N + \\ & + \gamma v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes v_i' \otimes v_{i+1} \otimes \cdots \otimes v_N. \end{aligned}$$

De la misma forma, considerando que considerando que F es una función multilineal, se puede probar que se cumple

$$\begin{aligned} & \lambda (v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes v_i \otimes v_{i+1} \otimes \cdots \otimes v_N) = \\ = \; & v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes (\lambda v_i) \otimes v_{i+1} \otimes \cdots \otimes v_N, \end{aligned}$$

que implica

$$\begin{aligned} & -(v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes v_i \otimes v_{i+1} \otimes \cdots \otimes v_N) = \\ = \; & v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes (-v_i) \otimes v_{i+1} \otimes \cdots \otimes v_N \end{aligned}$$

y

$$(v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes 0 \otimes v_{i+1} \otimes \cdots \otimes v_N) = 0.$$

36.5. Tensor

Supongamos que $V_1, V_2, \cdots, V_n$ son espacios vectoriales sobre K, entonces a las funciones

$$F : V_1 \times V_2 \times \cdots \times V_n \to K$$

multilineales se les llama tensores.

Para la geometría diferencial y la relatividad general son de particular importancia los tensores que involucran espacios vectoriales

$$V_1, V_2, \cdots, V_n$$

y sus espacios vectoriales duales

$$V_1^*, V_2^*, \cdots, V_n^*$$

En este caso, tendremos funciones multilineales de la fornma

$$F : V_1^* \times V_2^* \times \cdots \times V_n^* \times V_1 \times V_2 \times \cdots \times V_n \to K$$

En este con texto, a los vectores de los espacios V_i se les llama contravariantes y a los elementos de los espacios duales V_i^* se les llama covariantes.

Por ejemplo, si

$$V = V_1 = V_2 = \cdots = V_n$$

tenemos el espacio dual

$$V^* = V_1^* = V_2^* = \cdots = V_n^*.$$

Por lo que podemos construir el tensor

$$F : \underbrace{V^* \times V^* \times \cdots \times V^*}_{l} \times \underbrace{V \times V \times \cdots \times V}_{m} \to K \tag{36.10}$$

Ahora, supogamos que V tiene dimensión n y que

$$\{v_1, v_2, v_3, \cdots, v_n\}$$

es una de sus bases. En este caso el espacio dual V^* tiene dimensión n y supongamos que

$$\{v_1^*, v_2^*, v_3^*, \cdots, v_n^*\}$$

es una de sus bases. Entonces, los tensores de la forma (36.10) son generados por las funciones multilineales de la forma

$$v_{i_1}^* \otimes v_{i_2}^* \otimes \cdots \otimes v_{i_l}^* \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m} \tag{36.11}$$

con

$$i_1, i_2, \cdots, i_l, j_1, j_2, \cdots, j_m = 1, 2, \cdots, n.$$

Ahora, por simplicidad a los vectores que pertenecen a una base de V (contravariantes) se suelen escribir con el índice abajo. Así, una base de V se suele escribir como

$$\{v_1, v_2, v_3, \cdots, v_n\}$$

Mientras que los vectores que pertenecen a una base de V^* (covariantes) se suelen escribir con el índice arriba. Así, una base de V^* se suele escibir como

$$\{v_1^* = v^1, v_2^* = v^2, v_3^* = v^3, \ldots, v_n^* = v^n\}.$$

Por lo que las funciones (36.11) se pueden expresar de la forma

$$v^{i_1} \otimes v^{i_2} \otimes v^{i_3} \otimes \cdots \otimes v^{i_l} \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m}.$$

Así, cualquier tensor de la forma (36.10) se puede escribir como

$$F = F_{i_1 i_2 \cdots i_l}{}^{j_1 j_2 \cdots j_m} v^{i_1} \otimes v^{i_2} \otimes \cdots \otimes v^{i_l} \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m}. \tag{36.12}$$

con

$$F_{i_1 i_2 \cdots i_l}{}^{j_1 j_2 \cdots j_m} \in K.$$

Además, recordemos que si

$$v_1, v_2, v_3, \cdots, v_n$$

y

$$v_1', v_2', v_3', \cdots, v_n'$$

son base de V se cumple

$$v_i' = U_{ij} v_j, \tag{36.13}$$

con U la matriz de cambio de base. También recordemos que si

$$v_1^*, v_2^*, v_3^*, \cdots, v_n^*$$

y

$$v_1^{*\prime}, v_2^{*\prime}, v_3^{*\prime}, \cdots, v_n^{*\prime}$$

son base del espacio dual V^* se cumple

$$v_i^{*\prime} = \left(\left(U^{-1}\right)^T\right)_{ij} v_j^*. \tag{36.14}$$

Con la notación de índices arriba y abajo (vectores covariantes y contravariantes) la identidades (36.13) y (36.14) se suelen escribir como

$$\begin{aligned} v_i' &= U_i{}^j v_j, \\ v'^i &= U^i{}_j v^j. \end{aligned}$$

Ahora, con un cambio de base, se debe cumplir la igualdad

$$\begin{aligned} F &= F_{i_1 i_2 \cdots i_l}{}^{j_1 j_2 \cdots j_m} v^{i_1} \otimes v^{i_2} \otimes \cdots \otimes v^{i_l} \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m} \\ &= F'_{i_1' i_2' \cdots i_l'}{}^{j_1' j_2' \cdots j_m'} v'^{i_1'} \otimes v'^{i_2'} \otimes \cdots \otimes v'^{i_l'} \otimes v'_{j_1'} \otimes v'_{j_2'} \otimes \cdots \otimes v'_{j_m'}. \end{aligned}$$

También se tiene

$$\begin{aligned} F &= F'_{i_1' i_2' \cdots i_l'}{}^{j_1' j_2' \cdots j_m'} v'^{i_1'} \otimes v'^{i_2'} \otimes \cdots \otimes v'^{i_l'} \otimes v'_{j_1'} \otimes v'_{j_2'} \otimes \cdots \otimes v'_{j_m'} \\ &= F'_{i_1' i_2' \cdots i_l'}{}^{j_1' j_2' \cdots j_m'} \left(U^{i_1'}{}_{i_1} v^{i_1}\right) \otimes \left(U^{i_2'}{}_{i_2} v^{i_2}\right) \otimes \cdots \otimes \left(U^{i_l'}{}_{i_l} v^{i_l}\right) \otimes \\ &\quad \otimes \left(U_{j_1'}{}^{j_1} v_{j_1}\right) \otimes \left(U_{j_2'}{}^{j_2} v_{j_2}\right) \otimes \cdots \otimes \left(U_{j_m'}{}^{j_m} v_{j_m}\right) \\ &= U^{i_1'}{}_{i_1} \cdot U^{i_2'}{}_{i_2} \cdots U^{i_l'}{}_{i_l} \cdot U_{j_2'}{}^{j_2} \cdots U_{j_2'}{}^{j_2} U_{j_m'}{}^{j_m} F'_{i_1' i_2' \cdots i_l'}{}^{j_1' j_2' \cdots j_m'} \\ &\quad \left(v^{i_1} \otimes v^{i_2} \otimes \cdots \otimes v^{i_l} \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m}\right), \end{aligned}$$

esto es,

$$\begin{aligned} F &= U^{i_1'}{}_{i_1} \cdot U^{i_2'}{}_{i_2} \cdots U^{i_l'}{}_{i_l} \cdot U_{j_2'}{}^{j_2} \cdots U_{j_2'}{}^{j_2} U_{j_m'}{}^{j_m} F'_{i_1' i_2' \cdots i_l'}{}^{j_1' j_2' \cdots j_m'} \\ &\quad \left(v^{i_1} \otimes v^{i_2} \otimes \cdots \otimes v^{i_l} \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m}\right). \end{aligned}$$

Esta igualdad implica el resultado

$$\begin{aligned} F &= F_{i_1 i_2 \cdots i_l}{}^{j_1 j_2 \cdots j_m} v^{i_1} \otimes v^{i_2} \otimes \cdots \otimes v^{i_l} \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m} \\ F &= U^{i_1'}{}_{i_1} \cdot U^{i_2'}{}_{i_2} \cdots U^{i_l'}{}_{i_l} \cdot U_{j_2'}{}^{j_2} \cdots U_{j_2'}{}^{j_2} U_{j_m'}{}^{j_m} F'_{i_1' i_2' \cdots i_l'}{}^{j_1' j_2' \cdots j_m'} \\ &\quad \left(v^{i_1} \otimes v^{i_2} \otimes \cdots \otimes v^{i_l} \otimes v_{j_1} \otimes v_{j_2} \otimes \cdots \otimes v_{j_m}\right), \end{aligned}$$

que induce la regla de transformación

$$F_{i_1 i_2 \cdots i_l}{}^{j_1 j_2 \cdots j_m} = U^{i'_1}{}_{i_1} \cdots U^{i'_l}{}_{i_l} \cdots U_{j'_2}{}^{j_2} U_{j'_m}{}^{j_m} F'_{i'_1 i'_2 \cdots i'_l}{}^{j'_1 j'_2 \cdots j'_m}.$$

Este tipo de reglas de transformación fueron importantes para la formulación de los principios de la relatividad general y para establecer sus ecuaciones de movimiento.

Capítulo 37

Compuertas clásicas y cuánticas

En este capítulo veremos que diferentes compuertas de la computación clásica y cuántica se pueden expresar con herramientas del álgebra lineal.

