

Preguntas y respuestas de Álgebra Lineal

Domingo López

2020-05-25

Índice general

1	Introducción	1
2	Espacios Vectoriales	3
2.1	Qué es un subespacio vectorial	4
2.2	Qué es un sistema de vectores linealmente independiente	5
2.3	Qué es un sistema generador y el subespacio generado por un conjunto de vectores	6
2.4	Cómo extraemos una base a partir de un sistema generador de un (sub)espacio vectorial	7
2.5	Cómo encontrar la base (y la dimensión) para un subespacio vectorial	10
2.6	Cómo convertir entre ecuaciones cartesianas y paramétricas de un subespacio vectorial	12
2.7	Cómo hallar la intersección de dos subespacios vectoriales	13
2.8	¿Es la unión de subespacios un subespacio?	15
2.9	Cómo hallar la suma de dos subespacios vectoriales	15
2.10	Qué dice el teorema de la dimensión	17
2.11	Cómo calcular el subespacio suplementario de uno dado	17
2.12	Cómo calcular las coordenadas de un vector con respecto a una base dada	19
2.13	Cómo calcular la matriz de cambio de base entre dos bases de un mismo espacio vectorial	20
3	Aplicaciones lineales	25
3.1	Qué propiedades básicas tienen las aplicaciones lineales	26
3.2	Cuál es la matriz asociada a una aplicación lineal en unas bases dadas	26
3.3	Cómo influye un cambio de base en la matriz asociada a una aplicación lineal	30
3.4	Qué es y cómo determinar el núcleo de una aplicación lineal	32
3.5	Cómo podemos identificar una aplicación lineal inyectiva	34
3.6	Cómo calcular el subespacio imagen de una aplicación lineal	36
3.7	Cómo identificar si una aplicación lineal es sobreyectiva	37

3.8	Qué es un isomorfismo	38
3.9	Cómo determinar la imagen de un subespacio vectorial	39
3.10	A qué se llama rango y nulidad de la aplicación lineal	41
3.11	Qué dice el teorema de la dimensión para núcleo e imagen	42
4	Diagonalización	43
4.1	Qué son los autovalores de un endomorfismo y de una matriz	45
4.2	Cómo calculamos los autovectores asociados a un autovalor	48
4.3	Cómo sabemos si un endomorfismo o una matriz es diagonalizable	50
4.4	Si una matriz es diagonalizable, cómo podemos hallar sus potencias	53
4.5	Para qué podemos utilizar el Teorema de Cayley-Hamilton	55
5	Espacios Euclídeos	59
5.1	Qué son las normas y distancias	60
5.2	Qué significa el concepto de ortogonalidad	62
5.3	Qué es el complemento ortogonal de un subespacio	66
5.4	Cómo calculamos la proyección de un vector sobre un subespacio	69
5.5	Qué es y qué utilidad tiene una base ortonormal	72
5.6	Cómo construimos una base ortonormal	74
5.7	Qué es una matriz ortogonal y una aplicación ortogonal	78
5.8	Qué es la diagonalización ortogonal	80
5.9	Qué matrices son diagonalizables ortogonalmente	81
6	Problemas Resueltos	89
6.1	Espacios Vectoriales	89
6.2	Aplicaciones lineales	94
6.3	Diagonalización	98
6.4	Espacios Euclídeos	102

Capítulo 1

Introducción

En este libro vamos a recopilar algunas de las preguntas y dudas más frecuentes en los temas de Álgebra Lineal, con ejemplos resueltos, y acompañados de la teoría necesaria para comprender los resultados.

No pretende ser un libro de teoría exhaustivo, con demostraciones de teoremas, sino que pretende complementar a ese tipo de manuales.

Cada capítulo incluye preguntas y sus respuestas acerca de aquellos procedimientos más usuales relacionados con el tema correspondiente.

La idea de uso de este *libro* es que se pueda navegar libremente, yendo a aquellas preguntas sobre las que se tenga dudas, ya que en su respuesta se incluirán enlaces a todos los contenidos y procedimientos necesarios para poder resolver dicha cuestión.

Además, incluimos un [último capítulo](#) con problemas resueltos, siguiendo las indicaciones dadas en los capítulos correspondientes.

Capítulo 2

Espacios Vectoriales

Supongamos un cuerpo \mathbb{K} (generalmente $\mathbb{K} = \mathbb{R}$). Consideremos un conjunto V dotado de dos operaciones:

- Una operación interna, la *suma*, de forma que si $u, v \in V$ entonces su suma es $u + v \in V$.
- Una operación externa, *producto por un escalar* de \mathbb{K} : si $c \in \mathbb{K}$ y $v \in V$, su producto es $c \cdot v \in V$.

Si:

- V con la operación $+$ es grupo abeliano.
- El producto por un escalar verifica las propiedades distributiva $((c + d) \cdot v = c \cdot v + d \cdot v$ y $c \cdot (u + v) = c \cdot u + c \cdot v$, para $c, d \in \mathbb{K}$, $u, v \in V$), pseudoasociativa $(c \cdot (d \cdot v) = (c \cdot d) \cdot v$ para $c, d \in \mathbb{K}$, $v \in V$) y existencia de neutro ($1 \in \mathbb{K}$, tal que $1 \cdot v = v$ para todo $v \in V$).

Entonces a $(V, +, \cdot)$ se le denomina espacio vectorial, y a los elementos de V , vectores.

En este documento, nos referiremos normalmente a resultados sobre espacios vectoriales denominados *de dimensión finita*. Este concepto lo hablaremos más adelante.

Ejemplo

$V = \mathbb{R}^4$ es un espacio vectorial y $\begin{pmatrix} 2 \\ 1 \\ -2 \\ -1 \end{pmatrix}$ es un vector de V .

¿Qué preguntas vamos a responder en este capítulo?

- ¿Qué es un subespacio vectorial?
- ¿Qué es un sistema de vectores linealmente independiente?
- ¿Qué es un sistema generador y el subespacio generado por un conjunto de vectores?
- ¿Cómo extraemos una base a partir de un sistema generador de un (sub)espacio vectorial?
- ¿Cómo encontrar la base (y la dimensión) para un subespacio vectorial?
- ¿Cómo convertir entre ecuaciones cartesianas y paramétricas de un subespacio vectorial?
- ¿Cómo hallar la intersección de dos subespacios vectoriales?
- ¿Es la unión de subespacios un subespacio?
- ¿Cómo hallar la suma de dos subespacios vectoriales?
- ¿Qué dice el teorema de la dimensión?
- ¿Cómo calcular el subespacio suplementario de uno dado?
- ¿Cómo calcular las coordenadas de un vector con respecto a una base dada?
- ¿Cómo calcular la matriz de cambio de base entre dos bases de un mismo espacio vectorial?

2.1 Qué es un subespacio vectorial

Definición: Subespacio vectorial

Un subconjunto de $U \subset V$ es un **subespacio vectorial** de V si la restricción de las operaciones suma y producto por un escalar a ese conjunto son operaciones cerradas.

Eso equivale a decir que la suma de dos vectores cualesquiera de U debe estar en U , y si multiplicamos por un escalar un vector de U , el producto vuelve a estar en U .

Ejemplo

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : x + y - z = 0 \right\} \text{ es un subespacio de } V = \mathbb{R}^4.$$

2.2 Qué es un sistema de vectores linealmente independiente

Definición: Sistema linealmente independiente

Supongamos un conjunto de vectores $\{v_1, \dots, v_n\}$. Este conjunto es **linealmente independiente** si ninguno de los vectores se puede poner como combinación lineal de los demás vectores.

Esto es equivalente a que si tenemos una combinación lineal igualada a 0, de la forma

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0$$

entonces el sistema de los v_i será independiente si necesariamente todos los α_i valen 0.

Evidentemente, cualquier sistema de vectores que incluya al vector 0 no es linealmente independiente.

Ejemplo

Consideremos el sistema de vectores $\left\{ \begin{pmatrix} 2 \\ 1 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ -8 \\ -3 \end{pmatrix} \right\}$.

¿Cómo podemos saber si es linealmente independiente?

Comenzamos por ver que no está el vector 0. Estudiemos entonces qué pasa si escribimos una combinación lineal de los tres vectores igualada a cero:

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \\ -2 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ 3 \\ -8 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Esto lo podemos reescribir como un sistema de ecuaciones lineales donde las incógnitas son las α_i .

$$\begin{cases} 2\alpha_1 + \alpha_2 + 4\alpha_3 = 0 \\ \alpha_1 + 3\alpha_3 = 0 \\ -2\alpha_1 + \alpha_2 - 8\alpha_3 = 0 \\ -\alpha_1 - 3\alpha_3 = 0 \end{cases}$$

Fijaos en que la matriz de coeficientes está formada por los vectores puestos por *columnas*.

Pueden ocurrir dos casos:

- Si el sistema es compatible *determinado*, entonces su única solución es que todos los $\alpha_i = 0$. Esto, por el comentario anterior, nos dice que el sistema será linealmente independiente.
- Si el sistema es compatible *indeterminado*, entonces admite más soluciones aparte de la trivial. Por tanto, hay una combinación lineal, donde no todos los α_i son nulos, y que da igual al vector 0. Eso nos indica que el sistema no es linealmente independiente.

Lo más sencillo en este caso es usar el método de Gauss para encontrar la forma escalonada de la matriz de coeficientes. Del estudio de sistemas homogéneos, sabemos el siguiente resultado, que se aplica tras tener la forma escalonada:

Teorema

Si el número de ecuaciones (filas) es menor estrictamente que el número de incógnitas (columnas), entonces el sistema es compatible indeterminado. En otro caso (mismo número de ecuaciones que de incógnitas), será compatible determinado.

Usamos entonces Gauss:

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

En este caso, tenemos 2 filas no nulas y 3 incógnitas. Por lo anterior, el sistema es compatible indeterminado, tiene infinitas soluciones, y eso implica que el sistema no es linealmente independiente.

Si hubiéramos llegado a 3 ecuaciones no nulas, entonces estaríamos en la situación de un sistema compatible determinado, luego la única solución es la nula. Por lo anterior, el sistema sería linealmente independiente.

2.3 Qué es un sistema generador y el subespacio generado por un conjunto de vectores

Definición: Subespacio generado

Consideramos un sistema de vectores $\mathcal{S} = \{v_1, \dots, v_n\}$. El **subespacio generado** por \mathcal{S} es el conjunto de vectores que se calcula como combinaciones lineales de los de \mathcal{S} .

Es decir, si vamos dando valores a unos escalares $\alpha_1, \dots, \alpha_n$, entonces el vector $\alpha_1 v_1 + \dots + \alpha_n v_n$ es combinación lineal de los de \mathcal{S} , luego pertenece al subespacio generado por éste.

Como conclusión, el subespacio generado por \mathcal{S} es el menor subespacio de V que contiene a los elementos de \mathcal{S} .

Denotamos como $\mathcal{L}(\mathcal{S})$ al subespacio generado por \mathcal{S} .

Ejemplo

Consideremos el sistema de vectores $\mathcal{S} = \left\{ \begin{pmatrix} 2 \\ 1 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$.

Entonces

$$\mathcal{L}(\mathcal{S}) = \left\{ \alpha_1 \begin{pmatrix} 2 \\ 1 \\ -2 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} : \alpha_i \in \mathbb{R} \right\}$$

Dando valores a α_1 y α_2 tendremos todos los vectores generados en $\mathcal{L}(\mathcal{S})$. Por ejemplo, para

$$\alpha_1 = -1, \quad \alpha_2 = 2 \quad \Rightarrow \quad v = -1 \begin{pmatrix} 2 \\ 1 \\ -2 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 4 \\ 1 \end{pmatrix} \in \mathcal{L}(\mathcal{S})$$

2.4 Cómo extraemos una base a partir de un sistema generador de un (sub)espacio vectorial

Definición: Base

Para cualquier espacio vectorial, o subespacio, una **base** es un sistema de vectores que **lo genera** (mediante combinaciones lineales, como hemos visto antes) y que es **linealmente independiente**.

La **base canónica** \mathcal{C} en un espacio \mathbb{R}^n es la dada por los vectores e_i ($i = 1, \dots, n$), donde e_i es el vector que tiene todas sus componentes a 0 menos la i -ésima que vale 1.

Si tenemos un sistema generador de un (sub)espacio vectorial, siempre podremos encontrar una base, gracias al siguiente resultado:

Teorema

Todo sistema generador de un subespacio contiene un sistema linealmente independiente, que también genera el mismo subespacio.

El concepto de **dimensión** está intrínsecamente relacionado con el de base: es el número de vectores en *cualquier* base del espacio. Hay que tener en cuenta que todas las bases que se puedan encontrar en un mismo espacio vectorial tienen el mismo número de elementos.

Aparte, hay dos propiedades más que relacionan las bases y las propiedades de ser sistema generador o linealmente independiente. Si n es la dimensión de un espacio vectorial:

- Todo sistema de vectores con más de n elementos no puede ser linealmente independiente.
- Todo sistema generador de ese espacio debe tener, al menos, n elementos.

Dado un sistema de vectores S que se sepa que es sistema generador de un espacio o de un subespacio, la mejor estrategia es utilizar transformaciones elementales sobre S para eliminar aquellos vectores que sean combinaciones lineales de los demás.

Ejemplo

Consideremos el sistema de vectores

$$S = \left\{ \begin{pmatrix} -1 \\ 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -5 \\ -1 \\ -3 \\ 0 \end{pmatrix} \right\}$$

Vamos a encontrar un sistema de vectores que sea base de $\mathcal{L}(S)$. Para ellos, vamos a eliminar de S aquellos vectores que sean combinaciones lineales de los demás.

Esa tarea es justo la que realiza el método de Gauss con las filas de una matriz, al ponerla en forma escalonada. Si aplicamos Gauss a una matriz, las filas que acaban siendo de ceros son las que originalmente eran combinación lineal de las demás.

Por tanto, nos aprovechamos de esta propiedad para eliminar vectores linealmente dependientes. Vamos a construir la matriz M resultado de poner los vectores de S por filas (primer vector en primera fila, segundo en segunda fila, etc.).

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

Si aplicamos Gauss para encontrar la forma escalonada (omitimos los detalles intermedios):

$$\begin{pmatrix} -1 & 2 & -1 & -1 \\ -2 & -1 & -1 & -1 \\ -2 & -2 & -1 & 2 \\ -5 & -1 & -3 & 0 \end{pmatrix} \sim \begin{pmatrix} -1 & 2 & -1 & -1 \\ 0 & -5 & 1 & 1 \\ 0 & 0 & -\frac{1}{5} & \frac{14}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Como se puede ver, tenemos que la última fila es de ceros. Por tanto, el vector original del que procede esa fila es linealmente dependiente del resto de vectores de S .

Si hubiéramos tenido que intercambiar filas en algún momento, habría que llevar cuenta de cada intercambio para saber a qué vector original se corresponden aquellas filas que valen 0.

Por tanto, la base del subespacio generado por S , que consiste en únicamente los vectores linealmente independientes de S , es:

$$\mathcal{B}_{\mathcal{L}(S)} = \left\{ \begin{pmatrix} -1 \\ 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \\ -1 \\ 2 \end{pmatrix} \right\}$$

2.5 Cómo encontrar la base (y la dimensión) para un subespacio vectorial

En muchas ocasiones, necesitaremos encontrar una base para un subespacio vectorial, puesto que con las bases tenemos mucha información adicional sobre un subespacio. Además, cada subespacio queda completamente caracterizado una vez dada su base.

Sin embargo, existen otras formas en las que podemos definir un subespacio vectorial, y necesitamos conocer cómo extraer una base de dicho subespacio, para tener una forma *estándar* de describir los subespacios.

2.5.1 Partiendo de las ecuaciones paramétricas

La primera forma escrita en que podemos definir un subespacio vectorial es mediante sus ecuaciones paramétricas, que nos indican cómo determinar las diversas componentes de cada vector (x, y, \dots) a partir de unos parámetros libres $\alpha_1, \dots, \alpha_n$.

Realmente esto es equivalente a escribir un vector genérico v como combinación lineal de otros vectores. Esos otros vectores que forman la combinación lineal son, por tanto, un sistema generador del subespacio que queremos definir. Usando lo explicado en la [sección correspondiente](#), obtendríamos la base deseada.

Ejemplo

Consideremos el subespacio U dado en su forma paramétrica:

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in V : \begin{array}{lcl} x = & -\alpha - 2\beta - 2\gamma - 5\delta \\ y = & 2\alpha - \beta - 2\gamma - \delta \\ z = & -\alpha - \beta - \gamma - 3\delta \\ t = & -\alpha - \beta + 2\gamma \end{array}, \alpha, \beta, \gamma, \delta \in \mathbb{R} \right\}$$

En este caso, lo que debemos hacer es expresar esas ecuaciones paramétricas en forma de combinación lineal de vectores. Para ello, consideramos los coeficientes que acompañan a cada uno de los parámetros $\alpha, \beta, \gamma, \delta$.

Si lo hacemos, nos queda (ya poniéndolo en formato vectorial) que un vector v pertenece a U si se puede expresar como una combinación lineal de la forma:

$$v = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 2 \\ -1 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ -1 \\ -1 \\ -1 \end{pmatrix} + \gamma \begin{pmatrix} -2 \\ -2 \\ -1 \\ 2 \end{pmatrix} + \delta \begin{pmatrix} -5 \\ -1 \\ -3 \\ 0 \end{pmatrix}$$

Es decir, $\mathcal{S} = \left\{ \begin{pmatrix} -1 \\ 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -5 \\ -1 \\ -3 \\ 0 \end{pmatrix} \right\}$ es un sistema generador del

subespacio U . Este proceso no nos asegura que sean vectores linealmente independientes, así que no podemos decir, de momento, que sea una base de U .

A partir de aquí, basta aplicar lo descrito en [la sección sobre cómo extraer una base a partir de un sistema generador](#).

2.5.2 Partiendo de las ecuaciones cartesianas

Un subespacio vectorial puede venir definido mediante unas expresiones cartesianas que relacionan las componentes x, y, \dots entre sí.

Esas expresiones cartesianas realmente expresan un sistema de ecuaciones homogéneo cuyo conjunto solución es el propio subespacio que buscamos.

Por tanto, la forma de proceder es tomar las ecuaciones cartesianas, formar el sistema, y resolverlo. Para ello, haciendo Gauss, pasamos a forma escalonada o escalonada reducida y desde esa forma, calculamos las expresiones paramétricas del conjunto solución y aplicamos lo explicado en la sección anterior.

Ejemplo

Sea el subespacio

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : \begin{array}{l} 3x - y + t = 0 \\ 3x + y - t = 0 \end{array} \right\}$$

El sistema homogéneo dado por su ecuaciones cartesianas es:

$$\begin{cases} 3x & -y & +t & = & 0 \\ 3x & +y & -t & = & 0 \end{cases}$$

Este sistema, resuelto por Gauss, nos da las siguientes expresiones paramétricas:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

A partir de aquí, ya sabemos que estos dos vectores son un sistema generador de U , luego podemos encontrar una base para U si seguimos las explicaciones en [la sección correspondiente](#).

Podemos fijarnos que, en este ejemplo, ya los dos vectores son linealmente independientes, así que ya conforman una base, pero no es la situación general y conviene hacer, para estar seguros, lo mencionado en el párrafo anterior.

Luego, además, podemos saber que la dimensión de U es $\dim(U) = 2$.

2.6 Cómo convertir entre ecuaciones cartesianas y paramétricas de un subespacio vectorial

En la [sección anterior](#), ya hemos visto que para pasar de ecuaciones cartesianas a paramétricas, basta con usar el método de Gauss-Jordan para resolver el sistema de ecuaciones homogéneo.

Para pasar de ecuaciones paramétricas de un subespacio a las ecuaciones cartesianas, se suele utilizar el método de *eliminación de parámetros*.

Si tenemos m parámetros α_i ($i = 1, \dots, m$) y el espacio es de dimensión $n > m$, entonces tendremos, a lo sumo, $n - m$ ecuaciones cartesianas. Buscamos un sistema de ecuaciones lineales homogéneo (las ecuaciones cartesianas) cuya solución, en forma paramétrica sea las ya dadas.

Para ello, comenzamos con la expresión paramétrica con una ecuación por cada coordenada cartesiana del subespacio (una expresión de x en función de los α_i , una para la y , etc.).

Se toma uno de los parámetros, y se despeja de alguna de las ecuaciones en las que aparece. Se sustituye en las demás ecuaciones y deseamos la ecuación de la que hemos despejado el parámetro. En esta situación, tenemos una ecuación menos y un parámetro menos.

Con el nuevo sistema de ecuaciones, repetimos el proceso anterior: seleccionar un parámetro, despejarlo de una ecuación y sustituirlo en las demás, eliminando la ecuación de la que se ha despejado.

Repetimos todo el proceso hasta que no queden parámetros que eliminar.

Ejemplo

Consideramos el subespacio

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in V : \begin{array}{lcl} x = & -\alpha - 2\beta - 2\gamma - 5\delta \\ y = & 2\alpha - \beta - 2\gamma - \delta \\ z = & -\alpha - \beta - \gamma - 3\delta \\ t = & -\alpha - \beta + 2\gamma \end{array}, \alpha, \beta, \gamma, \delta \in \mathbb{R} \right\}$$

Vamos a aplicar el método de eliminación de los parámetros:

- Despejamos α de la ecuación 1, y sustituimos en las demás, lo cual nos da:

$$\begin{array}{rcl} 2x + y & = & -5\beta - 6\gamma - 11\delta \\ -x + z & = & \beta + \gamma + 2\delta \\ -x + t & = & \beta + 4\gamma + 5\delta \end{array}$$

- Despejamos β de la ecuación 1, y sustituimos en las demás, lo cual nos da:

$$\begin{array}{rcl} -\frac{3}{5}x + \frac{1}{5}y + z & = & -\frac{1}{5}\gamma - \frac{1}{5}\delta \\ -\frac{3}{5}x + \frac{1}{5}y + t & = & \frac{14}{5}\gamma + \frac{14}{5}\delta \end{array}$$

- Por último, despejando γ de la ecuación 1, nos queda:

$$-9x + 3y + 14z + t = 0$$

que es la expresión en cartesianas del espacio U .

2.7 Cómo hallar la intersección de dos subespacios vectoriales

Si tenemos dos subespacios, U y W de V , podemos calcular su intersección, que es el conjunto de vectores que pertenecen a ambos subespacios a la vez.