37.1. Compuerta NOT

Primero veamos la negación. Supongamos que A es una proposición, entonces la tabla de verdad de $\neg A$ es (37.1)

A	$\neg A$
V	F
F	V

Cuadro 37.1: Tabla de verdad de la negación NOT, $\neg A$

En este caso tenemos

$$\mathrm{NOT}\,(V) = F, \tag{37.1}$$
$$\mathrm{NOT}\,(F) = V. \tag{37.2}$$

Ahora, haremos la asignación

$$V = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{37.3}$$
$$F = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{37.4}$$

Entonces, podemos proponer la matriz

$$\mathrm{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{37.5}$$

con la cual se satisface

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{37.6}$$

Estos resultados coinciden con los resultados (37.1)-(37.2). Así, con los vectores (37.3) y (37.4) junto con la matriz (37.5) podemos representar la tabla de verdad de la negación NOT (37.1).

37.2. Compuerta OR

Supongamos que tenemos las proposiciones A y B entonces la tabla de verdad de A o B, está dado por (37.2)

A	B	$A \vee B$
V	V	V
V	F	V
F	V	V
F	F	F

Cuadro 37.2: Tabla de verdad OR, $A \vee B$.

Esta tabla se puede representar con las siguientes ecuaciones

$$\mathrm{OR}\,(V, V) = V, \tag{37.7}$$
$$\mathrm{OR}\,(V, F) = V, \tag{37.8}$$
$$\mathrm{OR}\,(F, V) = V, \tag{37.9}$$
$$\mathrm{OR}\,(F, F) = F, \tag{37.10}$$

Ahora, podemos hacer el producto tensorial

$$\begin{aligned}
V \otimes V &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \\
V \otimes F &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\
F \otimes V &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\
F \otimes F &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.
\end{aligned}$$

Entonces, podemos proponer la matriz OR

$$\mathrm{OR} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}, \tag{37.11}$$

con la cual se satisface

$$\begin{aligned}
\mathrm{OR}\,(V \otimes V) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = V, \\
\mathrm{OR}\,(V \otimes F) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = V, \\
\mathrm{OR}\,(F \otimes V) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = V, \\
\mathrm{OR}\,(F \otimes F) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = F,
\end{aligned}$$

es decir,

$$\begin{aligned}
\mathrm{OR}\,(V \otimes V) &= V,\\
\mathrm{OR}\,(V \otimes F) &= V,\\
\mathrm{OR}\,(F \otimes V) &= V,\\
\mathrm{OR}\,(F \otimes F) &= F,
\end{aligned}$$

que coincide con la tabla de verdad de $A \vee B$, dada por las expresiones (37.7)-(37.10).

37.3. Compuerta AND

Si A y B son dos proposiciones, entonces la tabla de verdad de A y B, está dado por (37.3)

A	B	$A \wedge B$
V	V	V
V	F	F
F	V	F
F	F	F

Cuadro 37.3: Tabla de verdad AND, $A \wedge B$.

Esta tabla se puede representar mediante las siguientes ecuaciones

$$\mathrm{AND}\,(V, V) = V, \tag{37.12}$$
$$\mathrm{AND}\,(V, F) = F, \tag{37.13}$$
$$\mathrm{AND}\,(F, V) = F, \tag{37.14}$$
$$\mathrm{AND}\,(F, F) = F. \tag{37.15}$$

Entonces podemos proponer la matriz AND de la forma

$$\mathrm{AND} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{37.16}$$

con la cual se satisface

$$
\begin{aligned}
\mathrm{AND}\,(V\otimes V) &= \begin{pmatrix}1&1&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}=V,\\
\mathrm{AND}\,(V\otimes F) &= \begin{pmatrix}1&1&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}=F,\\
\mathrm{AND}\,(F\otimes V) &= \begin{pmatrix}1&1&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}=F,\\
\mathrm{AND}\,(F\otimes F) &= \begin{pmatrix}1&1&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}=F,
\end{aligned}
$$

es decir,

$$
\begin{aligned}
\mathrm{AND}\,(V\otimes V) &= V,\\
\mathrm{AND}\,(V\otimes F) &= F,\\
\mathrm{AND}\,(F\otimes V) &= F,\\
\mathrm{AND}\,(F\otimes F) &= F,
\end{aligned}
$$

que coincide con la tabla de verdad de $A\wedge B$ dada por las expresiones (37.12)-(37.15).

37.4. Compuerta NAND

Ya tenemos la compuerta NOT para $\neg A$ y la compuerta AND para $A\wedge B$. Entonces para construir la compuerta NAND para $\neg\,(A\wedge B)$, de la forma

$$
\begin{aligned}
\mathrm{NAND} &= \mathrm{NOT}\cdot\mathrm{AND}\\
&= \begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&1&1&0\\0&0&0&1\end{pmatrix}\\
&= \begin{pmatrix}0&0&0&1\\1&1&1&0\end{pmatrix},
\end{aligned}
$$

es decir,

$$\mathrm{NAND} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}. \tag{37.17}$$

Con la cual se satisface

$$\begin{aligned}
\mathrm{NAND}\,(V \otimes V) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = F, \\
\mathrm{NAND}\,(V \otimes F) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = V, \\
\mathrm{NAND}\,(F \otimes V) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = V, \\
\mathrm{NAND}\,(F \otimes F) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = V,
\end{aligned}$$

esto es

$$\begin{aligned}
\mathrm{NAND}\,(V \otimes V) &= F, \\
\mathrm{NAND}\,(V \otimes F) &= V, \\
\mathrm{NAND}\,(F \otimes V) &= V, \\
\mathrm{NAND}\,(F \otimes F) &= V,
\end{aligned}$$

que coincide con la tabla de verdad de

$$\neg\,(A \wedge B)\,.$$

Cuando se aplican compuertas de manera sucesiva se dice que se aplican operaciones secuenciales.

37.5. Compuertas paralelas

No todas las operaciones con compuertas se pueden obtener con operaciones secuenciales.

En algunos casos se requieren realizar operaciones paralelas en las cuales una compuerta solo afecta a algunos de los vectores. Por ejemplo, si A y B son dos proposiciones y queremos realizar la tabla de verdad de

$$\neg A \wedge B \tag{37.18}$$

la compuerta NOT solo afecta a la proposición A, pero no a B. Para la representación de las operaciones paralelas podemos ocupar el producto tensorial en las compuertas.

Antes de continuar, veamos la tabla de verdad de (37.18), la cual está representada en la tabla (37.4)

A	B	$\neg A$	$\neg A \wedge B$
V	V	F	F
V	F	F	F
F	V	V	V
F	F	V	F

Cuadro 37.4: Tabla de verdad $\neg A \vee B$.

Ahora, para solo negar a A ocuparemos la compuerta

$$\text{NOT} \otimes I,$$

con la cual tendremos

$$(\text{NOT} \otimes I)(A \otimes B) = (\text{NOT}(A)) \otimes B,$$

después, debemos aplicar la compuerta AND, es decir,

$$\text{AND}((\text{NOT} \otimes I)(A \otimes B)) = \text{AND}((\text{NOT}(A)) \otimes B),$$

esto es, debemos aplicar la compuerta

$$\text{AND}(\text{NOT} \otimes I).$$

De forma explícita tenemos

$$\begin{aligned}
\mathrm{NOT}\otimes I &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\otimes\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\\
&= \begin{pmatrix}0\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} & 1\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\\ 1\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} & 0\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\end{pmatrix}\\
&= \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{pmatrix},
\end{aligned}$$

es decir,

$$\mathrm{NOT}\otimes I = \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{pmatrix},$$

por lo que

$$\begin{aligned}
\mathrm{AND}\left(\mathrm{NOT}\otimes I\right) &= \begin{pmatrix}1 & 1 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{pmatrix}\\
&= \begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 0 & 0\end{pmatrix}.
\end{aligned}$$

Así, tenemos la compuerta

$$\mathrm{AND}\left(\mathrm{NOT}\otimes I\right) = \begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 0 & 0\end{pmatrix}.$$

Con la cual se satisface

$$
\begin{aligned}
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(V\otimes V) &= \begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 0 & 0\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = F,\\
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(V\otimes F) &= \begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 0 & 0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = F,\\
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(F\otimes V) &= \begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 0 & 0\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix} = V,\\
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(F\otimes F) &= \begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = F,
\end{aligned}
$$

es decir,

$$
\begin{aligned}
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(V\otimes V) &= F,\\
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(V\otimes F) &= F,\\
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(F\otimes V) &= V,\\
(\mathrm{AND}\,(\mathrm{NOT}\otimes I))\,(F\otimes F) &= F,
\end{aligned}
$$

que coincide con la tabla de verdad de

$$\neg A \wedge B.$$

37.6. Leyes de De Morgan

Si A y B son dos proposiciones, una de las implicaciones de las leyes de De Morgan es la igualdad

$$\neg(\neg A \wedge \neg B) = A \vee B. \tag{37.19}$$

La parte izquierda de la igualdad se obtiene con la compuerta

$$\mathrm{NOT}\,(\mathrm{AND}\,(\mathrm{NOT}\otimes\mathrm{NOT}))\,,$$

mientras que la parte derecha es simplemnte la compuerta OR

$$\mathrm{OR} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

Así, la expresión (37.19) es equivalente a la igualdad

$$\mathrm{NOT}\,(\mathrm{AND}\,(\mathrm{NOT} \otimes \mathrm{NOT})) = \mathrm{OR}. \tag{37.20}$$

Para probar esta igualdad notemos que se cumple

$$\begin{aligned} \mathrm{NOT} \otimes \mathrm{NOT} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 0\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

esto es

$$\mathrm{NOT} \otimes \mathrm{NOT} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Por lo que,

$$\begin{aligned} \mathrm{NOT}\,(\mathrm{AND}\,(\mathrm{NOT} \otimes \mathrm{NOT})) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \\ &= \mathrm{OR}. \end{aligned}$$

Lo que prueba la igualdad (37.20).

37.7. Notación binomial y bits

En muchos casos es más común emplear la notación binomial o de bits

$$
\begin{aligned}
V &= 1, \\
F &= 0
\end{aligned}
$$

y para los vectores se suelen usar los kets, que originalmente surgieron en Mecánica Cuántica en donde se toma

$$
\begin{aligned}
|V\rangle &= |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\
|F\rangle &= |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.
\end{aligned}
$$

Usando esta notación se toma

$$
\begin{aligned}
|00\rangle &= |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\
|01\rangle &= |0\rangle \otimes |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\
|10\rangle &= |1\rangle \otimes |0\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\
|11\rangle &= |1\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
\end{aligned}
$$

esto es

$$\begin{aligned}
|00\rangle &= |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\
|01\rangle &= |0\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\
|10\rangle &= |1\rangle \otimes |0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\
|11\rangle &= |1\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
\end{aligned}$$

37.8. Qubits

Una computadora cuántica usa como qubits los estados de espín de un electrón. Los estados de espín tienen como base los vectores

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Entonces, cualquier estado del espín del electrón se puede escribir como combinación lineal de los estados $|0\rangle$ y $|1\rangle$, esto es

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle ,$$

donde

$$|\alpha_0|^2 + |\alpha_1|^2 = 1.$$

Aquí

$$|\alpha_0|^2$$

representa la probabilidad de que el electrón se encuentre en el estado $|0\rangle$. Mientras que

$$|\alpha_1|^2$$

representa la probabilidad de que el electrón se encuentre en el estado $|1\rangle$.

Los estados de espín de un electrón se pueden representar en la llamada esfera de Bloch, que se muestra en la figura (37.1).

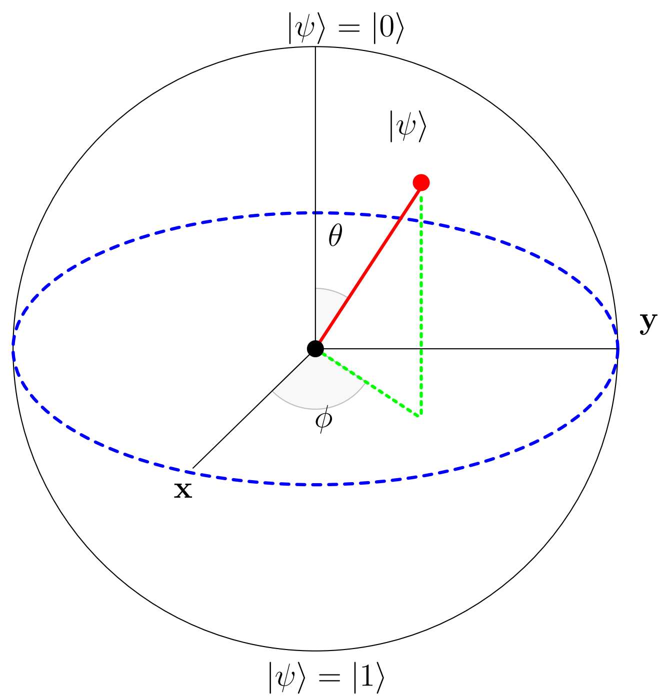

Figura 37.1: Representación del Qubit en la esfera de Bloch.

La diferencia entre la computación clásica y cuántica radica en que los bit solo tienen acceso a los polos de la esfera de Bloch, mientras que los qubits tienen acceso a toda la esfera de Bloch.

Para construir estados de varios qubits ocuparemos el producto tensorial, por ejemplo

$$|11\rangle = |1\rangle \otimes |1\rangle ,$$

o

$$|101\rangle = |1\rangle \otimes |0\rangle \otimes |1\rangle .$$

Podemos observar que en un sistema de n qubits, cada electrón aporta dos elementos a la base del sistema. Por lo que un sistema de n qubits tiene dimensión

$$2^n.$$

37.9. Compuertas cuánticas

Existen diferentes compuertas cuánticas, por ejemplo tenemos las matrices de Pauli, que en computación cuántica se denotan como

$$\begin{aligned} X &= \mathrm{NOT} = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ Y &= \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ Z &= \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Otro ejemplo es la compuerta de Hadamard

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

con la cual se obtiene

$$\begin{aligned} H|0\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} = |+\rangle, \\ H|1\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}} = |-\rangle. \end{aligned}$$

Se puede observar que las compuertas actuando en un qubit son transformaciones geométricas de puntos sobre la esfera de Bloch (37.1).

Ahora, para realizar operaciones secuenciales debemos aplicar de manera sucesivas las compuertas. Por ejemplo, si el sistema está en el estado $|\psi\rangle$ y tenemos las compuertas $M_1, M2, \cdots, M_{N-1}$ podemos realizar la operación secuencial

$$M_N M_{N-1} \cdots M_2 M_1 |\psi\rangle.$$

Para realizar operaciones paralelas se usa el producto tensorial. Por ejemplo, si el sistema se describe con el estado

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$$

y se tienen las dos compuertas

$$A_1, \qquad A_2$$

podemos construir la compuerta paralela

$$A = A_1 \otimes A_2,$$

con la cual se puede realizar la operación paralela

$$A\,|\psi\rangle = (A_1 \otimes A_2)\,(|\psi_1\rangle \otimes |\psi_2\rangle) = (A_1\,|\psi_1\rangle) \otimes (A_2\,|\psi_2\rangle)$$

esto es

$$A\,|\psi\rangle = (A_1\,|\psi_1\rangle) \otimes (A_2\,|\psi_2\rangle)\,.$$

37.10. Compuertas con control o controladas

Existe otro tipo de compuertas que se les llama C-Compuertas, las cuales actúan sobre sistemas de más de un qubit y si se cumple alguna condición particular sobre uno de los qubits.

Por ejemplo, supongamos que tenemos la compuerta A para sistemas de un qubit

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{37.21}$$

Entonces podemos construir la compuerta CA que actúa sobre un sistemas de dos qubits

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = |\psi_1\psi_2\rangle\,. \tag{37.22}$$

En donde la regla es que la compuerta CA no cambia a los estados en los cuales el primer qubit es $|0\rangle$, pero si el primer qubit es $|1\rangle$ la compuerta A actúa sobre el segundo qubit. Esto es

$$\begin{aligned} \mathrm{CA}\,|00\rangle &= |00\rangle\,, \\ \mathrm{CA}\,|01\rangle &= |01\rangle\,, \\ \mathrm{CA}\,|10\rangle &= |1\rangle \otimes A\,|0\rangle\,, \\ \mathrm{CA}\,|11\rangle &= |1\rangle \otimes A\,|1\rangle\,. \end{aligned}$$

Por ejemplo tenemos la compuerta CNOT es

$$\begin{aligned}
\mathrm{CNOT}\,|00\rangle &= |00\rangle\,, \\
\mathrm{CNOT}\,|01\rangle &= |01\rangle\,, \\
\mathrm{CNOT}\,|10\rangle &= |1\rangle \otimes \mathrm{NOT}\,|0\rangle = |1\rangle \otimes |1\rangle = |11\rangle\,, \\
\mathrm{CNOT}\,|11\rangle &= |1\rangle \otimes \mathrm{NOT}\,|1\rangle = |1\rangle \otimes |0\rangle = |10\rangle\,,
\end{aligned}$$

esto es

$$\begin{aligned}
\mathrm{CNOT}\,|00\rangle &= |00\rangle\,, \\
\mathrm{CNOT}\,|01\rangle &= |01\rangle\,, \\
\mathrm{CNOT}\,|10\rangle &= |11\rangle\,, \\
\mathrm{CNOT}\,|11\rangle &= |10\rangle\,.
\end{aligned}$$

Así, la compuerta CNOTse puede representar con la matriz

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Como otro ejemplo tenemos la compuerta CH, la cual nos da

$$\begin{aligned}
\mathrm{CH}\,|00\rangle &= |00\rangle\,, \\
\mathrm{CH}\,|01\rangle &= |01\rangle\,, \\
\mathrm{CH}\,|10\rangle &= |1\rangle \otimes \mathrm{H}\,|0\rangle = |1\rangle \otimes |+\rangle = |1+\rangle\,, \\
\mathrm{CH}\,|11\rangle &= |1\rangle \otimes \mathrm{H}\,|1\rangle = |1\rangle \otimes |-\rangle = |1-\rangle\,,
\end{aligned}$$

esto es

$$\begin{aligned}
\mathrm{CH}\,|00\rangle &= |00\rangle\,, \\
\mathrm{CH}\,|01\rangle &= |01\rangle\,, \\
\mathrm{CH}\,|10\rangle &= |1+\rangle\,, \\
\mathrm{CH}\,|11\rangle &= |1-\rangle\,.
\end{aligned}$$

Por lo que compuerta CH se puede representar con la matriz

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Para esete caso, una compuerta CA se puede representar con la matriz

$$\mathrm{CA} = \begin{pmatrix} I & 0 \\ 0 & A \end{pmatrix}.$$

37.11. Ejercicios

1.- Si A y B son dos proposiciones, obtener la representación matricial de la compuerta para la expresión

$$A \vee \neg B. \tag{37.23}$$

2.- Si A y B son dos proposiciones, usando la representación matricial de las compuertas mostrar que se cumplen la ley de De Morgan

$$\neg (A \vee B) = \neg A \wedge \neg B, \tag{37.24}$$

es decir, usando la representación matricial de compuertas mostrar que se cumple

$$\mathrm{NOT} \cdot (\mathrm{OR}) = \mathrm{AND} \cdot (\mathrm{NOT} \otimes \mathrm{NOT}) . \tag{37.25}$$

3.- Si A y B son dos proposiciones, usando la representación matricial de las compuertas mostrar que se cumplen la ley de De Morgan

$$\neg (A \wedge B) = \neg A \vee \neg B.$$

es decir, usando la representación matricial de compuertas mostrar que se cumple

$$\mathrm{NOT}\,(\mathrm{AND}) = \mathrm{OR}\,(\mathrm{NOT} \otimes \mathrm{NOT}) .$$

Capítulo 38

Transformada cuántica de Fourier

La Transformada Cuántica de Fourier está relacionada con la Transformada Discreta de Fourier, sin embargo, estas transformadas se aplican en contextos diferentes. Por esta razón, buscando que estos temas se puedan estudiar de forma autónoma, los capítulos de estos temas se presentan de forma independiente con su notación propia y sin hacer referencia entre ellos.