Teorema

La intersección de dos subespacios es un nuevo subespacio, aunque puede suceder que la intersección sea únicamente el vector 0.

Conviene tener ambos subespacios expresados en forma de ecuaciones cartesianas. Si no estuviera así (porque solo se dé su base o, lo que es equivalente, las ecuaciones paramétricas), basta ver la sección sobre [cómo convertir de ecuaciones paramétricas a cartesianas](#).

Si un vector v pertenece a la intersección, entonces verifica tanto las ecuaciones paramétricas de U como las de W . Por tanto, podemos construir un sistema de ecuaciones con todas las ecuaciones de U y de W , y la intersección $U \cap W$ será el conjunto solución de este sistema homogéneo *ampliado*.

A partir de ese sistema, ya hemos explicado [cómo encontrar la base](#) del subespacio intersección.

Ejemplo

Consideremos los subespacios siguientes:

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : \begin{array}{l} 3x - y + t = 0 \\ 3x + y - t = 0 \end{array} \right\}$$

$$W = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : 3x + z + 3t = 0 \right\}$$

El subespacio U ya lo hemos estudiado [antes](#). Para calcular la intersección $U \cap W$, construimos un sistema de ecuaciones con todas las ecuaciones de U y las de W . La intersección es el conjunto solución del sistema homogéneo resultante:

$$\begin{cases} 3x & -y & & +t & = & 0 \\ 3x & +y & & -t & = & 0 \\ 3x & & +z & +3t & = & 0 \end{cases}$$

Aplicamos lo que hemos visto [antes](#) para estudiar este sistema.

Este sistema, resuelto por Gauss, nos da las siguientes expresiones paramétricas:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 1 \\ -3 \\ 1 \end{pmatrix}$$

Por tanto, una base de $U \cap W$ será:

$$\mathcal{B}_{U \cap W} = \left\{ \begin{pmatrix} 0 \\ 1 \\ -3 \\ 1 \end{pmatrix} \right\}$$

Y, así, $\dim(U \cap W) = 1$.

2.8 ¿Es la unión de subespacios un subespacio?

En general, no. La única forma en que la unión de dos subespacios sea también un subespacio es que sean iguales o que, al menos, uno de ellos sea el espacio $\{0\}$ o el total, V .

2.9 Cómo hallar la suma de dos subespacios vectoriales

Aparte de la [intersección](#) y de la [unión](#) de subespacios vectoriales, podemos definir una nueva operación:

Definición: Suma de subespacios

Dados dos subespacios U y W de V , se define su **suma** $U + W$ como el conjunto de vectores que son combinación lineal de un vector de U y otro de W .

Si, además, $U \cap W = \{0\}$ y $U + W = V$, entonces se llama **suma directa** y se denota por $U \oplus W$. Se dice entonces que U es el **subespacio suplementario** de W y viceversa.

El resultado teórico esencial es el siguiente:

Teorema

Si \mathcal{B}_U y \mathcal{B}_W son bases de U y W respectivamente, entonces $\mathcal{G} = \mathcal{B}_U \cup \mathcal{B}_W$ es un sistema generador de $U + W$.

Por tanto, teniendo bases de U y de W , lo primero será unir las para obtener un sistema generador de la suma. Después, usaremos la [técnica para conseguir una base de un subespacio a partir de un sistema generador](#).

Ejemplo

Consideramos los mismos espacios U y W del ejemplo de la [sección sobre intersección](#).

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : \begin{array}{l} 3x - y + t = 0 \\ 3x + y - t = 0 \end{array} \right\}$$

$$W = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : 3x + z + 3t = 0 \right\}$$

Podemos aplicar la [técnica para conseguir una base a partir de las ecuaciones cartesianas](#), y llegamos a las siguientes bases:

$$\mathcal{B}_U = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

$$\mathcal{B}_W = \left\{ \begin{pmatrix} -\frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Ahora, tomando

$$\mathcal{G} = \mathcal{B}_U \cup \mathcal{B}_W = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

que es un sistema generador de la suma, podemos construir una base de $U + W$ [eliminando los vectores de \$\mathcal{G}\$ que sean combinaciones lineales de los demás](#), lo cual requiere usar el método de Gauss para obtener la forma escalonada reducida de ese sistema de vectores.

De esta forma, encontramos una base de $U + W$:

$$\mathcal{B}_{U+W} = \left\{ \begin{pmatrix} -\frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

Y, así, tenemos que $\dim(U + W) = 4$.

2.10 Qué dice el teorema de la dimensión

Hay un resultado teórico que relaciona las dimensiones de dos subespacios U y W con las de su intersección y su suma:

Teorema (Fórmula de Grassmann)

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Ejemplo

Si consideramos los espacios U y W que hemos usado como ejemplos en [las secciones anteriores](#), podemos comprobar fácilmente el teorema de la dimensión:

$$\begin{array}{ccccccc} \dim(U + W) & = & \dim(U) & + & \dim(W) & - & \dim(U \cap V) \\ 4 & = & 2 & + & 3 & - & 1 \end{array}$$

2.11 Cómo calcular el subespacio suplementario de uno dado

Como hemos comentado antes, el subespacio suplementario de un subespacio U es aquel W que verifica:

- $U \cap W = \{0\}$
- $U \oplus W = V$

Supongamos que tenemos una base \mathcal{B}_U de U . Para determinar W , basta con que demos una forma de construir una base suya.

El primer paso será encontrar una forma *sencilla* para la base de U . Realmente, queremos encontrar una base de U a partir de la que ya tenemos, donde los vectores formen un sistema escalonado reducido. El método más indicado es usar Gauss-Jordan sobre la matriz que forman los vectores de \mathcal{B}_U por *filas*.

Una vez que tenemos la forma escalonada reducida, nos fijamos en qué posiciones se encuentran los pivotes en esa matriz, es decir, las posiciones del primer elemento distinto de cero de cada fila.

Tomamos entonces, para formar la base de W , aquellos vectores de la base canónica que **no** tengan un 1 en esas posiciones. Los vectores así determinados nos dan la base de W , el subespacio suplementario, pues nos aseguramos las dos propiedades mencionadas más arriba.

Ejemplo

Consideremos el subespacio U cuya base es:

$$\mathcal{B}_U = \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Si no tuviéramos su base, sino que U viene dado de otra forma, [ya sabemos cómo encontrar la base](#).

Ponemos los vectores de la base de U en *filas* y aplicamos el método de Gauss (o Gauss-Jordan) para encontrar la forma escalonada de la matriz:

$$\begin{pmatrix} 2 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 1 & -1 \\ 0 & -1 & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}$$

Al inspeccionar la matriz escalonada, podemos ver que sus pivotes (primer elemento no nulo de cada fila) están en las columnas $\{1, 2\}$.

Consideramos la base canónica

$$\mathcal{C} = \{e_i\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Nos vamos a quedar con aquellos vectores de \mathcal{C} que no tengan un 1 en las posiciones de los pivotes ($\{1, 2\}$). De esta forma, lo que tenemos es una base de W , el subespacio suplementario de U :

$$\mathcal{B}_W = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

2.12 Cómo calcular las coordenadas de un vector con respecto a una base dada

Un concepto importante es el de **coordenadas** de un vector, asociadas a una base \mathcal{B} del espacio vectorial V .

Definición: Coordenadas

Supongamos una base $\mathcal{B} = \{v_1, \dots, v_n\}$ de V . Las **coordenadas** de un vector v son unos escalares $\alpha_1, \dots, \alpha_n$ tales que podemos expresar v como una combinación lineal de los vectores de la base con esos coeficientes:

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n$$

Generalmente, como venimos haciendo desde el principio, estamos denotando e identificando a un vector con sus coordenadas, de forma que escribimos:

$$v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}_{\mathcal{B}}$$

Cuando la base \mathcal{B} es la canónica, no es necesario indicar el nombre de la base con respecto a la cual tomamos las coordenadas, a menos que lleve a confusión.

Dado un vector con sus coordenadas en la base canónica, ¿cómo puedo saber sus coordenadas en otra base \mathcal{B} ? Tenemos dos vías, una que exploramos aquí y otra que veremos en [la sección siguiente](#).

El determinar las coordenadas de un vector en una base distinta de la canónica es realmente equivalente a la resolución de un sistema de ecuaciones lineales, donde la matriz de coeficientes es la formada por los vectores de la nueva base *por columnas*.

Lo vemos directamente con un ejemplo.

Ejemplo

Consideremos el vector v dado por sus coordenadas en la base canónica $\begin{pmatrix} -5 \\ 1 \\ 3 \\ -2 \end{pmatrix}$.

Consideremos también la base:

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

Unos escalares a, b, c, d serán las coordenadas de v en la base \mathcal{B}_1 si se verifica:

$$\begin{pmatrix} -5 \\ 1 \\ 3 \\ -2 \end{pmatrix} = a \begin{pmatrix} -1 \\ 1 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix} + c \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} + d \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

que se puede reescribir como el siguiente sistema de ecuaciones lineales (donde la matriz de coeficientes resulta de poner los vectores de la base \mathcal{B}_1 por columnas):

$$\begin{cases} -a & & +d & = & -5 \\ a & -b & & +d & = & 1 \\ & b & +c & -d & = & 3 \\ -a & -b & -c & -d & = & -2 \end{cases}$$

Este sistema se puede resolver usando métodos Gaussianos, y nos da:

$$v = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}_{\mathcal{B}_1} = \begin{pmatrix} 3 \\ 0 \\ 1 \\ -2 \end{pmatrix}_{\mathcal{B}_1}$$

2.13 Cómo calcular la matriz de cambio de base entre dos bases de un mismo espacio vectorial

Supongamos dos bases \mathcal{B}_1 y \mathcal{B}_2 de V , y que tenemos un vector v , expresado en coordenadas de \mathcal{B}_1 . Nos planteamos una forma sencilla de calcular las coordenadas

del mismo vector v respecto a la otra base \mathcal{B}_2 .

El principal resultado teórico en el que nos basamos (y del que alguna pincelada hemos dado en [la sección anterior](#)) es el que sigue:

Teorema

La transformación de las coordenadas para un mismo vector v entre dos bases \mathcal{B}_1 y \mathcal{B}_2 viene dada por la siguiente expresión:

$$[v]_{\mathcal{B}_2} = A[v]_{\mathcal{B}_1}$$

donde A es una matriz cuadrada formada por las coordenadas de cada vector de \mathcal{B}_1 en la base \mathcal{B}_2 y $[v]_{\mathcal{B}_i}$ indica las coordenadas de v en la base \mathcal{B}_i .

Notación: A la matriz de cambio entre la base \mathcal{B}_1 y \mathcal{B}_2 , se suele denotar por $P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$.

Para calcular $P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$, para cada vector de \mathcal{B}_1 vamos a encontrar sus coordenadas en la base \mathcal{B}_2 .

Siguiendo las explicaciones acerca de [cómo calcular coordenadas de un vector respecto de una base](#), queremos expresar cada v_i como combinación lineal de los v'_i :

$$v_i = a_{1,i}v'_1 + \dots + a_{n,i}v'_n \quad i = 1, \dots, n$$

En este caso, los valores $a_{1,i}, \dots, a_{n,i}$ son las coordenadas de v_i en la base \mathcal{B}_2 , y podemos escribir, para todo $i = 1, \dots, n$:

$$[v_i]_{\mathcal{B}_2} = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{n,i} \end{pmatrix}$$

Según el resultado anterior, la matriz de cambio de base $A = P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$ tendrá por columnas estas coordenadas previamente calculadas:

$$A = ([v_1]_{\mathcal{B}_2} | \dots | [v_n]_{\mathcal{B}_2})$$

Llamemos $B_1 = (v_1 | v_2 | \dots | v_n)$ y $B_2 = (v'_1 | v'_2 | \dots | v'_n)$ a las dos matrices que resultan de poner los vectores de las dos bases que tenemos, en forma de columna.

Entonces, resulta que la combinación lineal de más arriba que relaciona un v_i con los vectores de \mathcal{B}_2 mediante sus coordenadas se puede reescribir como:

$$v_i = B_2 [v_i]_{\mathcal{B}_2}$$

Si vamos concatenando los v_i en la izquierda, hasta formar la matriz B_1 , en la derecha debemos ir haciendo lo mismo con sus coordenadas en la base \mathcal{B}_2 , y entonces tenemos:

$$(v_1|v_2|\dots|v_n) = B_2 ([v_1]_{\mathcal{B}_2}|[v_2]_{\mathcal{B}_2}|\dots|[v_n]_{\mathcal{B}_2})$$

es decir,

$$B_1 = B_2 P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$$

Como nos interesa una forma fácil de calcular $P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$, podemos despejarla de la expresión anterior, luego tenemos $P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2} = B_2^{-1} B_1$.

El método práctico para calcular esta matriz de cambio de base, por tanto, será usar Gauss-Jordan, comenzando por la matriz $(B_2|B_1)$ hasta dejarla en la forma $(I_n|A)$. En ese momento, en la derecha, tendremos $A = P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$.

Una vez que tenemos determinada la matriz $A = P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$ que nos proporciona el cambio de base entre \mathcal{B}_1 y \mathcal{B}_2 , el cambio inverso, es decir, entre \mathcal{B}_2 y \mathcal{B}_1 , tiene como matriz asociada A^{-1} . Recordemos que la inversa de una matriz A se puede calcular mediante el método de Gauss-Jordan, comenzando por $(A|I_n)$ y reduciéndola a la forma $(I_n|A^{-1})$.

Caso particular: Cambio de base a y desde la base canónica

Supongamos que una de las bases, por ejemplo $\mathcal{B}_2 = \mathcal{C}$ es la base canónica. La matriz que forman sus vectores por columna en este caso sería $B_2 = I_n$, la identidad.

Aplicando lo desarrollado, tendríamos que la matriz de cambio de base de \mathcal{B}_1 a la canónica es: $P_{\mathcal{B}_1 \rightarrow \mathcal{C}} = P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2} = B_2^{-1} B_1 = I_n B_1 = B_1$.

Y entonces $P_{\mathcal{C} \rightarrow \mathcal{B}_1} = P_{\mathcal{B}_2 \rightarrow \mathcal{B}_1} = B_1^{-1} B_2 = B_1^{-1} I_n = B_1^{-1}$.

Ejemplo

Consideremos dos bases:

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix} \right\}$$

Vamos a hallar la matriz del cambio de base entre ambas bases, en ambos sentidos.

Comenzamos por hallar la matriz A de cambio de base de \mathcal{B}_1 a \mathcal{B}_2 .

Construimos las matrices que tienen los vectores de \mathcal{B}_1 y \mathcal{B}_2 por columnas:

$$B_1 = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ -1 & -1 & -1 & -1 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} 2 & 0 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix}$$

Como se ha mencionado antes, para hallar A , podemos hacer Gauss-Jordan sobre $(B_2|B_1)$ hasta llegar a $(I_n|A)$:

$$\left(\begin{array}{cccc|cccc} 2 & 0 & 2 & 1 & -1 & 0 & 0 & 1 \\ 2 & 2 & 2 & 1 & 1 & -1 & 0 & 1 \\ 0 & 0 & 2 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 2 & 2 & -1 & -1 & -1 & -1 \end{array} \right) \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -\frac{1}{4} & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & -1 & -1 & 0 \end{array} \right)$$

Luego

$$P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2} = \begin{pmatrix} -\frac{1}{4} & 0 & 0 & 1 \\ 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -1 & -1 & 0 \end{pmatrix}$$

Para el camino contrario, pasar de coordenadas en \mathcal{B}_2 a coordenadas en \mathcal{B}_1 , hacemos la inversa de la matriz anterior (aunque podríamos haber aplicado el mismo razonamiento inicial):

$$\left(\begin{array}{cccc|cccc} -\frac{1}{4} & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & -\frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 1 & 0 \\ -\frac{1}{2} & -1 & -1 & 0 & 0 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -\frac{4}{3} & 0 & -\frac{8}{3} & -\frac{4}{3} \\ 0 & 1 & 0 & 0 & -\frac{8}{3} & -2 & -\frac{16}{3} & -\frac{8}{3} \\ 0 & 0 & 1 & 0 & \frac{10}{3} & 2 & \frac{20}{3} & \frac{7}{3} \\ 0 & 0 & 0 & 1 & \frac{2}{3} & 0 & -\frac{2}{3} & -\frac{1}{3} \end{array} \right)$$

De esta forma

$$P_{\mathcal{B}_2 \rightarrow \mathcal{B}_1} = \begin{pmatrix} -\frac{4}{3} & 0 & -\frac{8}{3} & -\frac{4}{3} \\ -\frac{8}{3} & -2 & -\frac{16}{3} & -\frac{8}{3} \\ \frac{10}{3} & 2 & \frac{20}{3} & \frac{7}{3} \\ \frac{2}{3} & 0 & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

Para completar el ejemplo, vemos el cambio de base entre \mathcal{B}_1 y \mathcal{C} :

$$P_{\mathcal{B}_1 \rightarrow \mathcal{C}} = B_1 = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ -1 & -1 & -1 & -1 \end{pmatrix}$$

Y entre \mathcal{C} y \mathcal{B}_1 (calculando la inversa de B_1 usando Gauss-Jordan):

$$P_{\mathcal{C} \rightarrow \mathcal{B}_1} = B_1^{-1} = \begin{pmatrix} -\frac{2}{3} & 0 & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -1 & -\frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & 1 & \frac{4}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}$$

Capítulo 3

Aplicaciones lineales

Consideremos dos [espacios vectoriales](#) V y W sobre el mismo cuerpo \mathbb{K} , y una aplicación $f : V \rightarrow W$ entre ellos.

Diremos que f es una **aplicación lineal** si verifica las siguientes dos condiciones:

- $f(v_1 + v_2) = f(v_1) + f(v_2)$ para todo $v_1, v_2 \in V$.
- $f(c \cdot v) = c \cdot f(v)$ para todo $c \in \mathbb{K}$, $v \in V$.

Una aplicación lineal cuyo dominio y codominio sean iguales ($V = W$) se llama **endomorfismo**.

Ejemplo

Existen muchos ejemplos de aplicaciones lineales:

- La aplicación 0, la identidad... son aplicaciones lineales de forma trivial.
- Las rotaciones de vectores en \mathbb{R}^n son aplicaciones lineales.
- Sobre el espacio vectorial $\mathcal{C}([a, b])$ de funciones continuas y derivables en un intervalo $[a, b]$, tanto el operador derivada $f \rightarrow f'$ como el operador integral $f \rightarrow \int_a^b f$, son aplicaciones lineales.

¿Qué preguntas vamos a responder en este capítulo?

- ¿Qué propiedades básicas tienen las aplicaciones lineales?
- ¿Cuál es la matriz asociada a una aplicación lineal en unas bases dadas?
- ¿Cómo influye un cambio de base en la matriz asociada a una aplicación lineal?
- ¿Qué es y cómo determinar el núcleo de una aplicación lineal?
- ¿Cómo podemos identificar una aplicación lineal inyectiva?

- ¿Cómo calcular el subespacio imagen de una aplicación lineal?
- ¿Cómo identificar si una aplicación lineal es sobreyectiva?
- ¿Qué es un isomorfismo?
- ¿Cómo determinar la imagen de un subespacio vectorial?
- ¿A qué se llama rango y nulidad de la aplicación lineal?
- ¿Qué dice el teorema de la dimensión para núcleo e imagen?

3.1 Qué propiedades básicas tienen las aplicaciones lineales

Vamos a revisar aquí algunas de las propiedades básicas de las aplicaciones lineales entre espacios vectoriales.

Consideremos una aplicación lineal $f : V \rightarrow W$.

- $f(0_V) = 0_W$
- $f(-v) = -f(v)$, $v \in V$
- $f(v_1 - v_2) = f(v_1) - f(v_2)$, $v_1, v_2 \in V$
- $f(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 f(v_1) + \dots + \alpha_n f(v_n)$, $\alpha_i \in \mathbb{K}, v_i \in V$

Además:

Teorema

- Si U es un subespacio de V , entonces $f(U)$ es un subespacio de W . Análogamente, si S es subespacio de W , su preimagen $f^{-1}(S)$ es también un subespacio de V .
- Si $\mathcal{G} = \{u_1, \dots, u_m\}$ es un sistema generador de un subespacio U de V , entonces $f(\mathcal{G}) = \{f(u_1), \dots, f(u_m)\}$ es un sistema generador de $f(U)$.

Y, por último,

Si $f : V \rightarrow W$ y $g : W \rightarrow U$ son dos aplicaciones lineales, entonces su composición $g \circ f : V \rightarrow U$ también es una aplicación lineal.

3.2Cuál es la matriz asociada a una aplicación lineal en unas bases dadas

Teorema

Si tenemos una matriz $A \in \mathcal{M}_{mn}(\mathbb{R})$, podemos definir una aplicación lineal f_A entre \mathbb{R}^n y \mathbb{R}^m de la forma:

$$f_A(v) = A \cdot v, \quad v \in \mathbb{R}^n$$

En este resultado, podemos cambiar el cuerpo \mathbb{R} de los números reales por cualquier otro cuerpo.

Este resultado nos habla de una forma de construir aplicaciones lineales, mediante el producto por una matriz que tenga las dimensiones adecuadas.

Un resultado más interesante es el siguiente, que nos dice que TODA aplicación lineal viene definida por el producto con una matriz:

Teorema (Matriz asociada a una aplicación lineal)

Sea $f : V \rightarrow W$ una aplicación lineal. Fijemos dos bases \mathcal{B}_V y \mathcal{B}_W en dichos espacios, respectivamente. Entonces existe una matriz $A \in \mathcal{M}_{mn}(\mathbb{R})$ tal que si $w = f(v)$ entonces $[w]_{\mathcal{B}_W} = A \cdot [v]_{\mathcal{B}_V}$ para todo $v \in V$.

Esto significa que, fijando las bases, podemos encontrar una matriz A de forma que para calcular la imagen de un vector $v \in V$, basta con tomar sus **coordenadas** en la base \mathcal{B}_V , que representamos por $[v]_{\mathcal{B}_V}$, y multiplicarlas por A , para obtener las coordenadas de $f(v)$ en la base correspondiente del *codominio* W , las cuales se denotan por $[f(v)]_{\mathcal{B}_W}$.

Definición: Matriz asociada a aplicación lineal

A la matriz A que nos proporciona este resultado la llamamos **matriz asociada** a f en las **bases** \mathcal{B}_V y \mathcal{B}_W .

Hay que notar que la matriz A está determinada siempre y cuando fijemos las dos bases antes dichas.

Pero, ¿cómo determinamos la matriz A ?

Necesitamos fijar las dos bases de antemano:

$$\mathcal{B}_V = \{v_1, \dots, v_n\}$$

$$\mathcal{B}_W = \{w_1, \dots, w_m\}$$

El procedimiento general para construir la matriz A es el que sigue. Para determinar la columna j -ésima de A :

- Tomamos el vector j -ésimo de la base de V : v_j .
- Calculamos su imagen $y_j = f(v_j)$.

- Calculamos las coordenadas de y_j en la base \mathcal{B}_W .

$$[y_j]_{\mathcal{B}_W} = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{pmatrix} \Leftrightarrow y_j = a_{1,j}w_1 + a_{2,j}w_2 + \dots + a_{m,j}w_m$$

- Esas coordenadas son la columna j de la matriz A .

Vamos a pasar todo esto a notación matricial, para encontrar una forma cómoda de resolver este problema.

Llamemos B_2 a la matriz formada, por columnas, por las coordenadas de los vectores de \mathcal{B}_W en la base canónica: $B_2 = (w_1|w_2|\dots|w_m)$.

Llamamos $Y = (y_1|y_2|\dots|y_n) = (f(v_1)|f(v_2)|\dots|f(v_n))$.

Entonces, la expresión anterior se puede escribir matricialmente como $Y = B_2 \cdot A$, siendo A la matriz asociada a la aplicación lineal. Por tanto, podemos despejar $A = B_2^{-1}Y$.

Por tanto, para calcular A de forma directa, podemos realizar Gauss-Jordan partiendo de $(B_2|Y)$ hasta lograr $(I_m|A)$.

Caso particular: Base canónica en V y W

En este caso, la matriz asociada a la aplicación lineal es

$$A = Y = (f(v_1)|f(v_2)|\dots|f(v_n))$$

puesto que los vectores de la base canónica puestos por columnas forman la matriz identidad.

De hecho, en la situación de la base canónica, lo más rápido es fijarse en los coeficientes de cada una de las variables x, y, z, \dots o x_1, x_2, \dots , según sea el caso, y ponerlos por columnas. Así, para formar la matriz A asociada a la aplicación lineal, tomaríamos como primera columna los coeficientes de la x , en la segunda, los coeficientes de la y , y así sucesivamente con el resto de variables.

Ejemplo

Consideremos la aplicación $f : V = \mathbb{R}^3 \rightarrow W = \mathbb{R}^3$ dada por:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + z \\ x - 2y \\ 2x - y + z \end{pmatrix}$$

Como hemos visto, para encontrar su matriz asociada en las bases canónicas de v y W , basta con mirar los coeficientes de x, y, \dots , y ponerlos por columnas:

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

Esto significa que si tenemos un vector $v \in V$ expresado en coordenadas en la base canónica, entonces $f(v) = A v$, también expresado en las coordenadas canónicas de W .

Consideremos ahora dos bases distintas de la canónica tanto en V como W :

$$\mathcal{B}_1 = \{v_i\} = \left\{ \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

$$\mathcal{B}_2 = \{w_i\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Vamos a calcular la matriz A' asociada a f en estas dos bases. Debemos calcular:

$$Y = (f(v_1) | \dots | f(v_n))$$

$$B_2 = (w_1 | \dots | w_m)$$

Aplicando f sucesivamente a los vectores de \mathcal{B}_1 , tenemos que

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

y

$$B_2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Luego podemos obtener A' como $B_2^{-1}Y$, que resolvemos por Gauss-Jordan:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & -1 & 0 & -3 \\ 0 & 2 & 0 & -3 & 2 & 1 \\ 2 & 0 & 1 & -4 & 2 & -2 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{9}{2} & 3 & \frac{3}{2} \\ 0 & 1 & 0 & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & 5 & -4 & -5 \end{array} \right)$$

Y de aquí que la nueva matriz, asociada a f en las nuevas bases, sea:

$$A' = \begin{pmatrix} -\frac{9}{2} & 3 & \frac{3}{2} \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 5 & -4 & -5 \end{pmatrix}$$

3.3 Cómo influye un cambio de base en la matriz asociada a una aplicación lineal

Supongamos que tenemos una aplicación lineal $f : V \rightarrow W$ y que tenemos prefijadas unas bases \mathcal{B}_V y \mathcal{B}_W y, por tanto, tenemos determinada la matriz A , asociada a f en esas bases.

Consideremos unas nuevas bases \mathcal{B}'_V y \mathcal{B}'_W y nos preguntamos cuál será la matriz A' asociada a f en dichas bases.

Para saber hallar A' , nos basamos en dos resultados teóricos.

El primero de ellos nos relaciona los [cambios de base](#) con las aplicaciones lineales:

Teorema (Cambio de base como aplicación lineal)

Un **cambio de base** entre dos bases \mathcal{B}_V y \mathcal{B}'_V en un espacio vectorial V se corresponde con la aplicación lineal **identidad** $\text{id}_V : V \rightarrow V$, donde en el dominio hemos considerado la base \mathcal{B}_V y en el codominio la base \mathcal{B}'_V . Además, la matriz asociada a id_V en esta situación es justo la del cambio de base $P_{\mathcal{B}_V \rightarrow \mathcal{B}'_V}$.

Es decir, podemos considerar que un cambio de base no es más que la aplicación de la función identidad en el espacio vectorial, donde en dominio y codominio hemos considerado bases distintas.

El otro resultado teórico nos relaciona la composición de aplicaciones lineales (que [es de nuevo una aplicación lineal](#)) con sus matrices asociadas:

Teorema (Composición de aplicaciones lineales)

Si $f : V \rightarrow W$ y $g : W \rightarrow U$ son dos aplicaciones lineales, con matrices asociadas A_f y A_g , respectivamente, habiendo fijado bases en V , W y U , entonces la aplicación lineal composición de f y g , $g \circ f : V \rightarrow U$, tiene como matriz asociada a $A_{g \circ f} = A_g \cdot A_f$.

Luego si componemos dos aplicaciones lineales, *habiendo fijado bases en cada caso*, la matriz asociada a la composición es el producto de las matrices individuales.

Con estos dos resultados, ya estamos en condiciones de estudiar cómo influye un cambio de base en la matriz de la aplicación lineal.

Veamos el siguiente esquema, donde denotamos $V_{\mathcal{B}}$ que estamos considerando la base \mathcal{B} en el espacio V :

$$\begin{array}{ccccccc} V_{\mathcal{B}'_V} & \xrightarrow{\text{id}_V} & V_{\mathcal{B}_V} & \xrightarrow{f} & W_{\mathcal{B}_W} & \xrightarrow{\text{id}_W} & W_{\mathcal{B}'_W} \\ & & P_{\mathcal{B}'_V \rightarrow \mathcal{B}_V} & & A & & P_{\mathcal{B}_W \rightarrow \mathcal{B}'_W} \end{array}$$

Llamemos, por comodidad, $P = P_{\mathcal{B}'_V \rightarrow \mathcal{B}_V}$ y $Q = P_{\mathcal{B}_W \rightarrow \mathcal{B}'_W}$. Recordemos que tanto P como Q se pueden calcular [de forma sencilla](#).

De esta forma, vemos que si fijamos las nuevas bases, entonces la aplicación f resultante equivale a la f anterior, a la que precedemos y sucedemos con la aplicación identidad dentro de V y W , respectivamente, puesto que $f = \text{id}_W \circ f \circ \text{id}_V$.

Teniendo en cuenta los resultados teóricos anteriores, esto quiere decir que $A' = Q A P$.

Como consecuencia, tenemos además que

Teorema

Todas las matrices asociadas a una misma aplicación lineal f en distintas bases son equivalentes.

Ejemplo

Continuemos con el ejemplo de [la sección anterior](#). Sabemos que la aplicación f allí definida tenía por matriz asociada:

$$A = \begin{pmatrix} -\frac{9}{2} & 3 & \frac{3}{2} \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 5 & -4 & -5 \end{pmatrix}$$

Consideremos ahora dos bases distintas tanto en V como W :

$$\mathcal{B}'_1 = \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

$$\mathcal{B}'_2 = \left\{ \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Vamos a aplicar lo anterior para encontrar la matriz A' asociada a f en las nuevas bases.

Para ello, sabemos que $A' = Q A P$, donde P es la matriz de cambio de base de \mathcal{B}'_1 a \mathcal{B}_1 y Q es el cambio de \mathcal{B}_2 a \mathcal{B}'_2 .

Si llamamos B_1, B'_1, B_2 y B'_2 a las matrices que resultan de poner por columnas los vectores de $\mathcal{B}_1, \mathcal{B}'_1, \mathcal{B}_2$ y \mathcal{B}'_2 , respectivamente, podemos calcular P y Q de la siguiente manera, según [ya hemos visto](#):

$$P = B_1^{-1} B'_1, \quad Q = B_2'^{-1} B_2$$

Hacemos Gauss-Jordan partiendo de $(B_1|B'_1)$ hasta llegar a $(I_n|P)$, y partiendo de $(B_2|B'_2)$ hasta $(I_m|Q)$, obteniendo:

$$P = \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

De ahí:

$$\begin{aligned} A' &= Q A P = \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} -\frac{9}{2} & 3 & \frac{3}{2} \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 5 & -4 & -5 \end{pmatrix} \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{3}{2} & 1 & 1 \\ -2 & -4 & -2 \\ 5 & 6 & 4 \end{pmatrix} \end{aligned}$$

3.4 Qué es y cómo determinar el núcleo de una aplicación lineal

Definición: Núcleo de una aplicación lineal

Llamamos **núcleo** de una aplicación lineal $f : V \rightarrow W$, y lo denotamos por $\text{Ker } f$, al conjunto de aquellos vectores $v \in V$ tales que $f(v) = 0_W$:

$$\text{Ker } f = \{v \in V : f(v) = 0_W\}$$

El primer resultado teórico interesante es:

Teorema

$\text{Ker } f \neq \emptyset$, puesto que, al menos, $0_V \in \text{Ker } f$, como [hemos visto antes](#).

Podemos decir todavía más:

Teorema (El núcleo es un subespacio)

Dada una aplicación lineal $f : V \rightarrow W$, su núcleo $\text{Ker } f$ es un subespacio vectorial de V .

Por tanto, para operar con el núcleo, podemos hacerlo como hemos visto en [el capítulo acerca de espacios vectoriales](#): podemos calcular [un sistema generador](#), [determinar una base y su dimensión](#).

¿Qué relación tiene el núcleo de f con su matriz asociada?

Supongamos que fijamos las bases en V y W y que la matriz asociada a f en esas bases es A .

Entonces, si $v \in V$ está en el núcleo, será que $f(v) = 0_W$, luego las coordenadas de $f(v)$ en la base de W son todas cero: $[f(v)]_{\mathcal{B}_W} = 0$.

Pero como conocemos la matriz asociada a f , entonces sabemos que $[f(v)]_{\mathcal{B}_W} = A[v]_{\mathcal{B}_V}$.

En resumen, $v \in \text{Ker } f$ si, y sólo si $A[v]_{\mathcal{B}_V} = 0$, es decir, si y sólo si las coordenadas de v son solución del sistema de ecuaciones lineales homogéneo $Ax = 0$.

Teorema

$\text{Ker } f$ coincide con el subespacio de las soluciones del sistema $Ax = 0$.

Esto también implica que **las ecuaciones cartesianas de $\text{Ker } f$ son $Ax = 0$** . Por tanto, podemos aplicar [lo que ya hemos visto](#) para calcular una base del núcleo de f .

Ejemplo

Seguimos con el ejemplo de la aplicación $f : V = \mathbb{R}^3 \rightarrow W = \mathbb{R}^3$ dada por:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + z \\ x - 2y \\ 2x - y + z \end{pmatrix}$$

Como hemos visto, su matriz asociada es:

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

Usamos esa matriz A para construir el sistema homogéneo:

$$\begin{array}{rrcr} x & +y & +z & = & 0 \\ x & -2y & & = & 0 \\ 2x & -y & +z & = & 0 \end{array}$$

que se corresponde con las ecuaciones cartesianas de $\text{Ker } f$ (podíamos haber hecho directamente $f(v) = 0$ usando la definición de f).

Usaremos [lo que hemos visto ya](#) para convertir esas ecuaciones a forma paramétrica (usando eliminación Gaussiana), y así conseguir un sistema generador y [una base](#):

$$\begin{array}{rrcr} x & +y & +z & = & 0 \\ x & -2y & & = & 0 \\ 2x & -y & +z & = & 0 \end{array} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ 1 \end{pmatrix}$$

Luego

$$\mathcal{B}_{\text{Ker } f} = \left\{ \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ 1 \end{pmatrix} \right\}$$

3.5 Cómo podemos identificar una aplicación lineal inyectiva

Recordemos que una aplicación $f : X \rightarrow Y$ es **inyectiva** si y sólo si $f(x_1) \neq f(x_2)$ siempre que $x_1 \neq x_2$, con $x_1, x_2 \in X$.

Hay una forma de caracterizar las aplicaciones lineales inyectivas:

Teorema (Caracterización de la inyectividad)

La aplicación lineal $f : V \rightarrow W$ es inyectiva si, y sólo si, $\text{Ker } f = \{0\}$, es decir, si y sólo si $\dim(\text{Ker } f) = 0$.

Un resultado interesante nos presenta una relación entre la inyectividad de f y las dimensiones de los espacios V y W :

Teorema

Si la aplicación lineal $f : V \rightarrow W$ es inyectiva, entonces se verifica que $\dim(V) \leq \dim(W)$.

El recíproco no es cierto, no siempre que tengamos $\dim(V) \leq \dim(W)$ la aplicación puede ser inyectiva. Por ejemplo, la aplicación cero $0 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, que asocia a todo vector el vector 0 del codominio, no puede ser inyectiva, a pesar de que la dimensión del dominio es menor que la del codominio.

Podemos decir todavía algún resultado teórico que nos proporciona más información acerca de cómo identificar aplicaciones lineales inyectivas:

Teorema

Una aplicación $f : V \rightarrow W$ es inyectiva si, y sólo si, la imagen de todo **sistema linealmente independiente** es también un sistema linealmente independiente.

Teorema

Si V es de dimensión finita, entonces la aplicación lineal $f : V \rightarrow W$ es inyectiva si, y sólo si, $\dim(V) = \dim(\text{Im } f)$

Ejemplo

En el ejemplo de [la sección anterior](#), tenemos que $\dim(\text{Ker } f) = 1$, luego f **no es inyectiva**, ya que en ese caso debería ser $\dim(\text{Ker } f) = 0$.

Consideremos otra aplicación $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por:

$$g \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 3 & 0 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y \\ x + 3y \\ 2x + 4z \end{pmatrix}$$

Podemos estudiar su núcleo para comprobar si es inyectiva:

$$\text{Ker } g = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{array}{lcl} 2x - y & = & 0 \\ x + 3y & = & 0 \\ 2x + 4z & = & 0 \end{array} \right\}$$

Resolvemos el sistema homogéneo para llegar a que

$$\begin{cases} 2x - y & = & 0 \\ x + 3y & = & 0 \\ 2x & + & 4z = 0 \end{cases} \Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Luego $\text{Ker } g = \{0\}$, y así g es inyectiva.

3.6 Cómo calcular el subespacio imagen de una aplicación lineal

Definición: Imagen de una aplicación lineal

Llamamos **imagen de la aplicación lineal** $f : V \rightarrow W$ a:

$$\text{Im } f = \{w \in W : \text{existe } v \in V \text{ con } f(v) = w\} = f(V)$$

Análogamente a lo que pasaba con [el núcleo de una aplicación lineal](#), tenemos:

Teorema (La imagen es un subespacio vectorial)

Si $f : V \rightarrow W$ es una aplicación lineal, entonces $\text{Im } f$ es un subespacio vectorial de W .

Por tanto, podemos estudiarlo con todas las [herramientas que conocemos](#).

¿Cómo calculamos $\text{Im } f$?

[Recordemos](#) que si \mathcal{G} es un [sistema generador](#) de un subespacio U de V , entonces $f(\mathcal{G})$ es un sistema generador de $f(U)$.

En concreto, tomando una base \mathcal{B} de V , su imagen $f(\mathcal{B})$ es un sistema generador de $\text{Im } f$. Basta, por tanto, aplicar f a cada vector de \mathcal{B} y con eso construimos un sistema generador del que podemos [eliminar los vectores linealmente dependientes para obtener una base](#) de $\text{Im } f$.

En consecuencia:

Teorema (Sistema generador de $\text{Im } f$)

Sea f una aplicación lineal entre dos espacios V y W , y sea A la matriz asociada a f , una vez que hemos fijado dos bases en V y W . Entonces $\text{Im } f$ es el subespacio de W generado por los vectores columna que forman la matriz A .

Ejemplo

Vamos a estudiar la imagen de las dos aplicaciones f y g del [ejemplo anterior](#).

Para f , tenemos que $\text{Im } f$ está generado por los vectores columnas de la matriz A ,

luego un sistema generador de $\text{Im } f$ será:

$$\mathcal{G}_{\text{Im } f} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

A partir de él, [eliminando los vectores que sean linealmente dependientes](#), encontramos una base de $\text{Im } f$:

$$\mathcal{B}_{\text{Im } f} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} \right\}$$

Luego $\dim(\text{Im } f) = 2$.

Vayamos ahora con la aplicación g . Con el mismo razonamiento, sabemos que

$$\mathcal{G}_{\text{Im } g} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix} \right\}$$

es un sistema generador de $\text{Im } g$.

Igual que antes, podemos encontrar una base de $\text{Im } g$:

$$\mathcal{B}_{\text{Im } g} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix} \right\}$$

Entonces $\dim(\text{Im } g) = 3$. Esto nos dice que $\text{Im } g = W$.

3.7 Cómo identificar si una aplicación lineal es sobreyectiva

Recordemos que una aplicación $f : X \rightarrow Y$ se dice **sobreyectiva** si, para cada $y \in Y$, existe un $x \in X$ tal que $f(x) = y$, es decir, si $f(X) = Y$.

En cuanto a aplicaciones lineales, tenemos el siguiente resultado, consecuencia de lo anterior:

Teorema (Caracterización de la sobreyectividad)

La aplicación lineal $f : V \rightarrow W$ es sobreyectiva si $f(V) = \text{Im } f = W$, y esto es equivalente a decir que $\dim(\text{Im } f) = \dim(W)$.

Una propiedad adicional, con relación a las dimensiones de V y de W :

Teorema

Si la aplicación lineal $f : V \rightarrow W$ es sobreyectiva, entonces $\dim(V) \geq \dim(W)$.

Ejemplo

En el [ejemplo anterior](#), estudiamos las imágenes de las aplicaciones f y g de los ejemplos.

Si aplicamos el resultado teórico de antes, llegamos a que:

- f no es sobreyectiva, pues $\dim(\text{Im } f) = 2 \neq 3 = \dim(W)$.
- g sí es sobreyectiva, pues $\dim(\text{Im } g) = 3 = \dim(W)$.

3.8 Qué es un isomorfismo

Definición: Isomorfismo y automorfismo

Un **isomorfismo** es una aplicación lineal [inyectiva](#) y [sobreyectiva](#). Un **auto-morfismo** es un isomorfismo donde dominio y codominio son el mismo espacio V .

Algunas propiedades básicas de los isomorfismos

- La composición de isomorfismos es un isomorfismo.
- La aplicación inversa de un isomorfismo es también un isomorfismo

Cuando, además, estamos hablando de espacios de dimensión finita, tenemos los siguientes resultados:

Teorema

- Una aplicación lineal $f : V \rightarrow W$ es un isomorfismo si, y sólo si, $\dim(V) = \dim(W)$.
- Una aplicación $f : V \rightarrow V$ es un *automorfismo* si, y sólo si, o bien es inyectiva o bien es sobreyectiva.

Este último resultado nos dice que en un *endomorfismo*, basta con ver si la aplicación es [inyectiva](#) o [sobreyectiva](#) para tener que es un isomorfismo.

Ejemplo

Si miramos las aplicaciones f y g de [las secciones anteriores](#), entonces podemos confirmar que g es un automorfismo, mientras que f no lo es, ya que no es ni inyectiva ni sobreyectiva.

3.9 Cómo determinar la imagen de un subespacio vectorial

Este problema es realmente un *caso particular* de [cómo calcular la imagen de una aplicación lineal](#).

Tanto en esa sección como en la de [propiedades](#), vimos que si U es un [subespacio vectorial generado](#) por los vectores de $\mathcal{G} = \{u_1, \dots, u_k\}$, entonces $f(\mathcal{G}) = \{f(u_1), \dots, f(u_k)\}$ es un sistema generador de $f(U)$.

Por tanto, si deseamos calcular la imagen de un subespacio U , los pasos a seguir serán:

- A partir de sus ecuaciones (si se dan), [encontrar un sistema generador o una base](#) de U .
 - Calcular la imagen, mediante la aplicación f , de los vectores de la base de U . Por [lo anterior](#), esto es un sistema generador de $f(U)$.
 - [Eliminar los vectores linealmente dependientes](#) de ese sistema generador, para obtener una base de $f(U)$.
-

Ejemplo

Consideremos el subespacio U dado por

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : -x - \frac{1}{2}y + z = 0 \right\}$$

Queremos calcular $f(U)$ y $g(U)$, siendo f y g las de [los ejemplos anteriores](#).

El primer paso es [calcular una base](#) de U . Para ello, pasamos las ecuaciones cartesianas de U a paramétricas, que nos dan un sistema generador, resolviendo el sistema de ecuaciones por Gauss. A partir del sistema generador, ya podríamos proceder, o podemos calcular una base de U .

En nuestro caso,

$$\mathcal{B}_U = \left\{ \begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Comenzamos por calcular $f(U)$. Debemos calcular la imagen de los vectores de la base de U con la aplicación f :

$$f(\mathcal{B}_U) = \left\{ \begin{pmatrix} -1 \\ -6 \\ -7 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\}$$

Si [eliminamos los vectores linealmente dependientes](#) de $f(\mathcal{B}_U)$, nos queda una base de $f(U)$:

$$\mathcal{B}_{f(U)} = \left\{ \begin{pmatrix} -1 \\ -6 \\ -7 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\}$$

Repetimos el proceso con la aplicación g . Calculamos:

$$g(\mathcal{B}_U) = \left\{ \begin{pmatrix} -6 \\ 4 \\ -8 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 6 \end{pmatrix} \right\}$$

Como sabemos que g es un [isomorfismo](#), sabemos que la imagen de un sistema linealmente independiente es también un sistema linealmente independiente. Luego $\mathcal{B}_{g(U)} = g(\mathcal{B}_U)$. Si no recordáramos este resultado, deberíamos intentar eliminar los vectores linealmente dependientes en $g(\mathcal{B}_U)$ para así obtener la base.