38.1. Expansión binaria

Antes de iniciar, recordemos la expresión de los números naturales en la forma binanria.

Primero notemos que los números 0 y 1 se pueden escribir como

$$\begin{aligned} 0 &= 0 \cdot 2^0, \\ 1 &= 1 \cdot 2^0 \end{aligned}$$

es decir

$$x = a_0 2^0, \qquad a_0 = 0, 1.$$

Además, los 4 números

$$0, 1, 2, 3$$

se pueden escribir como

$$\begin{aligned}
0 &= 0 \cdot 2^1 + 0 \cdot 2^0, \\
1 &= 0 \cdot 2^1 + 1 \cdot 2^0, \\
2 &= 1 \cdot 2^1 + 0 \cdot 2^0, \\
3 &= 1 \cdot 2^1 + 1 \cdot 2^0,
\end{aligned}$$

de donde

$$x = a_1 2^1 + a_0 2^0, \qquad a_1, a_0 = 0, 1.$$

De la misma forma, los primeros 8 números

$$0, 1, 2, 3, 4, 5, 6, 7$$

se pueden escribir como

$$\begin{aligned}
0 &= 0 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0, \\
1 &= 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0, \\
2 &= 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0, \\
3 &= 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0. \\
4 &= 1 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0, \\
5 &= 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0, \\
6 &= 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0, \\
7 &= 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0,
\end{aligned}$$

así

$$x = a_2 2^2 + a_1 2^1 + a_0 2^0, \qquad a_1, a_0 = 0, 1.$$

En general, la expansión binaria de los primeros 2^n números

$$0, 1, 2, \cdots, 2^n - 1$$

se pueden escribir como

$$\boxed{x = a_{n-1} 2^{n-1} + a_{n-2} 2^{n-2} + \cdots + a_1 2^1 + a_0 2^0, \qquad a_{n-1}, \cdots, a_0 = 0, 1.}$$

38.1.1. Notación binaria fraccionaria

Ahora veremos la notación binaria fraccionaria.

Note que dado el conjunto de n números

$$x_1, x_2, \cdots, x_n$$

se puede construir el número decimal

$$x = 0.x_1 x_2 \cdots x_n.$$

Con esta expresión se puede se define la fracción binaria como

$$\boxed{[0.x_1 x_2 \cdots x_n] = \frac{x_1}{2} + \frac{x_2}{2^2} + \frac{x_3}{2^3} + \cdots + \frac{x_n}{2^n},}$$

en particular,

$$[0.x_1] = \frac{x_1}{2},$$

además,

$$[0.x_1 x_2] = \frac{x_1}{2} + \frac{x_2}{2^2}.$$

Por ejemplo,

$$[0{,}1] = \frac{1}{2},$$

además,

$$[0{,}11] = \frac{1}{2} + \frac{1}{2^2} = \frac{3}{4}$$

otro ejemplo es

$$[0{,}101] = \frac{1}{2} + \frac{0}{2^2} + \frac{1}{2^3} = \frac{5}{8}.$$

Los resultados de esta sección los usaremos para expresar de forma más sencilla los qubits y su transformada cuántica de Fourier.

38.2. Estados con notación binaria

Un qubit tienen como base los estados

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Para un sistema de un qubit cualquier estado se puede escribir como

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle, \qquad |\alpha_0|^2 + |\alpha_1|^2 = 1.$$

Además, note que si tenemos dos estados, $N = 2^1$, de forma binaria tenemos

$$x = a_0 2^0, \qquad a_0 = 0, 1,$$

por lo que la base de estados de un qubit se pueden escribir como

$$|a_0\rangle, \qquad a_0 = 0, 1.$$

Para dos qubits tenemos $N = 2^2 = 4$ estados, los cuales son

$$\boxed{\begin{aligned} |0\rangle \otimes |0\rangle &= |00\rangle, \\ |0\rangle \otimes |1\rangle &= |01\rangle, \\ |1\rangle \otimes |0\rangle &= |10\rangle, \\ |1\rangle \otimes |1\rangle &= |11\rangle. \end{aligned}}$$

Como podemos ver, estos estados se expresan en términos de los coeficientes de la expansión binaria de los primeros cuatro números. Así, inspirados en la expansión binaria, si tenemos 2 qubits, a los elementos de la base del sistema los enumeramos como

$$\boxed{\begin{aligned} |0\rangle &= |00\rangle, \\ |1\rangle &= |01\rangle, \\ |2\rangle &= |10\rangle, \\ |3\rangle &= |11\rangle. \end{aligned}}$$

Por lo que, cuando tenemos 2 qubits cualquier estado se puede escribir como

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \alpha_2 |2\rangle + \alpha_3 |3\rangle,$$

con

$$|\alpha_0|^2 + |\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 = 1.$$

Para tres qubits tenemos $N = 2^3 = 8$ estados, los cuales son

$$\boxed{\begin{aligned}
|0\rangle \otimes |0\rangle \otimes |0\rangle &= |000\rangle\,, \\
|0\rangle \otimes |0\rangle \otimes |1\rangle &= |001\rangle\,, \\
|0\rangle \otimes |1\rangle \otimes |0\rangle &= |010\rangle\,, \\
|0\rangle \otimes |1\rangle \otimes |1\rangle &= |011\rangle\,. \\
|1\rangle \otimes |0\rangle \otimes |0\rangle &= |100\rangle\,, \\
|1\rangle \otimes |0\rangle \otimes |1\rangle &= |101\rangle\,, \\
|1\rangle \otimes |1\rangle \otimes |0\rangle &= |110\rangle\,, \\
|1\rangle \otimes |1\rangle \otimes |1\rangle &= |111\rangle\,.
\end{aligned}}$$

Entonces, siguiendo esta expansión binaria de los primeros 8 números, si tenemos 3 qubits enumeramos los elementos de la base como

$$\boxed{\begin{aligned}
|0\rangle &= |000\rangle\,, \\
|1\rangle &= |001\rangle\,, \\
|2\rangle &= |010\rangle\,, \\
|3\rangle &= |011\rangle\,, \\
|4\rangle &= |100\rangle\,, \\
|5\rangle &= |101\rangle\,, \\
|6\rangle &= |110\rangle\,, \\
|7\rangle &= |111\rangle\,.
\end{aligned}}$$

Así, cuando tenemos 3 qubits, cualquier estado se puede escribir como

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \alpha_2 |2\rangle + \alpha_3 |3\rangle + \alpha_4 |4\rangle + \alpha_5 |5\rangle + \alpha_6 |6\rangle + \alpha_7 |7\rangle$$

con

$$|\alpha_0|^2 + |\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 + |\alpha_4|^2 + |\alpha_5|^2 + |\alpha_6|^2 + |\alpha_7|^2 = 1$$

Si tenemos un sistema de n qubits, su base consta de $N = 2^n$ estados, los cuales son

$$\boxed{\begin{aligned} &|0\rangle \otimes \cdots \otimes |0\rangle \otimes |0\rangle = |0\cdots 00\rangle\,, \\ &|0\rangle \otimes \cdots \otimes |0\rangle \otimes |1\rangle = |0\cdots 01\rangle\,, \\ &\qquad\qquad\vdots \\ &|1\rangle \otimes \cdots \otimes |1\rangle \otimes |1\rangle = |1\cdots 11\rangle\,. \end{aligned}}$$

Para enumerar estos estados emplearemos la expansión binaria de los primeros 2^n números. En este caso, si

$$\boxed{k \in \{0, 1, 2, \cdots, 2^n - 1\}}$$

definiremos el estado

$$\boxed{|k\rangle = |a_{n-1}a_{n-2}\cdots a_1 a_0\rangle\,,}$$

donde

$$\boxed{k = a_{n-1}2^{n-1} + a_{n-2}2^{n-2} + \cdots + a_1 2^1 + a_0 2^0, \qquad a_{n-1}, \cdots, a_0 = 0, 1.}$$

Así, cuando tenemos n qubits, cualquier estado se puede escribir como

$$\boxed{|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \cdots + \alpha_{2^n-1} |2^n - 1\rangle\,,}$$

con

$$\boxed{|\alpha_0|^2 + |\alpha_1|^2 + \cdots + |\alpha_{2^n-1}|^2 = 1.}$$

38.3. Transformada cuántica de Fourier

Supongamos que tenemos n qubits, entonces si

$$|\psi\rangle = \sum_{k=0}^{2^n-1} \gamma_k |k\rangle$$

es un estado del sistema, su Transformada Cuántica de Fourier se define como

$$\boxed{QFT\,|\psi\rangle = \sum_{k=0}^{2^n-1} \tilde{\gamma}_k\,|k\rangle}$$

donde

$$\boxed{\tilde{\gamma}_k = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} \gamma_j \omega^{kj}, \qquad \omega = e^{\frac{2\pi i}{2^n}} = e^{\frac{\pi i}{2^{n-1}}}.} \tag{38.1}$$

Por lo que

$$\begin{aligned} QFT\,|\psi\rangle &= \sum_{k=0}^{2^n-1} \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} \gamma_j \omega^{kj}\,|k\rangle \\ &= \sum_{j=0}^{2^n-1} \gamma_j \left(\frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} \omega^{kj}\,|k\rangle \right) \\ &= \sum_{j=0}^{2^n-1} \gamma_j QFT\,|j\rangle \end{aligned}$$

con

$$\boxed{QFT\,|j\rangle = \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} \omega^{jk}\,|k\rangle\,.}$$

De forma explícita tenemos la transformación

$$\begin{aligned} QFT\,|0\rangle &= \frac{1}{\sqrt{2^n}} \left(|0\rangle + |1\rangle + |2\rangle + \cdots + |2^n-1\rangle\right), \\ QFT\,|1\rangle &= \frac{1}{\sqrt{2^n}} \left(|0\rangle + \omega\,|1\rangle + \omega^2\,|2\rangle + \cdots + \omega^{2^n-1}\,|2^n-1\rangle\right), \\ QFT\,|2\rangle &= \frac{1}{\sqrt{2^n}} \left(|0\rangle + \omega^2\,|1\rangle + \omega^4\,|2\rangle + \cdots + \omega^{2(2^n-1)}\,|2^n-1\rangle\right), \\ &\vdots \\ QFT\,|2^n-1\rangle &= \frac{1}{\sqrt{2^n}} \left(|0\rangle + \omega^{2^n-1}\,|1\rangle + \omega^{2(2^n-1)}\,|2\rangle + \cdots + \omega^{(2^n-1)^2}\,|2^n-1\rangle\right). \end{aligned}$$

Por lo tanto, la transforama cuántica de Fourier se puede escribir como

$$\boxed{QFT\begin{pmatrix}|0\rangle\\|1\rangle\\|2\rangle\\\vdots\\|2^n-1\rangle\end{pmatrix}=\frac{1}{\sqrt{2^n}}\begin{pmatrix}1&1&1&\cdots&1\\1&\omega&\omega^2&\cdots&\omega^{2^n-1}\\1&\omega^2&\omega^4&\cdots&\omega^{2(2^n-1)}\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&\omega^{2^n-1}&\omega^{2(2^n-1)}&\cdots&\omega^{(2^n-1)^2}\end{pmatrix}\begin{pmatrix}|0\rangle\\|1\rangle\\|2\rangle\\\vdots\\|2^n-1\rangle\end{pmatrix}}.$$

A la matriz

$$\boxed{W=\frac{1}{\sqrt{2^n}}\begin{pmatrix}1&1&1&\cdots&1\\1&\omega&\omega^2&\cdots&\omega^{2^n-1}\\1&\omega^2&\omega^4&\cdots&\omega^{2(2^n-1)}\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&\omega^{2^n-1}&\omega^{2(2^n-1)}&\cdots&\omega^{(2^n-1)^2}\end{pmatrix}} \tag{38.2}$$

se le llama matriz de Fourier.

38.3.1. Un qubit

Si tenemos un qubit se cumple $2^1 = 2$, por lo que

$$\omega = e^{\pi i}$$

de donde

$$\boxed{W_2 = H = \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix},}$$

a la matriz H se le llama matriz de Hadamard y es utilizada como una compuerta lógica en diferentes algoritmos de la computación cuántica [15].

Por lo tanto,

$$QFT\begin{pmatrix}|0\rangle\\|1\rangle\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}|0\rangle\\|1\rangle\end{pmatrix},$$

de donde

$$\begin{aligned} QFT\left|0\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|0\right\rangle + \left|1\right\rangle\right), \\ QFT\left|1\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|0\right\rangle - \left|1\right\rangle\right). \end{aligned}$$

Usando la notación de la fracción binaria, podemos ver que

$$\begin{aligned} QFT\left|0\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|0\right\rangle + e^{2\pi i[0,0]}\left|1\right\rangle\right), \\ QFT\left|1\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|0\right\rangle + e^{2\pi i[0,1]}\left|1\right\rangle\right), \end{aligned}$$

es decir

$$\boxed{QFT\left|a_0\right\rangle = \frac{1}{\sqrt{2}}\left(\left|0\right\rangle + e^{2\pi i[0.a_0]}\left|1\right\rangle\right), \qquad a_1 = 0, 1.}$$

Como podemos ver, con la notación binaria la transformada cuántica de Fourier se puede escribir de forma compacta.

38.3.2. Dos qubits

Para dos qubits se tiene $2^2 = 4$, por lo que

$$\omega = e^{\frac{\pi i}{2}}$$

de donde

$$W_4 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}.$$

Por lo tanto

$$QFT\begin{pmatrix} \left|0\right\rangle \\ \left|1\right\rangle \\ \left|2\right\rangle \\ \left|3\right\rangle \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}\begin{pmatrix} \left|0\right\rangle \\ \left|1\right\rangle \\ \left|2\right\rangle \\ \left|3\right\rangle \end{pmatrix},$$

que implica

$$\begin{aligned}
QFT\,|0\rangle &= \frac{1}{2}\left(|0\rangle + |1\rangle + |2\rangle + |3\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle \otimes (|0\rangle + |1\rangle) + |1\rangle \otimes (|0\rangle + |1\rangle)\right) \\
&= \frac{1}{2}\left(|0\rangle + |1\rangle\right) \otimes \left(|0\rangle + |1\rangle\right),
\end{aligned}$$

es decir,

$$QFT\,|0\rangle = \frac{1}{2}\left(|0\rangle + |1\rangle\right) \otimes \left(|0\rangle + |1\rangle\right).$$

Usando la expansión binaria este resultado se puede escribir como

$$QFT\,|00\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,00]}\,|1\rangle\right).$$

Además,

$$\begin{aligned}
QFT\,|1\rangle &= \frac{1}{2}\left(|0\rangle + i\,|1\rangle - |2\rangle - i\,|3\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle \otimes |0\rangle + i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle - i\,|1\rangle \otimes |1\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle \otimes (|0\rangle + i\,|1\rangle) - |1\rangle \otimes (|0\rangle + i\,|1\rangle)\right) \\
&= \frac{1}{2}\left(|0\rangle - |1\rangle\right) \otimes \left(|0\rangle + i\,|1\rangle\right),
\end{aligned}$$

de donde

$$QFT\,|1\rangle = \frac{1}{2}\left(|0\rangle - |1\rangle\right) \otimes \left(|0\rangle + i\,|1\rangle\right).$$

Por lo que usando la expansión binaria, este último resultado se puede escribir como

$$QFT\,|01\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,01]}\,|1\rangle\right).$$

Adicionalmente,

$$\begin{aligned}
QFT\,|2\rangle &= \frac{1}{2}\left(|0\rangle - |1\rangle + |2\rangle - |3\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle \otimes |0\rangle - |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle - |0\rangle \otimes |1\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle \otimes (|0\rangle - |1\rangle) + |1\rangle \otimes (|0\rangle - |1\rangle)\right) \\
&= \frac{1}{2}\left(|0\rangle + |1\rangle\right) \otimes \left(|0\rangle - |1\rangle\right),
\end{aligned}$$

entonces,

$$QFT\,|2\rangle = \frac{1}{2}\left(|0\rangle + |1\rangle\right) \otimes \left(|0\rangle - |1\rangle\right).$$

Usando la expansión binaria, este estado se puede escribir como

$$QFT\,|10\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,10]}\,|1\rangle\right).$$

De la misma forma

$$\begin{aligned}
QFT\,|3\rangle &= \frac{1}{2}\left(|0\rangle - i\,|1\rangle - |2\rangle + i\,|3\rangle\right)\\
&= \frac{1}{2}\left(|0\rangle \otimes |0\rangle - i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle + i\,|1\rangle \otimes |1\rangle\right)\\
&= \frac{1}{2}\left(|0\rangle \otimes (|0\rangle - i\,|1\rangle) - |1\rangle \otimes (|0\rangle - i\,|1\rangle)\right)\\
&= \frac{1}{2}\left(|0\rangle - |1\rangle\right) \otimes \left(|0\rangle - i\,|1\rangle\right),
\end{aligned}$$

así,

$$QFT\,|3\rangle = \frac{1}{2}\left(|0\rangle - |1\rangle\right) \otimes \left(|0\rangle - i\,|1\rangle\right).$$

Con la expansión binaria este resultado se puede escribir como

$$QFT\,|11\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,11]}\,|1\rangle\right).$$

En resumen tenemos las transformaciones

$$\begin{aligned}
QFT\,|0\rangle &= QFT\,|00\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,00]}\,|1\rangle\right),\\
QFT\,|1\rangle &= QFT\,|01\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,01]}\,|1\rangle\right),\\
QFT\,|2\rangle &= QFT\,|10\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,10]}\,|1\rangle\right),\\
QFT\,|3\rangle &= QFT\,|11\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0,11]}\,|1\rangle\right),
\end{aligned}$$

que de forma compacta se puede expresar como

$$\boxed{QFT\,|a_1a_0\rangle = \frac{1}{2}\left(|0\rangle + e^{2\pi i[0.a_0]}\,|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i[0.a_1a_0]}\,|1\rangle\right), \quad a_0, a_1 = 0, 1.}$$

De esta forma la notación binaria nos permite escribri a la transformada cuántica de Fourier de forma simple.