3.10 A qué se llama rango y nulidad de la aplicación lineal

Definición: Rango

Se llama **rango** de una aplicación lineal $f : V \rightarrow W$ a la dimensión de la **imagen** de f : $\text{rg}(f) = \dim(\text{Im } f)$.

Definición: Nulidad

Se llama **nulidad** de una aplicación lineal $f : V \rightarrow W$ a la dimensión del **núcleo** de f : $\text{nul}(f) = \dim(\text{Ker } f)$.

De forma análoga, se pueden definir rango y nulidad para **matrices**, puesto que tenemos **una importante relación entre matrices y aplicaciones lineales**.

Definición: Rango y nulidad de una matriz

Dada una matriz A , su **rango** $\text{rg}(A)$ es la dimensión del subespacio generado por los vectores que forman sus columnas, y su **nulidad** $\text{nul}(A)$ es la dimensión del subespacio vectorial de las soluciones del sistema de ecuaciones lineales $A v = 0$.

Un resultado importante, debido a que **todas las matrices asociadas a una misma aplicación lineal son equivalentes**, es que todas tienen el mismo rango y la misma nulidad, y coinciden con el rango y nulidad de f .

Por tanto, para calcular rango y nulidad de f , bastará con estudiar el rango y nulidad de cualquiera de sus expresiones matriciales en unas bases fijadas de V y W .

Ejemplo

Vamos a calcular el rango y la nulidad de las aplicaciones f y g que estamos estudiando en los **ejemplos anteriores**.

El caso fácil es el de g , que era un **isomorfismo**. Luego:

- $\text{nul}(g) = \dim(\text{Ker } g) = \dim(\{0\}) = 0$, por ser g **inyectiva**.
- $\text{rg}(g) = \dim(\text{Im } g) = \dim(W) = 3$, por ser g **sobreyectiva**.

Y lo mismo se aplica para sus matrices asociadas.

En cuanto a f , recopilando la información que tenemos:

- Su nulidad es la dimensión de su **núcleo**: $\text{nul}(f) = \dim(\text{Ker } f) = 1$.
- Su rango es la dimensión de su **imagen**: $\text{rg}(f) = \dim(\text{Im } f) = 2$.

3.11 Qué dice el teorema de la dimensión para núcleo e imagen

El resultado teórico que nos relaciona las dimensiones de [núcleo](#) e [imagen](#) de una aplicación lineal $f : V \rightarrow W$, es decir, su nulidad y su rango, es el siguiente:

Teorema (Teorema de la dimensión)

Si $f : V \rightarrow W$ es una aplicación lineal, entonces

$$\dim(V) = \dim(\text{Ker } f) + \dim(\text{Im } f) = \text{nul}(f) + \text{rg}(f)$$

Este resultado, reescrito en términos matriciales, queda como:

Teorema

Si A es una matriz con m filas y n columnas, entonces

$$n = \text{nul}(A) + \text{rg}(A)$$

Ejemplo

Vamos a *comprobar* el teorema de la dimensión para las aplicaciones lineales f y g de [los ejemplos anteriores](#).

- Para f :

$$\begin{array}{rclclcl} \dim(\mathbb{R}^3) & = & \dim(\text{Ker } f) & + & \dim(\text{Im } f) \\ 3 & = & 1 & + & 2 \end{array}$$

- Para g :

$$\begin{array}{rclclcl} \dim(\mathbb{R}^3) & = & \dim(\text{Ker } g) & + & \dim(\text{Im } g) \\ 3 & = & 0 & + & 3 \end{array}$$

Capítulo 4

Diagonalización

En este capítulo nos vamos a centrar en el estudio de *endomorfismos*, es decir, de aplicaciones lineales $f : V \rightarrow V$ donde el dominio y el codominio son el mismo *espacio vectorial* sobre un cuerpo \mathbb{K} (generalmente $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$).

Una de las propiedades que hemos estudiado es que si tenemos una base \mathcal{B} fija en V , entonces *podemos encontrar una matriz* A , asociada a f , tal que podemos expresar las *coordenadas* de $f(v)$ en dicha base como $A v$, sea cual sea $v \in V$.

¿Qué ocurre si fijamos otra base distinta \mathcal{B}' en V ?

En ese caso, la matriz asociada al endomorfismo f será $A' = P^{-1} A P$, donde P es la *matriz de cambio de base* de \mathcal{B}' a \mathcal{B} .

Esta relación especial entre A y A' tiene nombre: son *matrices semejantes*.

Definición: Matrices semejantes

Dos matrices $A, A' \in \mathcal{M}_n(\mathbb{K})$ se dicen **semejantes** si existe $P \in \mathcal{M}_n(\mathbb{K})$ regular tal que $A' = P^{-1} A P$.

Esto tiene implicaciones como que los cálculos que podamos necesitar hacer con un endomorfismo, como calcular su núcleo y su imagen, o los que veremos en este tema, lo podemos hacer con cualquiera de sus matrices asociadas, respecto a diferentes bases, ya que son todas semejantes.

Sin embargo, para ciertas operaciones, podremos necesitar tener una matriz asociada a f lo más simple posible. ¿Qué entendemos por una matriz simple? Una matriz diagonal.

La diagonalización es el proceso por el cual pretendemos encontrar una base \mathcal{B} de V tal que la matriz asociada a un endomorfismo f sea diagonal.

Definición: Endomorfismo diagonalizable

Un endomorfismo $f : V \rightarrow V$ se dice **diagonalizable** si existe una base de V donde la matriz asociada a f es diagonal.

No todos los endomorfismos serán **diagonalizables**, así que veremos cuándo podremos asegurar que sí lo son, y el mecanismo para encontrar la forma diagonal de la matriz asociada al endomorfismo.

Nota: A menudo, abusaremos de la notación y usaremos indistintamente v para designar a un vector o a sus coordenadas con respecto a la base canónica.

Ejemplo

Consideremos el siguiente endomorfismo $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, dado por

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 9x + 6y + 6z \\ -2x + y - 2z \\ -10x - 10y - 7z \end{pmatrix}$$

Es claro que su matriz asociada en la base canónica es:

$$A = \begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix}$$

Si consideramos la siguiente base:

$$\mathcal{B} = \left\{ \begin{pmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

entonces la matriz de cambio de \mathcal{B} a la canónica \mathcal{C} es:

$$P = \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

y la matriz asociada a f en la nueva base es:

$$\begin{aligned} A' &= \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix} \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \end{aligned}$$

En este capítulo, veremos el método para hallar esa descomposición, la base y la matriz diagonal.

¿Qué preguntas vamos a responder en este capítulo?

- ¿Qué son los autovalores de un endomorfismo y de una matriz?
- ¿Cómo calculamos los autovectores asociados a un autovalor?
- ¿Cómo sabemos si un endomorfismo o una matriz es diagonalizable?
- ¿Si una matriz es diagonalizable, cómo podemos hallar sus potencias?
- ¿Para qué podemos utilizar el Teorema de Cayley-Hamilton?

4.1 Qué son los autovalores de un endomorfismo y de una matriz

Definición: Autovalor y autovector

Consideremos un endomorfismo $f : V \rightarrow V$. Diremos que un valor $\lambda \in \mathbb{K}$ es un **autovalor** (o **valor propio**) de f si existe un vector no nulo $v \in V$ tal que $f(v) = \lambda v$. A ese v se le denomina **autovector** (o **vector propio**) asociado a λ .

Nota: Eliminamos la posibilidad de que v sea nulo en la definición, porque **sabemos que $f(0) = 0$** en cualquier aplicación lineal, así que no nos proporcionaría información alguna.

Intuitivamente, un *autovector* v es un vector no nulo que, al aplicarle f , su imagen $f(v)$ es proporcional a él mismo, y la constante de proporcionalidad es el *autovalor* λ .

Si consideramos una matriz A asociada a f , la condición $f(v) = \lambda v$ se puede reescribir como $A v = \lambda v$.

Podemos generalizar entonces esta definición a matrices cuadradas:

Definición: Autovalor y autovector de una matriz cuadrada

Un escalar λ se llama **autovalor** o **valor propio** de una matriz $A \in \mathcal{M}_n(\mathbb{K})$ si existe $v \in \mathbb{K}^n$ no nulo, llamado **autovector** o **vector propio** asociado a λ , tal que $A v = \lambda v$.

Nos falta comprobar si hay alguna incoherencia en las definiciones: si distintas matrices asociadas a un mismo endomorfismo tuvieran distintos *autovalores*, eso sería un problema de la definición. Sin embargo, el siguiente resultado nos asegura que esto no puede ser así:

Teorema

Si dos matrices A y A' son *semejantes*, entonces tienen los mismos autovalores.

Podemos decir entonces que:

Teorema

Si λ es un autovalor de f , entonces lo es de cualquiera de sus matrices asociadas en diferentes bases, y esto es debido a su relación de *semejanza*.

Esto significa que si queremos calcular los autovalores de un endomorfismo f , basta con calcularlos para cualquiera de sus matrices asociadas, en particular para su matriz asociada en la *base canónica*, la más sencilla de encontrar.

¿Cómo podemos encontrar los autovalores de un endomorfismo o de su matriz asociada?

Para encontrar los autovalores de un endomorfismo, siempre partiremos de una matriz asociada al mismo. Por tanto, explicaremos cómo calcular los autovalores de una matriz cuadrada, y eso nos bastará para encontrar los del endomorfismo, por lo comentado anteriormente.

La condición para ser $\lambda \in \mathbb{K}$ un autovalor es que exista $v \in V$ tal que $A v = \lambda v$.

Esto es equivalente a $A v - \lambda v = 0$, es decir, $A v - \lambda I v = 0$, siendo I la matriz identidad del mismo tamaño que v . En forma compacta, podemos decir que λ es autovalor de A si existe v tal que $(A - \lambda I) v = 0$.

Es decir, v debe ser solución del siguiente sistema homogéneo: $(A - \lambda I) x = 0$.

Recordemos que un sistema homogéneo tiene siempre solución. Estamos interesados en soluciones distintas del vector 0, puesto que hemos impuesto en la definición de *autovector* que sea un vector no nulo.

¿Qué condición tenemos para saber que un sistema homogéneo es compatible in-

determinado? Pues que la matriz de coeficientes del sistema sea *singular*, es decir, tenga determinante igual a 0.

En este caso, la matriz de coeficientes del sistema es $A - \lambda I$, así que tendremos soluciones distinta de la trivial si $\det(A - \lambda I) = 0$.

Hay que notar que la expresión $\det(A - \lambda I)$ depende de λ , así que realmente estamos buscando **los valores de λ que hacen que $\det(A - \lambda I) = 0$** .

Definición: Polinomio característico

Llamamos **polinomio característico** de la matriz A a $p(\lambda) = \det(A - \lambda I)$, y **ecuación característica** a $\det(A - \lambda I) = 0$.

Por tanto, **los autovalores de una matriz A son** soluciones a la ecuación característica, es decir, **las raíces del polinomio característico**.

Ejemplo

Vamos a calcular los autovalores del endomorfismo $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, dado por

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 9x + 6y + 6z \\ -2x + y - 2z \\ -10x - 10y - 7z \end{pmatrix}$$

del ejemplo anterior.

Sabemos que su matriz asociada en la base canónica es:

$$A = \begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix}$$

Definimos entonces el polinomio característico como:

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) = \\ &= \det \left(\begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) = \\ &= \det \begin{pmatrix} 9 - \lambda & 6 & 6 \\ -2 & 1 - \lambda & -2 \\ -10 & -10 & -7 - \lambda \end{pmatrix} \end{aligned}$$

Si desarrollamos el determinante, nos queda como *polinomio característico*:

$$p(\lambda) = \lambda^3 - 3\lambda^2 - 9\lambda + 27$$

Igualándolo a 0, tenemos la ecuación característica $\lambda^3 - 3\lambda^2 - 9\lambda + 27 = 0$, que podemos resolver (generalmente se podrá factorizar de forma simple o usando Ruffini) y nos quedan las siguientes soluciones: $\lambda = -3, 3$. Esta última solución es raíz doble, ya que:

$$p(\lambda) = (\lambda + 3) \cdot (\lambda - 3)^2$$

Por tanto, $\lambda = -3, 3$ son autovalores de la matriz A y del endomorfismo f .

4.2 Cómo calculamos los autovectores asociados a un autovalor

[Acabamos de ver](#) que los autovectores asociados a un autovalor λ son soluciones no triviales del sistema homogéneo $(A - \lambda I)x = 0$.

También sabemos que las soluciones de un sistema homogéneo forman un subespacio vectorial. Por tanto, **el conjunto de autovectores asociado a un autovalor λ es un subespacio vectorial que llamamos subespacio asociado al autovalor λ** . Generalmente, denotaremos U_λ al subespacio asociado al autovalor λ .

En ocasiones, se denomina **subespacio invariante por f** a U_λ , puesto que $f(U_\lambda) \subseteq U_\lambda$.

Como con cualquier subespacio vectorial, para tener perfectamente determinado a U_λ nos basta con dar una base suya que, evidentemente, estará formada por autovectores asociados a λ .

En este caso, podemos [partir de las ecuaciones cartesianas](#) de U_λ , que son el sistema $(A - \lambda I)x = 0$, y encontrar una base suya, como ya hemos visto en el capítulo de [espacios vectoriales](#).

Ejemplo

Retomamos el endomorfismo f de los ejemplos anteriores. Calculemos los subespacios asociados a los distintos autovalores:

- Para el autovalor $\lambda = -3$:

$$\begin{aligned} (A + 3I) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} 12 & 6 & 6 \\ -2 & 4 & -2 \\ -10 & -10 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \Leftrightarrow \\ &\Leftrightarrow \begin{cases} 12x + 6y + 6z = 0 \\ -2x + 4y - 2z = 0 \\ -10x - 10y - 4z = 0 \end{cases} \end{aligned}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de su conjunto solución, que es U_{-3} :

$$\begin{aligned} \left(\begin{array}{ccc|c} 12 & 6 & 6 & 0 \\ -2 & 4 & -2 & 0 \\ -10 & -10 & -4 & 0 \end{array} \right) &\sim \left(\begin{array}{ccc|c} 12 & 0 & \frac{36}{5} & 0 \\ 0 & 5 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \\ &\Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{pmatrix} \end{aligned}$$

De aquí que una base de U_{-3} sea:

$$\mathcal{B}_{U_{-3}} = \left\{ \begin{pmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{pmatrix} \right\}$$

- Para el autovalor $\lambda = 3$:

$$\begin{aligned} (A - 3I) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} 6 & 6 & 6 \\ -2 & -2 & -2 \\ -10 & -10 & -10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \Leftrightarrow \\ &\Leftrightarrow \begin{cases} 6x + 6y + 6z = 0 \\ -2x - 2y - 2z = 0 \\ -10x - 10y - 10z = 0 \end{cases} \end{aligned}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de su conjunto solución, que es U_3 :

$$\left(\begin{array}{ccc|c} 6 & 6 & 6 & 0 \\ -2 & -2 & -2 & 0 \\ -10 & -10 & -10 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 6 & 6 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

De aquí que una base de U_3 sea:

$$\mathcal{B}_{U_3} = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Los vectores de las bases encontradas, así como cualquiera de sus múltiplos, son autovectores asociados al correspondiente autovalor λ .

4.3 Cómo sabemos si un endomorfismo o una matriz es diagonalizable

Una vez que sabemos cómo hallar los autovalores y los autovectores de una matriz y de un endomorfismo, vamos a plantearnos cómo utilizarlos para saber si podemos encontrar una base \mathcal{B} del espacio V tal que la matriz asociada al endomorfismo f es diagonal. Es decir, nos planteamos el estudio de si f es **diagonalizable**.

El principal resultado teórico que necesitamos es el siguiente:

Teorema (Caracterización de diagonalizable)

Un endomorfismo $f : V \rightarrow V$ es *diagonalizable* si y sólo si existe una base \mathcal{B} de V formada únicamente por autovectores de f . En ese caso, la **matriz asociada** D en esa base es diagonal y tiene en su diagonal principal los autovalores de f . Además, la **matriz de cambio de base** $P_{\mathcal{B} \rightarrow \mathcal{C}}$ (formada por los vectores de \mathcal{B} por columnas) verifica:

$$P_{\mathcal{B} \rightarrow \mathcal{C}}^{-1} A P_{\mathcal{B} \rightarrow \mathcal{C}} = D$$

o, equivalentemente,

$$A = P_{\mathcal{C} \rightarrow \mathcal{B}}^{-1} D P_{\mathcal{C} \rightarrow \mathcal{B}}$$

Este resultado proporciona mucha información: nos dice que para que sea diagonalizable un endomorfismo (o una matriz), necesitamos una base de V formada por únicamente autovectores, y que la matriz diagonal asociada tendrá a los autovalores en la diagonal principal.

Nos preguntamos entonces bajo qué condiciones podemos asegurar que tenemos una base de V formada enteramente por autovectores.

El siguiente resultado es de gran ayuda:

Teorema

Autovectores asociados a distintos autovalores son linealmente independientes.

Supongamos que llamamos $\lambda_1, \dots, \lambda_k$ a los autovalores de un endomorfismo o de una matriz, y $U_{\lambda_1}, \dots, U_{\lambda_k}$ a los subespacios asociados a los autovalores. Supongamos que todos los λ_i son distintos unos de otros.

El resultado anterior me asegura que si $i \neq j$, entonces los vectores de las bases de U_{λ_i} y de U_{λ_j} son linealmente independientes.

Llamamos $\mathcal{B} = \mathcal{B}_{U_{\lambda_1}} \cup \dots \cup \mathcal{B}_{U_{\lambda_k}}$, es un sistema **linealmente independiente** por lo que acabamos de comentar. Si, además, \mathcal{B} tuviera n elementos, puesto que $\dim(V) = n$, entonces \mathcal{B} sería una **base** de V (recordemos que una base es un sistema linealmente independiente con tantos elementos como la dimensión del espacio vectorial).

Por tanto, si unimos todas las bases de los subespacios asociados a los autovalores, y tenemos tantos elementos como la dimensión del espacio V , entonces se verifica todo lo anterior: el endomorfismo (y sus matrices asociadas) son diagonalizables. Esto es así porque entonces se verificaría:

$$V = U_{\lambda_1} + \dots + U_{\lambda_k}$$

La primera consecuencia de esto es:

Teorema

Si una matriz o un endomorfismo tienen exactamente n autovalores distintos, entonces es diagonalizable.

Esto es así porque cada uno de los n autovalores tendría un subespacio asociado cuya base es linealmente independiente de la base de los demás subespacios. Esto solo podría suceder si cada una de las n bases $\mathcal{B}_{U_{\lambda_i}}$, ($i = 1, \dots, n$) solo tiene un elemento, es decir, si $\dim(U_{\lambda_i}) = 1$ para todo $i = 1, \dots, n$. Si no fuera así, acabaríamos con más de n vectores linealmente independientes en V , lo cual es imposible.

Entonces \mathcal{B} , definida como la unión de las bases mencionadas, tiene exactamente n elementos. Esto nos demuestra que la matriz o el endomorfismo con n autovalores distintos es diagonalizable.

Pero ¿qué pasa cuando no tenemos n autovalores distintos?

Es de gran importancia conocer, para todo $i = 1, \dots, k$, la dimensión del subespacio asociado al autovalor λ_i .

Definición: Multiplicidad geométrica

Llamaremos **multiplicidad geométrica del autovalor** λ a $m_g(\lambda) = \dim(U_{\lambda_i})$.

Se tiene el siguiente resultado:

Teorema

Para todo autovalor λ de un endomorfismo o de una matriz, se tiene que

$$1 \leq m_g(\lambda) \leq m_a(\lambda)$$

donde $m_g(\lambda)$ es su multiplicidad geométrica y $m_a(\lambda)$ es la **multiplicidad algebraica de** λ , es decir, su multiplicidad como raíz del **polinomio característico**.

Es decir, la *multiplicidad geométrica* siempre es mayor o igual que 1 (indicando que el subespacio U_{λ} no es el trivial $\{0\}$), y menor que la *multiplicidad algebraica*.

Un corolario de todos los resultados anteriores es:

Teorema

Si, para todo autovalor λ de un endomorfismo o de una matriz cuadrada, se tiene que $m_g(\lambda) = m_a(\lambda)$, entonces es diagonalizable. Y el recíproco es también cierto.

Aquí tenemos un criterio sencillo para determinar si un endomorfismo o una matriz son diagonalizables:

- Si tenemos n autovalores distintos, entonces es diagonalizable.
- Si tenemos menos de n autovalores distintos, debemos comprobar que las multiplicidades geométricas de **todos** los autovalores coincidan con sus multiplicidades algebraicas.

En ambos casos, la **base** de V definida por $\mathcal{B} = \mathcal{B}_{U_{\lambda_1}} \cup \dots \cup \mathcal{B}_{U_{\lambda_k}}$ es una base formada por autovectores y la matriz asociada al endomorfismo f en dicha base es diagonal con los autovalores en la diagonal principal (repetidos según su multiplicidad).

Si P es la **matriz de paso** de \mathcal{B} a la canónica \mathcal{C} , entonces la **expresión matricial del endomorfismo f en dicha base \mathcal{B}** es $D = P^{-1} A P$.

Ejemplo

Consideramos el endomorfismo f de los ejemplos anteriores. Si tenemos en cuenta lo que se ha determinado de sus **autovalores** y de los **subespacios asociados**, tenemos

la siguiente tabla resumen:

Autovalor	Mult. Algebraica	Mult. Geométrica
-3	1	1
3	2	2

Con esto, queda demostrado que el endomorfismo f **sí** es diagonalizable.

Además, la base \mathcal{B} de autovectores es:

$$\mathcal{B} = \mathcal{B}_{-3} \cup \mathcal{B}_3 = \left\{ \begin{pmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

La matriz de cambio de base de \mathcal{B} a \mathcal{C} es la que tiene los elementos de \mathcal{B} por columnas:

$$P = \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Y la matriz diagonal es la que tiene los autovalores en la diagonal, que se calcula como:

$$\begin{aligned} D &= P^{-1} A P = \\ &= \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix} \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \end{aligned}$$

4.4 Si una matriz es diagonalizable, cómo podemos hallar sus potencias

Consideremos una matriz cuadrada [diagonalizable](#). En concreto, eso significa que existe otra matriz cuadrada P regular (cuyas columnas son los vectores de la base de [autovectores](#)) tal que

$$A = P^{-1} D P$$

donde D es la matriz diagonal con los [autovalores](#) de A en la diagonal principal.