38.3.3. Tres qubits

Para 3 qubits se tiene $2^3 = 8$, entonces

$$\omega = e^{\frac{\pi i}{4}}$$

por lo que la matriz de Fourier es

$$W_8 = \frac{1}{\sqrt{8}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & e^{\frac{\pi i}{4}} & i & e^{\frac{3\pi i}{4}} & -1 & -e^{\frac{\pi i}{4}} & -i & -e^{\frac{3\pi i}{4}} \\ 1 & i & -1 & -i & 1 & i & -1 & -i \\ 1 & e^{\frac{3\pi i}{4}} & -i & e^{\frac{\pi i}{4}} & -1 & -e^{\frac{3\pi i}{4}} & i & -e^{\frac{\pi i}{4}} \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -e^{\frac{\pi i}{4}} & i & -e^{\frac{3\pi i}{4}} & -1 & e^{\frac{\pi i}{4}} & -i & e^{\frac{3\pi i}{4}} \\ 1 & -i & -1 & i & 1 & -i & -1 & i \\ 1 & -e^{\frac{3\pi i}{4}} & -i & -e^{\frac{\pi i}{4}} & -1 & e^{\frac{3\pi i}{4}} & i & e^{\frac{\pi i}{4}} \end{pmatrix}.$$

De esta matriz se obtiene

$$\begin{aligned} QFT\,|0\rangle &= \frac{1}{\sqrt{8}}\left(|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle + |7\rangle\right) \\ &= \frac{1}{\sqrt{8}}\Big(|0\rangle \otimes |0\rangle \otimes |0\rangle + |0\rangle \otimes |0\rangle \otimes |1\rangle + |0\rangle \otimes |1\rangle \otimes |0\rangle \\ &\quad + |0\rangle \otimes |1\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle \otimes |0\rangle + |1\rangle \otimes |0\rangle \otimes |1\rangle \\ &\quad + |1\rangle \otimes |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \otimes |1\rangle \Big) \\ &= \frac{1}{\sqrt{8}}\Big(|0\rangle \otimes (|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle) \\ &\quad + |1\rangle \otimes (|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle) \Big) \\ &= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \Big) \\ &= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle \otimes (|0\rangle + |1\rangle) + |1\rangle \otimes (|0\rangle + |1\rangle) \Big) \\ &= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle + |1\rangle \Big), \end{aligned}$$

es decir,

$$QFT\,|0\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle + |1\rangle \Big).$$

Usando la expansión binaria, este resultado se puede escribir como

$$QFT\,|000\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{2\pi i[0,00]}\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{2\pi i[0,000]}\,|1\rangle \Big).$$

Además,

$$\begin{aligned}
QFT\,|1\rangle &= \frac{1}{\sqrt{8}}\Big(|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle + i\,|2\rangle + e^{\frac{3\pi i}{4}}\,|3\rangle - |4\rangle - e^{\frac{\pi i}{4}}\,|5\rangle - i\,|6\rangle - e^{\frac{3\pi i}{4}}\,|7\rangle\Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle \otimes |0\rangle \otimes |0\rangle + e^{\frac{\pi i}{4}}\,|0\rangle \otimes |0\rangle \otimes |1\rangle + i\,|0\rangle \otimes |1\rangle \otimes |0\rangle \\
&\quad + e^{\frac{3\pi i}{4}}\,|0\rangle \otimes |1\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle \otimes |0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle \otimes |0\rangle \otimes |1\rangle \\
&\quad - i\,|1\rangle \otimes |1\rangle \otimes |0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle \otimes |1\rangle \otimes |1\rangle \Big) \\
&= \frac{1}{2}\Big(|0\rangle \otimes \Big(|0\rangle \otimes |0\rangle + e^{\frac{\pi i}{4}}\,|0\rangle \otimes |1\rangle + i\,|1\rangle \otimes |0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle \otimes |1\rangle\Big) \\
&\quad - |1\rangle \otimes \Big(|0\rangle \otimes |0\rangle + e^{\frac{\pi i}{4}}\,|0\rangle \otimes |1\rangle + i\,|1\rangle \otimes |0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle \otimes |1\rangle\Big)\Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle - |1\rangle \Big) \otimes \Big(|0\rangle \otimes |0\rangle + e^{\frac{\pi i}{4}}\,|0\rangle \otimes |1\rangle + i\,|1\rangle \otimes |0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle \otimes |1\rangle \Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle - |1\rangle \Big) \otimes \Big(|0\rangle \otimes \Big(|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle\Big) + |1\rangle \otimes \Big(i\,|0\rangle + ie^{\frac{\pi i}{4}}\,|1\rangle\Big)\Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle - |1\rangle \Big) \otimes \Big(|0\rangle + i\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle \Big),
\end{aligned}$$

de donde

$$QFT\,|1\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle - |1\rangle \Big) \otimes \Big(|0\rangle + e^{\frac{\pi i}{2}}\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle \Big).$$

que usando la expansión binaria, se puede escribir como

$$QFT\,|001\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + e^{2\pi i[0,1]}\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{2\pi i[0,01]}e^{\frac{\pi i}{2}}\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{2\pi i[0,001]}\,|1\rangle \Big).$$

De la misma forma

$$
\begin{aligned}
QFT\,|2\rangle &= \frac{1}{\sqrt{8}}\left(|0\rangle + i\,|1\rangle - |2\rangle - i\,|3\rangle + |4\rangle + i\,|5\rangle - |6\rangle - i\,|7\rangle\right) \\
&= \frac{1}{\sqrt{8}}\Bigg(|0\rangle \otimes |0\rangle \otimes |0\rangle + i\,|0\rangle \otimes |0\rangle \otimes |1\rangle - |0\rangle \otimes |1\rangle \otimes |0\rangle - i\,|0\rangle \otimes |1\rangle \otimes |1\rangle \\
&\quad + |1\rangle \otimes |0\rangle \otimes |0\rangle + i\,|1\rangle \otimes |0\rangle \otimes |1\rangle - |1\rangle \otimes |1\rangle \otimes |0\rangle - i\,|1\rangle \otimes |1\rangle \otimes |1\rangle \Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(|0\rangle \otimes (|0\rangle \otimes |0\rangle + i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle - i\,|1\rangle \otimes |1\rangle) \\
&\quad + |1\rangle \otimes (|0\rangle \otimes |0\rangle + i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle - i\,|1\rangle \otimes |1\rangle) \Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(|0\rangle + |1\rangle \Bigg) \otimes \Bigg(|0\rangle \otimes |0\rangle + i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle - i\,|1\rangle \otimes |1\rangle \Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(|0\rangle + |1\rangle \Bigg) \otimes \Bigg(|0\rangle \otimes (|0\rangle + i\,|1\rangle) - |1\rangle \otimes (|0\rangle + i\,|1\rangle) \Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(|0\rangle + |1\rangle \Bigg) \otimes \Bigg(|0\rangle - |1\rangle \Bigg) \otimes \Bigg(|0\rangle + i\,|1\rangle \Bigg),
\end{aligned}
$$

por lo que

$$QFT\,|2\rangle = \frac{1}{\sqrt{8}}\Bigg(|0\rangle + |1\rangle \Bigg) \otimes \Bigg(|0\rangle - |1\rangle \Bigg) \otimes \Bigg(|0\rangle + i\,|1\rangle \Bigg),$$

que usando la expansión binaria se puede escribir como

$$QFT\,|010\rangle = \frac{1}{\sqrt{8}}\Bigg(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle \Bigg) \otimes \Bigg(|0\rangle + e^{2\pi i[0,10]}\,|1\rangle \Bigg) \otimes \Bigg(|0\rangle + e^{2\pi i[0,010]}\,|1\rangle \Bigg).$$

Como otro caso tenemos

$$
\begin{aligned}
QFT\,|3\rangle &= \frac{1}{\sqrt{8}}\Big(|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle - i\,|2\rangle + e^{\frac{\pi i}{4}}\,|3\rangle - |4\rangle - e^{\frac{3\pi i}{4}}\,|5\rangle + i\,|6\rangle - e^{\frac{\pi i}{4}}\,|7\rangle\Big) \\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle\otimes|0\rangle\otimes|0\rangle + e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|0\rangle\otimes|1\rangle - i\,|0\rangle\otimes|1\rangle\otimes|0\rangle + \\
&\quad + e^{\frac{\pi i}{4}}\,|0\rangle\otimes|1\rangle\otimes|1\rangle - |1\rangle\otimes|0\rangle\otimes|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\otimes|0\rangle\otimes|1\rangle + \\
&\quad + i\,|1\rangle\otimes|1\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\otimes|1\rangle\,\Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle\otimes\Big(|0\rangle\otimes|0\rangle + e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|1\rangle - i\,|1\rangle\otimes|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\Big) \\
&\quad - |1\rangle\otimes\Big(|0\rangle\otimes|0\rangle + e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|1\rangle - |1\rangle\otimes|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\Big)\Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle\otimes|0\rangle + e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|1\rangle - i\,|1\rangle\otimes|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\,\Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle\otimes\Big(|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle\Big) - i\,|1\rangle\otimes\Big(|0\rangle + ie^{\frac{\pi i}{4}}\,|1\rangle\Big)\Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle\otimes\Big(|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle\Big) - i\,|1\rangle\otimes\Big(|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle\Big)\Bigg) \\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle - i\,|1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle\Bigg),
\end{aligned}
$$

entonces,

$$
QFT\,|3\rangle = \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle - i\,|1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle\Bigg),
$$

que usando la expansión binaria se puede escribir como

$$
QFT\,|011\rangle = \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\Bigg)\otimes\Bigg(\,|0\rangle + e^{2\pi i[0,11]}\,|1\rangle\Bigg)\otimes\Bigg(\,|0\rangle + e^{2\pi i[0,011]}\,|1\rangle\Bigg).
$$

Ahora,

$$\begin{aligned}
QFT\,|4\rangle &= \frac{1}{\sqrt{8}}\left(|0\rangle - |1\rangle + |2\rangle - |3\rangle + |4\rangle - |5\rangle + |6\rangle - |7\rangle\right)\\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle\otimes|0\rangle\otimes|0\rangle - |0\rangle\otimes|0\rangle\otimes|1\rangle + |0\rangle\otimes|1\rangle\otimes|0\rangle - |0\rangle\otimes|1\rangle\otimes|1\rangle\\
&\quad + |1\rangle\otimes|0\rangle\otimes|0\rangle - |1\rangle\otimes|0\rangle\otimes|1\rangle + |1\rangle\otimes|1\rangle\otimes|0\rangle - |1\rangle\otimes|1\rangle\otimes|1\rangle \Big)\\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle\otimes(|0\rangle\otimes|0\rangle - |0\rangle\otimes|1\rangle + |1\rangle\otimes|0\rangle - |1\rangle\otimes|1\rangle)\\
&\quad + |1\rangle\otimes(|0\rangle\otimes|0\rangle - |0\rangle\otimes|1\rangle + |1\rangle\otimes|0\rangle - |1\rangle\otimes|1\rangle) \Big)\\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big)\otimes\Big(|0\rangle\otimes|0\rangle - |0\rangle\otimes|1\rangle + |1\rangle\otimes|0\rangle - |1\rangle\otimes|1\rangle \Big)\\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big)\otimes\Big(|0\rangle\otimes(|0\rangle - |1\rangle) + |1\rangle\otimes(|0\rangle - |1\rangle) \Big)\\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big)\otimes\Big(|0\rangle + |1\rangle \Big)\otimes\Big(|0\rangle - |1\rangle \Big),
\end{aligned}$$

por lo tanto,

$$QFT\,|4\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big)\otimes\Big(|0\rangle + |1\rangle \Big)\otimes\Big(|0\rangle - |1\rangle \Big),$$

que usando la expansión binaria se puede escribir como

$$QFT\,|100\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle \Big)\otimes\Big(|0\rangle + e^{2\pi i[0,00]}\,|1\rangle \Big)\otimes\Big(|0\rangle + e^{2\pi i[0,100]}\,|1\rangle \Big).$$

Otro caso es

$$
\begin{aligned}
QFT\,|5\rangle &= \frac{1}{\sqrt{8}}\Big(|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle + i\,|2\rangle - e^{\frac{3\pi i}{4}}\,|3\rangle - |4\rangle + e^{\frac{\pi i}{4}}\,|5\rangle - i\,|6\rangle + e^{\frac{3\pi i}{4}}\,|7\rangle\Big)\\
&= \frac{1}{\sqrt{8}}\bigg(\,|0\rangle\otimes|0\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|0\rangle\otimes|0\rangle\otimes|1\rangle + i\,|0\rangle\otimes|1\rangle\otimes|0\rangle\\
&\quad -e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|1\rangle\otimes|1\rangle - |1\rangle\otimes|0\rangle\otimes|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle\otimes|0\rangle\otimes|1\rangle\\
&\quad -i\,|1\rangle\otimes|1\rangle\otimes|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle\otimes|1\rangle\otimes|1\rangle\,\bigg)\\
&= \frac{1}{\sqrt{8}}\bigg(\,|0\rangle\otimes\Big(|0\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|0\rangle\otimes|1\rangle + i\,|1\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\Big)\\
&\quad -|1\rangle\otimes\Big(|0\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|0\rangle\otimes|1\rangle + i\,|1\rangle\otimes|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\otimes|1\rangle\Big)\bigg)\\
&= \frac{1}{\sqrt{8}}\bigg(\,|0\rangle - |1\rangle\,\bigg)\otimes\bigg(\,|0\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|0\rangle\otimes|1\rangle + i\,|1\rangle\otimes|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\otimes|1\rangle\,\bigg)\\
&= \frac{1}{\sqrt{8}}\bigg(\,|0\rangle - |1\rangle\,\bigg)\otimes\bigg(\,|0\rangle\otimes\Big(|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\Big) + i\,|1\rangle\otimes\Big(|0\rangle + ie^{\frac{3\pi i}{4}}\,|1\rangle\Big)\bigg)\\
&= \frac{1}{\sqrt{8}}\bigg(\,|0\rangle - |1\rangle\,\bigg)\otimes\bigg(\,|0\rangle\otimes\Big(|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\Big) + i\,|1\rangle\otimes\Big(|0\rangle + ie^{\frac{\pi i}{2}}e^{\frac{\pi i}{4}}\,|1\rangle\Big)\bigg)\\
&= \frac{1}{\sqrt{8}}\bigg(\,|0\rangle - |1\rangle\,\bigg)\otimes\bigg(\,|0\rangle\otimes\Big(|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\Big) + i\,|1\rangle\otimes\Big(|0\rangle + iie^{\frac{\pi i}{4}}\,|1\rangle\Big)\bigg)\\
&= \frac{1}{\sqrt{8}}\bigg(\,|0\rangle - |1\rangle\,\bigg)\otimes\bigg(\,|0\rangle + i\,|1\rangle\,\bigg)\otimes\bigg(\,|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\,\bigg),
\end{aligned}
$$

así,

$$QFT\,|5\rangle = \frac{1}{\sqrt{8}}\bigg(\,|0\rangle - |1\rangle\,\bigg)\otimes\bigg(\,|0\rangle + i\,|1\rangle\,\bigg)\otimes\bigg(\,|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\,\bigg),$$

que usando la expansión binaria se puede escribir como

$$QFT\,|101\rangle = \frac{1}{\sqrt{8}}\bigg(\,|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\,\bigg)\otimes\bigg(\,|0\rangle + e^{2\pi i[0,01]}\,|1\rangle\,\bigg)\otimes\bigg(\,|0\rangle + e^{2\pi i[0,101]}\,|1\rangle\,\bigg).$$

También se encuentra

$$\begin{aligned}
QFT\,|6\rangle &= \frac{1}{\sqrt{8}}\left(|0\rangle - i\,|1\rangle - |2\rangle + i\,|3\rangle + |4\rangle - i\,|5\rangle - |6\rangle + i\,|7\rangle\right) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle \otimes |0\rangle \otimes |0\rangle - i\,|0\rangle \otimes |0\rangle \otimes |1\rangle - |0\rangle \otimes |1\rangle \otimes |0\rangle + i\,|0\rangle \otimes |1\rangle \otimes |1\rangle \\
&\quad + |1\rangle \otimes |0\rangle \otimes |0\rangle - i\,|1\rangle \otimes |0\rangle \otimes |1\rangle - |1\rangle \otimes |1\rangle \otimes |0\rangle + i\,|1\rangle \otimes |1\rangle \otimes |1\rangle \Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle \otimes (|0\rangle \otimes |0\rangle - i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle + i\,|1\rangle \otimes |1\rangle) \\
&\quad + |1\rangle \otimes (|0\rangle \otimes |0\rangle - i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle + i\,|1\rangle \otimes |1\rangle) \Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle \otimes |0\rangle - i\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle + i\,|1\rangle \otimes |1\rangle \Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle \otimes (|0\rangle - i\,|1\rangle) - |1\rangle \otimes (|0\rangle - i\,|1\rangle) \Big) \\
&= \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle - |1\rangle \Big) \otimes \Big(|0\rangle - i\,|1\rangle \Big),
\end{aligned}$$

por lo cual

$$QFT\,|6\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + |1\rangle \Big) \otimes \Big(|0\rangle - |1\rangle \Big) \otimes \Big(|0\rangle - i\,|1\rangle \Big),$$

que usando la expansión binaria, se puede escribir como

$$QFT\,|110\rangle = \frac{1}{\sqrt{8}}\Big(|0\rangle + e^{2\pi i[0,0]}\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{2\pi i[0,10]}\,|1\rangle \Big) \otimes \Big(|0\rangle + e^{2\pi i[0,110]}\,|1\rangle \Big).$$

De la misma forma

$$
\begin{aligned}
QFT\,|7\rangle &= \frac{1}{\sqrt{8}}\Big(|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle - i\,|2\rangle - e^{\frac{\pi i}{4}}\,|3\rangle - |4\rangle + e^{\frac{3\pi i}{4}}\,|5\rangle + i\,|6\rangle + e^{\frac{\pi i}{4}}\,|7\rangle\Big)\\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle\otimes|0\rangle\otimes|0\rangle - e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|0\rangle\otimes|1\rangle - i\,|0\rangle\otimes|1\rangle\otimes|0\rangle\\
&\quad -e^{\frac{\pi i}{4}}\,|0\rangle\otimes|1\rangle\otimes|1\rangle - |1\rangle\otimes|0\rangle\otimes|0\rangle + e^{\frac{3\pi i}{4}}\,|1\rangle\otimes|0\rangle\otimes|1\rangle +\\
&\quad +i\,|1\rangle\otimes|1\rangle\otimes|0\rangle + e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\otimes|1\rangle\,\Bigg)\\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle\otimes\Big(|0\rangle\otimes|0\rangle - e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|1\rangle - i\,|1\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\Big)\\
&\quad - |1\rangle\otimes\Big(|0\rangle\otimes|0\rangle - e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|1\rangle - i\,|1\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\Big)\Bigg)\\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle\otimes|0\rangle - e^{\frac{3\pi i}{4}}\,|0\rangle\otimes|1\rangle - i\,|1\rangle\otimes|0\rangle - e^{\frac{\pi i}{4}}\,|1\rangle\otimes|1\rangle\,\Bigg)\\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle\otimes\Big(|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\Big) - i\,|1\rangle\otimes\Big(|0\rangle - ie^{\frac{\pi i}{4}}\,|1\rangle\Big)\Bigg)\\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle\otimes\Big(|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\Big) - i\,|1\rangle\otimes\Big(|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\Big)\Bigg)\\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle - i\,|1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\,\Bigg),
\end{aligned}
$$

en consecuencia

$$
QFT\,|7\rangle = \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle - |1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle - i\,|1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle - e^{\frac{3\pi i}{4}}\,|1\rangle\,\Bigg),
$$

que usando la expansión binaria se puede escribir como

$$
\begin{aligned}
QFT\,|7\rangle &= QFT\,|111\rangle\\
&= \frac{1}{\sqrt{8}}\Bigg(\,|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle + e^{2\pi i[0,11]}\,|1\rangle\,\Bigg)\otimes\Bigg(\,|0\rangle + e^{2\pi i[0,111]}\,|1\rangle\,\Bigg).
\end{aligned}
$$