Supongamos que queremos hallar A^k , con $k > 1$ entero. Si usamos la expresión anterior, nos queda:

$$A^k = A \cdot A \cdot \dots \cdot A = (P^{-1} D P) \cdot (P^{-1} D P) \cdot \dots \cdot (P^{-1} D P)$$

donde los productos se repiten k veces. Si agrupamos cada par $P \cdot P^{-1}$, que es igual a la identidad, entonces nos queda:

$$A^k = P^{-1} D^k P$$

Como D es una matriz diagonal con los autovalores, es fácil calcular su potencia k -ésima:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix}$$

De esta forma, calcular la potencia de A se reduce a calcular la potencia k -ésima de la matriz D . Esto es especialmente útil para calcular potencias *altas* de A .

Ejemplo

Calculemos A^{100} donde A es la matriz del [ejemplo anterior](#).

Conocemos su descomposición:

$$\begin{aligned} A &= P D P^{-1} = \\ &= \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \end{aligned}$$

Entonces, por lo comentado anteriormente, $A^{100} = P D^{100} P^{-1}$, es decir:

$$\begin{aligned} A^{100} &= P D^{100} P^{-1} = \\ &= \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} (-3)^{100} & 0 & 0 \\ 0 & 3^{100} & 0 \\ 0 & 0 & 3^{100} \end{pmatrix} \begin{pmatrix} -\frac{3}{5} & -1 & -1 \\ \frac{1}{5} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \end{aligned}$$

4.5 Para qué podemos utilizar el Teorema de Cayley-Hamilton

El **Teorema de Cayley-Hamilton** nos dice que:

Teorema (Cayley-Hamilton)

Toda matriz cuadrada A verifica su ecuación característica, es decir, si $p(\lambda)$ es el polinomio característico de A , entonces $p(A) = 0$.

Supongamos que el **polinomio característico** de una matriz A es

$$p(\lambda) = c_n \lambda^n + \dots + c_1 \lambda + c_0$$

El teorema nos dice que A hace cero su polinomio característico:

$$p(A) = c_n A^n + \dots + c_1 A + c_0 I = 0$$

donde A^k representa el producto de A por sí misma k veces.

¿Qué utilidades tiene este teorema?

Supongamos que el término independiente del polinomio característico, c_0 es no nulo (esto equivale a que 0 no sea un autovalor, lo cual es también equivalente a que A sea regular). Si de la igualdad superior despejamos la identidad, I , nos queda lo siguiente:

$$\begin{aligned} I &= -\frac{1}{c_0} (c_n A^n + \dots + c_1 A) = \\ &= -\frac{1}{c_0} (c_n A^{n-1} + \dots + c_1 I) A \end{aligned}$$

De aquí deducimos que

$$A^{-1} = -\frac{1}{c_0} (c_n A^{n-1} + \dots + c_1 I)$$

Por tanto, **el teorema de Cayley-Hamilton nos da una forma de calcular la inversa de una matriz cuadrada A , únicamente a partir de sumas de potencias de la matriz.**

Esta misma estrategia nos proporciona un **método para calcular potencias de la matriz A** , **aparte del ya explicado**, reduciendo el problema a potencias de A de grado menor que el tamaño de la propia matriz.

Si, de la igualdad del teorema, despejamos el término A^n , nos queda

$$A^n = -\frac{1}{c_n} (-c_{n-1} A^{n-1} - \dots - c_1 A - c_0 I)$$

A partir de A^n , siendo n el número de filas o columnas de la matriz A , todas sus potencias se pueden poner como suma de potencias de A de grado menor estrictamente que n :

$$\begin{aligned}
 A^{n+1} &= A \cdot A^n = \\
 &= A \cdot \left(-\frac{1}{c_n} (-c_{n-1}A^{n-1} - \dots - c_1A - c_0I) \right) = \\
 &= -\frac{1}{c_n} (-c_{n-1}A^n - \dots - c_1A^2 - c_0A) = \\
 &= -\frac{1}{c_n} \left(-c_{n-1} \left(-\frac{1}{c_n} (-c_{n-1}A^{n-1} - \dots - c_1A - c_0I) \right) - \dots - c_1A^2 - c_0A \right) = \\
 &= \dots
 \end{aligned}$$

Siguiendo con este patrón, se puede poner cualquier potencia A^m , con $m \geq n$, como suma de A, A^2, \dots, A^{n-1} .

Ejemplo

Vamos a considerar la matriz

$$A = \begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix}$$

asociada al endomorfismo f de los ejemplos anteriores.

Vamos a calcular A^{-1} , A^4 y A^5 usando el Teorema de Cayley-Hamilton.

Conocemos su [polinomio característico](#):

$$p(\lambda) = \lambda^3 - 3\lambda^2 - 9\lambda + 27$$

El teorema nos proporciona la siguiente igualdad:

$$p(A) = A^3 - 3A^2 - 9A + 27I = 0$$

Despejando la identidad, nos queda:

$$\begin{aligned}
 I &= \frac{1}{27} (A^3 - 3A^2 - 9A) = \\
 &= \left(-\frac{1}{27}A^2 + \frac{1}{9}A + \frac{1}{3}I \right) A
 \end{aligned}$$

Luego

$$\begin{aligned}
 A^{-1} &= -\frac{1}{27}A^2 + \frac{1}{9}A + \frac{1}{3}I = \\
 &= -\frac{1}{27} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} + \frac{1}{9} \begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \\
 &= \begin{pmatrix} 1 & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{9} & \frac{1}{9} & -\frac{2}{9} \\ -\frac{10}{9} & -\frac{10}{9} & -\frac{7}{9} \end{pmatrix}
 \end{aligned}$$

Vamos ahora a calcular A^4 y A^5 como suma de potencias de A de grado menor que su tamaño, es decir, potencias de A de grado menor que 3.

En todos los casos, tenemos que partir de despejar el término A^3 de la igualdad que nos proporciona el Teorema de Cayley-Hamilton:

$$A^3 = 3A^2 + 9A - 27I$$

Para A^4 , tenemos:

$$A^4 = A \cdot A^3 = A(3A^2 + 9A - 27I) = 3A^3 + 9A^2 - 27A$$

Si queremos dejarlo como suma de potencias de A de grado menor que 3, debemos sustituir aquí A^3 por la expresión que hemos despejado hace un momento.

$$\begin{aligned}
 A^4 &= 3(3A^2 + 9A - 27I) + 9A^2 - 27A = \\
 &= 18A^2 - 81I = \\
 &= 18 \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} - 81 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \\
 &= \begin{pmatrix} 81 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 81 \end{pmatrix}
 \end{aligned}$$

Evidentemente, podríamos haber hecho el cálculo bien directamente bien haciendo $A^4 = A^2 \cdot A^2$, pero esa forma no está expresada como suma de potencias de grado menor que n .

Repetimos el proceso para calcular A^5 .

$$A^5 = A \cdot A^4 = A(18A^2 - 81I) = 18A^3 - 81A$$

Sustituimos de nuevo A^3 por su expresión:

$$\begin{aligned} A^5 &= 18(3A^2 + 9A - 27I) - 81A = \\ &= 54A^2 + 81A - 486I = \\ &= 54 \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} + 81 \begin{pmatrix} 9 & 6 & 6 \\ -2 & 1 & -2 \\ -10 & -10 & -7 \end{pmatrix} - 486 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 729 & 486 & 486 \\ -162 & 81 & -162 \\ -810 & -810 & -567 \end{pmatrix} \end{aligned}$$

Capítulo 5

Espacios Euclídeos

En este capítulo, damos un paso para acercarnos a los conceptos geométricos clásicos de distancias y ángulos, abstrayéndolos en el contexto de los espacios vectoriales.

Para ello, precisaremos de la noción de *producto escalar*: Una aplicación $g : V \times V \rightarrow \mathbb{R}$, donde V es un espacio vectorial sobre \mathbb{R} , es un **producto escalar** si verifica:

- Es simétrica: $g(u, v) = g(v, u)$ para todo $u, v \in V$.
- Es bilineal: $g(\alpha u + \beta v, w) = \alpha g(u, w) + \beta g(v, w)$ para todos los $\alpha, \beta \in \mathbb{R}$, $u, v, w \in V$. Por simetría, se tiene lo mismo para la linealidad en la segunda componente.
- Es definida positiva: $g(v, v) > 0$ para todo $v \in V \setminus \{0\}$, $g(v, v) = 0$ si, y sólo si $v = 0$.

En resumen, se suele definir un producto escalar como una *forma bilineal, simétrica y definida positiva*.

Notación: El producto escalar de dos vectores u y v suele recibir diversas notaciones, aunque las más usadas son $u \cdot v$ y $\langle u, v \rangle$.

A un espacio vectorial V donde podemos definir un producto vectorial lo llamamos **espacio euclídeo**, y lo solemos denotar mediante el par $(V, \langle \cdot, \cdot \rangle)$.

Ejemplo

En \mathbb{R}^n , disponemos de un *producto escalar usual*, de forma que si $u, v \in \mathbb{R}^n$, entonces

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \dots + u_n v_n$$

siendo u_i, v_i , $i = 1, \dots, n$ las coordenadas de u y de v , respectivamente, en una base dada (la canónica como caso particular). A este producto vectorial es al que nos referiremos cuando usemos la notación $u \cdot v$.

Podemos generalizar este producto escalar, definiendo unos pesos $\alpha_i \in \mathbb{R}^+$ para cada componente:

$$\langle u, v \rangle = \alpha_1 u_1 v_1 + \alpha_2 u_2 v_2 + \dots + \alpha_n u_n v_n$$

Podemos construir productos escalares de la siguiente forma: Tomemos una matriz $A \in \mathcal{M}_n(\mathbb{R})$ simétrica y definida positiva. Entonces la aplicación $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$\langle u, v \rangle = u^t A v$$

es un producto escalar. A esta matriz A se le denomina *matriz de Gram*, y todo producto escalar tiene asociada una. Los dos ejemplos anteriores se correspondían con tomar como A la matriz identidad o una matriz diagonal con elementos positivos $\alpha_i \in \mathbb{R}^+$ en la diagonal.

¿Qué preguntas vamos a responder en este capítulo?

- ¿Qué son las normas y distancias?
- ¿Qué significa el concepto de ortogonalidad?
- ¿Qué es el complemento ortogonal de un subespacio?
- ¿Cómo calculamos la proyección de un vector sobre un subespacio?
- ¿Qué es y qué utilidad tiene una base ortonormal?
- ¿Cómo construimos una base ortonormal?
- ¿Qué es una matriz ortogonal y una aplicación ortogonal?
- ¿Qué es la diagonalización ortogonal?
- ¿Qué matrices son diagonalizables ortogonalmente

5.1 Qué son las normas y distancias

Definición: Norma

Una **norma** vectorial sobre un espacio V (sobre un cuerpo \mathbb{K}) es un operador que sirve para *medir* la longitud de un vector (es decir, su distancia al origen), y se define formalmente como una aplicación $\| \cdot \| : V \rightarrow \mathbb{R}$ que verifica:

- Para todo $v \in V$, $\|v\| \geq 0$, y $\|v\| = 0$ si, y solo si, $v = 0$.
- Para todo $c \in \mathbb{K}$, $v \in V$, se tiene $\|cv\| = |c| \cdot \|v\|$.
- Para todo $u, v \in V$, $\|u + v\| \leq \|u\| + \|v\|$. Esta propiedad se llama *desigualdad triangular* (en un triángulo, la suma de la longitud de dos de sus lados $\|u\|$ y $\|v\|$ es siempre menor que la longitud de su otro lado $\|u + v\|$).

Cualquier operador $\| \cdot \|$ que verifique lo anterior es una norma vectorial. Sin embargo, podemos construir normas a partir de un producto escalar:

Teorema

Si $\langle \cdot, \cdot \rangle$ es un producto escalar en el espacio V , entonces la aplicación $\|\cdot\| : V \rightarrow \mathbb{R}$ dada por $\|v\| = \sqrt{\langle v, v \rangle}$ es una **norma derivada del producto escalar** sobre V .

Al definir la norma vectorial a partir de un producto escalar, además, tenemos la siguiente propiedad que será de gran interés en la [siguiente sección](#):

Teorema (Desigualdad de Cauchy-Schwarz)

Si $\|\cdot\|$ es una norma derivada del producto escalar $\langle \cdot, \cdot \rangle$, entonces

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|$$

para todo $u, v \in V$.

Un concepto implícitamente relacionado con el de norma vectorial es el de *distancia*.

Definición: Distancia

Una **distancia** (también llamada *métrica*) es un operador $d : V \times V \rightarrow \mathbb{R}$ que verifica las siguientes propiedades:

- $d(u, v) = 0$ si, y solo si, $u = v$.
- Es simétrica: $d(u, v) = d(v, u)$.
- Cumple la desigualdad triangular: $d(u, v) \leq d(u, w) + d(w, v)$.

Además, de esta definición se deduce que $d(u, v) \geq 0$ para todo $u, v \in V$.

Podemos deducir una distancia a partir de una norma, para así completar las relaciones existentes entre los conceptos aquí explicados:

Teorema

Si $\|\cdot\|$ es una norma sobre un espacio vectorial V , entonces la aplicación $d : V \times V \rightarrow \mathbb{R}$ definida por

$$d(u, v) = \|u - v\|$$

para todo $u, v \in V$ es una *distancia*.

Nota: Existen muchas más normas de las que se deducen a partir de un producto escalar, al igual que más *distancias*.

Ejemplo

Podemos deducir la expresión de la norma (a veces llamada *módulo*) de un vector $v \in \mathbb{R}^n$.

Si $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$, entonces

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

Y así, además, deducimos la expresión para la distancia euclídea usual en \mathbb{R}^n :

$$d(u, v) = \|u - v\| = \sqrt{(u_1 - v_1)^2 + \dots + (u_n - v_n)^2}$$

Como ejemplo, podemos calcular la norma del vector

$$v = \begin{pmatrix} -2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\|v\| = \sqrt{(-2)^2 + 1^2 + 1^2 + 1^2} = \sqrt{7}$$

5.2 Qué significa el concepto de ortogonalidad

Si partimos de la *desigualdad de Cauchy-Schwarz* que [hemos visto antes](#), podemos llegar a que

$$-1 \leq \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|} \leq 1$$

para todo $u, v \in V$.

Por otro lado, la función coseno restringida a $[0, \pi]$, es decir, $\cos : [0, \pi] \rightarrow [-1, 1]$, es biyectiva. Eso quiere decir que existe un único $\theta \in [0, \pi]$ tal que $\cos \theta = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$.

Esto nos permite la siguiente definición:

Definición: Ángulo entre vectores

Llamamos **ángulo entre los vectores** $u, v \in V$ al único $\theta \in [0, \pi]$ tal que $\cos \theta = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$.

Definición: Vectores ortogonales

Diremos que dos vectores u y v de V son **ortogonales** (o **perpendiculares**) si el ángulo que forman es $\frac{\pi}{2}$, es decir, si $\langle u, v \rangle = 0$.

La noción de *ortogonalidad* en el espacio euclídeo se corresponde con la perpendicularidad de vectores que todos tenemos de forma intuitiva. Lo podemos extender también a ortogonalidad con respecto a un espacio vectorial: Un vector $v \in V$ es **ortogonal al subespacio** U de V si, y sólo si, es ortogonal a todos y cada uno de los vectores de U , es decir, $\langle v, u \rangle = 0$ para todo $u \in U$.

El siguiente resultado nos proporciona una ayuda a la hora de determinar la ortogonalidad entre subespacios vectoriales.

Teorema

Sea U un subespacio de V , y sea $\mathcal{B} = \{u_1, \dots, u_k\}$ una base de U . Un vector $v \in V$ es ortogonal a U si, y solo si, es ortogonal a cada u_i : $\langle v, u_i \rangle = 0$ para todo $i = 1, \dots, k$.

Un **sistema ortogonal** es un conjunto de vectores $\{v_1, \dots, v_m\} \subset V$ donde los v_i son ortogonales dos a dos, es decir, $\langle v_i, v_j \rangle = 0$ para todo $i \neq j$.

Algunas propiedades y caracterizaciones de los vectores y sistemas ortogonales**Teorema (Generalización del teorema de Pitágoras)**

Dos vectores u y w son ortogonales si, y solo si, $\|u + w\|^2 = \|u\|^2 + \|w\|^2$

Teorema

Si un sistema de vectores es ortogonal, entonces es linealmente independiente. Por tanto, en un espacio V de dimensión finita n , cualquier conjunto de n vectores ortogonales entre sí es una base.

Este último resultado es interesante, pues nos constata que cualquier sistema ortogonal con tantos vectores como la dimensión del espacio vectorial es una base.

Ejemplo

Consideremos el espacio U dado por:

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : \begin{array}{rcl} -x + 4y + z & = & 0 \\ y + t & = & 0 \end{array} \right\}$$

y los vectores

$$v = \begin{pmatrix} 1 \\ -2 \\ -1 \\ 2 \end{pmatrix} \quad w = \begin{pmatrix} -1 \\ 2 \\ -1 \\ -1 \end{pmatrix}$$

Calculemos el ángulo entre v y w :

$$\theta = \arccos \frac{v \cdot w}{\|v\| \cdot \|w\|}$$

$$v \cdot w = 1 \cdot (-1) + (-2) \cdot 2 + (-1) \cdot (-1) + 2 \cdot (-1) = -6$$

$$\|v\| = \sqrt{1^2 + (-2)^2 + (-1)^2 + 2^2} = \sqrt{10}$$

$$\|w\| = \sqrt{(-1)^2 + 2^2 + (-1)^2 + (-1)^2} = \sqrt{7}$$

Luego

$$\theta = \arccos \frac{-6}{\sqrt{10}\sqrt{7}} = \arccos \frac{-3 \cdot \sqrt{70}}{35}$$

Vamos a comprobar ahora que v es ortogonal al subespacio U . Para ello, a partir de las ecuaciones cartesianas de U , vamos a [hallar una base](#) suya, resolviendo el sistema por Gauss-Jordan y pasando por las ecuaciones paramétricas:

$$\begin{pmatrix} -1 & 4 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & 4 \\ 0 & 1 & 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -4 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

De aquí que la base de U sea:

$$\mathcal{B}_U = \{u_i\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Como hemos visto antes, para comprobar que v es ortogonal a U , nos basta con comprobar que lo es a cada vector de \mathcal{B}_U , hallando su producto escalar:

$$\langle v, u_1 \rangle = \left\langle \begin{pmatrix} 1 \\ -2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 1 \cdot 1 + (-2) \cdot 0 + (-1) \cdot 1 + 2 \cdot 0 = 0$$

$$\langle v, u_2 \rangle = \left\langle \begin{pmatrix} 1 \\ -2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle = 1 \cdot (-4) + (-2) \cdot (-1) + (-1) \cdot 0 + 2 \cdot 1 = 0$$

Consideremos ahora el siguiente sistema de vectores:

$$S = \{s_i\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Podemos comprobar fácilmente que es un sistema ortogonal, hallando los productos escalares de los vectores de S dos a dos:

$$\langle s_1, s_2 \rangle = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix} \right\rangle = 1 \cdot (-1) + 1 \cdot 2 + 1 \cdot (-1) + 0 \cdot 0 = 0$$

$$\langle s_1, s_3 \rangle = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix} \right\rangle = 1 \cdot 1 + 1 \cdot 0 + 1 \cdot (-1) + 0 \cdot 2 = 0$$

$$\langle s_1, s_4 \rangle = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle = 1 \cdot (-1) + 1 \cdot 0 + 1 \cdot 1 + 0 \cdot 1 = 0$$

$$\langle s_2, s_3 \rangle = \left\langle \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix} \right\rangle = (-1) \cdot 1 + 2 \cdot 0 + (-1) \cdot (-1) + 0 \cdot 2 = 0$$

$$\langle s_2, s_4 \rangle = \left\langle \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle = (-1) \cdot (-1) + 2 \cdot 0 + (-1) \cdot 1 + 0 \cdot 1 = 0$$

$$\langle s_3, s_4 \rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle = 1 \cdot (-1) + 0 \cdot 0 + (-1) \cdot 1 + 2 \cdot 1 = 0$$

Tenemos, por tanto, un sistema de 4 vectores ortogonales en \mathbb{R}^4 , y, como hemos comentado antes, son linealmente independientes, luego forman una base del espacio vectorial.

5.3 Qué es el complemento ortogonal de un subespacio

Consideremos un espacio vectorial V con producto escalar $\langle \cdot, \cdot \rangle$, y sea U un subespacio de V .

Definición: Complemento ortogonal

Llamamos **complemento ortogonal de U** al conjunto de vectores que son ortogonales a todos los de U :

$$U^\perp = \{v \in V : \langle v, u \rangle = 0 \text{ para todo } u \in U\}$$

Tenemos las siguientes propiedades del complemento ortogonal de un subespacio:

Teorema (Propiedades del complemento ortogonal)

- U^\perp es un subespacio vectorial de V .
- $U \cap U^\perp = \{0\}$.
- Si V es de dimensión finita, $\dim(U) + \dim(U^\perp) = \dim(V)$.

Como consecuencia, este resultado nos dice que V **se puede descomponer como la suma directa de U y de su complemento ortogonal**: $U \oplus U^\perp = V$.

¿Cómo calculamos el complemento ortogonal de un subespacio?

Hemos de hallar los vectores que son ortogonales a un subespacio U . Como hemos [visto antes](#), un vector v es ortogonal a U si, y solo si, es ortogonal a todos los vectores de una base de U .

Así pues, partimos de que tenemos una base $\mathcal{B} = \{u_1, \dots, u_m\}$ de U y tomemos un

vector genérico $v \in V$. Para que $v \in U^\perp$, debe cumplirse que $\langle v, u_i \rangle = 0$ para todo $i = 1, \dots, m$.

Cada una de las restricciones $\langle v, u_i \rangle = 0$ nos proporciona una ecuación lineal que debe verificar un vector para poder pertenecer a U^\perp .

Tomamos, por tanto, todas las ecuaciones lineales que se deducen de las restricciones anteriores: ese sistema de ecuaciones nos representa las ecuaciones cartesianas del espacio U^\perp .

En el caso concreto de un espacio $V = \mathbb{R}^n$ con el *producto escalar euclídeo*, podemos ver qué aspecto tienen las ecuaciones del tipo $\langle v, u_i \rangle = 0$. Supongamos que los vectores v y u_i tienen por coordenadas, en la base canónica, las siguientes:

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad u_i = \begin{pmatrix} \alpha_{i,1} \\ \alpha_{i,2} \\ \vdots \\ \alpha_{i,n} \end{pmatrix}$$

Entonces

$$\langle v, u_i \rangle = 0 \Leftrightarrow \left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} \alpha_{i,1} \\ \alpha_{i,2} \\ \vdots \\ \alpha_{i,n} \end{pmatrix} \right\rangle = 0 \Leftrightarrow \alpha_{i,1}x_1 + \alpha_{i,2}x_2 + \dots + \alpha_{i,n}x_n = 0$$

Si nos fijamos, nos queda una ecuación lineal cuyas incógnitas son las coordenadas canónicas de v y cuyos coeficientes son las coordenadas de u_i . Si repetimos el proceso para todo $i = 1, \dots, m$, tenemos el siguiente sistema de ecuaciones:

$$Av = 0 \quad \text{donde} \quad A = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \dots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \dots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \dots & \alpha_{m,n} \end{pmatrix}$$

Es decir, las *ecuaciones cartesianas* de U^\perp son las que tienen por matriz de coeficientes a las coordenadas de los vectores de la base de U puestos *por filas*. De hecho, si llamamos B a la matriz que resulta de poner por columnas los vectores de \mathcal{B} , entonces $A = B^t$.

A partir de esas ecuaciones cartesianas, ya podríamos [encontrar una base](#) del subespacio U^\perp .

Ejemplo

Consideremos el espacio U generado por la base siguiente:

$$\mathcal{B} = \{u_i\} = \left\{ \begin{pmatrix} -2 \\ 0 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -2 \\ 2 \end{pmatrix} \right\}$$

Para hallar su complemento ortogonal, suponemos un vector genérico

$$v = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4$$

y hacemos su producto escalar por todos los vectores de \mathcal{B} , imponiendo que valga 0:

$$\langle v, u_1 \rangle = 0 \Leftrightarrow \left\langle \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ -2 \\ 2 \end{pmatrix} \right\rangle = 0 \Leftrightarrow -2x - 2z + 2t = 0$$

$$\langle v, u_2 \rangle = 0 \Leftrightarrow \left\langle \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -2 \\ 2 \end{pmatrix} \right\rangle = 0 \Leftrightarrow -x - y - 2z + 2t = 0$$

Hemos llegado, por tanto, al siguiente sistema de ecuaciones que representan las ecuaciones cartesianas de U^\perp :

$$\begin{array}{rrcr} -2x & & -2z & +2t & = & 0 \\ -x & -y & -2z & +2t & = & 0 \end{array}$$

Nos podemos dar cuenta de que la matriz del sistema es la que resulta de poner los vectores de la base \mathcal{B} *por filas*.

Como siempre, podemos resolver el sistema por Gauss-Jordan para [deducir la forma paramétrica](#) de la solución, y, por tanto, un [sistema generador](#) de U^\perp :

$$\begin{pmatrix} -2 & 0 & -2 & 2 \\ -1 & -1 & -2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & -1 \end{pmatrix} \Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

De aquí que una base del complemento ortogonal U^\perp sea:

$$\mathcal{B}' = \left\{ \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Podemos comprobar fácilmente que los vectores de \mathcal{B}' son ortogonales a los de \mathcal{B} .

5.4 Cómo calculamos la proyección de un vector sobre un subespacio

Definición: Proyección ortogonal

Consideremos un espacio euclídeo $(V, \langle \cdot, \cdot \rangle)$. Consideremos un subespacio U , del que conocemos una **base ortogonal** $\mathcal{B}_U = \{u_1, \dots, u_m\}$. Definimos la **proyección ortogonal de un vector v sobre el subespacio U** , como:

$$\text{proy}_U(v) = \frac{\langle v, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle v, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 + \dots + \frac{\langle v, u_m \rangle}{\langle u_m, u_m \rangle} u_m$$

Es decir, la proyección $\text{proy}_U(v)$ es un vector del subespacio U , ya que es combinación lineal de los elementos de su base, y además esta combinación lineal tiene como coeficientes el producto escalar de v por cada elemento de la base, dividido por la norma de u_i al cuadrado (**recordemos que $\|x\| = \sqrt{\langle x, x \rangle}$**).

Evidentemente, si los elementos de la base de U tienen norma 1, el denominador desaparece.

La proyección de v sobre el subespacio U tiene la propiedad siguiente:

Teorema

Si u es un vector del subespacio U , y $v \in V$, entonces $d(u, v) \geq d(\text{proy}_U(v), v)$, y la única forma para que se dé la igualdad es que $u = \text{proy}_U(v)$.

Esto significa que la proyección ortogonal de v sobre U es el vector de U más cercano (según la distancia d) a v .

Como caso particular, podríamos definir la **proyección ortogonal de un vector v sobre otro vector u** como

$$\text{proy}_u(v) = \frac{\langle v, u \rangle}{\langle u, u \rangle} u$$

(realmente es la proyección del vector v sobre el **subespacio generado** por el vector u).

Podemos entonces concluir que, si la base ortogonal de U es $\mathcal{B}_U = \{u_1, \dots, u_m\}$, entonces

$$\text{proy}_U(v) = \text{proy}_{u_1}(v) + \dots + \text{proy}_{u_m}(v)$$

A partir del hecho de que la base de U es *ortogonal*, podemos deducir que:

Teorema

Si U es un subespacio de V , $v \in V$, y una base ortogonal de U es $\mathcal{B}_U = \{u_1, \dots, u_m\}$, entonces el vector $v - \text{proy}_U(v)$ es ortogonal a U , es decir, $v - \text{proy}_U(v) \in U^\perp$ está en el **complemento ortogonal** de U .

Esto, junto con que $\text{proy}_U(v) \in U$, nos justifican el siguiente resultado, llamado **teorema de la descomposición ortogonal**:

Teorema (Descomposición ortogonal)

En un espacio euclídeo $(V, \langle \cdot, \cdot \rangle)$, dado un subespacio U , y $v \in V$, existe una única descomposición de v como suma de dos vectores ortogonales, $v = v_1 + v_2$, tales que $v_1 \in U$ y $v_2 \in U^\perp$. Concretamente, $v_1 = \text{proy}_U(v)$ y $v_2 = v - v_1 = \text{proy}_{U^\perp}(v)$.

Podemos siempre descomponer un vector de V como suma de dos vectores, uno de U y otro de U^\perp , que coinciden con las proyecciones sobre ambos subespacios.

Ejemplo

Consideremos el subespacio U cuya base es

$$\mathcal{B} = \{u_i\} = \left\{ \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ -5 \\ 1 \\ -3 \end{pmatrix} \right\}$$

Notemos que \mathcal{B} es una base *ortogonal*.

Consideremos el vector

$$v = \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}$$

Buscamos la *descomposición ortogonal* de v con respecto al subespacio U . Sabemos que podemos escribir $v = v_1 + v_2$, donde $v_1 = \text{proy}_U(v)$ y $v_2 = v - v_1 \in U^\perp$.

$$v_1 = \text{proy}_U(v) = \text{proy}_{u_1}(v) + \text{proy}_{u_2}(v) = \frac{\langle v, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle v, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2$$

Si calculamos esas cantidades

$$\langle v, u_1 \rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \right\rangle = (-1) \cdot (-1) + 1 \cdot 1 + (-2) \cdot (-1) + 0 \cdot (-1) = 4$$

$$\langle v, u_2 \rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ -5 \\ 1 \\ -3 \end{pmatrix} \right\rangle = (-1) \cdot (-3) + 1 \cdot (-5) + (-2) \cdot 1 + 0 \cdot (-3) = -4$$

$$\langle u_1, u_1 \rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \right\rangle = (-1)^2 + 1^2 + (-1)^2 + (-1)^2 = 4$$

$$\langle u_2, u_2 \rangle = \left\langle \begin{pmatrix} -3 \\ -5 \\ 1 \\ -3 \end{pmatrix}, \begin{pmatrix} -3 \\ -5 \\ 1 \\ -3 \end{pmatrix} \right\rangle = (-3)^2 + (-5)^2 + 1^2 + (-3)^2 = 44$$

podemos sustituirlas en la expresión de más arriba y llegamos a

$$v_1 = 1 \cdot \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix} + \left(-\frac{1}{11}\right) \cdot \begin{pmatrix} -3 \\ -5 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -\frac{8}{11} \\ \frac{16}{11} \\ -\frac{12}{11} \\ -\frac{8}{11} \end{pmatrix}$$

Por tanto,

$$v_2 = \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix} - \begin{pmatrix} -\frac{8}{11} \\ \frac{16}{11} \\ -\frac{12}{11} \\ -\frac{8}{11} \end{pmatrix} = \begin{pmatrix} -\frac{3}{11} \\ -\frac{5}{11} \\ -\frac{10}{11} \\ \frac{8}{11} \end{pmatrix}$$

Por lo comentado anteriormente, sabemos que $v_2 = \text{proy}_{U^\perp}(v) \in U^\perp$, lo cual se podría comprobar haciendo el producto escalar $\langle v_2, u_i \rangle$ y viendo que su valor es 0 para todo i .

5.5 Qué es y qué utilidad tiene una base ortonormal

Definición: Vector unitario

Un vector $v \in V$ se dice **unitario** si $\|v\| = 1$.

Es fácil crear vectores unitarios: dado $u \in V$, si definimos $v = \frac{1}{\|u\|}u$, este vector tiene norma 1. A este proceso se le llama *normalizar*.

Un sistema de vectores $\{v_1, \dots, v_k\}$ se llama **sistema ortonormal** si es un [sistema ortogonal](#) y cada v_i es un vector unitario ($i = 1, \dots, k$).

En relación a lo comentado de las propiedades de los sistemas ortogonales, en nuestro caso podemos decir que un sistema ortonormal $\{v_1, \dots, v_n\}$, donde $\dim(V) = n$, es una **base ortonormal**.

El interés de una base ortonormal es que es sencillo calcular las [coordenadas](#) de cualquier vector con respecto a dicha base:

Teorema (Coordenadas en base ortonormal)

Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base *ortonormal* del espacio vectorial V . Entonces, para cada $v \in V$, tenemos:

$$v = \langle v, v_1 \rangle v_1 + \dots \langle v, v_n \rangle v_n$$

luego las coordenadas de v en la base \mathcal{B} las podemos calcular como:

$$[v]_{\mathcal{B}} = \begin{pmatrix} \langle v, v_1 \rangle \\ \langle v, v_2 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{pmatrix}_{\mathcal{B}}$$

Ejemplo

Consideremos ahora el sistema de vectores ortogonales del [ejemplo anterior](#):

$$S = \{s_i\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

A partir de S , formemos un sistema ortonormal S' , donde cada vector de S' es un vector de S , que se ha normalizado.

$$S' = \left\{ \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ \frac{-\sqrt{6}}{6} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{6}}{6} \\ 0 \\ \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{3}}{3} \\ 0 \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix} \right\}$$

S' forma una base *ortonormal* de \mathbb{R}^4 . Si consideramos el vector

$$v = \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}$$

podemos hallar sus coordenadas en la base S' , ya que cada coordenada de v será el producto escalar de v por el correspondiente vector de S' :

$$\langle v, s'_1 \rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \end{pmatrix} \right\rangle = \frac{-2 \cdot \sqrt{3}}{3}$$

$$\langle v, s'_2 \rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ \frac{-\sqrt{6}}{6} \\ 0 \end{pmatrix} \right\rangle = \frac{5 \cdot \sqrt{6}}{6}$$

$$\langle v, s'_3 \rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{6}}{6} \\ 0 \\ \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \end{pmatrix} \right\rangle = \frac{\sqrt{6}}{6}$$

$$\langle v, s'_4 \rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{3}}{3} \\ 0 \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix} \right\rangle = \frac{-\sqrt{3}}{3}$$

Entonces, podremos escribir:

$$v = \begin{pmatrix} \frac{-2\sqrt{3}}{3} \\ \frac{5\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \\ \frac{-\sqrt{3}}{3} \end{pmatrix}_{S'}$$

5.6 Cómo construimos una base ortonormal

En las secciones anteriores hemos usado bases ortogonales y ortonormales, pero cabe preguntarnos si la existencia de tales bases está asegurada. La respuesta nos la da el siguiente resultado:

Teorema

Sea V un espacio vectorial euclídeo de *dimensión finita*. Entonces V tiene una base ortonormal.

La demostración de este resultado teórico realmente nos presenta un método constructivo que permite hallar una base ortonormal del espacio V a partir de una base cualquiera.

Método de Ortonormalización de Gram-Schmidt

Consideremos una base $\mathcal{B} = \{v_1, \dots, v_n\}$ en V . Vamos a utilizar el método de Gram-Schmidt para construir una base ortonormal a partir de ella.

El proceso se divide en dos fases:

- En la primera fase, construiremos una base ortogonal \mathcal{B}' .
- En un segundo paso, cada vector de \mathcal{B}' se normaliza, de forma que obtendremos una base ortonormal.

Primera fase

El proceso del primer paso es incremental.

Llamamos $u_1 = v_1$, el primer vector de la base \mathcal{B} .

Ahora consideramos el segundo vector v_2 . Por un [resultado teórico](#) sobre las proyecciones que vimos antes, si llamamos $U_1 = \mathcal{L}(\{u_1\})$, entonces $v_2 - \text{proy}_{U_1}(v_2)$ es un vector ortogonal a U_1 , en particular es ortogonal a u_1 .

Llamemos entonces $u_2 = v_2 - \text{proy}_{U_1}(v_2)$. Entonces, el sistema $\{u_1, u_2\}$ es ortogonal.

Llamamos ahora $U_2 = \mathcal{L}(\{u_1, u_2\})$. Entonces, por el mismo motivo que antes, si llamamos $u_3 = v_3 - \text{proy}_{U_2}(v_3)$, tenemos que u_3 es ortogonal a U_2 , luego $\{u_1, u_2, u_3\}$ es un sistema ortogonal.

De esta forma, sucesivamente, en el paso m , vamos construyendo $U_m = \mathcal{L}(\{u_1, \dots, u_m\})$ y consideramos $u_{m+1} = v_{m+1} - \text{proy}_{U_m}(v_{m+1})$, que es un vector ortogonal a U_m y, por tanto, $\{u_1, \dots, u_{m+1}\}$ es un sistema ortogonal.

Una vez se haya completado el proceso, se habrá construido el conjunto $\mathcal{B}' = \{u_1, \dots, u_n\}$. Como estamos en un espacio V de dimensión n , y tenemos n vectores ortogonales, luego **linealmente independientes**, éstos forman una base. Luego \mathcal{B}' es una base ortogonal.

Segunda fase

La segunda fase es más sencilla, ya que implica únicamente normalizar cada vector de \mathcal{B}' .

De esta forma, la base ortonormal es la formada por

$$\left\{ \frac{u_1}{\|u_1\|}, \dots, \frac{u_n}{\|u_n\|} \right\}$$

Ejemplo

Consideremos la base \mathcal{B} en \mathbb{R}^4 dada por

$$\mathcal{B} = \{v_i\} = \left\{ \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \right\}$$

Vamos a usar el método de Gram-Schmidt para encontrar una base ortonormal, a partir de \mathcal{B} .

Fase 1

Llamamos

$$u_1 = v_1 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

A partir de este momento, procedemos de forma recursiva:

- Construimos:

$$U_1 = \mathcal{L}(\{u_1\}) = \mathcal{L}\left(\left\{\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}\right\}\right)$$

Entonces

$$\begin{aligned} \text{proy}_{U_1}(v_2) &= \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 \\ &= \frac{0}{4} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

Y llamamos

$$u_2 = v_2 - \text{proy}_{U_1}(v_2) = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

- Construimos:

$$U_2 = \mathcal{L}(\{u_1, u_2\}) = \mathcal{L}\left(\left\{\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}\right\}\right)$$

Entonces

$$\begin{aligned} \text{proy}_{U_2}(v_3) &= \frac{\langle v_3, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle v_3, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 \\ &= \frac{3}{4} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} \\ -1 \\ -\frac{1}{2} \\ -1 \end{pmatrix} \end{aligned}$$

Y llamamos

$$u_3 = v_3 - \text{proy}_{U_2}(v_3) = \begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \end{pmatrix} - \begin{pmatrix} -\frac{1}{2} \\ -1 \\ -\frac{1}{2} \\ -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}$$

- Construimos:

$$U_3 = \mathcal{L}(\{u_1, u_2, u_3\}) = \mathcal{L}\left(\left\{\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}\right\}\right)$$

Entonces

$$\begin{aligned} \text{proy}_{U_3}(v_4) &= \frac{\langle v_4, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle v_4, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 + \frac{\langle v_4, u_3 \rangle}{\langle u_3, u_3 \rangle} u_3 \\ &= \frac{0}{4} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} + \frac{0}{4} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} + \frac{\frac{1}{2}}{\frac{1}{2}} \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

Y llamamos

$$u_4 = v_4 - \text{proy}_{U_3}(v_4) = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

Por tanto, la base ortogonal a la que llegamos es

$$\mathcal{B}' = \{u_i\} = \left\{ \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\}$$

Fase 2

Ahora debemos multiplicar cada vector u_i en \mathcal{B}' por el inverso de su norma, y llegamos a los vectores $\frac{1}{\|u_i\|}u_i$, que nos forman la base ortonormal siguiente:

$$\begin{aligned}\mathcal{B}'' &= \left\{ \frac{1}{2} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{-1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\} = \\ &= \left\{ \begin{pmatrix} \frac{-1}{2} \\ \frac{-1}{2} \\ \frac{-1}{2} \\ \frac{-1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{-1}{2} \\ \frac{1}{2} \\ \frac{-1}{2} \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{\sqrt{2}}{2} \\ 0 \\ \frac{-\sqrt{2}}{2} \end{pmatrix} \right\}\end{aligned}$$

5.7 Qué es una matriz ortogonal y una aplicación ortogonal

Definición: Matriz ortogonal

Una matriz cuadrada Q se llama **ortogonal** si $Q^t Q = I$ o, equivalentemente, $Q^{-1} = Q^t$.

Es decir, una matriz ortogonal es aquella cuya inversa es su propia traspuesta.

Teorema

Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ es ortogonal si, y sólo si, sus columnas forman una base *ortonormal* de \mathbb{R}^n .

Teorema

Sea A una matriz ortogonal. Entonces $\det(A) = \pm 1$.

Definición: Endomorfismo ortogonal

Un endomorfismo $f : V \rightarrow V$ se llama **ortogonal** si respeta el producto escalar, es decir, si $\langle \cdot, \cdot \rangle$ es el producto escalar en V , entonces $\langle f(u), f(v) \rangle = \langle u, v \rangle$ para todo $u, v \in V$.

Como la aplicación ortogonal conserva el producto escalar, también conserva todas las cantidades que hemos definido a partir de él: **la norma**, **la distancia** y **el ángulo** entre vectores de V : $\|f(v)\| = \|v\|$, $d(f(u), f(v)) = d(u, v)$ y $\text{ang}(f(u), f(v)) = \text{ang}(u, v)$ para todo $u, v \in V$.

Teorema

Una aplicación $f : V \rightarrow V$ es ortogonal si, y sólo si, la imagen de toda base ortonormal es de nuevo una base ortonormal de V .

¿Qué relación existe entre matrices ortogonales y aplicaciones ortogonales?

Teorema

Sea $f : V \rightarrow V$ una aplicación lineal y sea A su matriz asociada respecto a una base ortonormal de V . Entonces f es ortogonal si, y solo si, la matriz A es ortogonal.

Ejemplo

Como hemos comentado, si tomamos una base ortonormal y ponemos sus vectores por columnas, nos forman una matriz *ortogonal*. Por tanto, podemos considerar la base \mathcal{B}'' del [ejemplo anterior](#), y construir la matriz ortogonal asociada:

$$Q = \begin{pmatrix} \frac{-1}{2} & \frac{1}{2} & \frac{-\sqrt{2}}{2} & 0 \\ \frac{-1}{2} & \frac{-1}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{-1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{-1}{2} & \frac{-1}{2} & 0 & \frac{-\sqrt{2}}{2} \end{pmatrix}$$

Es fácil comprobar que, en este caso, $Q^t Q = I$.

Hay otros ejemplos de matrices ortogonales. En particular, una tipología concreta de matrices ortogonales en \mathbb{R}^2 es muy utilizada en la práctica. Llamemos:

$$R(\alpha) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Son las denominadas **matrices de rotación**. Están asociadas a la aplicación lineal que representa un giro o rotación de α radianes en sentido positivo (*antihorario*) alrededor del origen de coordenadas.

Así

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R(\alpha) \begin{pmatrix} x \\ y \end{pmatrix}$$

son las coordenadas del vector girado, en función de las coordenadas del vector original.

Estas matrices de rotación se pueden extender a dimensiones superiores. Por ejemplo, en \mathbb{R}^3 , la rotación de α radianes alrededor del eje X se representa mediante la matriz ortogonal

$$R(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

5.8 Qué es la diagonalización ortogonal

Recordemos [la definición de endomorfismo o de matriz diagonalizable](#). Una aplicación lineal $f : V \rightarrow V$ es diagonalizable si y solo si existe una base de V formada por autovectores de f , lo cual equivale a que la representación matricial de f en dicha base sea una matriz diagonal D . En representación matricial, esto se traduce en que una matriz $A \in \mathcal{M}_n(\mathbb{R})$ es diagonalizable si y solo si, existe una matriz regular P tal que $D = P^{-1} A P$. Las columnas de P están formadas por los autovectores de la base de la definición.

Veamos cómo influye el hecho de tener bases ortonormales en un espacio euclídeo en la diagonalización, lo que da lugar a la denominada *diagonalización ortogonal*.

Definición: Aplicación diagonalizable ortogonalmente

Diremos que una aplicación lineal $f : V \rightarrow V$ es **diagonalizable ortogonalmente** si y solo si existe una base ortonormal en V formada por autovectores de f .

Definición: Matriz diagonalizable ortogonalmente

En matrices, una matriz A se dice **diagonalizable ortogonalmente** si existe una matriz P *ortogonal* tal que $D = P^t A P$, donde D , la matriz asociada a f en la base de autovectores, es diagonal con autovalores en la diagonal principal.

Ejemplo

Consideremos la matriz A dada por

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

asociada a un endomorfismo $f : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ en la base canónica.

Si tomamos la base

$$\mathcal{B} = \left\{ \begin{pmatrix} \frac{-\sqrt{6}}{6} \\ \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{6}}{6} \\ \frac{-\sqrt{6}}{6} \\ 0 \\ \frac{\sqrt{6}}{3} \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \\ \frac{\sqrt{3}}{3} \end{pmatrix} \right\}$$

resulta ser una [base ortonormal](#) formada por autovectores de f (en [la siguiente sección](#) detallaremos el proceso completo).

Por tanto, esta matriz (y el endomorfismo f) es diagonalizable ortogonalmente. De hecho, si consideramos como matriz P la del [cambio de base](#) de \mathcal{B} a la canónica, podemos calcular la expresión diagonal, teniendo en cuenta que P es una matriz ortogonal:

$$\begin{aligned} D &= P^t A P = \\ &= \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & \frac{-\sqrt{3}}{3} \\ \frac{-\sqrt{6}}{6} & \frac{-\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} & 0 \\ 0 & \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \end{pmatrix}^t \begin{pmatrix} 0 & -2 & -1 & -1 \\ -2 & 0 & -1 & 1 \\ -1 & -1 & -1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & \frac{-\sqrt{3}}{3} \\ \frac{-\sqrt{6}}{6} & \frac{-\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} & 0 \\ 0 & \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \end{aligned}$$

donde podemos comprobar que los elementos de la diagonal de D son autovalores de A .

5.9 Qué matrices son diagonalizables ortogonalmente

Supongamos una matriz $A \in \mathcal{M}_n(\mathbb{R})$ que sea [diagonalizable ortogonalmente](#), es decir, tal que existe P ortogonal con $P^t A P = D$ una matriz diagonal.

Como D es diagonal, en particular tenemos que $D^t = D$, luego:

$$P^t A P = D = D^t = (P^t A P)^t = P^t A^t (P^t)^t = P^t A^t P$$

Luego, multiplicando por P a la izquierda en los extremos de la igualdad, y por P^t en la derecha, recordando que P es una [matriz ortogonal](#), llegamos a que $A = A^t$.

Esto significa que:

Teorema

Si una matriz $A \in \mathcal{M}_n(\mathbb{R})$ es *diagonalizable ortogonalmente*, entonces es simétrica.

Por tanto, una matriz *no simétrica* podrá ser *diagonalizable* (dependiendo de si cumple los criterios ya estudiados), pero *nunca* puede ser *diagonalizable ortogonalmente*.

Podemos estudiar entonces algunas de las propiedades de las matrices simétricas en lo que se refiere a diagonalización:

Teorema

Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica. Entonces:

- Todos los autovalores de A son reales.
- Si λ_1 y λ_2 son dos autovalores distintos, y sus subespacios asociados son U_{λ_1} y U_{λ_2} , respectivamente, entonces U_{λ_1} es ortogonal a U_{λ_2} .
- La multiplicidad geométrica de cada autovalor de A coincide con su multiplicidad algebraica.

En consecuencia

Teorema

Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ es diagonalizable ortogonalmente si y solo si es simétrica.

¿Cómo diagonalizamos ortogonalmente una matriz simétrica?

El procedimiento es idéntico al que se conoce para *diagonalizar* una matriz cualquiera, con las siguientes puntualizaciones:

- Al ser una matriz simétrica, tenemos asegurada que es diagonalizable (incluso ortogonalmente).
- Al calcular la base del subespacio asociado a un autovalor λ , podemos hacer un paso extra, usando el *método de Gram-Schmidt*, para tener dicha base en forma ortonormal.
- Al finalizar, al igual que hacíamos en la diagonalización *estándar*, podemos unir todas las bases de los distintos subespacios asociados a los autovalores y tenemos una base ortonormal del espacio vectorial V .

De forma esquemática, incluyendo todo lo que conocemos de diagonalización, queda:

1. Calcular el *polinomio característico* $p(\lambda)$ de A , y sus raíces, los autovalores de A . Sean $\lambda_1, \dots, \lambda_k$ los autovalores.
2. Para cada autovalor λ :

1. Determinar una base \mathcal{B}_λ del subespacio propio U_λ .
2. Aplicar el método de Gram-Schmidt a \mathcal{B}_λ para obtener una base _ortonormal de U_λ a la que llamaremos \mathcal{B}'_λ .
3. Llamamos $\mathcal{B} = \mathcal{B}'_{\lambda_1} \cup \dots \cup \mathcal{B}'_{\lambda_k}$. Entonces \mathcal{B} es una base ortonormal de V . Además, la matriz P de cambio de base de \mathcal{B} a la canónica es una matriz ortogonal, tiene por columnas las coordenadas de los vectores de \mathcal{B} , y verifica que $D = P^t A P$ es una matriz diagonal con los autovalores en la diagonal principal (con multiplicidades incluidas).

Ejemplo

Consideremos la matriz

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

Como es simétrica, es diagonalizable ortogonalmente. Vamos a seguir los pasos vistos para encontrar su forma diagonal y la base ortonormal de autovectores.

Paso 1: Encontrar los autovalores de A

Para calcular los autovalores de A , comenzamos por construir su polinomio característico:

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) = \\ &= \det \left(\begin{pmatrix} 0 & -2 & -1 & -1 \\ -2 & 0 & -1 & 1 \\ -1 & -1 & -1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) = \\ &= \det \begin{pmatrix} -\lambda & -2 & -1 & -1 \\ -2 & -\lambda & -1 & 1 \\ -1 & -1 & -1-\lambda & 0 \\ -1 & 1 & 0 & 1-\lambda \end{pmatrix} \\ &= \lambda^4 - 9\lambda^2 \end{aligned}$$

Igualándolo a 0 y resolviendo la ecuación que queda, llegamos a que los autovalores son: $\lambda = 0, -3, 3$, ya que:

$$p(\lambda) = \lambda^2 \cdot (\lambda + 3) \cdot (\lambda - 3)$$

Paso 2: Encontrar una base ortonormal de cada subespacio propio

Recorremos ahora cada autovalor, encontrando una base del subespacio propio asociado, y ortonormalizándola usando el [método de Gram-Schmidt](#).

- Para el autovalor $\lambda = 0$:

$$\begin{aligned} (A - 0I) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} 0 & -2 & -1 & -1 \\ -2 & 0 & -1 & 1 \\ -1 & -1 & -1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow \\ &\Leftrightarrow \begin{cases} -2y - z - t = 0 \\ -2x - z + t = 0 \\ -x - y - z = 0 \\ -x + y + t = 0 \end{cases} \end{aligned}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de U_0 :

$$\begin{aligned} \left(\begin{array}{cccc|c} 0 & -2 & -1 & -1 & 0 \\ -2 & 0 & -1 & 1 & 0 \\ -1 & -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \end{array} \right) &\sim \left(\begin{array}{cccc|c} -2 & 0 & -1 & 1 & 0 \\ 0 & -2 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \\ \Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} &= \alpha \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

Luego una base de U_0 será:

$$\mathcal{B}_{U_0} = \left\{ \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right\}$$

El siguiente paso es *ortonormalizar* esta base, mediante el método de Gram-Schmidt. Llamemos v_i a los vectores de esta base hallada. Llamamos

$$u_1 = v_1 = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}$$

Procedemos con el resto de vectores:

Construimos:

$$U_1 = \mathcal{L}(\{u_1\}) = \mathcal{L}\left(\left\{\begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}\right\}\right)$$

Entonces

$$\text{proy}_{U_1}(v_2) = \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 = \frac{0}{\frac{3}{2}} \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Y llamamos

$$u_2 = v_2 - \text{proy}_{U_1}(v_2) = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}$$

Y así hemos llegado a un sistema ortogonal formado por los vectores u_i . Debemos normalizar ahora los vectores de este sistema ortogonal que tenemos, multiplicando cada uno por el inverso de su norma, llegando a la siguiente base ortonormal de U_0 .

$$\mathcal{B}'_{U_0} = \left\{ \frac{1}{\frac{\sqrt{6}}{2}} \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\frac{\sqrt{6}}{2}} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \frac{-\sqrt{6}}{6} \\ \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{6}}{6} \\ \frac{-\sqrt{6}}{6} \\ 0 \\ \frac{\sqrt{6}}{3} \end{pmatrix} \right\}$$

- Para el autovalor $\lambda = -3$:

$$\begin{aligned} (A + 3I) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} 3 & -2 & -1 & -1 \\ -2 & 3 & -1 & 1 \\ -1 & -1 & 2 & 0 \\ -1 & 1 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow \\ &\Leftrightarrow \begin{cases} 3x - 2y - z - t = 0 \\ -2x + 3y - z + t = 0 \\ -x - y + 2z = 0 \\ -x + y + 4t = 0 \end{cases} \end{aligned}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de U_{-3} :

$$\left(\begin{array}{cccc|c} 3 & -2 & -1 & -1 & 0 \\ -2 & 3 & -1 & 1 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ -1 & 1 & 0 & 4 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} 3 & 0 & -3 & 0 & 0 \\ 0 & \frac{5}{3} & -\frac{5}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{18}{5} & 0 \end{array} \right)$$

$$\Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Luego una base de U_{-3} será:

$$\mathcal{B}_{U_{-3}} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

Como solo tenemos un vector en la base, ya forma un sistema ortogonal. Debemos normalizar ahora los vectores de este sistema ortogonal que tenemos, multiplicando cada uno por el inverso de su norma, llegando a la siguiente base ortonormal de U_{-3} .

$$\mathcal{B}'_{U_{-3}} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \end{pmatrix} \right\}$$

- Para el autovalor $\lambda = 3$:

$$(A - 3I) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow \begin{pmatrix} -3 & -2 & -1 & -1 \\ -2 & -3 & -1 & 1 \\ -1 & -1 & -4 & 0 \\ -1 & 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} -3x - 2y - z - t = 0 \\ -2x - 3y - z + t = 0 \\ -x - y - 4z = 0 \\ -x + y - 2t = 0 \end{cases}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de U_3 :

$$\left(\begin{array}{cccc|c} -3 & -2 & -1 & -1 & 0 \\ -2 & -3 & -1 & 1 & 0 \\ -1 & -1 & -4 & 0 & 0 \\ -1 & 1 & 0 & -2 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} -3 & 0 & 0 & -3 & 0 \\ 0 & -\frac{5}{3} & 0 & \frac{5}{3} & 0 \\ 0 & 0 & -\frac{18}{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$$\Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Luego una base de U_3 será:

$$\mathcal{B}_{U_3} = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Como solo tenemos un vector en la base, ya forma un sistema ortogonal. Debemos normalizar ahora los vectores de este sistema ortogonal que tenemos, multiplicando cada uno por el inverso de su norma, llegando a la siguiente base ortonormal de U_3 .

$$\mathcal{B}'_{U_3} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \\ \frac{\sqrt{3}}{3} \end{pmatrix} \right\}$$

Paso 3: Hallar una base ortonormal formada por autovectores

Definimos

$$\mathcal{B} = \mathcal{B}'_{U_0} \cup \mathcal{B}'_{U_{-3}} \cup \mathcal{B}'_{U_3} = \left\{ \begin{pmatrix} \frac{-\sqrt{6}}{6} \\ \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{6}}{6} \\ \frac{-\sqrt{6}}{6} \\ 0 \\ \frac{\sqrt{6}}{3} \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \\ \frac{\sqrt{3}}{3} \end{pmatrix} \right\}$$

Es una base ortonormal de \mathbb{R}^4 , formada únicamente por autovectores de la matriz A .

Además, la matriz P del [cambio de base](#) de esta base \mathcal{B} a la canónica, formada por los vectores de \mathcal{B} puestos por columnas, es *ortogonal* y nos proporciona la relación

entre A y la matriz diagonal D de los autovalores:

$$\begin{aligned}
 D &= P^t A P = \\
 &= \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{-\sqrt{6}}{6} & \frac{\sqrt{6}}{3} & 0 \\ \frac{\sqrt{6}}{6} & \frac{-\sqrt{6}}{6} & 0 & \frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & 0 \\ \frac{-\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & 0 & \frac{\sqrt{3}}{3} \end{pmatrix} \begin{pmatrix} 0 & -2 & -1 & -1 \\ -2 & 0 & -1 & 1 \\ -1 & -1 & -1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & \frac{-\sqrt{3}}{3} \\ \frac{-\sqrt{6}}{6} & \frac{-\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} & 0 \\ 0 & \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \end{pmatrix} \\
 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}
 \end{aligned}$$

Recordemos que en la diagonal de D se encuentran los autovalores (en el mismo orden en que se han puesto los autovectores en las columnas de P), con sus respectivas multiplicidades.

Capítulo 6

Problemas Resueltos

En este capítulo planteamos varios problemas que recorren la mayoría de los contenidos de este libro y que intentan poner en práctica todas las técnicas de cada tema.

En cada problema resuelto, se hará referencia a la sección donde se ha dado la explicación que ayuda a la resolución del ejercicio.

6.1 Espacios Vectoriales

Ejercicio

Consideremos como espacio vectorial $V = \mathbb{R}^4$ y los subespacios siguientes:

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : 3x + y - z + 2t = 0 \right\}$$
$$W = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 : \begin{array}{rcl} 3x + 4y & = & 0 \\ 9x + 4z + 4t & = & 0 \end{array} \right\}$$

- Hallar la dimensión y una base de U y de W .
- Hallar la dimensión y una base de $U \cap W$ y de $U + W$.
- Comprobar el teorema de la dimensión.
- Encontrar el subespacio suplementario de U .

e. Consideremos el vector v dado por sus coordenadas en la base canónica

$$\begin{pmatrix} -5 \\ -9 \\ 1 \\ -5 \end{pmatrix}. \text{ Consideremos también la base:}$$

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Encontrar las coordenadas de v en la nueva base.

f. Encontrar las matrices de cambio de base entre \mathcal{B}_1 y \mathcal{C} (en ambos sentidos), siendo \mathcal{C} la base canónica.

Solución

a. Para encontrar unas bases de U y de W podemos seguir las indicaciones sobre [cómo encontrar un sistema generador y una base a partir de la descripción de un subespacio](#).

En el caso del subespacio U , el sistema homogéneo dado por su ecuaciones cartesianas es:

$$\begin{cases} 3x + y - z + 2t = 0 \end{cases}$$

Este sistema, resuelto por Gauss, nos da las siguientes expresiones paramétricas:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} -\frac{1}{3} \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} \frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} -\frac{2}{3} \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

De estas paramétricas podemos [extraer un sistema generador](#) y, [a partir de él, la base de \$U\$](#) :

$$\mathcal{B}_U = \left\{ \begin{pmatrix} -\frac{1}{3} \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{2}{3} \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Y, así, $\dim(U) = 3$.

Repetimos todo el proceso para el subespacio W . Tomamos su sistema homogéneo asociado:

$$\begin{cases} 3x + 4y & = 0 \\ 9x & + 4z + 4t = 0 \end{cases}$$

Lo resolvemos por Gauss para obtener las ecuaciones paramétricas del conjunto solución:

$$\left(\begin{array}{cccc|c} 3 & 4 & 0 & 0 & 0 \\ 9 & 0 & 4 & 4 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} 3 & 4 & 0 & 0 & 0 \\ 0 & -12 & 4 & 4 & 0 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} -\frac{4}{3} \\ 1 \\ 1 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 0 \\ 2 \\ -2 \end{pmatrix}$$

Por tanto, una base de W será:

$$\mathcal{B}_W = \left\{ \begin{pmatrix} -\frac{4}{3} \\ 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ -2 \end{pmatrix} \right\}$$

Y, así, $\dim(W) = 2$.

- b. El subespacio intersección $U \cap W$ es **el conjunto de vectores que verifican tanto las ecuaciones cartesianas de U como las de W** . Es decir, son el conjunto solución del sistema homogéneo:

$$\begin{cases} 3x + y - z + 2t = 0 \\ 3x + 4y = 0 \\ 9x + 4z + 4t = 0 \end{cases}$$

Este sistema, resuelto por Gauss, nos da las siguientes expresiones paramétricas:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} -\frac{2}{3} \\ \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

Por tanto, una base de $U \cap W$ será:

$$\mathcal{B}_{U \cap W} = \left\{ \begin{pmatrix} -\frac{2}{3} \\ \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} \right\}$$

Y, así, $\dim(U \cap W) = 1$.

Para la suma $U + W$, sabemos que $\mathcal{B}_U \cup \mathcal{B}_W$ es un sistema generador:

$$\mathcal{B}_U \cup \mathcal{B}_W = \left\{ \begin{pmatrix} -\frac{1}{3} \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{2}{3} \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{4}{3} \\ 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ -2 \end{pmatrix} \right\}$$

Basta con usar eliminación gaussiana en ese conjunto de vectores para eliminar aquellos que sean linealmente dependientes.

$$\begin{pmatrix} -\frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{3} & 0 & 1 & 0 \\ -\frac{2}{3} & 0 & 0 & 1 \\ -\frac{4}{3} & 1 & 1 & 2 \\ 0 & 0 & 2 & -2 \end{pmatrix} \sim \begin{pmatrix} -\frac{1}{3} & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}$$

Aquellas filas que hayan acabado siendo enteras de ceros, se corresponden con vectores que originalmente eran dependientes linealmente de los demás.

Por tanto, eliminándolos, tenemos una base de $U + W$:

$$\mathcal{B}_{U+W} = \left\{ \begin{pmatrix} -\frac{1}{3} \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{2}{3} \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ -2 \end{pmatrix} \right\}$$

Y, así, $\dim(U + W) = 4$.

c. Podemos comprobar el teorema de la dimensión:

$$\begin{array}{ccccccc} \dim(U + W) & = & \dim(U) & + & \dim(W) & - & \dim(U \cap W) \\ 4 & = & 3 & + & 2 & - & 1 \end{array}$$

d. Calculemos el espacio suplementario de U .

Tomamos una base de U , como la calculada antes, y le aplicamos Gauss para obtener un sistema de vectores en forma escalonada.

$$\begin{pmatrix} -\frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{3} & 0 & 1 & 0 \\ -\frac{2}{3} & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} -\frac{1}{3} & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 2 & 1 \end{pmatrix}$$

Siguiendo las instrucciones de la [sección acerca del espacio suplementario](#), vamos a fijarnos en qué columnas están los pivotes: $\{1, 2, 3\}$.

De la base canónica \mathcal{C} eliminamos aquellos vectores que tienen un 1 en las mismas posiciones que los pivotes. Lo que nos quede, ésa es la base del espacio suplementario de U :

$$\mathcal{B}_{\text{supl}} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

e. Unos escalares a, b, c, d serán [las coordenadas de \$v\$ en la base \$\mathcal{B}_1\$](#) si se verifica:

$$\begin{pmatrix} -5 \\ -9 \\ 1 \\ -5 \end{pmatrix} = a \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} + b \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix} + c \begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \end{pmatrix} + d \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

lo que equivale al siguiente sistema de ecuaciones lineales, que se puede resolver usando métodos Gaussianos, y nos da:

$$\begin{cases} a - b - d = -5 \\ -a - b - d = -9 \\ -a + b - c + d = 1 \\ -a - b - c + d = -5 \end{cases} \Rightarrow v = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}_{\mathcal{B}_1} = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 4 \end{pmatrix}_{\mathcal{B}_1}$$

f. El procedimiento para construir la [matriz de cambio de base](#) de una base \mathcal{B} a otra \mathcal{B}' es el siguiente:

- Calcular las coordenadas de cada vector en \mathcal{B} en la base \mathcal{B}' . Eso supone la resolución de n sistemas de ecuaciones lineales, y el procedimiento es igual que el visto en la sección anterior.
- Construir la matriz poniendo, en forma de columna, las coordenadas calculadas: la primera columna se corresponde con las calculadas para el primer vector de \mathcal{B} , la segunda columna, para las del segundo, etc.

Otro aspecto teórico importante es que la matriz de \mathcal{B} a \mathcal{B}' es la inversa de la del cambio de base de \mathcal{B}' a \mathcal{B} . Si una de ellas es fácil de calcular, la otra se puede construir usando, por ejemplo, el método de Gauss-Jordan para el cálculo de inversas.

Hay un caso fácil: el cambio de base desde \mathcal{B}_1 a la base canónica \mathcal{C} es la propia

matriz cuyas columnas son los vectores de \mathcal{B}_1 :

$$P_{\mathcal{B}_1 \rightarrow \mathcal{C}} = \begin{pmatrix} 1 & -1 & 0 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 \end{pmatrix}$$

La matriz del cambio de base opuesto la vamos a calcular [buscando la matriz inversa](#) de la de \mathcal{B}_1 a \mathcal{C} .

Para ello, se puede usar la eliminación de Gauss-Jordan:

$$\left(\begin{array}{cccc|cccc} 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 & 0 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & -1 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array} \right)$$

Luego

$$P_{\mathcal{C} \rightarrow \mathcal{B}_1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & -1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

6.2 Aplicaciones lineales

Ejercicio

Consideremos la aplicación $f : V = \mathbb{R}^3 \rightarrow W = \mathbb{R}^2$ dada por:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y - z \\ 2x + 2z \end{pmatrix}$$

- a. Encontrar la matriz asociada a f en las bases canónicas y en las bases:

$$\mathcal{B}_1 = \{v_i\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right\}$$

$$\mathcal{B}_2 = \{w_i\} = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$$

- b. Dar una base y la dimensión del núcleo y de la imagen de f .
- c. Identificar si f es inyectiva o sobreyectiva.
- d. Comprobar el teorema de la dimensión de rango y nulidad.
- e. Calcular una base de $f(U)$, donde U es el subespacio de V dado por:

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : 2x + 2y + z = 0 \right\}$$

Solución

- a. Para hallar la **matriz asociada** a la aplicación f en las bases canónicas, basta con mirar los coeficientes de las variables x, y, \dots en la expresión de f y ponerlos por columnas. De esta forma:

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

Para calcular la matriz A' asociada a f **con respecto a las nuevas bases**, tenemos dos opciones:

- 1. Calcular la imagen de los vectores de \mathcal{B}_1 y encontrar sus coordenadas con respecto a la base \mathcal{B}_2 , y ponerlas por columnas.
- 2. Calcular las matrices P y Q , de **cambio de base** desde la canónica en V y en W a \mathcal{B}_1 y \mathcal{B}_2 , respectivamente, y usarlas para calcular $A' = Q^{-1} A P$.

Ambos caminos son completamente equivalentes y, de hecho, uno se deduce del otro. Por tanto, en cada situación debemos ver cuál de los dos nos puede resultar más sencillo.

En este caso, optamos por el **camino 1**.

Llamemos $Y = (f(v_1)|f(v_2)|\dots|f(v_n))$ a la matriz que tiene por columnas las imágenes de los vectores de \mathcal{B}_1 mediante f :

$$Y = \begin{pmatrix} 2 & -1 & 1 \\ 2 & -2 & -2 \end{pmatrix}$$

y $B_2 = (w_1|w_2|\dots|w_m)$ a la matriz que tiene por columnas los vectores de \mathcal{B}_2 .

Entonces $A' = B_2^{-1} Y$, que podemos calcular mediante Gauss-Jordan, partiendo de $(B_2|Y)$ y llegando a $(I_m|A')$:

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

Luego

$$A' = \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{4} & -\frac{5}{4} \end{pmatrix}$$

b. Comenzamos por estudiar el **núcleo** de f .

Sabemos que el núcleo $\text{Ker } f$ es el conjunto de vectores $v \in V$ tal que $f(v) = 0$, y que eso coincide con el conjunto de soluciones del sistema de ecuaciones homogéneo cuya matriz de coeficientes es A .

Todo esto se resume en que las ecuaciones cartesianas de $\text{Ker } f$ son aquellas cuya matriz de coeficientes es A :

$$\text{Ker } f = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V : \begin{array}{rcl} 2x + y - z & = & 0 \\ 2x + 2z & = & 0 \end{array} \right\}$$

Resolvemos este sistema por Gauss para **encontrar las ecuaciones paramétricas** y, a partir de ahí, una **base** del núcleo:

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

Luego nos quedan las siguientes ecuaciones paramétricas para $\text{Ker } f$:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$$

Entonces la base del núcleo de f es:

$$\mathcal{B}_{\text{Ker } f} = \left\{ \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix} \right\}$$

y $\dim(\text{Ker } f) = 1$.

Con respecto a la imagen de f , hemos de recordar que **$\text{Im } f$ es el subespacio de W generado por los vectores que forman las columnas de la matriz asociada A .**

Por tanto, un **sistema generador** de $\text{Im } f$ será:

$$\mathcal{G} = \left\{ \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\}$$

Desde este sistema generador, **eliminamos los vectores linealmente dependientes** para obtener una base de $\text{Im } f$. Usamos Gauss con los vectores de \mathcal{G} por *filas*, y eliminamos aquellos que, al final, se hayan convertido en filas de ceros:

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

Luego

$$\mathcal{B}_{\text{Im } f} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

y $\dim(\text{Im } f) = 2$.

- c. Para determinar si f es **inyectiva** o **sobreyectiva**, nos fijaremos en las dimensiones de su núcleo y de su imagen:
 - f será inyectiva si y solo si se cumple $\dim(\text{Ker } f) = 0$. Sabemos que $\dim(\text{Ker } f) = 1$, luego f no es inyectiva.
 - f será sobreyectiva si y solo si se cumple que $\dim(\text{Im } f) = \dim(W) = \dim(\mathbb{R}^2)$, y esto se cumple, como hemos visto en el apartado anterior, luego f es sobreyectiva.
- d. El **teorema de la dimensión para núcleo e imagen** dice que $\dim(V) = \dim(\text{Ker } f) + \dim(\text{Im } f)$. Lo podemos comprobar fácilmente:

$$\begin{array}{rccccccc} \dim(\mathbb{R}^3) & = & \dim(\text{Ker } f) & + & \dim(\text{Im } f) \\ 3 & = & 1 & + & 2 \end{array}$$

- e. Dado el subespacio U , debemos encontrar un sistema de vectores \mathcal{G} que **lo genere**, calcular su imagen $f(\mathcal{G})$, y eso será un sistema generador de $f(U)$. A partir de ahí, podemos **extraer una base** de $f(U)$.

Para calcular un sistema generador de U , pasamos sus **ecuaciones cartesianas** a paramétricas, usando métodos Gaussianos:

$$\left\{ \begin{array}{l} 2x + 2y + z = 0 \\ \end{array} \right. \Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}$$

De aquí podemos determinar una base de U , si eliminamos sus vectores linealmente dependientes:

$$\mathcal{B}_U = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right\}$$

Ahora procedemos a calcular $f(\mathcal{B}_U)$:

$$f(\mathcal{B}_U) = \left\{ \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$$

que es un sistema generador de $f(U)$. Eliminamos los vectores linealmente dependientes (en este caso es sencillo) para obtener una base de $f(U)$:

$$\mathcal{B}_{f(U)} = \left\{ \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$$

y así $\dim(f(U)) = 2$.

6.3 Diagonalización

Ejercicio

Consideremos el siguiente endomorfismo $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, dado por

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5x - 3z \\ x + 8y + z \\ -x + 7z \end{pmatrix}$$

- Hallar todos los autovalores de f .
- Hallar los subespacios de autovectores de f .
- Comprobar si f es diagonalizable y dar una base \mathcal{B} de \mathbb{R}^3 tal que la matriz asociada a f en dicha base sea diagonal.
- Expresar y calcular, como suma de potencias de grado menor que 3, tanto A^{-1} como A^4 , siendo A la matriz asociada a f en la base canónica.

Solución

Como paso previo a todos los cálculos, debemos determinar [la matriz asociada](#) a f en las bases canónicas:

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

- Para [calcular los autovalores](#) de f , o equivalentemente, los de A , debemos construir el *polinomio característico*:

$$\begin{aligned}p(\lambda) &= \det(A - \lambda I) = \\&= \det \left(\begin{pmatrix} 5 & 0 & -3 \\ 1 & 8 & 1 \\ -1 & 0 & 7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) = \\&= \det \begin{pmatrix} 5-\lambda & 0 & -3 \\ 1 & 8-\lambda & 1 \\ -1 & 0 & 7-\lambda \end{pmatrix} = \\&= \lambda^3 - 20\lambda^2 + 128\lambda - 256\end{aligned}$$

Factorizamos el polinomio para hallar sus raíces (los autovalores) más fácilmente:

$$p(\lambda) = (\lambda - 4) \cdot (\lambda - 8)^2$$

Por tanto, los autovalores de f son $\{4, 8\}$.

b. Para calcular el [subespacio asociado a un autovector](#) λ , debemos hallar la solución del sistema homogéneo $(A - \lambda I)x = 0$:

- Para el autovalor $\lambda = 4$:

$$\begin{aligned}(A - 4I) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} 1 & 0 & -3 \\ 1 & 4 & 1 \\ -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \\&\Leftrightarrow \begin{cases} x & -3z & = & 0 \\ x & +4y & +z & = & 0 \\ -x & & +3z & = & 0 \end{cases}\end{aligned}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de su conjunto solución, que es U_4 :

$$\left(\begin{array}{ccc|c} 1 & 0 & -3 & 0 \\ 1 & 4 & 1 & 0 \\ -1 & 0 & 3 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -3 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$

De aquí que una base de U_4 sea:

$$\mathcal{B}_{U_4} = \left\{ \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \right\}$$

- Para el autovalor $\lambda = 8$:

$$\begin{aligned} (A - 8I) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} -3 & 0 & -3 \\ 1 & 0 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \\ &\Leftrightarrow \begin{cases} -3x & -3z = 0 \\ x & +z = 0 \\ -x & -z = 0 \end{cases} \end{aligned}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de su conjunto solución, que es U_8 :

$$\left(\begin{array}{ccc|c} -3 & 0 & -3 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} -3 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

De aquí que una base de U_8 sea:

$$\mathcal{B}_{U_8} = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

c. Estudiamos si f es [diagonalizable](#). Para ello, tenemos en cuenta la multiplicidad algebraica y la geométrica de cada autovalor, como presentamos en la siguiente tabla:

Autovalor	Mult. Algebraica	Mult. Geométrica
4	1	1
8	2	2

Con esto, queda demostrado que el endomorfismo f **sí** es diagonalizable.

Además, la base \mathcal{B} de autovectores es:

$$\mathcal{B} = \mathcal{B}_4 \cup \mathcal{B}_8 = \left\{ \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

La matriz de cambio de base de \mathcal{B} a \mathcal{C} es la que tiene los elementos de \mathcal{B} por columnas:

$$P = \begin{pmatrix} 3 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Y la matriz diagonal, asociada a f en la nueva base, es la que tiene los autovalores en la diagonal, que se calcula como:

$$\begin{aligned} D &= P^{-1} A P = \\ &= \begin{pmatrix} 3 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 5 & 0 & -3 \\ 1 & 8 & 1 \\ -1 & 0 & 7 \end{pmatrix} \begin{pmatrix} 3 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} \end{aligned}$$

- d. Ahora aplicaremos el [Teorema de Cayley-Hamilton](#) para poder expresar la inversa de A y A^4 como suma de potencias de grado menor que 3.

El teorema me dice que la matriz A anula su polinomio característico, es decir:

$$p(A) = A^3 - 20A^2 + 128A - 256I = 0$$

Despejando la identidad, nos queda:

$$\begin{aligned} I &= -\frac{1}{256} (A^3 - 20A^2 + 128A) = \\ &= \left(\frac{1}{256} A^2 - \frac{5}{64} A + \frac{1}{2} I \right) A \end{aligned}$$

Luego tenemos

$$\begin{aligned} A^{-1} &= \frac{1}{256} A^2 - \frac{5}{64} A + \frac{1}{2} I = \\ &= +\frac{1}{256} \begin{pmatrix} 28 & 0 & -36 \\ 12 & 64 & 12 \\ -12 & 0 & 52 \end{pmatrix} - \frac{5}{64} \begin{pmatrix} 5 & 0 & -3 \\ 1 & 8 & 1 \\ -1 & 0 & 7 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{7}{32} & 0 & \frac{3}{32} \\ -\frac{1}{32} & \frac{1}{8} & -\frac{1}{32} \\ \frac{1}{32} & 0 & \frac{5}{32} \end{pmatrix} \end{aligned}$$

Calculemos ahora A^4 . Tenemos que partir de despejar el término A^3 de la igualdad que nos proporciona el Teorema de Cayley-Hamilton:

$$A^3 = 20A^2 - 128A + 256I$$

Para A^4 , tenemos:

$$A^4 = A \cdot A^3 = A(20A^2 - 128A + 256I) = 20A^3 - 128A^2 + 256A$$

Si queremos dejarlo como suma de potencias de A de grado menor que 3, debemos sustituir aquí A^3 por la expresión que hemos despejado hace un momento.

$$\begin{aligned}
 A^4 &= 20(20A^2 - 128A + 256I) - 128A^2 + 256A = \\
 &= 272A^2 - 2304A + 5120I = \\
 &= 272 \begin{pmatrix} 28 & 0 & -36 \\ 12 & 64 & 12 \\ -12 & 0 & 52 \end{pmatrix} - 2304 \begin{pmatrix} 5 & 0 & -3 \\ 1 & 8 & 1 \\ -1 & 0 & 7 \end{pmatrix} + 5120 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \\
 &= \begin{pmatrix} 1216 & 0 & -2880 \\ 960 & 4096 & 960 \\ -960 & 0 & 3136 \end{pmatrix}
 \end{aligned}$$

6.4 Espacios Euclídeos

Ejercicio

- a. Consideremos los vectores

$$u = \begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \end{pmatrix} \quad v = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix}$$

Hallar el ángulo que forman u y v , la norma de cada uno de ellos, $d(u, v)$, y normalizarlos.

- b. Encontrar el complemento ortogonal del subespacio generado por el siguiente sistema de vectores:

$$\mathcal{B} = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ -2 \end{pmatrix} \right\}$$

- c. Diagonalizar ortogonalmente la matriz

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

hallando una base ortonormal de autovectores y expresando de forma explícita la matriz diagonal en función de la matriz A y de la matriz de cambio de base.

Solución

a. Vamos a calcular tanto el **producto escalar** de u y v como sus **normas**:

$$\langle u, v \rangle = \left\langle \begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix} \right\rangle = (-1) \cdot 1 + 2 \cdot 1 + 0 \cdot 1 + (-2) \cdot 2 = -3$$

$$\|u\| = \sqrt{\langle u, u \rangle} = \sqrt{(-1)^2 + 2^2 + 0^2 + (-2)^2} = 3$$

$$\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{1^2 + 1^2 + 1^2 + 2^2} = \sqrt{7}$$

Entonces

$$\text{ang}(u, v) = \arccos \frac{\langle u, v \rangle}{\|u\|\|v\|} = \arccos \frac{-3}{3\sqrt{7}} = \arccos \frac{-\sqrt{7}}{7}$$

Pasamos a normalizar cada vector:

$$\frac{u}{\|u\|} = \frac{1}{3} \begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \\ -\frac{2}{3} \end{pmatrix}, \quad \frac{v}{\|v\|} = \frac{\sqrt{7}}{7} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{7}}{7} \\ \frac{\sqrt{7}}{7} \\ \frac{\sqrt{7}}{7} \\ \frac{2\sqrt{7}}{7} \end{pmatrix}$$

Por último, calculamos la **distancia**:

$$d(u, v) = \|u - v\| = \left\| \begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -2 \\ 1 \\ -1 \\ -4 \end{pmatrix} \right\| = \sqrt{22}$$

b. Para calcular el **complemento ortogonal** del subespacio U , partimos de la base \mathcal{B} , y buscamos aquellos vectores v de V cuyo producto escalar por los vectores de la base sea 0.

Concretamente, usando el producto escalar usual en \mathbb{R}^4 , esto equivale a buscar las soluciones del sistema de ecuaciones lineales homogéneo:

$$\begin{array}{rrcr} -x & +y & & -t & = & 0 \\ & & 2z & -2t & = & 0 \end{array}$$

donde los coeficientes de cada fila son las coordenadas del vector correspondiente de \mathcal{B} . De esta forma, este sistema representa las ecuaciones cartesianas del espacio U^\perp .

Podemos resolver el sistema usando Gauss-Jordan, y despejando para encontrar las ecuaciones paramétricas del conjunto solución:

$$\begin{pmatrix} -1 & 1 & 0 & -1 \\ 0 & 0 & 2 & -2 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 0 & -1 \\ 0 & 0 & 2 & -2 \end{pmatrix} \Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

de donde podemos deducir una [base](#) para U^\perp :

$$\mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

- c. Comenzamos por hallar el [polinomio característico](#) de A y calculando sus [autovalores](#).

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) = \\ &= \det \left(\begin{pmatrix} -1 & 1 & -2 & 1 \\ 1 & -1 & 2 & -1 \\ -2 & 2 & -4 & 2 \\ 1 & -1 & 2 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) = \\ &= \det \begin{pmatrix} -1-\lambda & 1 & -2 & 1 \\ 1 & -1-\lambda & 2 & -1 \\ -2 & 2 & -4-\lambda & 2 \\ 1 & -1 & 2 & -1-\lambda \end{pmatrix} \\ &= \lambda^4 + 7\lambda^3 \end{aligned}$$

Igualándolo a 0 y resolviendo la ecuación que queda, llegamos a que los autovalores son: $\lambda = 0, -7$, ya que:

$$p(\lambda) = \lambda^3 \cdot (\lambda + 7)$$

Recorremos ahora cada autovalor, encontrando una base del subespacio propio asociado, y ortonormalizándola usando el [método de Gram-Schmidt](#).

- Para el autovalor $\lambda = 0$:

$$(A - 0I) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow \begin{pmatrix} -1 & 1 & -2 & 1 \\ 1 & -1 & 2 & -1 \\ -2 & 2 & -4 & 2 \\ 1 & -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} -x + y - 2z + t = 0 \\ x - y + 2z - t = 0 \\ -2x + 2y - 4z + 2t = 0 \\ x - y + 2z - t = 0 \end{cases}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de U_0 :

$$\left(\begin{array}{cccc|c} -1 & 1 & -2 & 1 & 0 \\ 1 & -1 & 2 & -1 & 0 \\ -2 & 2 & -4 & 2 & 0 \\ 1 & -1 & 2 & -1 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} -1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$$\Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Luego una base de U_0 es:

$$\mathcal{B}_{U_0} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

El siguiente paso es *ortonormalizar* esta base, mediante el método de Gram-Schmidt. Llamemos v_i a los vectores de esta base hallada. Llamamos

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Procedemos con el resto de vectores:

Construimos:

$$U_1 = \mathcal{L}(\{u_1\}) = \mathcal{L}\left(\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\}\right)$$

Entonces

$$\text{proy}_{U_1}(v_2) = \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 = \frac{-2}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

Y llamamos

$$u_2 = v_2 - \text{proy}_{U_1}(v_2) = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Construimos:

$$U_2 = \mathcal{L}(\{u_1, u_2\}) = \mathcal{L}\left(\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}\right\}\right)$$

Entonces

$$\text{proy}_{U_2}(v_3) = \frac{\langle v_3, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle v_3, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{-1}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{5}{6} \\ \frac{1}{6} \\ -\frac{1}{3} \\ 0 \end{pmatrix}$$

Y llamamos

$$u_3 = v_3 - \text{proy}_{U_2}(v_3) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} \frac{5}{6} \\ \frac{1}{6} \\ -\frac{1}{3} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ -\frac{1}{6} \\ \frac{1}{3} \\ 1 \end{pmatrix}$$

Y así hemos llegado a un sistema ortogonal formado por los vectores u_i . Debemos normalizar ahora los vectores de este sistema ortogonal que tenemos, multiplicando

cada uno por el inverso de su norma, llegando a la siguiente base ortonormal de U_0 .

$$\begin{aligned}\mathcal{B}'_{U_0} &= \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\frac{\sqrt{42}}{6}} \begin{pmatrix} \frac{1}{6} \\ \frac{-1}{6} \\ \frac{1}{3} \\ 1 \end{pmatrix} \right\} = \\ &= \left\{ \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{42}}{42} \\ \frac{-\sqrt{42}}{42} \\ \frac{\sqrt{42}}{21} \\ \frac{\sqrt{42}}{7} \end{pmatrix} \right\}\end{aligned}$$

- Para el autovalor $\lambda = -7$:

$$\begin{aligned}(A + 7I) \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 &\Leftrightarrow \begin{pmatrix} 6 & 1 & -2 & 1 \\ 1 & 6 & 2 & -1 \\ -2 & 2 & 3 & 2 \\ 1 & -1 & 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \Leftrightarrow \\ &\Leftrightarrow \begin{cases} 6x + y - 2z + t = 0 \\ x + 6y + 2z - t = 0 \\ -2x + 2y + 3z + 2t = 0 \\ x - y + 2z + 6t = 0 \end{cases}\end{aligned}$$

Resolvemos este sistema por Gauss-Jordan, encontrando la forma paramétrica de U_{-7} :

$$\begin{aligned}\left(\begin{array}{cccc|c} 6 & 1 & -2 & 1 & 0 \\ 1 & 6 & 2 & -1 & 0 \\ -2 & 2 & 3 & 2 & 0 \\ 1 & -1 & 2 & 6 & 0 \end{array} \right) &\sim \left(\begin{array}{cccc|c} 6 & 0 & 0 & 6 & 0 \\ 0 & \frac{35}{6} & 0 & -\frac{35}{6} & 0 \\ 0 & 0 & \frac{7}{5} & \frac{14}{5} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \\ &\Rightarrow \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 1 \\ -2 \\ 1 \end{pmatrix}\end{aligned}$$

Luego una base de U_{-7} es:

$$\mathcal{B}_{U_{-7}} = \left\{ \begin{pmatrix} -1 \\ 1 \\ -2 \\ 1 \end{pmatrix} \right\}$$

Como solo tenemos un vector en la base, ya forma un sistema ortogonal. Debemos normalizar ahora los vectores de este sistema ortogonal que tenemos, multiplicando cada uno por el inverso de su norma, llegando a la siguiente base ortonormal de U_{-7} .

$$\mathcal{B}'_{U_{-7}} = \left\{ \frac{1}{\sqrt{7}} \begin{pmatrix} -1 \\ 1 \\ -2 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \frac{-\sqrt{7}}{7} \\ \frac{\sqrt{7}}{7} \\ \frac{-2\sqrt{7}}{7} \\ \frac{\sqrt{7}}{7} \end{pmatrix} \right\}$$

Definimos

$$\mathcal{B} = \mathcal{B}'_{U_0} \cup \mathcal{B}'_{U_{-7}} = \left\{ \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{42}}{42} \\ \frac{-\sqrt{42}}{42} \\ \frac{\sqrt{42}}{21} \\ \frac{\sqrt{42}}{7} \end{pmatrix}, \begin{pmatrix} \frac{-\sqrt{7}}{7} \\ \frac{\sqrt{7}}{7} \\ \frac{-2\sqrt{7}}{7} \\ \frac{\sqrt{7}}{7} \end{pmatrix} \right\}$$

Es una base ortonormal de \mathbb{R}^4 , formada únicamente por autovectores de la matriz A .

Además, la matriz P del [cambio de base](#) de esta base \mathcal{B} a la canónica, formada por los vectores de \mathcal{B} puestos por columnas, es *ortogonal* y nos proporciona la relación entre A y la matriz diagonal D de los autovalores:

$$\begin{aligned} D &= P^t A P = \\ &= \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 \\ \frac{-\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & 0 \\ \frac{\sqrt{42}}{42} & \frac{-\sqrt{42}}{42} & \frac{\sqrt{42}}{21} & \frac{\sqrt{42}}{7} \\ \frac{-\sqrt{7}}{7} & \frac{\sqrt{7}}{7} & \frac{-2\sqrt{7}}{7} & \frac{\sqrt{7}}{7} \end{pmatrix} \begin{pmatrix} -1 & 1 & -2 & 1 \\ 1 & -1 & 2 & -1 \\ -2 & 2 & -4 & 2 \\ 1 & -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} & \frac{\sqrt{42}}{42} & \frac{-\sqrt{7}}{7} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & \frac{-\sqrt{42}}{42} & \frac{\sqrt{7}}{7} \\ 0 & \frac{\sqrt{3}}{3} & \frac{\sqrt{42}}{21} & \frac{-2\sqrt{7}}{7} \\ 0 & 0 & \frac{\sqrt{42}}{7} & \frac{\sqrt{7}}{7} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -7 \end{pmatrix} \end{aligned}$$