Por lo tanto,

$$\begin{aligned}
QFT\,|0\rangle &= QFT\,|000\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,00]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,000]}\,|1\rangle\Big),\\
QFT\,|1\rangle &= QFT\,|001\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,01]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,001]}\,|1\rangle\Big),\\
QFT\,|2\rangle &= QFT\,|010\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,10]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,010]}\,|1\rangle\Big),\\
QFT\,|3\rangle &= QFT\,|011\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,11]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,011]}\,|1\rangle\Big),\\
QFT\,|4\rangle &= QFT\,|100\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,00]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,100]}\,|1\rangle\Big),\\
QFT\,|5\rangle &= QFT\,|101\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,01]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,101]}\,|1\rangle\Big),\\
QFT\,|6\rangle &= QFT\,|110\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,0]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,10]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,110]}\,|1\rangle\Big),\\
QFT\,|7\rangle &= QFT\,|111\rangle\\
&= \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0,1]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,11]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0,111]}\,|1\rangle\Big).
\end{aligned}$$

Con la notación binaria todos estos estados se pueden escribir de la forma compacta

$$\boxed{\begin{array}{l} QFT\,|a_2a_1a_0\rangle = \\ \frac{1}{\sqrt{8}}\Big(\,|0\rangle + e^{2\pi i[0.a_0]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0.a_1a_0]}\,|1\rangle\Big)\otimes\Big(\,|0\rangle + e^{2\pi i[0.a_2a_1a_0]}\,|1\rangle\Big). \end{array}}$$

38.4. Caso general

En general, si tenemos n qubits tendremos 2^n estados en la base del sistema, además los primeros 2^n números se puede escribir con la expansión binariar como

$$\boxed{k = a_{n-1}2^{n-1} + a_{n-2}2^{n-2} + \cdots + a_1 2^1 + a_0 2^0, \qquad a_{n-1}, \cdots, a_0 = 0, 1.}$$

Entonces la transformada cuántica de Fourier se puede escribir como

$$
\begin{aligned}
&QFT\,|k\rangle = QFT\,|a_{n-1}\cdots a_1 a_0\rangle \\
&= \frac{1}{\sqrt{2^n}}\Big(|0\rangle + e^{2\pi i[0.a_0]}\,|1\rangle\Big) \otimes \Big(|0\rangle + e^{2\pi i[0.a_1 a_0]}\,|1\rangle\Big) \otimes \cdots \otimes \Big(|0\rangle + e^{2\pi i[0.a_n\cdots a_1 a_0]}\,|1\rangle\Big).
\end{aligned}
$$

38.5. La inversa de la matriz de Fourier

Se puede observar que las componentes de la matriz de Fourier W (38.2) se pueden escribir como

$$W_{ij} = \frac{1}{\sqrt{2^n}}\omega^{ij}, \qquad i, j = 0, 1, 2, 3, \cdots, 2^n - 1.$$

donde ω está definida en (38.1). Ahora, note que en esta matriz los índices empiezan en cero y teminan en $2^n - 1$, por lo que tenemos una matriz de $(2^n - 1) \times (2^n - 1)$.

Además, considerando la definición de ω dada en (38.1), las componentes de la matriz adjunta de W son

$$
\begin{aligned}
(W^\dagger)_{ij} &= \frac{1}{\sqrt{2^n}}\left(\omega^{ji}\right)^* \\
&= \frac{1}{\sqrt{2^n}}\omega^{-ji},
\end{aligned}
$$

es decir,

$$(W^\dagger)_{ij} = \frac{1}{\sqrt{2^n}}\omega^{-ji}.$$

De donde

$$
\begin{aligned}
(WW^\dagger)_{ij} &= W_{ik}W^\dagger_{kj} \\
&= \frac{1}{\sqrt{2^n}}\omega^{ik}\frac{1}{\sqrt{2^n}}\omega^{-jk} \\
&= \frac{1}{2^n}\omega^{ik-jk} \\
&= \frac{1}{2^n}\omega^{k(i-j)},
\end{aligned}
$$

esto es,

$$(WW^{\dagger})_{ij} \;=\; \frac{1}{2^n}\omega^{k(i-j)}.$$

Recordemos que, de acuerdo con la convención de la suma de Einstein, en esta expresión tenemos una suma sobre k. Así, por razones de claridad, escribiremos

$$(WW^{\dagger})_{ij} \;=\; \frac{1}{2^n}\sum_{k=0}^{2^n-1}\omega^{k(i-j)}. \tag{38.3}$$

Para el caso en que

$$i = j$$

tenemos

$$\begin{aligned}(WW^{\dagger})_{ii} &= \frac{1}{2^n}\sum_{k=0}^{2^n-1}\omega^{k0}\\ &= \frac{1}{2^n}\sum_{k=0}^{2^n-1}1\\ &= \frac{2^n}{2^n},\end{aligned}$$

por lo que

$$(WW^{\dagger})_{ii} \;=\; 1.$$

Mientras que en el caso

$$i \neq j$$

podemos definir

$$z = \omega^{i-j} \tag{38.4}$$

y la suma (38.3) se puede escribir como

$$(WW^{\dagger})_{ij} \;=\; \frac{1}{2^n}\sum_{k=0}^{2^n-1}z^k.$$

Por lo tanto, usando la suma geométrica, tenemos

$$\begin{aligned}(WW^{\dagger})_{ij} &= \frac{1}{N}\sum_{k=0}^{2^n-1}z^k\\ &= \frac{1}{2^n}\frac{z^{2^n}-1}{z-1},\end{aligned}$$

por lo cual

$$(WW^\dagger)_{ij} = \frac{1}{2^n}\frac{z^{2^n}-1}{z-1},$$

Ahora, considerando la definición de z dada en (38.4) y la definición ω dada en (38.1) se encuentra

$$\begin{aligned} z^{2^n} &= \left(\omega^{i-j}\right)^{2^n} \\ &= \left(e^{\frac{2\pi}{2^n}}\right)^{(i-j)2^n} \\ &= e^{\frac{2\pi(i-j)2^n}{2^n}} \\ &= e^{2\pi(i-j)} = 1 \end{aligned}$$

es decir,

$$z^{2^n} = 1.$$

Por lo tanto, si $i \neq j$, se encuentra

$$(WW^\dagger)_{ij} = 0.$$

Estos resultados nos indican que se cumple la igualdad

$$(WW^\dagger)_{ii} = \delta_{ij},$$

en otras palabras

$$WW^\dagger = I.$$

Por lo tanto, se cumple la igualdad

$$\boxed{W^\dagger = W^{-1}}$$

y la matriz de Fourier W es una matriz unitaria.

Así

$$\boxed{W^{-1} = W^\dagger = \frac{1}{\sqrt{2^n}}\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \cdots & \omega^{-(2^n-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \cdots & \omega^{-2(2^n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{-(2^n-1)} & \omega^{-2(2^n-1)} & \cdots & \omega^{-(2^n-1)(2^n-1)} \end{pmatrix}.}$$

Con esta matriz se construye la Inversa de la Tranformada Cuántica de Fourier.

Referencias

[1] E. Lluis-Puebla, *Álgebra lineal,* Primera edición, Sistemas Técnicos de Edición, México (1997).

[2] S. H. Friedberg, A. J. Insel y L. E. Spence, *Álgebra lineal,* Primera edición, Publicaciones Cultural S.A, México (1982).

[3] P. Halmos, *Finite-dimensional vector spaces,* Segunda edición, Dover, New York (2017).

[4] H. A. Rincón, *Álgebra lineal,* Facultad de Ciencias UNAM, México (2015).

[5] J. A. de la Peña, *Álgebra lineal avanzada,* Fondo de Cultura Económica, México (1996).

[6] H. Anton, *Introducción al álgebra lineal,* Limusa, México (2001).

[7] A. M. Sales-Cruz y O. González-Gaxiola, *Introducción al álgebra lineal con Mathematica,* UAM Cuajimalpa, México (2014).

[8] R. Courant y F. John, *Introducción al cálculo y al análisis matemático,* Volumen I y II, Editorial Limusa, México (1990).

[9] M. Reed y B. Simon, *Methods of modern mathematical physics,* Academic Press Inc, New York (1972).

[10] A. N. Kolmogorov, *Elementos de la teoría de funciones y del análisis funcional,* Mir Moscú, México (1975).

[11] D. Luzardo y A. L. Peña, *Historia del álgebra lineal hasta los albores del siglo XX,* Divulgaciones Matemáticas **14** Vol. 4, 153 (2006).

[12] N. Bourbaki, *Elementos de historia de las matemáticas,* Alianza Editorial, Madrid (1976).

[13] M. J. Crowe, *A History of vector analysis,* Dover, New York (1967).

[14] A. D. Aczel, *El último teorema de Fermat, el secreto de un antiguo problema matemático,* Fondo de Cultura Económica, México (2003).

[15] J. D. Hidary, *Quantum computing: an applied approach,* Springer, Switzerland (2019).

[16] P. V. O' Neil, *Matemáticas avanzadas para la ingeniería.* Quinta edición. International Thomson Editores, México (2004).

[17] H. Gardner, *Mentes creativas: una anatomía de la creatividad.* Ediciones Paidós Ibérica, Madrid (1995).

[18] H. Weyl, *Space, Time, Matter.* Dover, USA (1955).

[19] L. F. Magaña, *Las matemáticas y los mayas,* Ciencias 19, (1990). (1955).

Índice Alfabético

Álgebra lineal, con ejercicios en Python, aplicaciones y notas históricas, de Juan Manuel Romero Sanpedro, es una obra que se puede encontrar en la página web del repositorio de la UAM Cuajimalpa Concéntric@.

La asistencia editorial estuvo a cargo de Denise Ocaranza.

Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA