Álgebra
(Lineal Básica)

Ana Mª Díaz Hernández
Vicente Bargueño Fariñas
Carlos Romera Carrión
Luis Manuel Ruiz Virumbrales
Luis Tejero Escribano

ÁLGEBRA (LINEAL BÁSICA)

Pinos Alta, 49 - 28029 Madrid
Teléfs.: 902 400 415 - 91 314 07 82
www.sanzytorres.com
editorial@sanzytorres.com
librería@sanzytorres.com

ISBN: 84-88667-91-4
84-88667-93-0 (Obra completa)
Depósito legal: M-37.703-2002

Compuesto por FER Fotocomposición, S.A.
C/ Bocángel, 45. 28028 Madrid
Impreso por Edigrafos.
C/ Volta, 2. Polígono Industrial San Marcos
28906 Getafe (Madrid)

SÍMBOLOS	LECTURA
$\forall$	Para todo
$\exists$	Existe
$\nexists$	No existe
$\in$	Pertenece
$\notin$	No pertenece
$\{x / x \in A\}$	El conjunto de los elementos x, tales que x pertenece al conjunto A
$\subset$	Contenido
$\not\subset$	No contenido
$\emptyset$	Conjunto vacío
$A \cup B$	A unión B: Conjunto de todos los elementos que pertenecen al conjunto A o al conjunto B
$A \cap B$	A intersección B: Conjunto de todos los elementos que pertenecen a la vez al conjunto A y al conjunto B
$\bar{x}$	El elemento $\bar{x}$ es un vector
$\lVert\bar{x}\rVert$	La norma del vector $\bar{x}$
$\mathbb{N}$	El conjunto de los números naturales
$\mathbb{Z}$	El conjunto de los números enteros
$\mathbb{Q}$	El conjunto de los números racionales
$\mathbb{R}$	El conjunto de los números reales
I	La matriz identidad
$\oplus$	Suma directa

ÍNDICE GENERAL

Capítulo 1
ESPACIOS VECTORIALES

Capítulo 2
APLICACIONES LINEALES

Capítulo 3
DETERMINANTE DE UNA MATRIZ CUADRADA

Capítulo 4
SISTEMAS DE ECUACIONES LINEALES

Capítulo 5
PRODUCTO ESCALAR DE VECTORES Y ESPACIO EUCLÍDEO

Capítulo 6
MATRICES SEMEJANTES

Capítulo 7
FORMAS BILINEALES

Capítulo 8
EL PROBLEMA DE LA PROGRAMACIÓN LINEAL

PRÓLOGO

En la declaración del Año Mundial de las Matemáticas, que fue el año 2000, se afirma que *«...las Matemáticas constituyen un pilar fundamental de la cultura, no sólo por ser el lenguaje de la Ciencia, sino por lo que suponen como bagaje necesario para entender el mundo en que vivimos...»*.

El mundo en que vivimos es básicamente tecnológico, y como tal, necesitado de las Matemáticas, y de entre las disciplinas matemáticas, el Álgebra es herramienta de uso ineludible para el estudio de otras muchas materias que forman el diseño curricular del estudiante de Informática.

Este equipo de autores aborda la confección de un texto moderno para explicar una ciencia antigua, que tiene la doble faceta de ciencia autónoma y herramienta.

Para conseguir nuestro objetivo primordial: *«que sea un texto útil al estudiante de Informática»*, lo hemos dotado de las siguientes características:

El texto completo está desarrollado en dos soportes diferentes:

En papel se han confeccionado dos tomos, el primero son las Unidades Didácticas propiamente dichas, contienen teoría, problemas, gráficos, aplicaciones, etc.

- El criterio para la distribución el contenido en capítulos es la homogeneidad del objeto de estudio, independientemente de su longitud. Dicho objeto de estudio, se indica brevemente en la introducción que le precede.
- En cada capítulo hemos incluido un diagrama de flujo que permite al lector ver rápidamente que objetivo se persigue en él, así como el discurso lógico que fija las directrices seguidas para conseguirlo.

- Hemos incorporado, en la medida de lo posible, aplicaciones que alienten al estudio del Álgebra por su carácter instrumental.

El segundo tomo contiene:

- Notas útiles para trabajar con DERIVE.
- Introducción a las herramientas de DERIVE necesarias en cada capítulo a través de ejemplos.
- Ejercicios y problemas propuestos en exámenes de Álgebra (Informática Gestión y Sistemas) resueltos.
- Los mismos ejercicios y problemas resueltos paso a paso con DERIVE.
- Glosario de términos algebraicos utilizados en el texto.

Soporte C. D: Es un soporte familiar para el estudiante de Informática, y con él, tratamos tanto de acercar los contenidos plasmados en el papel a su ámbito natural, como de incluir posibilidades de actuación interactiva que en papel no existen.

Madrid, agosto 2002.
Ana Díaz Hernández
Coordinadora,

CAPÍTULO 1

ESPACIOS VECTORIALES

INTRODUCCIÓN

El concepto de vector como segmento dirigido en el plano o en el espacio para representar fuerzas, velocidades o aceleraciones, se remonta a Aristóteles, quien también conocía algunas de sus propiedades algebraicas.

El desarrollo de la Geometría Analítica a partir del siglo XVII lleva a los físicos a la interpretación de los vectores como parejas o ternas de números (sus coordenadas) y a investigar cómo se traducen las operaciones vectoriales en estos términos (composición de vectores, dilataciones, rotaciones, producto escalar, producto vectorial, etc.). El desarrollo y formalización completa de estas ideas no se lleva a cabo, sin embargo, hasta bien entrado el siglo XIX.

El estudio de los sistemas de ecuaciones propició, de modo natural, el paso del plano y el espacio ordinario al espacio de n dimensiones (como conjunto de entes determinados por n coordenadas, o bien directamente como conjunto de n-uplas), que puede adivinarse ya en Gauss y aparece claramente en los matemáticos de la siguiente generación. En todo caso, Cayley y Grassmann manejan libremente en 1846 las propiedades algebraicas del espacio n-dimensional: subespacios, generadores, dimensión, suma, intersección, así como las fórmulas del cambio de coordenadas.

En el último tercio del siglo XIX aparece la noción de estructura y toma carta de naturaleza el método axiomático.

Los vectores son utilizados en la forma actual desde 1888, cuando Peano (1858-1932) definió los espacios vectoriales axiomáticamente. Antes hay gran cantidad de precursores, citaremos entre ellos a Descartes (1596-1650) que habla de coordenadas, Gauss (1777-1855) que utiliza la suma de vectores de $\mathbb{R}^2$, Hamilton (1805-1865) que describe cómo operar con pares de números reales, y Grassman (1809-1877), quien sin utilizar la nomenclatura actual obtiene resultados importantes.

GIUSEPE PEANO (1858-1932)

Matemático y lógico italiano, influyó activamente en el tratamiento axiomático de Hilbert de la geometría plana y los trabajos de Whitehead y Russel sobre lógica matemática.

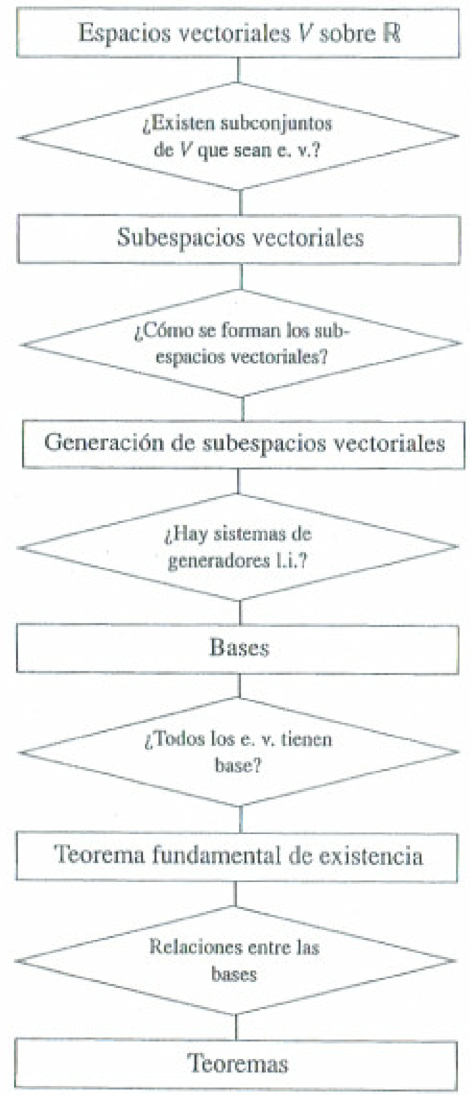
Es uno de los creadores del método axiomático. Dio en el año 1888 la definición axiomática de los espacios vectoriales (de dimensión finita o no) sobre los números reales. También dio la definición de aplicación lineal entre espacios vectoriales.

Sus postulados sobre los enteros positivos han hecho que generaciones enteras de estudiantes se pregunten si el álgebra moderna no es alguna especie de conspiración encaminada a hacer que lo obvio resulte oscuro.

En 1890, sorprendió al mundo de las matemáticas con su notable construcción de una curva continua en el plano que llena completamente el cuadrado $0 \leq x \leq 1, 0 \leq y \leq 1$.

Desgraciadamente para un hombre que atribuía tanto valor a la lógica, su prueba de 1886 del teorema de existencia para soluciones de $y' = f(x, y)$ (teorema de Peano) fue inadecuada y no se encontró una prueba satisfactoria sino hasta varios años después.

CAPÍTULO 1


Espacios vectoriales V sobre $\mathbb{R}$
¿Existen subconjuntos de V que sean e. v.?
Subespacios vectoriales
¿Cómo se forman los subespacios vectoriales?
Generación de subespacios vectoriales
¿Hay sistemas de generadores l.i.?
Bases
¿Todos los e. v. tienen base?
Teorema fundamental de existencia
Relaciones entre las bases
Teoremas

1.1 El espacio vectorial real ($\mathbb{R}^n$, +, $\mathbb{R}$)

Algunos espacios vectoriales son conocidos por todos los estudiantes, así ocurre cuando hablamos de $\mathbb{R}$, $\mathbb{R}^2$, o $\mathbb{R}^3$, y en general de $\mathbb{R}^n$, por ello, comenzaremos el capítulo haciendo un breve repaso de los elementos de dichos espacios, o sea, vectores en la recta, en el plano o en el espacio.

1.1.1 Definición de vector $\overline{u}$ de $\mathbb{R}^n$

Los vectores fila $\overline{u}$ que forman los espacios $\mathbb{R}^n$ son n-uplas $(u_1, u_2, ..., u_n)$ de números reales u_i, donde cada u_i (coordenada i del vector) es un escalar.

Las n-uplas se pueden escribir en columnas, y entonces hablaremos de vectores columna del mismo espacio vectorial $\mathbb{R}^n$: $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$.

Nuestra familiaridad con los espacios vectoriales $\mathbb{R}$, $\mathbb{R}^2$ y $\mathbb{R}^3$, se debe a que modelizan fielmente el espacio en que vivimos, siendo cada uno de ellos apto para describir la posición de un punto cuando se ha fijado un sistema de referencia (sistema de coordenadas con un origen). En estas condiciones, los vectores se pueden representar geométricamente uniendo mediante una flecha el origen de coordenadas 0 con el punto u_1 en $\mathbb{R}$, (0,0) con (u_1,u_2) en $\mathbb{R}^2$, y (0,0,0) con (u_1,u_2,u_3) en $\mathbb{R}^3$, y reciben el nombre de <u>vectores fijos</u>.

En física, es muy útil el uso de los vectores de $\mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}$, porque al añadir a los de $\mathbb{R}^3$ una cuarta coordenada, permiten situar el tiempo (una vez fijado un origen para u_4).

Los espacios vectoriales $\mathbb{R}^n$ juegan un papel importante en el mundo empresarial, ya que, sus vectores permiten almacenar datos ordenados, gestionarlos, y extraer resultados, de forma eficiente. Aunque este es el principal uso que hacen las empresas de los sistemas informáticos, también es muy frecuente, hacer un modelo abstracto, introducir distintos juegos de datos reales y prever los posibles resultados antes de tomar decisiones.

Ejemplo 1.1.1

El vector de $\mathbb{R}^7$ (día, mes, año, vendedor, producto, porcentaje, destino) puede dar información de la fecha en que se realiza una operación de venta, el vendedor que la realiza, qué producto vende, qué tanto por ciento lleva, y en qué ciudad se ha producido.

Hay que observar que, las variables de este vector deben ser escalares para que el vector pertenezca a $\mathbb{R}^7$, pero un vector puede ser definido de forma que las variables pertenezcan a conjuntos distintos de $\mathbb{R}$.

Para hacer la gestión de los datos es necesario operar con ellos, y las operaciones más frecuentes son las que se pueden hacer con los vectores.

OPERACIONES CON VECTORES: SUMA

Los conceptos que se pueden modelizar mediante elementos de $\mathbb{R}$, $\mathbb{R}^2$ y $\mathbb{R}^3$, como distancias, fuerzas, velocidades, etc., se pueden sumar, y además, podemos representar gráficamente dichas sumas (recordemos la regla del paralelogramo).

Ejemplo 1.1.2

La velocidad real de un barco se puede expresar mediante un vector, dicho vector resulta de sumar el vector velocidad que lleva la máquina y otros vectores velocidad ajenos al barco, como pueda ser la velocidad de una corriente determinada.

La suma de vectores sirve para modelizar la velocidad real o corregida, y llevar el rumbo y la velocidad deseados.

$\mathbb{R}$, $\mathbb{R}^2$ y $\mathbb{R}^3$, son casos particulares de $\mathbb{R}^n$; de forma general, definimos analíticamente la suma de vectores de $\mathbb{R}^n$ a través de la suma de sus coordenadas, que son escalares.

1.1.2 Definición de suma de vectores de $\mathbb{R}^n$

Vector $\overline{u}+\overline{v}$ suma de los vectores $\overline{u}=(u_1,u_2,\ldots,u_n)$ y $\overline{v}=(v_1,v_2,\ldots,v_n)$ de $\mathbb{R}^n$ es la imagen en la aplicación:

$\mathbb{R}^n\times\mathbb{R}^n \xrightarrow{+} \mathbb{R}^n$

$(\overline{u},\overline{v}) \xrightarrow{+} \overline{u}+\overline{v}=(u_1,u_2,\ldots,u_n)+(v_1,v_2,\ldots,v_n)=(u_1+v_1,u_2+v_2,\ldots,u_n+v_n)$

Cuando sumamos dos vectores cualesquiera $\overline{u},\overline{v}\in\mathbb{R}^n$ asignamos al par $(\overline{u},\overline{v})$, otro vector $\overline{u}+\overline{v}$ del mismo conjunto $\mathbb{R}^n$ al que llamamos vector suma, que se obtiene sumando las coordenadas de $\overline{u}$ y $\overline{v}$ que ocupan el mismo lugar.

1.1.3 Propiedades de la suma de los vectores de $\mathbb{R}^n$

La suma de vectores de $\mathbb{R}^n$ tiene las propiedades:

1. Asociativa: $(\overline{u}+\overline{v})+\overline{w}=\overline{u}+(\overline{v}+\overline{w})$
2. Conmutativa: $\overline{u}+\overline{v}=\overline{v}+\overline{u}$
3. Existe un vector neutro $\overline{0}\in\mathbb{R}^n$, que verifica $\overline{u}+\overline{0}=\overline{u}$, siendo $\overline{0}=(0,0,\ldots,0)$
4. Cada vector $\overline{u}=(u_1,u_2,\ldots,u_n)\in\mathbb{R}^n$ tiene un vector simétrico, $-\overline{u}=(-u_1,-u_2,\ldots,-u_n)\in\mathbb{R}^n$, tal que, $\overline{u}+(-\overline{u})=\overline{0}$

MULTIPLICACIÓN DE VECTORES POR UN ESCALAR

En $\mathbb{R}$, $\mathbb{R}^2$ y $\mathbb{R}^3$, tiene sentido hablar de tamaños, la posibilidad de hacer más grande o más pequeño algo que se puede representar con un vector, se modeliza en álgebra multiplicando dicho vector por un escalar.

Si llamamos λ al escalar, $\lambda\overline{u}$ representa una dilatación o una contracción de $\overline{u}$ (según que $|\lambda|>1$ o $|\lambda|<1$). Además si $\lambda<0$, el sentido de $\lambda\overline{u}$ es el opuesto al de $\overline{u}$.

Se utilizan indistintamente las expresiones escalar por vector $(\lambda\overline{u})$ o vector por escalar $(\overline{u}\lambda)$, y su definición es la misma.

1.1.4 Definición de multiplicación de vectores de $\mathbb{R}^n$ por un escalar

Vector λu, producto del escalar $\lambda\in\mathbb{R}$ por el vector $\overline{u}=(u_1,u_2,\ldots,u_n)$ de $\mathbb{R}^n$ es la imagen en la aplicación:

$\mathbb{R}^n\times\mathbb{R} \xrightarrow{\times} \mathbb{R}^n$

$(\overline{u},\lambda) \xrightarrow{\times} \lambda\overline{u}=(\lambda u_1,\lambda u_2,\ldots,\lambda u_n)$

Al multiplicar cualquier vector $\overline{u}\in\mathbb{R}^n$ por un escalar $\lambda\in\mathbb{R}$, asignamos al par $(\overline{u},\lambda)$ otro vector $\lambda\overline{u}\in\mathbb{R}^n$, por tanto, no es una operación interna entre vectores de un conjunto, sino una operación externa que involucra elementos de dos conjuntos diferentes: $\mathbb{R}^n$ (vectores) y de $\mathbb{R}$ (escalares).

1.1.5 Propiedades de la multiplicación de vectores de $\mathbb{R}^n$ por un escalar

La multiplicación de vectores de $\mathbb{R}^n$ por un escalar tiene las propiedades:

1. Distributiva de escalares respecto a vectores: $\lambda(\overline{u}+\overline{v})=\lambda\overline{u}+\lambda\overline{v}$
2. Distributiva de vectores respecto a escalares: $(\lambda+\mu)\,\overline{u}=\lambda\overline{u}+\mu\overline{u}$
3. Asociativa: $(\lambda\mu)\,\overline{u}=\lambda(\mu\overline{u})$
4. Existe el escalar unidad, 1, que cumple $1\overline{u}=\overline{u}$

Hemos enunciado las propiedades que tienen las operaciones con los vectores de $\mathbb{R}^n$, pero alguna de las cosas que hemos dicho podría no ser correcta, aunque lo parezca.

Para tener la certeza de que alguna proposición que enunciamos es cierta, necesita ser demostrada.

En cada demostración se pueden utilizar axiomas (verdades evidentes), u otras verdades demostradas con anterioridad, es el método lógico-deductivo utilizado desde los griegos.

Un enunciado y su demostración forman un teorema o una proposición.

1.1.6 Teorema

En $(\mathbb{R}^n, + \mathbb{R})$ se verifica:

a) La operación de sumar vectores de $\mathbb{R}^n$ tiene las propiedades *[1.1.3]*.

b) La operación de multiplicar vectores de $\mathbb{R}^n$ por escalares tiene las propiedades *[1.1.5]*

Demostración:

a) La operación de sumar vectores de $\mathbb{R}^n$ tiene las propiedades ***[1.1.3]*** porque las tiene la suma de los escalares que forman las n-uplas.

b) Para demostrar que la operación de multiplicar vectores de $\mathbb{R}^n$ por escalares cumple las propiedades ***[1.1.5]***, hay que ir comprobando cada una de las igualdades.

Vamos a hacerlo con una de ellas, con las demás, se hace de la misma manera.

Una forma de demostrar la veracidad de una igualdad es, desarrollar los dos miembros de manera que se llegue en ambos al mismo resultado.

Vamos a utilizar esta técnica para demostrar la propiedad distributiva de vectores respecto a escalares.

Queremos demostrar que: $(\lambda + \mu)\overline{u} = \lambda\overline{u} + \mu\overline{u}$.

Desarrollamos el primer miembro aplicando la definición *[1.1.4]*:

$$(\lambda + \mu)\overline{u} = (\lambda + \mu)(u_1, u_2, \ldots, u_n) = ((\lambda + \mu)u_1, (\lambda + \mu)u_2, \ldots, (\lambda + \mu)u_n) = \\ = (\lambda u_1, \mu u_1, \lambda u_2 + \mu u_2, \ldots, \lambda u_n + \mu u_n)$$

Desarrollamos el segundo miembro aplicando la definición *[1.1.4]* a cada uno de los dos sumandos:

$$\lambda\overline{u} + \mu\overline{u} = \lambda(u_1, u_2, \ldots, u_n) + \mu(u_1, u_2, \ldots, u_n) = \\ = (\lambda u_1, \lambda u_2, \ldots, \lambda u_n) + (\mu u_1, \mu u_2, \ldots, \mu u_n)$$

Si utilizamos la definición *[1.1.2]* para sumar los vectores resultantes, comprobamos que es lo que queríamos demostrar.

Es muy fácil repetir este procedimiento para cada una de las propiedades, cuando lo hayamos hecho podremos decir que hemos demostrado el teorema *[1.1.6]*.

1.2 La estructura algebraica de espacio vectorial

Es muy frecuente que, al realizar una determinada operación en un conjunto y otra operación en otro conjunto, ambas operaciones tengan las mismas propiedades. La razón es que tienen una estructura común.

La importancia que tiene el estudio de las estructuras radica en que es mucho más eficiente analizar de una sola vez las propiedades que derivan de una estructura abstracta y aplicarlas a los conjuntos concretos que la tienen, independientemente del tipo de elementos que forman los conjuntos y de la operación que se haya definido en ellos.

Las estructuras se utilizan en Informática con mucha frecuencia, por ejemplo, al hacer un programa se asigna a cada variable la estructura a que pertenece (INTEGER, REAL, CHARACTER, LOGICAL, ...).

1.2.1 Definición de estructura algebraica

$(A,*,\circ,...)$ es una estructura algebraica $\Leftrightarrow$ A es un conjunto en el que hay definidas una o más operaciones: $*$, $\circ$,...

Ejemplo 1.2.1

$(\mathbb{R}^n, +)$ es una estructura algebraica.

Los elementos con que se opera son vectores $\in \mathbb{R}^n$, y la operación "+" es la suma de vectores.

Ejemplo 1.2.2

Si $A = \{1,2,5\}$ y la operación "$*$" asigna a (a, b) el resto de $a^b / 3$, $(A,*)$, es una estructura algebraica.

Cuando la operación se realiza en un conjunto finito, se pueden dar los resultados de la operación entre todos sus elementos, es decir, se puede tabular:

	1	2	5
1	1	1	1
2	2	1	2
5	2	1	2

Hay algunas estructuras que se repiten con frecuencia, incluso formando parte de otras; este es el caso de los grupos.

1.2.2. Definición de grupo

$(A,*)$ tiene estructura de grupo $\Leftrightarrow$ La operación $*$ en el conjunto A (no vacío) tiene las propiedades:

1. Asociativa: $(a * b) * c = a * (b * c)$, $\forall a, b, c \in A$
2. Existe un elemento neutro $e \in A$, que verifica $a * e = a$, $\forall a \in A$
3. $\forall a \in A$, $\exists a' \in A$ que verifica $a * a' = e$; a' es el elemento simétrico o inverso de a
4. Es un grupo conmutativo si además tiene la propiedad conmutativa $\forall a, b \in A$ se verifica $a * b = b * a$

Ejemplo 1.2.3

$(\mathbb{R}^n, +)$ tiene estructura algebraica de grupo.

La operación + en $\mathbb{R}^n$, cumple las propiedades: asociativa, existe un vector neutro y todo vector tiene un simétrico, como hemos visto en *[1.1.6]*.

Se suele llamar opuesto de a al elemento simétrico para la operación de sumar, y se representa como $-a$.

El estudio de las propiedades de $(\mathbb{R}^n, +, \mathbb{R})$ se ha ido incorporando a las matemáticas desde el siglo XIX, y ha sido tan significativo, que ha dado nombre a la estructura algebraica que tiene: espacio vectorial.

Hay muchos conjuntos que tienen las mismas propiedades, es decir, la misma estructura que definimos de forma general.

1.2.3. Definición de espacio vectorial

$(V, *, \cdot \mathbb{R})$ tiene estructura de espacio vectorial sobre $\mathbb{R} \Leftrightarrow (V, *)$ es grupo conmutativo y $(V, *, \cdot\mathbb{R})$ tiene las propiedades:

- Distributiva de las escalares respecto a los vectores: $\lambda \cdot (\overline{u} * \overline{v}) = \lambda \cdot \overline{u} * \lambda \cdot \overline{v}$
- Distributiva de los vectores respecto a la suma de escalares: $(\lambda + \mu) \cdot \overline{u} = \lambda \cdot \overline{u} * \mu \cdot \overline{u}$
- Asociativa respecto al producto de escalares: $(\lambda\mu) \cdot \overline{u} = \lambda \cdot (\mu \cdot \overline{u})$
- Existe un escalar unidad, $1 \in \mathbb{R}$, que verifica $1 \cdot \overline{u} = \overline{u}$

Los elementos de cualquier conjunto con estructura de espacio vectorial se llaman vectores y normalmente se representan con una barra sobre una letra latina.

Los elementos de $\mathbb{R}$ se llaman escalares y se suelen representar con una letra griega minúscula.

El lugar de $\mathbb{R}$ puede ser ocupado por un cuerpo diferente, el de los complejos, o cualquier otro cuerpo $\mathbb{K}$, entonces $(V, *, \cdot\mathbb{K})$ es un espacio vectorial sobre $\mathbb{K}$.

En este curso, si no decimos lo contrario, nos referiremos siempre a $(V, *, \cdot\mathbb{R})$ y omitiremos el símbolo "$\cdot$" para la ley de composición externa que indicaremos mediante una simple yustaposición.

Ejemplo 1.2.4

El conjunto de polinomios de grado menor o igual que 2 con coeficientes reales y una variable $(P_2(x), +, \mathbb{R})$ es un espacio vectorial sobre $\mathbb{R}$.

La operación interna es la suma de polinomios y la operación externa es la multiplicación de un polinomio por un escalar. Los polinomios son los vectores y los números reales los escalares.

Ambas operaciones nos resultan familiares, las utilizamos con frecuencia y conocemos sus propiedades.

Ejemplo 1.2.5

El conjunto de polinomios de grado 2 con coeficientes reales y una variable $(P_2(x), +, \mathbb{R})$ no es un espacio vectorial sobre $\mathbb{R}$.

La suma de dos polinomios de grado 2, puede no ser un polinomio de grado 2.

Por ejemplo: $(x^2 + 2x - 1) + (-x^2 + x + 3) = 3x + 2$; $3x + 2 \notin P_2(x)$, ya que, es de grado 1.

En el mismo conjunto se pueden definir distintas operaciones y según las propiedades que verifiquen, tendrán una estructura diferente.

Ejemplo 1.2.6

Si en $(\mathbb{R}^2, +, \mathbb{R})$ definimos $\begin{cases} (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2) \\ \lambda \overline{u} = (\lambda u_1, u_2) \end{cases}$, siendo $\overline{u} = (u_1, u_2)$, $\overline{v} = (v_1, v_2) \in \mathbb{R}^2$, $\lambda \in \mathbb{R}$, entonces $(\mathbb{R}^2, +, \mathbb{R})$ no es un espacio vectorial.

Veamos que no tiene la propiedad $(\lambda + \mu)\overline{u} = \lambda\overline{u} + \mu\overline{u}$:

$$(\lambda + \mu)\overline{u} = ((\lambda + \mu)u_1, u_2) = (\lambda u_1 + \mu u_1, u_2)$$
$$\lambda\overline{u} + \mu\overline{u} = (\lambda u_1, u_2) + (\mu u_1, u_2) = (\lambda u_1 + \mu u_1, 2u_2)$$

Los resultados obtenidos en ambos miembros son diferentes, es decir, $(\lambda + \mu)\overline{u} \neq \lambda\overline{u} + \mu\overline{u}$.

Hay propiedades que se deducen directamente de la definición de espacio vectorial, y es más eficiente demostrarlas una sola vez de forma general, que comprobar su veracidad para cada caso concreto. Algunas de las más utilizadas son las siguientes:

1.2.4. Consecuencias

En cualquier espacio vectorial $(V, *, \mathbb{R})$ se verifica:

1. $\lambda\overline{0} = \overline{0}$
2. $0\overline{u} = \overline{0}$
3. $(-\lambda)\overline{u} = \lambda(-\overline{u}) = -(\lambda\overline{u})$
4. Si $\lambda\overline{u} = \overline{0} \Rightarrow \lambda = 0$ ó $\overline{u} = \overline{0}$
5. Si $\lambda\overline{u} = \mu\overline{u}$, $\overline{u} \neq \overline{0} \Rightarrow \lambda = \mu$
6. Si $\lambda\overline{u} = \lambda\overline{v}$, $\lambda \neq 0 \Rightarrow \overline{u} = \overline{v}$

Demostraciones:

1. $$\lambda\,(\overline{u} + \overline{0}) = \begin{cases} = \lambda\overline{u} + \lambda\overline{0} \\ = \lambda\overline{u} \end{cases} \Rightarrow \lambda\overline{0} = \overline{0}$$

Para obtener $\lambda\overline{u} + \lambda\overline{0}$ hemos aplicado la propiedad distributiva de los vectores respecto a la suma de escalares y para obtener $\lambda\overline{u}$ hemos realizado primero la suma de vectores.

2. $$(\lambda + 0)\overline{u} = \begin{cases} = \lambda\overline{u} + 0\overline{u} \\ = \lambda\overline{u} \end{cases} \Rightarrow 0\overline{u} = \overline{0}$$

3. $$\overline{0} = \begin{cases} = (-\lambda + \lambda)\overline{u} = (-\lambda)\overline{u} + \lambda\overline{u} \Rightarrow (-\lambda)\overline{u} = -(\lambda\overline{u}) \\ = \lambda(-\overline{u} + \overline{u}) = \lambda(-\overline{u}) + \lambda\overline{u} \Rightarrow \lambda(-\overline{u}) = -(\lambda\overline{u}) \end{cases}$$

4. En $\lambda\overline{u} = \overline{0}$, puede ocurrir $\lambda = 0$ ó $\lambda \neq 0$; si $\lambda \neq 0$, $\exists\lambda^{-1} \in \mathbb{R}$,

$$\lambda^{-1}(\lambda\overline{u}) = \begin{cases} = \lambda^{-1}(\overline{0}) = \overline{0} \\ (\lambda^{-1}\lambda)\overline{u} = 1\overline{u} = \overline{u} \end{cases} \Rightarrow \overline{u} = \overline{0}$$

5. Si $\lambda\overline{u} = \mu\overline{u} \Rightarrow \lambda\overline{u} - \mu\overline{u} = (\lambda - \mu)\overline{u} = \overline{0}$, como $\overline{u} \neq \overline{0}$, aplicando la propiedad 4 $\Rightarrow \lambda - \mu = 0 \Rightarrow \lambda = \mu$.
6. Si $\lambda\overline{u} = \lambda\overline{v} \Rightarrow \lambda\overline{u} - \lambda\overline{v} = \lambda(\overline{u} - \overline{v}) = \overline{0}$, como $\lambda \neq 0$, aplicando la propiedad 4 $\Rightarrow \overline{u} - \overline{v} = \overline{0} \Rightarrow \overline{u} = \overline{v}$.

Subespacio

- Conservar la suma interna
- La operación interna con escalares
- El vector $\overline{0}$ tiene q pertenecer

1.3. Subespacios vectoriales

En el organigrama de este capítulo, después de saber qué es un espacio vectorial, nos hacemos la pregunta: ¿Hay subconjuntos de un espacio vectorial V que sean a su vez espacios vectoriales?

La respuesta siempre es sí.

Cualquier espacio vectorial V tiene, al menos, dos subconjuntos que con las mismas operaciones tienen también estructura de espacio vectorial; son el elemento neutro $(\overline{0})$, y el mismo conjunto V se llaman subespacios impropios. Si algún otro subconjunto de V es espacio vectorial, diremos que es un subespacio propio de V.

1.3.1. Definición de subespacio vectorial

U es un subespacio vectorial del espacio $(V, *, \mathbb{R}) \Leftrightarrow U$ es subconjunto no vacío de V y $(U, *, \mathbb{R})$ es espacio vectorial

Para ver si un subconjunto U de V es un subespacio, o no, se puede ir comprobando que verifica todas y cada una de las propiedades que definen un espacio vectorial, o bien, buscar una caracterización, es decir, una o varias condiciones equivalentes, el símbolo que se utiliza para la equivalencia es $\Leftrightarrow$.

Con $\Leftrightarrow$ debemos entender:

Tener las propiedades $\Rightarrow$ Cumplir la caracterización

y

Cumplir la caracterización $\Rightarrow$ Tener las propiedades

1.3.2. Teorema: Caracterización de subespacio vectorial

U es subespacio vectorial de $V \Leftrightarrow \begin{cases} \forall \overline{u}, \overline{v} \in U \Rightarrow \overline{u} * \overline{v} \in U \\ \forall \overline{u} \in U, \forall \lambda \in \mathbb{R} \Rightarrow \lambda\overline{u} \in U \end{cases}$ y siendo U un subjunto no vacío de V.

Demostración:

$\Rightarrow$) Es evidente; por ser $(U, *, \mathbb{R})$ espacio, se verifica $\overline{u} * \overline{v} \in U$ y $\lambda\overline{u} \in U$.

$\Leftarrow$) Para los elementos de U se cumplen automáticamente las propiedades de la multiplicación por elementos de $\mathbb{R}$, ya que los elementos de U lo son también de V.

Lo mismo ocurre con las propiedades asociativa y conmutativa de $(U, *)$.

El vector neutro $\overline{0}$ pertence a U. En efecto, $\forall \overline{u} \in U$, $0\overline{u} = \overline{0} \in U$.

Para cada elemento $\overline{u}$ de U, hay un simétrico $-\overline{u}$ que también pertenece a U: $\forall \overline{u} \in U$, $(-1)\overline{u} = -\overline{u} \in U$.

Ejemplo 1.3.1

El conjunto de soluciones de la ecuación $2x_1 - x_2 + x_3 = 0$, es un subespacio vectorial de $(\mathbb{R}^3, +, \mathbb{R})$.

El conjunto de soluciones de la ecuación dada es $S = \{(s_1, s_2, s_3)/2s_1 - s_2 + s_3 = 0\}$, que es un subconjunto no vacío de $\mathbb{R}^3$. Comprobemos que se cumple *[1.3.2]*:

$$(s_1, s_2, s_3) \in S \Leftrightarrow 2s_1 - s_2 + s_3 = 0; \qquad (t_1, t_2, t_3) \in S \Leftrightarrow 2t_1 - t_2 + t_3 = 0;$$

$$(s_1, s_2, s_3) + (t_1, t_2, t_3) \in S \text{ porque } 2(s_1 + t_1) - (s_2 + t_2) + (s_3 + t_{13}) =$$
$$= (2s_1 - s_2 + s_3) + (2t_1 - t_2 + t_3) = 0$$
$$\lambda(s_1, s_2, s_3) \in S \text{ porque } \lambda(2s_1 - s_2 + s_3) = 0$$

Ejemplo 1.3.2

$U = \{(u_1, u_2, u_3)/u_1 = 1\}$ no es subespacio de $\mathbb{R}^3$.

Es un subconjunto no vacío, en el que la suma de dos elementos de U como $(1, 2, 3)$ y $(1, 0, 0)$ es $(2, 2, 3)$ que no pertenece a U porque su primera coordenada es $\neq 1$.

En este ejemplo es evidente que U no puede ser espacio vectorial, ya que $\overline{0} = (0, 0, 0) \notin U$, pero el elemento neutro debe pertenecer a U porque U es un grupo.

La dos condiciones que caracterizan los subespacios se pueden agrupar en una sola.

1.3.3. Teorema: otra caracterización de subespacio vectorial

U (subconjunto no vacío de V) es subespacio vectorial de $V \Leftrightarrow \forall \overline{u}, \overline{v} \in U; \forall \lambda, \mu \in \mathbb{R} \Rightarrow \lambda\overline{u} * \mu\overline{v} \in U$

Ambas caracterizaciones son equivalentes:

Demostración:

$\Rightarrow$) Por cumplirse *[1.3.2.]*, podemos afirmar que $\forall \overline{u}, \overline{v} \in U; \forall \lambda, \mu \in \mathbb{R} \Rightarrow \Rightarrow \lambda\overline{u} * \mu\overline{v} \in U$.

$\Leftarrow$) Si en *[1.3.3.]*, hacemos $\forall \overline{u}, \overline{v} \in U; \lambda = 1, \mu = 1 \Rightarrow \overline{u} * \overline{v} \in U$.

Si en *[1.3.3]*, hacemos $\forall \overline{u}, \overline{v} \in U; \forall \lambda \in \mathbb{R}, \mu = 0 \Rightarrow \lambda\overline{u} \in U$.

En *[1.3.3]* hemos utilizado un tipo especial de expresiones: las que son de la forma $\lambda\overline{u} * \mu\overline{v} \in U$; son muy importantes y reciben el nombre de combinaciones lineales

1.3.4. Definición de combinación lineal

$\overline{v}$ es una combinación lineal o depende linealmente de los vectores de $S \subset V$
$\overline{v} = \lambda_1\overline{u}_1 * \lambda_2\overline{u}_2 * \ldots * \lambda_p\overline{u}_p$; $S = \{\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_p\} \subset$ espacio vectorial V; $\lambda_1, \lambda_2, \ldots, \lambda_p \in \mathbb{R}$

El conjunto formado por todas las combinaciones lineales de los elementos de S, recibe diversos nombres, como envolvente lineal de S, o clausura de S y se representa generalmente como $\langle S \rangle$.

Los múltiples colores de la pantalla de un monitor se definen como combinación lineal de los colores primarios rojo, verde y azul.

1.3.5. Definición de sistema de generadores de un espacio vectorial

S es un sistema de generadores de U o U está generado por $S \Leftrightarrow U = \langle S \rangle$

Ejemplo 1.3.3

Cualquier vector de la forma $\overline{v} = \lambda_1 (1, 0, 1) + \lambda_2 (2, 0, 1)$ es combinación lineal de los vectores de $S = \{(1, 0, 1), (2, 0, 1)\} \subset \mathbb{R}^3$. S es el sistema de generadores de $\langle S \rangle$.

El subespacio de $\mathbb{R}^3$ generado por S es $\langle S \rangle = \{(\lambda_1 + 2\lambda_2, 0, \lambda_1 + \lambda_2)/\lambda_1, \lambda_2 \in \mathbb{R}\}$.

Gráficamente, vemos que, cualquier vector del plano que contiene los vectores dados, es una combinación lineal de ellos. Dicho plano es el subespacio más pequeño que los contiene.

Ejemplo 1.3.4

Compruébese que $S' = \{(1, 0, 1), (3, 0, 2)\}$ genera el mismo subespacio de $\mathbb{R}^3$ que el generado por S del ejemplo anterior.

La ecuación del plano generado por S es $x_2 = 0$.
La ecuación del plano generado por S' es $x_2 = 0$.
$\langle S \rangle$, $= \langle S' \rangle$ se forman con todas las combinaciones lineales de dos vectores del plano $x_2 = 0$.
Ambos sistemas de vectores son sistemas equivalentes porque generan el mismo subespacio.

Consecuencia:

Para generar un subespacio no hay un sistema único.

1.3.6. Definición de vectores linealmente dependientes

$S = \{\overline{u}_1, \overline{u}_2, \dots, \overline{u}_p\} \subset V$ es un sistema linealmente dependiente o ligado $\Leftrightarrow$
$\Leftrightarrow$ Al menos uno de los $\overline{u}_i \in S$ depende linealmente de los otros vectores de S

1.3.7. Caracterización de vectores linealmente dependientes

$S = \{\overline{u}_1, \overline{u}_2, \dots, \overline{u}_p\} \subset V$ es un sistema linealmente dependiente $\Leftrightarrow$ Si $\overline{0} = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + \dots + \lambda_p\overline{u}_p;\ \lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R} \Rightarrow \exists$ algún $\lambda_i \neq 0$

Demostración

$\Rightarrow$) Si $S = \{\overline{u}_1, \overline{u}_2, \dots, \overline{u}_p\}$ es un sistema linealmente dependiente, la definición *[1.3.6]* nos asegura que, al menos uno de los $\overline{u}_i \in S$ depende linealmente de los otros vectores de S. Podemos suponer sin pérdida de generalidad (bastaría reordenar) que es $\overline{u}_1$ el vector que depende linealmente de los otros, entonces, $\overline{u}_1 = \lambda_2\overline{u}_2 + \dots + \lambda_p\overline{u}_p$, y por tanto $\overline{0} = -1\overline{u}_1 + \lambda_2\overline{u}_2 + \dots + \lambda_p\overline{u}_p$ $(\lambda_1 \neq 0)$.
$\Leftarrow$) Si $\overline{0} = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + \dots + \lambda_p\overline{u}_p;\ \lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R} \Rightarrow \exists$ algún $\lambda_1 \neq 0$. Podemos suponer sin pérdida de generalidad (bastaría reordenar) que es $\lambda_1 \neq 0$ *como* $\lambda_1 \in \mathbb{R}$, $\exists\, \lambda_1^{-1}$, $\overline{0} = \lambda_1, \lambda_1^{-1}\,\overline{u}_1 + \lambda_2\,\lambda_1^{-1}\,\overline{u}_2 + \dots + \lambda_p\,\lambda_1^{-1}\,\overline{u}_p$; $\overline{u}_1 = \mu_2\overline{u}_2 + \dots + \mu_p\overline{u}_p$.

1.3.8. Definición de vectores linealmente independientes

$S = \{\overline{u}_1, \overline{u}_2, \dots, \overline{u}_p\} \subset V$ es un sistema linealmente independiente o libre $\Leftrightarrow$
$\Leftrightarrow \overline{0} = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + \dots + \lambda_p\overline{u}_p;\ \lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R} \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_p = 0$

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

Ejemplo 1.3.5

Compruébese que, de entre los siguientes conjuntos de vectores, son libres o linealmente independientes B y C:

$A = \{(1, 2, 0), (0, 0, 0), (4, 5, 6)\}$; $B = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$;
$C = \{(1, 4, 7), (5-3, 4)\}$

a) $$\lambda\begin{pmatrix}1\\2\\0\end{pmatrix} + \mu\begin{pmatrix}0\\0\\0\end{pmatrix} + \gamma\begin{pmatrix}4\\5\\6\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \Rightarrow \begin{cases}\lambda + 4\gamma = 0\\ 2\lambda + 5\gamma = 0 \Rightarrow \lambda = 0;\ \gamma = 0;\\ 6\gamma = 0\end{cases}$$

$\langle S - \{v\} \rangle$

La igualdad es cierta $\forall \mu \in \mathbb{R} \Rightarrow A$ es un sistema linealmente dependiente o ligado.

Si $\overline{0}$ pertenece a un sistema de generadores, éste es linealmente dependiente.

Los vectores deberían escribirse verticalmente para distinguirlos de los puntos, pero está admitido que se escriban también horizontalmente por comodidad, nosotros lo haremos indistintamente dependiendo de la situación, cuando haga falta especificar más, haremos la precisión correspondiente.

b) $\lambda \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} \lambda = 0 \\ \mu = 0 \\ \gamma = 0 \\ \delta = 0 \end{cases}; \Rightarrow B$ es un sistema libre.

B es la base canónica o estándar de $\mathbb{R}^4$.

c) $\lambda \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} + \mu \begin{pmatrix} 5 \\ -3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} \lambda = 5\mu = 0 \\ 4\lambda - 3\mu = 0 \Rightarrow \lambda = 0; \mu = 0; \\ 7\lambda + 4\mu = 0 \end{cases} \Rightarrow C$ es un sistema libre.

1.3.9. Proposición

Si $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_p\}$ es un sistema linealmente independiente $\subset V$, y $\overline{v} \notin \langle S \rangle$, y $\overline{v} \in V$, entonces, $S \cup \{\overline{v}\}$ es un sistema linealmente independiente $\subset V$

Demostración:

$\Rightarrow$) Si $\overline{0} = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + ... + \lambda_p\overline{u}_p + \mu\overline{v}$; tiene que ser $\mu = 0$, porque si $\mu \neq 0$, se podría escribir $\overline{v}$ como combinación lineal de los elementos de S, y eso va contra la hipótesis de que $\overline{v} \notin \langle S \rangle$. Todos los λ son 0 porque si no, iríamos contra la hipótesis de que $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_p\}$ es un sistema linealmente independiente.

Por tanto, $\overline{0} = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + ... + \lambda_p\overline{u}_p + \mu\overline{v} \Rightarrow \lambda_1 = \lambda_2 = ... = \lambda_p = \mu = 0$, que es la condición necesaria para poder decir que $S \cup \{\overline{v}\}$ es un sistema libre contenido en V.

Se pueden escribir más proposiciones relativas a la dependencia e independencia lineal, pero son consecuencia directa de lo que ya hemos aprendido, y se demuestran utilizando las mismas técnicas.

1.4. Transformaciones en un sistema de generadores. Espacios vectoriales finitos. Bases

En **1.3** hemos visto que los espacios vectoriales se pueden generar mediante las combinaciones lineales de un sistema de generadores cualquiera, pero en general, es más interesantes considerar los sistemas de generadores linealmente independientes.

Las proposiciones siguientes expresan propiedades relacionadas con la dependencia y la independencia lineal de los vectores $\overline{u}_i$, que forman el sistema de generadores $G = \{\overline{u}_1, \overline{u}_2, ..., u_p\}$ de un espacio vectorial, V.

Además de la importancia intrínseca que tienen las propiedades que vamos a enunciar, tienen un valor añadido porque serán utilizadas en las demostraciones de teoremas posteriores.

La proposición siguiente expresa que, la cantidad de vectores de un sistema de generadores ligado, se puede reducir sin que varíe el espacio vectorial generado.

1.4.1. Proposición

Si $\overline{u}_i \in G$ es combinación lineal de vectores de G, ($G = \{\overline{u}_1, \overline{u}_2, ..., u_p\}$ generador de V), entonces, el espacio generado por $G - \{\overline{u}_i\}$ es V

Demostración:

Supongamos sin perder generalidad que $\overline{u}_1 = \mu_2\overline{u}_2, + ... + \mu_p\overline{u}_p$. Si $\overline{v} \in V$, $\overline{v} = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + ... + \lambda_p\overline{u}_p = \lambda_1(\mu_2\overline{u}_2 + ... + \mu_p\overline{u}_p) + \lambda_2\overline{u}_2 + ... + \lambda_p\overline{u}_p = \gamma_2\overline{u}_2 + ... + \gamma_p\overline{u}_p$.

Es decir, cualquier vector de V se puede generar prescindiendo de los vectores de G que dependan linealmente de otros vectores de G.

Ejemplo 1.4.1

En el sistema $G = \{(0, 1), (3, 0), (0, 2)\}$, el vector $(0, 2) = 2(0, 1) + 0\,(3, 0)$.

El espacio que genera G es el mismo que genera $G_1 = G - \{(0, 2)\} = \{(0, 1), (3, 0)\}$, ambos generan $\mathbb{R}^2$.

Obsérvese que cualquier vector de $\mathbb{R}^2$ se puede expresar con los vectores de G, pero no de forma única.

Los colores de la pantalla de un monitor se pueden definir como una combinación lineal de rojo, azul y verde, pero no se pueden definir con ninguna combinación lineal de rojo, azul y morado, porque el morado depende linealmente de rojo y azul.

La proposición siguiente expresa que se puede sustituir un vector del sistema de generadores G, por cualquier combinación lineal de los vectores de G donde intervenga dicho vector con coeficiente no nulo.

1.4.2. Proposición

Si $\overline{v} \in V$ es combinación lineal de vectores $\overline{u}_i$ de G, ($G = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_p\}$ generador de V), entonces, al sustituir un $\overline{u}_i$ (con coeficiente $\neq 0$ en la combinación lineal que expresa $\overline{v}$) por $\overline{v}$ en G, el sistema obtenido G_1 sigue siendo generador de V

Demostración:

Sea $\overline{v} = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + ... + \lambda_p\overline{u}_p$, y supongamos que es $\lambda_1 \neq 0$, entonces, $\overline{u}_1 = \lambda_1^{-1}\overline{v} - \lambda_1^{-1}\lambda_2\overline{u}_2 - ... - \lambda_1^{-1}\lambda_p\overline{u}_p$.

Cualquier vector $\overline{x} \in V$ verifica $\overline{x} = \mu_1\overline{u}_1, \mu_2\overline{u}_2, + ... + \mu_p\overline{u}_p = \mu_1(\lambda_1^{-1}\overline{v} - \lambda_1^{-1}\lambda_2\overline{u}_2 - ... - \lambda_1^{-1}\lambda_p\overline{u}_p) + \mu_2\overline{u}_2 + ... + \mu_p\overline{u}_p = \gamma_1\overline{v} + \gamma_2\overline{u}_2 + ... + \gamma_p\overline{u}_p$.

Ejemplo 1.4.2

Sea $G = \{(1, 0, 0), (0, 1, 0)\}$ generador del plano de $\mathbb{R}^3$, $z = 0$.

$\overline{v} = 2(1, 0, 0) + 0(0, 1, 0) = (2, 0, 0)$ puede sustituir a $(1, 0, 0)$ en $G = \{(1, 0, 0), (0, 1, 0)\}$, obteniendo $G_1 = \{(2, 0, 0), (0, 1, 0)\}$, generador del mismo plano $z = 0$.

Sin embargo, (2, 0, 0) no puede sustituir a (0, 1, 0) porque el sistema $G_2 = \{(1, 0, 0), (2, 0, 0)\}$ no genera el plano $z = 0$, sino la recta $y = 0$; $z = 0$.

El teorema siguiente expresa que el número de vectores de un sistema generador libre es menor o igual al de cualquier otro sistema generador.

1.4.3. Teorema fundamental de la independencia lineal

Si $I = \{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_h\}$ es sistema libre de generadores de V y $G = \{\overline{u}_1, \overline{u}_2, \dots, \overline{u}_p\}$ un sistema generador de V, entonces, $h \leq p$.

Demostración:

$\Rightarrow$) $\overline{v}_i$ son vectores de un sistema libre de V, G es generador de V, por tanto $\overline{v}_1 = \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + \dots + \lambda_p\overline{u}_p$ con algún $\lambda_i \neq 0$, supongamos $\lambda_1 \neq 0$, utilizando la proposición anterior podemos asegurar que $G_1 = \{\overline{v}_1, \overline{u}_2, \dots, \overline{u}_p\}$ es generador de V.

(Observemos que hemos sustituido un vector del sistema de generadores por un vector del sistema libre).

Por ser $G_1 = \{\overline{v}_1, \overline{u}_2, \dots, \overline{u}_p\}$ generador de V y $\overline{v}_2$ vector de un sistema libre de V, $\overline{v}_2 = \mu_1\overline{v}_1, \mu_2\overline{u}_2 + \dots + \mu_p\overline{u}_p$ con algún $\mu_i \neq 0$, supongamos $\mu_2 \neq 0$, utilizando la proposición anterior podemos asegurar que $G_2 = \{\overline{v}_1, \overline{u}_2, \dots, \overline{u}_p\}$ es generador de V.

Este proceso lo podemos repetir, como máximo, p veces, obteniendo $G_p = \{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_p\}$, si $p < h$ todavía quedan vectores $\{\overline{v}_{i+1}, \overline{v}_{i+2}, \dots, \overline{v}_h\}$ de I, que se pueden expresar como combinaciones lineales de $G_p = \{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_p\}$, que es un sistema generador, la suposición de que $p < h$ nos ha llevado a una contradicción, porque $I = \{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_h\}$ es un sistema libre, como consecuencia, la suposición era falsa y $p \geq h$.

Para hacer esta demostración, hemos utilizado una forma de razonar muy frecuente, es la reducción al absurdo. Consiste en suponer que es cierto lo contrario a aquello que queremos demostrar (queremos demostrar que $p \geq h$ y suponemos que $p < h$), seguir el razonamiento lógico, y llegar a una contradicción con la hipótesis (concluimos que $I = \{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_h\}$ no es libre, en contra de la hipótesis de que partimos).

Al repetir el proceso de sustitución de un vector del sistema de generadores, por un vector del sistema libre, hemos sustituido parcialmente los vectores de un sistema de generadores por los de un sistema libre cualquiera, sin que varíe el espacio generado.

1.4.4. Definición

Espacio vectorial finito es el que está generado por un número finito de vectores

En lo sucesivo, si no decimos lo contrario, nos referiremos a este tipo de espacios.

Un espacio vectorial finito puede tener infinitos elementos, sólo es necesario que tenga un sistema generador finito.

En el ejemplo 1.3.4 vimos que un espacio vectorial se puede generar a partir de distintos sistemas de generadores. Entre ellos, hay algunos que presentan ventajas indudables, por ejemplo, que cualquier vector sea una combinación lineal única de los vectores del sistema de generadores, o que contengan el menor número posible de vectores, los sistemas de generadores que cumplen esta condición reciben el nombre de bases.

1.4.5. Definición de base

$B = \{\overline{e}_1, \overline{e}_2, \dots, \overline{e}_n\}$ es base del espacio vectorial $V \Leftrightarrow B$ es sistema libre que genera V

Es muy frecuente y útil utilizar la siguiente definición de base:

1.4.6. Otra definición de base

$B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\}$ es base del espacio vectorial $V \Leftrightarrow$ cualquier $\overline{v} \in V$, se puede expresar de manera única como combinación lineal de los elementos de B.

1.4.7. Equivalencia entre las dos definiciones

B es sistema libre que genera $V \Leftrightarrow$ cualquier $\overline{v} \in V$, se puede expresar de manera única como combinación lineal de los elementos de B.

Demostración

$\Rightarrow$) Como $\overline{v} \in V$, y B es sistema generador de $V \Rightarrow \overline{v} = \lambda_1\overline{e}_1 + \lambda_2\overline{e}_2 + ... + \lambda_n\overline{e}_n$.
Si la expresión de $\overline{v}$ no fuera única, $\overline{v} = \mu_1\overline{e}_1 + \mu_2\overline{e}_2 + ... + \mu_n\overline{e}_n$.
Restando miembro a miembro las dos expresiones, se obtiene:
$\overline{0} = (\lambda_1 - \mu_1)\overline{e}_1 + (\lambda_2 - \mu_2)\overline{e}_2 + ... + (\lambda_n - \mu_n)\overline{e}_n$; por ser B un sistema libre,
$0 = \lambda_1 - \mu_1 = \lambda_2 - \mu_2 = ... = \lambda_n - \mu_n$; por tanto, $\lambda_1 = \mu_1$; $\lambda_2 = \mu_2$; ... ; $\lambda_n = \mu_n$.

$\Rightarrow$) Cualquier $\overline{v} \in V$, se puede expresar como combinación lineal de los elementos de B. Por tanto, B genera V.

$\overline{0} \in V$ se puede expresar como combinación lineal de los elementos de B, $\overline{0} = \lambda_1\overline{e}_1 + \lambda_2\overline{e}_2 + ... + \lambda_n\overline{e}_n$; dado que de $\overline{0} = 0\overline{e}_1 + 0\overline{e}_2 + ... + 0\overline{e}_n$ y los vectores se pueden expresar de manera única, $0 = \lambda_1 = \lambda_2 = ... = \lambda_n$, es decir, B es un sistema libre.

Ejemplo 1.4.2

El ejemplo más sencillo de base en $\mathbb{R}^n$ es la base canónica: $B = \{(1, 0, ..., 0), (0, 1, ..., 0), ..., (0, 0, ..., 1)\}$.

Cualquier vector $\overline{x} = (x_1, x_2, ..., x_n)$ de $\mathbb{R}^n$ es combinación lineal de los elementos de B, $(x_1, x_2, ..., x_n) = x_1 (1, 0, ..., 0) + x_2 (0, 1, ..., 0) + ... + x_n (0, 0, ..., 1)$, luego B genera $\mathbb{R}^n$.

Si $x_1 (1, 0, ..., 0) + x_2 (0, 1, ..., 0) + ... + x_n (0, 0, ..., 1) = (0, 0, ..., 0) \Rightarrow x_1 = x_2 = ... = x_n = 0$, por tanto, B es libre $\Rightarrow$ B generador y libre, es base.

Consecuencia:

La segunda definición de base nos permite decir que, $\overline{x} = (x_1, x_2, ..., x_n)$ se puede expresar de forma única mediante los coeficientes $x_1, x_2, ..., x_n$, respecto a la base canónica, dichos coeficientes reciben el nombre de coordenadas del vector $\overline{x}$ respecto a la base dada.

Las relaciones que existen entre las distintas bases de un espacio vectorial dan lugar a diversos teoremas que reciben el nombre genérico de **teoremas de la base.**

1.4.8. Teorema de existencia de la base

Cualquier espacio vectorial finito $V \neq \overline{0}$ tiene base

Demostración

Sea $G = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_p\}$ un sistema finito de generadores de $V \neq \overline{0}$.

Si sus vectores son independientes, G es base y está demostrado el teorema,

Si sus vectores no son independientes, es porque alguno depende linealmente de otros, es decir, se cumple la hipótesis de la proposición *[1.4.1]*, y por tanto, es cierta la conclusión de dicha proposición: Se genera el mismo V sustituyendo G por G_1. (G_1 se ha obtenido quitando en G los vectores que no son independientes). El nuevo sistema de generadores linealmente independientes, G_1, es una base.

1.4.9. Teorema de la dimensión

Todas las bases de un espacio vectorial V tienen el mismo número de elementos

Demostración

Sea $B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\}$ una base de V y sea $B' = \{\overline{e}'_1, \overline{e}'_2, ..., \overline{e}'_m\}$ otra base de V.

B es sistema libre de generadores, B' es sistema de generadores, por *[1.4.3]* $\Rightarrow$ $\Rightarrow n \leq m$.

Además, B' es sistema libre de generadores, B es sistema de generadores, por *[1.4.3]* $\Rightarrow m \leq n$.

Por tanto, $m = n$.

Todas las bases de un espacio tienen el mismo número de elementos, que recibe el nombre de dimensión del espacio.

1.4.10. Definición de dimensión de un espacio vectorial

Dimensión de un espacio vectorial $V \neq \{\overline{0}\}$, finito, Dim V, es el número de vectores de cualquiera de sus bases

Ejemplo 1.4.3

La base canónica de $\mathbb{R}^3$ es $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$, tiene tres elementos, los vectores $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, por tanto, cualquier base de $\mathbb{R}^3$ tiene tres elementos y tres es la dimensión del espacio vectorial $\mathbb{R}^3$.

Si el espacio es $V = \overline{0}$, por convenio se toma Dim $V = 0$.

1.4.11. Consecuencias

1. Si Dim $V = n \Rightarrow$ No puede haber en V más de n vectores linealmente independientes.

Demostración:

$$\left.\begin{array}{l} B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\} \text{ es base, por tanto, es sistema de generadores} \\ S = \{\overline{v}_1, \overline{v}_2, ..., \overline{v}_p\} \text{ es un conjunto de vectores linealmente} \\ \quad \text{independientes} \end{array}\right\} \Rightarrow p \leq n$$

Hemos utilizado *[1.4.3]*.

2. Si Dim $V = n \Rightarrow$ En cualquier sistema de generadores de V no puede haber menos de n vectores.

Demostración:

$$\left.\begin{array}{l} B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\} \text{ Sus vectores } \overline{e}_i \text{ son linealmente independientes} \\ \quad \text{(es base)} \\ G = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_p\} \text{ es sistema de generadores} \end{array}\right\} \Rightarrow n \leq p$$

3. Si Dim $V = n \Rightarrow$ Cualquier conjunto S de n vectores linealmente independientes forma una base.

Demostración:

Recordemos que en la proposición *[1.4.3]*, habíamos concluido que, los vectores de un sistema de generadores pueden ser sustituidos parcialmente por los de un sistema libre cualquiera. Por tanto, los n vectores linealmente independientes de S, pueden sustituir a los n vectores de cualquier base (son generadores), y el resultado sigue siendo generador y linealmente independiente, es decir, es otra base.

4. Si Dim $V = n \Rightarrow$ Cualquier conjunto de n generadores forma una base.

Demostración:

Si $G = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_n\}$ es sistema de generadores de V, y en G hay un sistema G' con p vectores l. i. ($p \leq n$), entonces, el espacio que genera G' también es V, por tanto, G' es base de V. Como la dimensión de V es n, el número de vectores de G' es $n \Rightarrow p = n$.
$\Rightarrow G = G'$ es base de V.

5. Si Dim $V = n \Rightarrow$ Todo conjunto de vectores $S = \{\overline{v}_1, \overline{v}_2, ..., \overline{v}_p\}$ de V con $p < n$, linealmente independiente, se puede completar con $p - n$ vectores de V hasta convertirse en base (Teorema de la base incompleta).

Demostración:

Por *[1.4.8]* podemos afirmar que en V existe una base de dimensión n. En B hay al menos un vector que no pertenece a $\langle S \rangle$, porque si no, S generaría una base de V y como consecuencia V, y esto no es posible porque $p < n$ como sabemos por *[1.4.11-1]*. Formemos S' añadiendo a S el vector encontrado en B.
Si $n = p + 1 \Rightarrow S'$ es la base buscada porque es un sistema de n vectores linealmente independientes en un espacio de dimensión n.
Si $n > p + 1 \Rightarrow$ Partiendo del sistema generador S' que tiene $p + 1$ vectores, obtenemos S'' con $p + 2$ vectores repitiendo todos los pasos que hemos dado para S.
Será necesario repetir este proceso $n - p$ veces.

1.4.12. Definición de coordenadas de un vector

$(x_1, x_2, ..., x_n)$ son las coordenadas de $\overline{x} \in V$ respecto a la base $B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\} \Leftrightarrow$
$\Leftrightarrow \overline{x} = x_1\overline{e}_1 + x_2\overline{e}_2 + ... + x_n\overline{e}_n$

Ejemplo 1.4.4

Cuando decimos que las coordenadas de un vector de $\mathbb{R}^2$ respecto a la base canónica $B = \{(1, 0), (0, 1)\}$ son $(2, 3)$, estamos expresando que $(2, 3) = 2(1, 0) + 3(0, 1)$.
Si hubiéramos tomado $B' = \{(2, 0), (0, 1)\}$ como base de $\mathbb{R}^2$, el mismo vector se expresaría como $1(2, 0) + 3(0, 1)$, y diríamos que sus coordenadas respecto a B' son $(1, 3)$.

1.5. Dimensión de los subespacios de un espacio vectorial finito

Los subespacios vectoriales son a su vez espacios vectoriales, y por tanto, tienen dimensión. Vamos a estudiar la relación que existe entre la dimensión de distintos subespacios de un espacio vectorial V de dimensión n.

1.5.1. Proposición. La intersección de subespacios es subespacio

La intersección de subespacios U_i con $i = 1, ..., n$ de V, es un subespacio de V.

Demostración:

Si $\overline{u}, \overline{v} \in \bigcap_{i=1}^{n} U_i \Rightarrow \lambda\overline{v} + \mu\overline{u} \in \bigcap_{i=1}^{n} U_i$, ya que todos los U_i son subespacios de V.

Además, la intersección no es vacía porque $\overline{0}$ pertenece a todos los U_i.

Hemos demostrado que la intersección de subespacios cumple la condición necesaria y suficiente para ser un subespacio.

1.5.2. Teorema

El conjunto $\langle S\rangle$ es el menor de todos los subespacios de V que contienen S

Demostración:

Es claro que es el menor, porque para ser subespacio que contenga S debe contener todas las combinaciones lineales de los elementos de S, es decir, debe

contener $\langle S\rangle$, y $\langle S\rangle$ es espacio vectorial porque su definición evidencia que se cumple *[1.1.3]*.

1.5.3. Proposición. La unión de subespacios puede no ser subespacio

La unión de subespacios U_i de V, puede no ser un subespacio de V

Demostración:

Es suficiente encontrar un ejemplo en el que la unión de subespacios no sea un subespacio, eso es lo que haremos en el ejemplo siguiente.

Ejemplo 1.5.1

Sean los subespacios del espacio $\mathbb{R}^2$, U_1 = eje de las y, U_2 = eje de las x, si $U_1 \cup U_2$ fuera espacio, $\lambda(0, 1) + \mu(1, 0) = (\lambda, \mu)$ debería pertenecer a $U_1 \cup U_2$, porque $(0, 1)$, $(1, 0)$ pertenecen a $U_1 \cup U_2$ y $\lambda, \mu, \in \mathbb{R}$, pero, es evidente que $(\lambda, \mu) \notin U_1 \cup U_2$ porque si $\lambda \neq 0$ y $\mu \neq 0$ (λ, μ) no está en ningún eje.

Es importante observar que, para comprobar que una afirmación es cierta, no es suficiente verificarla en un caso particular o ejemplo, hay que hacer una demostración, es decir, ver que la afirmación es cierta para todos los casos.

Sin embargo, para ver que una afirmación no es correcta, es suficiente que no lo sea, al menos, en un caso, y por tanto, vale con encontrar un ejemplo que no verifique la afirmación. Dicho ejemplo se llama contraejemplo.

Sabemos que $U_1 \cup U_2$ es el conjunto más pequeño que contiene U_1 y U_2, pero con frecuencia es interesante utilizar, no el conjunto mínimo, sino el subespacio mínimo de V, que contiene U_1 y U_2, ese subespacio se llama subespacio suma.

1.5.4. Definición de suma de subespacios

Suma de los subespacios U_1 y U_2, de V es el conjunto $U_1 + U_2$:

$U_1 + U_2 = \{\overline{u}_1 + \overline{u}_2 \,/\, \overline{u}_1 \in U_1,\ \overline{u}_2 \in U_2\}$

1.5.5. Proposición. La suma de subespacios es subespacio

La suma de subespacios de V es un subespacio de V

Demostración:

Sean $\overline{u}, \overline{v} \in U_1 + U_2$, entonces

$$\left\{\begin{array}{l}\overline{u} = \overline{u}_1 + \overline{u}_2 \,/\, \overline{u}_1 \in U_1, \overline{u}_2 \in U_2 \\ \overline{v} = \overline{v}_1 + \overline{v}_2 \,/\, \overline{v}_1 \in U_1, \overline{v}_2 \in U_2\end{array}\right. \Rightarrow \lambda\overline{u} + \mu\overline{u} = \lambda(\overline{u}_1 + \overline{u}_2) + \mu(\overline{v}_1 + \overline{v}_2) =$$

$$= (\lambda\overline{u}_1 + \mu\overline{v}_1) + (\lambda\overline{u}_2 + \mu\overline{v}_2) \in U_1 + U_2.$$

Ejemplo 1.5.2

Dados los subespacios $U_1 = \{(\alpha, \beta, 0)\} \subset \mathbb{R}^3$ y $U_2 = \{(0, \lambda, \mu)\} \subset \mathbb{R}^3$, ($U_1$ es el plano $z = 0$; (plano XY) y U_2 es el plano $x = 0$; (plano YZ)).

Los vectores que forman $U_1 + U_2$ son de la forma $(\alpha, \beta + \lambda, \mu)$, es decir, son todos los vectores de $\mathbb{R}^3$.

La descomposición de un vector $\mathbb{R}^3$ en suma de uno del plano $z = 0$ y otro del plano $x = 0$ no es única.

Ejemplo 1.5.3

$(2, 5, 3) = (2, 0, 0) + (0, 5, 3) = (2, 2, 0) + (0, 3, 3) = (2, 1, 0) + (0, 4, 3) = \ldots$

La dimensión de los subespacios está relacionada con la dimensión del subespacio suma mediante la expresión conocida como fórmula de dimensión o fórmula de Grassmann.

1.5.6. Teorema (fórmula de Grassmann)

$\text{Dim}\,(U_1 + U_2) + \text{Dim}\,(U_1 \cap U_2) = \text{Dim}\, U_1 + \text{Dim}\, U_2$

En sumas Directas: $\text{Dim}(U_1 + U_2) + 0 = \text{Dim}\,U_1 + \text{Dim}\,U_2$

Demostración:

Si $B_0 = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_p\}$ es base de $U_1 \cap U_2$), sabemos que se puede ampliar hasta conseguir una base $B_1 = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_p, \overline{u}_1, \ldots, \overline{u}_r\}$ de U_1, y por otra parte, se puede ampliar hasta conseguir otra base $B_2 = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_p, \overline{v}_1, \ldots, \overline{v}_s\}$ de U_2.

Podemos construir $B = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_p, \overline{u}_1, \ldots, \overline{u}_r, \overline{v}_1, \ldots, \overline{v}_s\}$, y demostraremos que es base de $U_1 + U_2$.

Para que B sea base hace falta que sea un sistema libre y que genere $U_1 + U_2$.

B genera $U_1 + U_2$:

$\forall \overline{u} \in U_1 + U_2, \overline{u} = \overline{u}_1 + \overline{u}_2, \overline{u}_1 \in U_1, \overline{u}_2 \in U_2$

Como B_1 es base de U_1 y B_2 es base de U_2, $\overline{u}_1$ se puede escribir como combinación lineal de $B_1 = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_p, \overline{u}_1, \ldots, \overline{u}_r\}$, y $\overline{u}_2$ como combinación lineal de los elementos de $B_2 = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_p, \overline{v}_1, \ldots, \overline{v}_s\}$, por tanto, $\overline{u}$ se puede escribir como combinación lineal de los elementos de B.

B es un sistema libre:

Sea $\overline{w}_0 + \overline{w}_1 + \overline{w}_2 = \overline{0}$, una combinación lineal de los vectores $\overline{w}_0 = \lambda_1\overline{e}_1 + \ldots + \lambda_p\overline{e}_p$, $\overline{w}_1 = \alpha_1\overline{u}_1 + \ldots + \alpha_r\overline{u}_r$, $\overline{w}_2 = \beta_1\overline{v}_1 + \ldots + \beta_s\overline{v}_s$; podemos escribir $\overline{w}_0 + \overline{w}_1 = -\overline{w}_2$, como $\overline{w}_0 + \overline{w}_1$ depende linealmente de B_1 y $\overline{w}_2$ depende linealmente de B_2 se deduce que, $\overline{w}_0 + \overline{w}_1 \in U_1 \cap U_2$, y como consecuencia, se puede expresar como combinación lineal de vectores de B_o: $\overline{w}_0 + \overline{w}_1 = \delta_1\overline{e}_1 + \ldots + \delta_p\overline{e}_p$.

Si formamos la combinación $\overline{w}_0 + \overline{w}_1 + \overline{w}_2 = \delta_1\overline{e}_1 + \ldots + \delta_p\overline{e}_p + \beta_1\overline{v}_1 + \ldots + \beta_s\overline{v}_s = \overline{0}$, hemos tomado un vector de $U_1 \cap U_2$, y otro de U_2, es decir, un vector generado por B_2, que es un sistema libre por ser base $\Rightarrow \delta_1 = \ldots = \delta_p = \beta_1 = \ldots = \beta_s = 0$.

Por ser $\beta_1 = \ldots = \beta_s = 0$ la expresión $\overline{w}_0 + \overline{w}_1 + \overline{w}_2 = \overline{0}$, se transforma en $\overline{w}_o + \overline{w}_1 = \overline{0} \Leftrightarrow \lambda_1\overline{e}_1 + \ldots + \lambda_p\overline{e}_p + \alpha_1\overline{u}_1 + \ldots + \alpha_r\overline{u}_r = \overline{0} \Rightarrow \lambda_1 = \ldots = \lambda_p = \alpha_1 = \ldots = \alpha_r = 0$, con lo que hemos demostrado que B es un sistema libre.

Como antes habíamos demostrado que B es un sistema generador podemos afirmar que B es base de $U_1 + U_2$.

Para que la descomposición de un vector sea única debemos movernos en una clase especial de sumas: las sumas directas, de ellas hablaremos muy pronto.

1.5.7. Definición de suma directa de subespacios

Suma de directa de los subespacios U_1 y U_2, de V es el conjunto $U_1 \oplus U_2 = \{\overline{u} = \overline{u}_1 + \overline{u}_2 \,/\, \overline{u}_1 \in U_1,\ \overline{u}_2 \in U_2\}$, tal que la descomposición de cada vector $\overline{u}$ es única.

Ejemplo 1.5.4

Dados los subespacios $U_1 = \{(\alpha, \beta, 0)\} \subset \mathbb{R}^3$ y $U_2 = \{(0, 0, \mu)\} \subset \mathbb{R}^3$ (U_1 es el plano $z = 0$; (plano XY) y U_2 es la recta $\begin{cases} x = 0 \\ y = 0 \end{cases}$; (eje Z)), los vectores que forman $U_1 + U_2$ son de la forma (α, β, μ), es decir, son todos los vectores de $\mathbb{R}^3$.

La descomposición de un vector de $\mathbb{R}^3$ en suma de uno del plano XY y otro del eje Z es única.

Así $(2, 5, 3) = (2, 5, 0) + (0, 0, 3)$, siendo $(2, 5, 0) \in$ plano $z = 0$ y $(0, 0, 3) \in \begin{cases} x = 0 \\ y = 0 \end{cases}$

1.5.8. Caracterización I de la suma directa

$U_1 \oplus U_2$ es suma de directa de los subespacios $U_1 + U_2 \Leftrightarrow \overline{u} = \overline{u}_1 + \overline{u}_2 = \overline{0} \Rightarrow \overline{u}_1 = \overline{u}_2 = \overline{0}$, $\forall\, \overline{u} \in U_1 \oplus U_2$

Demostración

$\Rightarrow$) $\overline{u} = \overline{u}_1 + \overline{u}_2 = \overline{0} + \overline{0}$, como la descomposición de $\overline{u}$ es única, $\overline{u}_1 = \overline{u}_2 = \overline{0}$.

$\Leftarrow$) Supongamos que se verifica la segunda condición.

Si la descomposición no fuera única, $\overline{u} = \overline{u}_1 + \overline{u}_2 = \overline{u}'_1 + \overline{u}'_2 \Rightarrow (\overline{u}'_1 - \overline{u}_1) + (\overline{u}'_2 - \overline{u}_2) = \overline{0} \Rightarrow \overline{u}'_1 - \overline{u}_1 = \overline{u}'_2 - \overline{u}_2 = \overline{0} \Rightarrow \overline{u}'_1 = \overline{u}_1$; $\overline{u}'_2 = \overline{u}_2$, es decir, la descomposición es única, que es una contradicción con la hipótesis.

Otra forma de ver que una suma es directa es comprobar que la intersección de los dos subespacios es $\{\overline{0}\}$.

1.5.9. Caracterización II de la suma directa

$U_1 \oplus U_2$ es suma de directa de los subespacios U_1 y $U_2 \Leftrightarrow \overline{u} \in U_1 \cap U_2 \Rightarrow \overline{u} = \overline{0}$

Demostración:

$\Rightarrow$) $\overline{u} \in U_1 \cap U_2 \Rightarrow \overline{u} \in U_1$ y $\overline{u} \in U_2$; U_1, U_2 son espacios vectoriales $-\overline{u} \in U_2$.
Como $(\overline{u}) + (-\overline{u}) = \overline{0}$, aplicando *[1.5.8.]* $\Rightarrow \overline{u} = \overline{0}$.
$\Leftarrow$) Ahora, la hipótesis es $\overline{u} \in U_1 \cap U_2 \Rightarrow \overline{u} = \overline{0}$.
Sea $\overline{u} = \overline{u}_1 + \overline{u}_2 = \overline{0}$, un vector de $U_1 + U_2$, $\overline{u}_1 \in U_1$, $\overline{u}_2 \in U_2$.
$\overline{u}_1 \in U_1$, como $\overline{u}_1 = -\overline{u}_2 \Rightarrow -\overline{u}_2 \in U_1$, $\Rightarrow \overline{u}_2 \in U_1$.
$\overline{u}_2 \in U_2$, como $\overline{u}_2 = -\overline{u}_1 \Rightarrow -\overline{u}_1 \in U_1$, $\Rightarrow \overline{u}_1 \in U_2$.
Por tanto, los vectores $\overline{u}_1$, $\overline{u}_2$, son de $U_1 \cap U_2$, cuyo único vector es $\overline{0}$.
Hemos demostrado que la suma es suma directa.

1.5.10. Definición y existencia de subespacio suplementario

> Dado un subespacio finito U_1, existe un subespacio U_2, tal que $V = U_1 \oplus U_2$
> U_2 es el *subespacio suplementario* de U_1

Demostración de la existencia de U_2:

Si $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_p\}$ es base de U_1, y $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_p, \overline{e}_{p+1}, ..., \overline{e}_n\}$ es base de V, entonces podemos construir $U_2 = \langle \overline{e}_{p+1}, ..., \overline{e}_n \rangle$, que cumple la condición requerida.

Una buena forma de demostrar que un conjunto existe, es construirlo, eso es lo que hemos hecho para demostrar que existe U_2.

En el ejemplo 1.5.4, el complementario del plano XY (Subespacio U_1) es el eje Z (Subespacio U_2).
Es evidente que si U_1 y U_2 son suplementarios, sus dimensiones están relacionadas por la expresión: $\mathrm{Dim}U_1 + \mathrm{Dim}U_2 = \mathrm{Dim}V$.

1.5.11. Consecuencias

Si U_2 es el subespacio suplementario de U_1 en V, se deducen las siguientes consecuencias de forma inmediata al aplicar la definición de subespacio suplementario dada en *[1.5.10]*:

1. Relación entre las dimensiones de subespacios suplementarios:

$$\mathrm{Dim}\ U_1 + \mathrm{Dim}\ U_2 = \mathrm{Dim}\ V$$

Demostración:

Como $\mathrm{Dim}\ (U_1 \cap U_2) = 0$, el resultado es consecuencia directa de *[1.5.6.]*.

2. *Caracterización de subespacio suplementario de un subespacio:*

$$U_1 + U_2 = V \text{ y } U_1 \cap U_2 = \overline{0}$$

Demostración:

La primera afirmación es consecuencia de la definición de suma directa y la segunda de la caracterización II de suma directa.

3. Si $B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\}$ es base del espacio $V \Rightarrow V = \langle\overline{e}_1\rangle \oplus \langle\overline{e}_2\rangle \oplus ... \oplus \oplus \langle\overline{e}_n\rangle$.

Demostración:

Es consecuencia inmediata de las definiciones de base y suma directa.

EJERCICIOS

Ejercicio 1.1

Demuéstrese que el conjunto de funciones $\mathfrak{F} : \mathbb{R} \rightarrow \mathbb{R}$ con las siguientes operaciones:

1. Suma de funciones, definida como $(f+g)(x) + f(x) + g(x)$, siendo $f, g \in \Im$.
2. Multiplicación por escalares, definida como $(\lambda f)(x) = \lambda f(x)$, siendo $\lambda \in \mathbb{R}, f \in \Im$, es un espacio vectorial sobre $\mathbb{R}$.

Solución:

Tenemos que demostrar que las funciones tienen una serie de propiedades; para ello, sabiendo que dos funciones son iguales si sus imágenes son iguales para cada elemento del dominio de definición, trasladamos las comprobaciones al campo de las imágenes, cuyas propiedades conocemos por ser números reales.

$(\Im, +)$ es grupo conmutativo. La operación + en el conjunto $\Im$ verifica todas las propiedades necesarias:

- Asociativa: $(f+g)+h = f+(g+h)$, $\forall f, g, h \in \Im$.

Primer miembro	Segundo miembro
$((f+g)+h)(x) = (f+g)(x) + h(x) =$ $= f(x) + g(x) + h(x)$	$(f+(g+h))(x) = f(x) + (g+h)(x) =$ $= f(x) + g(x) + h(x)$

- Propiedad conmutativa : $\forall f, g \in \Im$ se verifica $f+g = g+f$

Primer miembro	Segundo miembro
$(f+g)(x) = f(x) + (g)(x)$	$(g+f)(x) = g(x) + f(x)$

- Existe un elemento neutro: $0 \in \Im$, que verifica $f + 0 = f$, $\forall f \in \Im$
La función 0 es la función nula, es decir, $0(x) = 0$, $\forall x \in \mathbb{R}$

Primer miembro	Segundo miembro
$(f+0)(x) = f(x) + 0(x) = f(x) + 0 = f(x)$	$f(x)$

- Existe un elemento simétrico para cada elemento del conjunto: $\forall f \in \Im$, $\exists(-f) \in \Im$ definido como $(-f)(x) = -f(x)$, que verifica $f + (-f) = 0$

Primer miembro	Segundo miembro
$(f+(-f))(x) = f(x) + (-f)(x) = 0$	0

$(\Im, +\mathbb{R})$ tiene las propiedades:

- Distributiva de escalares respecto a elementos de $\Im$: $\lambda(f+g) = \lambda f + \lambda g$

Primer miembro	Segundo miembro
$\lambda(f+g)(x) = \lambda(f(x) + g(x)) =$ $= \lambda(f(x)) + \lambda(g(x))$	$(\lambda f + \lambda g)(x) = (\lambda f)(x) + (\lambda g)(x) =$ $= \lambda f(x) + \lambda g(x)$

- Distributiva de elementos de $\Im$ respecto a escaleres: $(\lambda + \mu)f = \lambda f + \mu f$

Primer miembro	Segundo miembro
$((\lambda + \mu)f)(x) = (\lambda + \mu)f(x) =$ $= \lambda f(x) + \mu f(x)$	$(\lambda f + \mu f)(x) = (\lambda f)(x) + (\mu f)(x) =$ $= \lambda f(x) + \mu f(x)$

- Asociativa respecto al producto de escalares: $(\lambda\mu)f = \lambda(\mu f)$

Primer miembro	Segundo miembro
$((\lambda\mu)f)(x) = (\lambda\mu)f(x) = \lambda\mu f(x)$	$(\lambda(\mu f))(x) = \lambda(\mu f)(x) + \lambda\mu f(x)$

♦ Existe un escalar unidad, 1, que verifica $1f = f$

Primer miembro	Segundo miembro
$(1f)(x) = 1f(x) = f(x)$	$f(x)$

Podemos decir que las funciones son vectores por ser elementos de un espacio vectorial, y los elementos de $\mathbb{R}$ son los escalares.

Ejercicio 1.2

$(A, *, \mathbb{R})$ es una estructura algebraica, tal que, en el conjunto $A = \{x \,/\, x$ son dígitos impares$\}$ se han definido las siguientes operaciones:

1. $*$ es una operación, tal que, $a * b = c$, siendo c la cifra de las unidades del producto ab.
2. Multiplicación estándar por elementos de $\mathbb{R}$.

Se pide hacer la tabla de $(A, *)$ y comprobar si $(A, *, \mathbb{R})$ es, o no es, espacio vectorial.

Solución:

♦ $(A, *)$ no es grupo, porque no tiene alguna de las propiedades necesarias.

*	1	3	5	7	9
1	1	3	5	7	9
3	3	9	5	1	7
5	5	5	5	5	5
7	7	1	5	9	3
9	9	7	5	3	1

— Es ley de composición interna: $a * b \in A, \forall\, a, b \in A$
— Es asociativa por serlo el producto de números enteros
— Es conmutativa $a * b = b * a, \forall\, a, b \in A$
— Existe elemento neutro: $1 \in A$; $1 * a = a, \forall\, a \in A$
— No todos los elementos de A tienen simétrico
— Poseen simétrico 1, 3, 7, 9. No tiene simétrico 5

♦ $(A, *, \mathbb{R})$ no es un espacio vectorial, ya que no es grupo.

♦ Observación:

En la tabla se pueden ver algunas de las propiedades, por ejemplo:
Es ley de composición interna: No aparece ningún número distinto a los elementos de A.
Es conmutativa: La tabla es simétrica.
Existe elemento neutro: $1 \in A$; La 1ª fila y la primera columna, correspondientes a la multiplicación por 1, permanecen invariables.
Poseen simétrico 1, 3, 7, 9. No tiene simétrico 5: En cada fila (no hace falta especificar columna porque es simétrica) hay un 1, excepto en las correspondientes a 5.

Ejercicio 1.3

Si en el conjunto $\mathbb{R}^2$ se definen las operaciones:

$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$
$\lambda(x_1, x_2) = (\lambda^2 x_1, \lambda x_2)$
no tiene estructura de espacio vectorial. ¿Por qué?

Solución:

$(\mathbb{R}^2, +)$: Hemos visto en teoría que $(\mathbb{R}^2, +)$ es un grupo conmutativo. (La operación de sumar estándar en $\mathbb{R}^2$ es la que hemos definido.)
$(\mathbb{R}^2, +, \mathbb{R})$: Veamos las propiedades que tiene esta operación de multiplicar escalar por vector.

1: No cumple la distributiva de los vectores respecto a los escalares: $(\lambda + \mu)\bar{u} = \lambda\bar{u} * \mu\bar{u}$:

$$(\lambda + \mu)(x_1, x_2) = ((\lambda + \mu)^2 x_1, (\lambda + \mu)x_2) = (\lambda^2 x_1 + 2\lambda\mu x_1 + \mu^2 x_1, \lambda x_2 + \mu x_2)$$
$$\lambda(x_1, x_2) + \mu(x_1, x_2) = (\lambda^2 x_1, \lambda x_2) + (\mu^2 x_1, \mu x_2) = (\lambda^2 x_1 + \mu^2 x_1, \lambda x_2 + \mu x_2)$$

Ya podemos afirmar que no es un espacio vectorial, puesto que, no cumple una de las condiciones necesarias.

Aunque no es necesario, vamos a ver si tiene el resto de las propiedades, o no.

2: Sí cumple la distributiva de los escalares respecto a los vectores: $\lambda(\overline{u} * \overline{v}) = \lambda\overline{u} * \lambda\overline{v}$:

$$\lambda((x_1, x_2) + (y_1, y_2)) = \lambda(x_1 + y_1, x_2 + y_2) = (\lambda^2(x_1 + y_1), \lambda(x_2 + y_2))$$
$$\lambda(x_1, x_2) + \lambda(y_1, y_2) = (\lambda^2 x_1 + \lambda x_2) + (\lambda^2 y_1 + \lambda y_2) = (\lambda^2 x_1 + \lambda^2 y_1, \lambda x_2 + \lambda y_2)$$

3: Sí cumple la asociativa: $(\lambda\mu)\overline{u} = \lambda(\mu\overline{u})$:

$$\lambda(\mu(x_1, x_2)) = \lambda(\mu^2 x_1, \mu x_2) = (\lambda^2\mu^2 x_1, \lambda\mu x_2)$$
$$(\lambda\mu)(x_1, x_2) = ((\lambda\mu)^2 x_1, \lambda\mu x_2)$$

4: También verifica la existencia de un escalar unidad, 1, tal que $1\overline{u} = \overline{u}$:

$$1(x_1, x_2) = (1^2 x_1, 1x_2) = (x_1, x_2)$$

$\mathbb{R}^2$ con las operaciones estándar tiene estructura de espacio vectorial, pero no la tiene con las operaciones que hemos definido en este ejercicio.

Hemos comprobado, por tanto, que la estructura algebraica de un conjunto depende del conjunto y de la operación u operaciones que se definan en él.

Ejercicio 1.4

Sea $(\mathbb{R}^3, +, \mathbb{R})$ el espacio vectorial real estándar de dimensión tres. Se pide decidir cuáles de los siguientes subconjuntos son subespacios vectoriales suyos:

1. $U_1 = \{(x_1, x_2, x_3) \,/\, x_1 + x_2 + x_3 = 1\}$
2. $U_2 = \{(x_1, x_2, x_3) \,/\, x_1 + x_2 + x_3 = 0\}$

3. $U_3 = \{(x_1, x_2, x_3) \,/\, x_1 + x_2 \leq 1\}$
4. $U_4 = \{(x_1, x_2, x_3) \,/\, x_1 x_2 = 1\}$
5. $U_5 = \{(x_1, x_2, 0) \,/\, x_1^2 + x_2^2 \leq 1\}$

Solución:

La condición necesaria y suficiente para que U sea subespacio vectorial es:

$\forall\overline{u}, \overline{v} \in U, \forall\lambda, \mu, \in \mathbb{R} \Rightarrow \lambda\overline{u} + \mu\overline{v} \in U.$
Veamos qué subespacios la verifican.

1. Si $\overline{x}, \overline{y} \in U_1 \Rightarrow \overline{x} + \lambda\overline{y} = (x_1 + \lambda y_1, x_2 + \lambda y_2, x_3 + \lambda y_3)$;
$(x_1 + \lambda y_1) + (x_2 + \lambda y_2) + (x_3 + \lambda y_3) = (x_1 + x_2 + x_3) + \lambda(y_1 + y_2 + y_3) = 1 + 1\lambda$
Como $1 + \lambda$ no tiene por qué ser $1 \Rightarrow U_1$ no es subespacio vectorial.

2. Si $\overline{x}, \overline{y} \in U_2 \Rightarrow \overline{x} + \lambda\overline{y} = (x_1 + \lambda y_1, x_2 + \lambda y_2, x_3 + \lambda y_3)$;
$(x_1 + \lambda y_1) + (x_2 + \lambda y_2) + (x_3 + \lambda y_3) = (x_1 + x_2 + x_3) + \lambda(y_1 + y_2 + y_3) = 0 + 0\lambda = 0$
Como $\overline{x} + \lambda\overline{y} \in U_2 \Rightarrow U_2$ es subespacio vectorial.

3. Si $\overline{x}, \overline{y} \in U_3 \Rightarrow \overline{x} + \lambda\overline{y} = (x_1 + \lambda y_1, x_2 + \lambda y_2, x_3 + \lambda y_3)$;
$(x_1 + \lambda y_1) + (x_2 + \lambda y_2) = (x_1 + x_2) + \lambda(y_1 + y_2)$, que no tiene por qué verificar la condición que caracteriza a este subconjunto, como ocurre en el ejemplo siguiente:

$$(-1, 0, 1) + 5(0, 1, 2) = (-1, 5, 11) \Rightarrow U_3 \text{ no es un subespacio vectorial.}$$

4. Si $\overline{x}, \overline{y} \in U_4 \Rightarrow \overline{x} + \lambda\overline{y} = (x_1 + \lambda y_1, x_2 + \lambda y_2, x_3 + \lambda y_3)$;
$(x_1 + \lambda y_1) + (x_2 + \lambda y_2) = x_1 x_2 + \lambda x_1 y_2 + \lambda x_2 y_1 + \lambda^2 y_1 y_2 = 1 + \lambda x_1 y_2 + \lambda x_2 y_1 + \lambda^2$
Como $1 + \lambda x_1 y_2 + \lambda x_2 y_1 + \lambda^2$ no tiene por qué ser $1 \Rightarrow U_4$ no es subespacio vectorial.

5. Si $\overline{x}, \overline{y} \in U_5 \Rightarrow \overline{x} + \lambda\overline{y} = (x_1 + \lambda y_1, x_2 + \lambda y_2, x_3 + \lambda y_3)$;
$(x_1 + \lambda y_1)^2 + (x_2 + \lambda y_2)^2 = x_1^2 + 2\lambda x_1 y_1 + \lambda^2 y_1^2 + x_2^2 + 2\lambda x_2 y_2 + \lambda^2 y_2^2$
Como la expresión anterior no tiene por qué ser ≤ 1, $\forall\overline{x}, \overline{y} \in U_5 \Rightarrow U_5$ no es subespacio vectorial.

- En los conjuntos U_1, U_4, no hacía falta haber hecho nada más que observar que (0, 0, 0) no pertenece a ellos para afirmar que no son espacios vectoriales.
- (0, 0, 0) sí pertenece a U_3, U_5, que tampoco son subespacios.
- La no pertenencia del $\overline{0}$ al subconjunto garantiza que no es subespacio vectorial, pero su pertenencia no garantiza que lo sea.

Ejercicio 1.5

En un almacén se dispone de tres tipos diferentes de abonos A, B, C, envasados en pastillas de 300 gramos. Cada pastilla contiene fósforo, potasio, nitrógeno, calcio y hierro. Las cantidades en gramos de cada componente por pastilla, según la clase, son: 100, 90, 60 de fósforo, 50, 50, 50 de potasio, 100, 105, 120 de nitrógeno, 50, 45, 30 de calcio y 0, 10, 40 de hierro.

a) Un agricultor necesita para un cultivo determinado, una mezcla que contenga 500 de fósforo, 100 de potasio, 500 de nitrógeno, 2.500 de calcio y 150 de hierro. ¿Se puede conseguir con los productos que hay en el almacén sin romper las pastillas? ¿Y rompiéndolas?

b) La exigencia de otro cultivo es una mezcla suministrada en bloques, tales que, cada uno contenga 290 g de fósforo, 150 g de potasio, 10 g de hierro y no importa la cantidad de nitrógeno y calcio. ¿Es posible hacer el compuesto pedido con los productos A, B, C?

c) Determínese la cantidad de cada componente por bloque de compuesto, el peso de cada bloque y la cantidad de pastillas de A, B, C que intervienen en él.

	P	K	N	Ca	Fe
Abono A	100	50	100	50	0
Abono B	90	50	105	45	10
Abono C	60	50	120	30	40

Solución:

a) Las cantidades de cada elemento en los distintos compuestos básicos se pueden representar mediante los vectores $\overline{a} = (100, 50, 100, 50, 0)$; $\overline{b} = (90, 50, 105, 45, 10)$; $\overline{c} = (60, 50, 120, 30, 40)$.

La mezcla pedida debe formarse tomando una cantidad de cada compuesto A, B, C, es decir, para que se pueda formar debe haber una combinación lineal de los vectores dados.

La habrá si existen coeficientes reales α, β, γ que verifiquen la igualdad: $(500, 100, 500, 2.500, 150) = \alpha\,(100, 50, 100, 50, 0) + \beta\,(90, 50, 105, 45, 10) + \gamma\,(60, 50, 120, 30, 40)$.

Para que dos vectores sean iguales, lo deben ser todos sus componentes, por tanto, el sistema siguiente debe tener solución.

$$\begin{aligned} 500 &= 100\alpha + 90\beta + 60\gamma; \\ 100 &= 50\alpha + 50\beta + 50\gamma; \\ 500 &= 100\alpha + 105\beta + 120\gamma; \\ 2.500 &= 50\alpha + 45\beta + 30\gamma; \\ 150 &= 0\alpha + 10\beta + 40\gamma; \end{aligned}$$

El sistema no tiene solución porque $100\alpha + 90\beta + 60\gamma = 2(50\alpha + 45\beta + 30\gamma)$, mientras que $500 \neq 2 \times 2500$ (2.ª y 4.ª ecuaciones).

No se puede formar el abono pedido, ni partiendo ni sin partir las pastillas.

b) Igual que en el apartado a), será posible hacer la mezcla si existen coeficientes reales α, β, γ que verifiquen la igualdad:

$(290, 150, u_3, u_4, 10) = \alpha\,(100, 50, 100, 50, 0) + \beta\,(90, 50, 105, 45, 10) + \gamma\,(60, 50, 120, 30, 40)$.

Es decir, si tiene solución el sistema siguiente:

$$\begin{aligned} 290 &= 100\alpha + 90\beta + 60\,\gamma; \\ 150 &= 50\alpha + 50\beta + 50\gamma; \\ u_3 &= 100\alpha + 105\beta + 120\gamma; \end{aligned}$$

$u_4 = 50\alpha + 45\beta + 30\gamma;$
$10 = 0\alpha + 10\beta + 40\gamma;$

El sistema formado por las ecuaciones 1.ª, 2.ª y 5.ª tiene como solución $\alpha = 2$, $\beta = 1$, $\gamma = 0$.

Para que tenga solución el sistema formado por las cinco ecuaciones, los valores $\alpha = 2$, $\beta = 1$, $\gamma = 0$ deben ser solución también de las ecuaciones 3.ª y 4.ª, y por tanto, $u_3 = 305$, $u_4 = 145$.

Como consecuencia, se puede formar la mezcla porque el vector (290, 150, u_3, u_4, 10) es combinación lineal de los vectores (100, 50, 100, 50, 0), (90, 50, 105, 45, 10), (60, 50, 120, 30, 40).

c) Los valores $u_3 = 305$, $u_4 = 145$, indican las cantidades de nitrógeno y calcio que lleva cada bloque, el vector que indica la cantidad de cada componente por bloque de compuesto que se va a suministrar al agricultor es: (290, 150, 305, 145, 10), siendo el peso de cada bloque 290 + 150 + 305 + 145 + + 10 = 900 g.

Los coeficientes $\alpha = 2$, $\beta = 1$, $\gamma = 0$, indican el número de pastillas de tipo A, B, C respectivamente que hay que tomar para formar cada bloque del compuesto pedido.

El peso del bloque también lo podíamos haber calculado sabiendo que cada bloque se forma con dos pastillas del tipo A y una pastilla del tipo B, de 300 g de peso cada una, es decir, $3 \times 300 = 900$ g.

Ejercicio 1.6

Sea $V = \{f: \mathbb{R} \to \mathbb{R}\}$. Se pide determinar si el conjunto $W = \{f(x) = e^{3x};\ g(x) = x^3;\ h(x) = x\}$ es libre, o ligado.

Solución:

Sea $\lambda e^{3x} + \beta x^3 + \gamma x = 0$, esta igualdad debe ser cierta para todos los valores de x.

Para cada uno de ellos obtendremos una ecuación:

Si $x = 0 \Rightarrow \quad \lambda = 0$
Si $x = 1 \Rightarrow \quad \beta + \gamma = 0$
Si $x = 2 \Rightarrow \quad 8\beta + 2\gamma = 0$

Del sistema de ecuaciones formado por $\beta + \gamma = 0$ y $8\beta + 2\gamma = 0$, se deduce que $\beta = \gamma = 0$, es decir, es un sistema libre o linealmente independiente.

Ejercicio 1.7

Si P_2 es el espacio vectorial real de los polinomios de grado menor que 3. Se pide:

a) Decidir razonadamente si $\{x^2, x^2 + x, x + 1\}$ es una base de dicho espacio vectorial.

b) Calcular las coordenadas de $5x^2 + 3x + 1$ respecto a la base $\{x^2, x^2 + x, x + 1\}$.

c) Encontrar una base del espacio vectorial real de los polinomios de grado menor o igual que 3, que contenga a la base dada.

d) Calcular las coordenadas de $5x^2 + 3x + 1$ respecto a la base encontrada en el apartado anterior.

Solución:

a) Cualquier base B del conjunto P_2 $\{ax^2 + bx + c \,/\, a, b, c \in \mathbb{R}\}$ debe ser un sistema libre de vectores de P_2, tal que, todo polinomio de grado menor o igual que tres, se pueda escribir como combinación lineal de los vectores de B.

Vamos a comprobar si $\{x^2, x^2 + x, x + 1\}$ cumple estas dos condiciones:

a1) $\{x^2, x^2 + x, x + 1\}$ es un sistema libre:

$\alpha x^2 + \beta(x^2 + x) + \delta(x + 1) = 0 \Rightarrow \alpha = 0, \beta = 0, \delta = 0 \Rightarrow \{x^2, x^2 + x, x + 1\}$ es un sistema libre.

a2) $\{x^2, x^2 + x, x + 1\}$ es un sistema generador de P_2:

Cualquier elemento de P_2 es de la forma $ax^2 + bx + c$; para que $\{x^2, x^2 + x, x + 1\}$ sea generador la ecuación $ax^2 + bx + c = \alpha x^2 + \beta(x^2 + x) + \delta(x + 1)$ debe tener soluciones reales para α, β, δ.

En efecto, $ax^2 + bx + c = \alpha x^2 + \beta(x^2 + x) + \delta(x + 1) = x^2(\alpha + \beta) + x(\beta + \delta) + \delta \Rightarrow \alpha = a - b + c; \Rightarrow \beta = b - c; \delta = c$; α, β, δ existen y son reales para cualquier polinomio de P_2, porque existen y son reales, a, b, c.

b) En $5x^2 + 3x + 1$, los coeficientes son: $a = 5; b = 3; c = 1$
$\Rightarrow \alpha = 5 - 3 + 1 = 3; \beta = 3 - 1 = 2; c = 1 \Rightarrow 5x^2 + 3x + 1 = 3x^2 + 2(x^2 + x) + 1(x + 1)$.

Las coordenadas de $5x^2 + 3x + 1$ respecto a la base dada son $(3, 2, 1)$.

c) Para que $\{x^2, x^2 + x, x + 1\}$ sea base de P_3, nos falta añadir x^3; el conjunto $\{x^3, x^2, x^2 + x, x + 1\}$ puede generar P_3, y los vectores que lo forman son un conjunto de vectores linealmente independientes. Hemos añadido x^3, pero hay otras muchas posibilidades como por ejemplo, $x^3 - x^2, x^3 - x^2 + x, x^2 + 1, \ldots$

d) El polinomio $5x^2 + 3x + 1$ también pertenece a P_3, por tanto, se puede expresar como combinación lineal de los elementos de cualquiera de sus bases.

En particular, $5x^2 + 3x + 1 = \alpha x^3 + \beta x^2 + \delta(x^2 + x) + \gamma(x + 1) = \alpha x^3 + (\beta + \delta)x^2 + (\delta + \gamma)x + \gamma \Rightarrow \alpha = 0, \beta = 3, \delta = 2, \gamma = 1$.

Las coordenadas de $5x^2 + 3x + 1$ respecto a la base $\{x^3, x^2, x^2 + x, x + 1\}$ son $(0, 3, 2, 1)$.

Ejercicio 1.8

Sea $S = \{(2, -1, 2, 3), (1, 3, 2, 1)\} \subset \mathbb{R}^4$. Se pide:

a) Escribir unas ecuaciones paramétricas de $\langle S \rangle$;

b) Determinar si el vector $(11, -2, 12, 16)$ pertenece a $\langle S \rangle$, y si es así, encontrar la combinación lineal que lo genera;

c) Determinar el valor de los parámetros m y n para que el vector $\overline{u} = (1, m, 4, n)$ pertenezca al subespacio de $\mathbb{R}^4$ generado por los vectores, $\overline{v} = (2, -1, 2, 3)$ y $\overline{w} = (1, 3, 2, 1)$.

Solución:

a) Ecuaciones paramétricas:

$$\langle S \rangle = \{\lambda(2, -1, 2, 3) + \mu(1, 3, 2, 1) / \lambda, \mu \in \mathbb{R}^4\} \Rightarrow \begin{cases} x_1 = 2\lambda + \mu \\ x_2 = -\lambda + 3\mu \\ x_3 = 2\lambda + 2\mu \\ x_4 = 3\lambda + \mu \end{cases}$$

b) Para que $(11, -2, 12, 16)$ pertenezca a $\langle S \rangle$, deben existir soluciones en el sistema formado por las ecuaciones paramétricas:

$$\begin{cases} 11 = 2\lambda + \mu \\ -2 = -\lambda + 3\mu \\ 12 = 2\lambda + 2\mu \\ 16 = 3\lambda + \mu \end{cases}$$

Los valores de λ, μ obtenidos como solución de las dos primeras ecuaciones son $\lambda = 5, \mu = 1$; que convierten en identidades las dos últimas ecuaciones, por tanto, $(11, -2, 12, 16) \in \langle S \rangle$, y la combinación lineal que lo genera es: $(11, -2, 12, 16) = 5(2, -1, 2, 3) + 1(1, 3, 2, 1)$.

c) Del mismo modo que en el apartado anterior, para que $(1, m, 4, n)$ pertenezca a $\langle S \rangle$, deben existir soluciones en el sistema de ecuaciones formado por las ecuaciones paramétricas:

$$\begin{cases} 1 = 2\lambda + \mu \\ m = -\lambda + 3\mu \\ 4 = 2\lambda + 2\mu \\ n = 3\lambda + \mu \end{cases}$$

Los valores de λ, μ obtenidos con solución de las ecuaciones primera y tercera son $\lambda = -1, \mu = 3$; que al ser sustituidos en las dos ecuaciones restantes, proporcionan los valores de los parámetros $m = 10, n = 0$.

Ejercicio 1.9

Sean $L_1 = \{(x_1, x_2, x_3, x_4)\} \in \mathbb{R}^4 / x_1 = x_2 = x_3 = x_4\}$, $L_2 = \{(x_1, x_2, x_3, x_4)\} \in \in \mathbb{R}^4 / x_1 = x_2 = x_3 = x_4\}$, $L_3 = \{(x_1, x_2, x_3, x_4)\} \in \mathbb{R}^4 / x_2 = 2x_4, x_4 = 2x_3\}$ subespacios vectoriales de $\mathbb{R}^4$. Se pide:

a) Una base y unas ecuaciones paramétricas de $L_1 \cap L_2$, $L_2 \cap L_3$, $L_1 \cap L_3$.

b) Indicar cuáles de las siguientes sumas son directas y dar sus ecuaciones. $L_1 + L_2$, $L_2 + L_3$ y $L_1 + L_3$.

Solución:

Los vectores de $L_1 \cap L_2$ pertenecen a L_1 y a L_2, por tanto, deben verificar las ecuaciones de ambos.

Cualquier vector que cumpla las condiciones $\begin{cases} x_1 = x_2 \\ x_2 = x_3, \\ x_3 = x_4 \end{cases}$ es evidente que,

también cumple $\begin{cases} x_1 = x_3 \\ x_2 = x_4 \end{cases}$, es decir, $L_1 \cap L_2 \Rightarrow L_1 \cap L_2 = L_1$.

Las ecuaciones cartesianas de L_1, son: $\begin{cases} x_1 = x_2 \\ x_2 = x_3 \\ x_3 = x_4 \end{cases}$

Unas ecuaciones paramétricas son: $\begin{cases} x_1 = \alpha \\ x_2 = \alpha \\ x_3 = \alpha \\ x_4 = \alpha \end{cases} \Leftrightarrow \begin{pmatrix} x_2 \\ x_2 \\ x_2 \\ x_4 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

Una base de $L_1 \cap L_2$ está formada por el vector (1, 1, 1, 1).

$L_1 + L_2$ no es suma directa porque $L_1 \cap L_2 \neq \overline{0}$.

Cálculo de $L_1 + L_2$:

Conocemos una base de L_1, siguiendo el mismo proceso encontramos una base de L_2:

$$\begin{cases} x_1 = \delta \\ x_2 = \beta \\ x_3 = \delta \\ x_4 = \beta \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \delta \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Los vectores (1, 0, 1, 0) y (0, 1, 0, 1) son generadores de L_2 y son linealmente independientes, por tanto, son una base de L_2.

Cualquier vector $\overline{v}$ de $L_1 + L_2$ es suma de un vector $\overline{v}_1 \in L_1$ y un vector de $\overline{v}_2 \in \in L_2$; $\overline{v}_1$ está generado por los vectores de una base de L_1, y $\overline{v}_2$ por los vectores de una base de L_2.

Es decir, es de la forma: $\overline{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \delta \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

Los tres vectores generadores no son linealmente independientes, (1, 1, 1, 1) = (1, 0, 1, 0) + (0, 1, 0, 1), por tanto (1, 1, 1, 1) se puede eliminar del sistema de generadores de $L_1 + L_2$ sin que varíe el subespacio generado, que es L_2, como ya sabíamos porque $L_1 \subset L_2$.

Buscamos los vectores que verifican las ecuaciones $\begin{cases} x_1 = x_3 \\ x_2 = x_4 \end{cases}$ y $\begin{cases} x_2 = 2x_4 \\ x_4 = 2x_3 \end{cases}$

la única solución posible del sistema formado por las cuatro ecuaciones es el vector (0, 0, 0, 0), por tanto, $L_2 \cap L_3$ es el subespacio impropio de $\mathbb{R}^4 : \overline{0}$.

$L_2 + L_3$ es suma directa porque $L_2 \cap L_3 = \overline{0}$.

Cálculo de $L_2 \oplus L_3$:

Base de $L_2 = \{(1, 0, 1, 0), (0, 1, 0, 1)\}$

Base de $L_3 = \{(1, 0, 0, 0), (0, 4, 1, 2)\}$

Si $\overline{v} \in L_2 + L_3 \Rightarrow \overline{v} = \overline{v}_2 + \overline{v}_3 / \overline{v}_2 \in L_2$ y $\overline{v}_3 \in L_3$.

Es decir, es de la forma: $\overline{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \delta \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 4 \\ 1 \\ 2 \end{pmatrix}$

Los cuatro vectores que forman el sistema generador son linealmente independientes, por tanto, son base de $\mathbb{R}^4 \Rightarrow L_2 \oplus L_3 = \mathbb{R}^4$.

Cualquier vector de $\mathbb{R}^4$ se puede obtener como suma de un vector de L_2 y otro de L_3.

$L_1 \cap L_3$ está formado por los vectores que verifican las ecuaciones $\begin{cases} x_1 = x_2 \\ x_2 = x_3 \\ x_3 = x_4 \end{cases}$ y

$\begin{cases} x_2 = 2x_4 \\ x_4 = 2x_3 \end{cases}$

$\Rightarrow L_1 \cap L_3 \{(0, 0, 0, 0)\} \Rightarrow L_1 + L_3$ es suma directa.

Cálculo de $L_1 \oplus L_3$:

Base de $L_1 = \{(1, 1, 1, 1)\}$

Base de $L_3 = \{(1, 0, 0, 0), (0, 4, 1, 2)\}$

Si $\overline{v} \in L_1 \oplus L_3 \Rightarrow \overline{v} = \overline{v}_1 + \overline{v}_3 / \overline{v}_1 \in L_1$ y $\overline{v}_3 \in L_3$.

Es decir, es de la forma: $\overline{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} 0 \\ 4 \\ 1 \\ 2 \end{pmatrix}$

Unas ecuaciones paramétricas de $L_1 \oplus L_3$ son: $\begin{cases} x_1 = \alpha + \beta \\ x_2 = \alpha + 4\delta \\ x_3 = \alpha + \delta \\ x_4 = \alpha + 2\delta \end{cases}$, que eliminando los parámetros dan lugar a las ecuaciones cartesinas $x_2 + 2x_3 - 3x_4 = 0$.

Ejercicio 1.10

Sean $A = \{(1, 2, 1, 1), (-1, 0, 1, 0), (2, 2, 0, 1)\}$ y $B = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 / / 2x_1 - 2x_2 + 2x_3 - x_4 = 0\}$.

Se pide:

a) Obtener una base, la dimensión y ecuaciones de ambos subespacios de $\mathbb{R}^4$.

b) Determinar si $A + B$ es suma directa, y escribir sus ecuaciones paramétricas.

c) Obtener el conjunto complementario de A en $\mathbb{R}^4$.

Solución:

a) El sistema $\{(1, 2, 1, 1), (-1, 0, 1, 0), (2, 2, 0, 1)\}$ no es libre, ya que:

$(1, 2, 1, 1) - (-1, 0, 1, 0) = (2, 2, 0, 1)$; al quitar el vector dependiente queda el sistema $\{(1, 2, 1, 1), (-1, 0, 1, 0)\}$, que es independiente y forma una base del subespacio A, cuya dimensión es dos.

Las ecuaciones paramétricas son: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, y las cartesianas $\begin{cases} x_1 + x_3 - 2x_4 = 0 \\ x_2 - 2x_4 = 0 \end{cases}$

El subespacio B de $\mathbb{R}^4$ cuyas ecuaciones cartesianas son: $2x_1 - 2x_2 + 2x_3 - x_4 = 0$ tiene por ecuaciones paramétricas:

$\begin{cases} x_1 = \alpha \\ x_2 = \beta \\ x_3 = \delta \\ x_4 = 2\alpha - 2\beta + 2\delta \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ -2 \end{pmatrix} + \delta \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}$, los vectores $(1, 0, 0, 2)$, $(0, 1, 0, -2)$, $(0, 0, 1, 2)$ son un sistema libre, por tanto una base de B, y la dimensión de este subespacio es tres.

b) $A+B$ no puede ser directa porque A y B son subespacios de $\mathbb{R}^4$, si la suma fuera directa, la dimensión de $A \cap B$ sería 0, y como $A + B$ también es un subespacio de $\mathbb{R}^4$, el mayor valor que puede tener su dimensión es cuatro, y no podría cumplirse la igualdad:

$$\text{Dim}\,(A + B) + \text{Dim}\,(A \cap B) = \text{Dim}\,A + \text{Dim}\,B.$$

$A + B$ es subespacio de $\mathbb{R}^4$, $\forall \overline{v} \in A + B \Rightarrow$

$$\overline{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 1 \\ 0 \\ -2 \end{pmatrix} + \eta \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

Pero de entre los cinco vectores del sistema de generadores, hay sólo tres independientes como puede comprobar el lector por cualquiera de los procedimientos que conoce.

Podemos escribir $\overline{v} = \delta \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 1 \\ 0 \\ -2 \end{pmatrix} + \eta \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 = \delta \\ x_2 = \gamma \\ x_3 = \eta \\ x_4 = 2\delta - 2\gamma + 2\eta \end{cases}$

que son las ecuaciones paramétricas de $A + B$.

c) Para encontrar el subespacio suplementario de A, basta con:

Ampliar la base $\{(1, 2, 1, 1), (-1, 0, 1, 0)\}$ de A hasta obtener una base de $\mathbb{R}^4$, lo conseguimos con dos vectores independientes de ellos, por ejemplo: $\{(0, 0, 1, 0), (0, 0, 0, 1)\}$; el subespacio generado por $\{(0, 0, 1, 0), (0, 0, 0, 1)\}$ es el suplementario de A.

Sus ecuaciones paramétricas son: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = \alpha \\ x_4 = \beta \end{cases}$

Sus ecuaciones cartesianas son: $\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$

Aunque al terminar el capítulo uno el lector dispone de herramientas para ver la dependencia e independencia lineal, después de estudiar los capítulos dos y tres, le resultarán más fácil hacerlo, por lo que se recomienda una revisión cuando haya leído dichos capítulos.

CAPÍTULO

2 APLICACIONES LINEALES

INTRODUCCIÓN

Aunque los cambios de coordenadas cartesianas conducen a cálculos idénticos a los de la teoría de transformaciones lineales, éstas aparecen más claramente (bajo el nombre de sustitución lineal) en los trabajos sobre clasificación y reducción de formas cuadráticas de coeficientes enteros, de dos variables en Lagrange y de tres variables en Gauss.

Gauss para representar la sustitución lineal, utiliza por primera vez una notación matricial que designa con un único símbolo S, hace notar también que si después de una sustitución S se realiza otra U, el efecto es el mismo que realizar una sustitución cuya tabla de coeficientes es precisamente la correspondiente a lo que después se llamaría el producto de las matrices de U y S. También comentan brevemente que el resultado sería cierto para cualquier número de variables.

Einstein, en 1844, utiliza el mismo simbolismo en sus trabajos sobre teoría de números, y destaca que hay que distinguir claramente entre SU y US.

Es Cayley, en 1855, quien al estudiar el comportamiento de los invariantes bajo transformaciones lineales, introduce la notación matricial para representar las transformaciones con toda generalidad al objeto de simplificar la notación.

El nombre de matriz para designar una tabla rectangular de números, se debe a la fértil imaginación de Sylvester, llama así a una tabla "oblonga de elementos porque es la madre de la que se pueden extraer tablas cuadradas para hacer sus determinantes". En 1858 publica *A Memoir on the Theory of Matrices*, que se considera el trabajo definitivo que dio origen a la teoría de matrices. En él se define la adición y el producto por escalares. La definición de producto de matrices está motivada (como en el caso de Gauss) por su interpretación como representación analítica de una transformación lineal y la imposición de que la matriz producto represente la transformación composición. Cayley prueba todas las propiedades algebraicas de las operaciones con matrices. Define el polinomio característico de una matriz cuadrada A como y enuncia el llamado Teorema de Hamilton-Cayley, aunque sólo lo demuestra para $n = 2$ y $n = 3$.

SYLVESTER (1814-1897)

Empezó sus trabajos como matemático resolviendo un difícil problema de cálculo de probabilidades para una empresa de lotería.

Su origen judío puso dificultades para que siguiera el camino de la ciencia en Inglaterra, lo que le impulsó a emigrar a EE.UU.

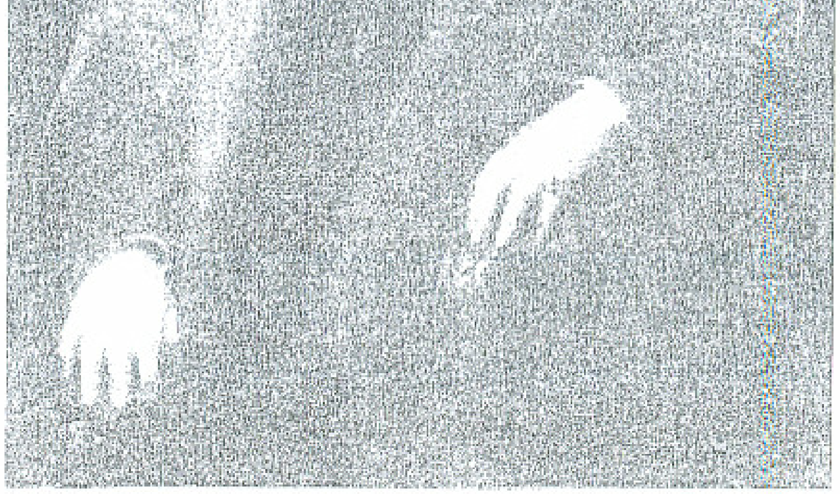
Tampoco allí se encontró a gusto, volvió a Inglaterra y se dedicó a la matemática aplicada desde su trabajo como actuario de seguros; más tarde abandonó este trabajo para dedicarse a la abogacía, actividad que volvió a abandonar para dedicarse a la matemática pura.

En 1850 conoció a Cayley y su relación cristalizó en la creación de la teoría de invariantes algebraicos y de las matrices, mientras era catedrático en Woolwich (Inglaterra).

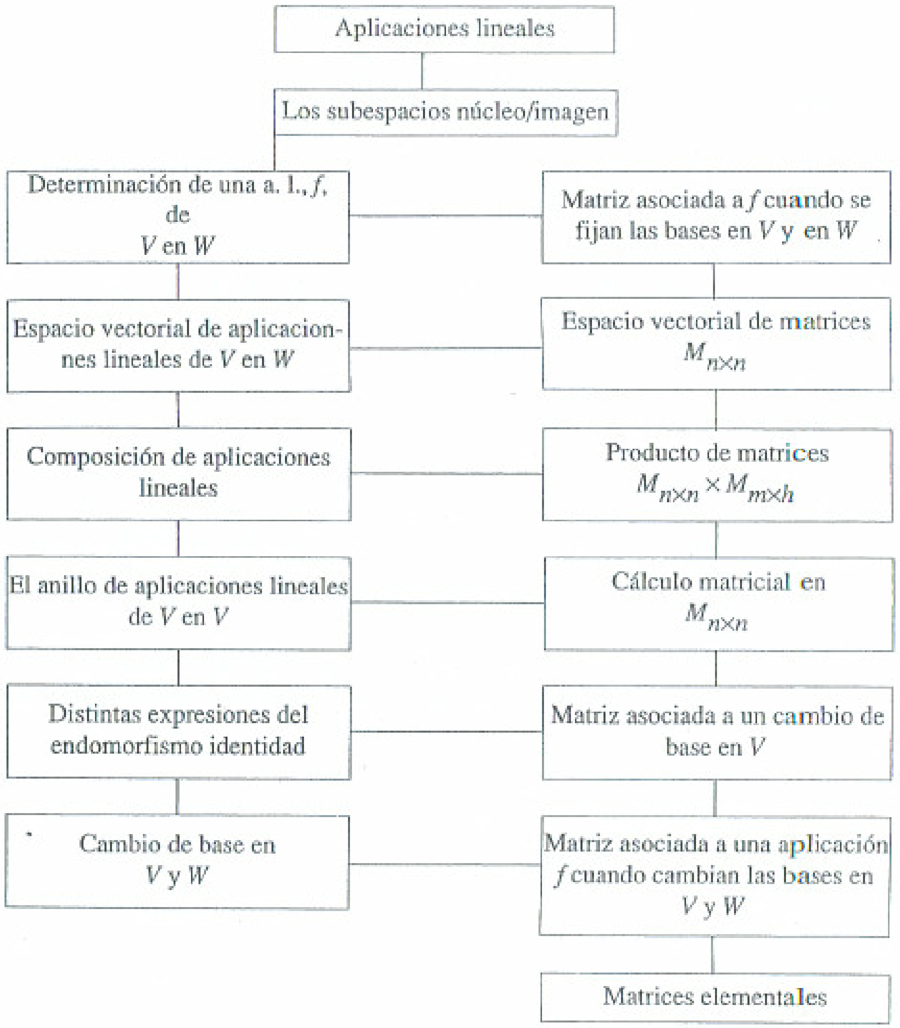
En 1876 fue como catedrático a la prestigiosa Universidad de John Hopleins de Baltimore (Estados Unidos).

Allí defendió la belleza de la abstracción matemática y provocó una revolución en la enseñanza de las matemáticas.

Su agitada vida puede resumirse en sus propias palabras: "*amo realmente mi ocupación*".



CAPÍTULO 2


Aplicaciones lineales
Los subespacios núcleo/imagen
Determinación de una a. l., f, de V en W
Matriz asociada a f cuando se fijan las bases en V y en W
Espacio vectorial de aplicaciones lineales de V en W
Espacio vectorial de matrices $M_{n\times n}$
Composición de aplicaciones lineales
Producto de matrices $M_{n\times n} \times M_{m\times h}$
El anillo de aplicaciones lineales de V en V
Cálculo matricial en $M_{n\times n}$
Distintas expresiones del endomorfismo identidad
Matriz asociada a un cambio de base en V
Cambio de base en V y W
Matriz asociada a una aplicación f cuando cambian las bases en V y W
Matrices elementales

2.1 Aplicaciones lineales

Al definir una aplicación entre dos espacios vectoriales, estamos utilizando muchos conceptos simultáneamente. Vamos a analizarlos sobre un ejemplo concreto.

Ejemplo 2.1.1

"Sea $f: (\mathrm{M}, +, \mathbb{R}) \to (\mathbb{R}^4, +, \mathbb{R})$ una aplicación definida como:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} \to (a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0)".$$

Tenemos:

- Estructura algebraica inicial: $(\mathrm{M}, +, \mathbb{R})$ es el espacio vectorial de las matrices de la forma $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix}$, entre las que hay definida la operación de suma estándar, y la operación de multiplicar estándar por un escalar.
- Estructura algebraica final: $(\mathbb{R}^4, +, \mathbb{R})$ es el espacio vectorial de los vectores de $\mathbb{R}^4$, entre los que hay definida la operación de suma estándar y la operación de multiplicar estándar por una escalar.
- Una aplicación f que transforma cada matriz $\begin{pmatrix} a_{11} & a_{12} \\ a_{11} & 0 \end{pmatrix} \in \mathrm{M}$ en el vector

$(a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0) \in \mathbb{R}^4$.

Es evidente que f es aplicación, porque para cada original la imagen es única.

- Además, es una aplicación especial que conserva la suma:

$$f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} = (a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0)$$

$$f\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{pmatrix} = (b_{11} + b_{12}, b_{12} - b_{11}, b_{21}, 0)$$

$$f\begin{pmatrix} a_{11} + b_{12} & a_{12} + b_{12} \\ a_{21} + b_{21} & 0 \end{pmatrix} = (a_{11} + b_{11} + a_{12} + b_{12}, a_{12} + b_{12} - a_{11} - b_{11}, a_{21} + b_{21}, 0)$$

$$f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} + f\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{pmatrix} = (a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0) + (b_{11} + b_{12}, b_{12} - b_{11}, b_{21}, 0)$$

$$\boxed{f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} + f\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{pmatrix} = f\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{pmatrix}\right)}$$

- También conserva el producto por escalares:

$$f\left(\lambda\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix}\right) = f\begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & 0 \end{pmatrix} = (\lambda a_{11} + \lambda a_{12}, \lambda a_{12} - \lambda a_{11}, \lambda a_{21}, 0)$$

$$\lambda f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} = \lambda\,(a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0)$$

$$\boxed{f\left(\lambda\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix}\right) = \lambda f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix}}$$

Podemos decir que, "Mediante la aplicación f, la imagen de la suma de matrices es la suma de las imágenes de dichas matrices, y la imagen del producto de una matriz por un escalar, es el escalar por la imagen de la matriz".

Las aplicaciones que, como ésta, tienen la propiedad de conservar la estructura de espacio vectorial porque la imagen de la suma es la suma de las imágenes, y la imagen del producto de un original por un escalar, es el escalar por la imagen del original, son de un tipo especial y reciben el nombre de aplicaciones lineales o transformaciones lineales

2.1.1. Definición de aplicación lineal

$f: V \to W$ es una aplicación lineal entre los espacios vectoriales V, W, si se verifica: $f(\lambda\overline{u} + \mu\overline{v}) = \lambda f(\overline{u}) + \mu f(\overline{v}) + \mu f(\overline{v})$; $\forall \overline{u}, \overline{v} \in V$; $\forall \lambda, \mu \in \mathbb{R}$

Las aplicaciones, lineales o no, se usan continuamente en Informática, citaremos algunos ejemplos claros:

- Cada dirección en la memoria se considera como un vector y se le asigna su posición mediante una aplicación no lineal.
- En una base de datos se relacionan (se establece una aplicación) varios elementos de diversos campos (conjuntos sobre los que se define la aplicación).
- En la programación orientada a objetos, los conjuntos y subconjuntos son las clases y las aplicaciones la relación entre clases.

De entre las aplicaciones lineales, algunas nos resultan conocidas, es el caso geométrico de las proyecciones que se ilustra en el ejemplo siguiente.

Ejemplo 2.1.2.

Compruébese que $f_1: \mathbb{R}^3 \to \mathbb{R}^3 / f_1(x_1, x_2, x_3) = (x_1, x_2, 0)$ es una aplicación lineal.

Veamos que verifica la condición $f_1(\lambda\overline{x} + \mu\overline{y}) = \lambda f_1(\overline{x}) + \mu f_1(\overline{y})$.

En efecto, al desarrollar ambos miembros obtenemos el mismo resultado.

$f_1(\lambda(x_1, x_2, x_3) + \mu(y_1, y_2, y_3)) = f_1(\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, \lambda x_3 + \mu y_3) = (\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, 0)$.

$f_1(\lambda(x_1, x_2, x_3)) + f_1(\mu(y_1, y_2, y_3)) = f_1(\lambda x_1, \lambda x_2, \lambda x_3) + f_1(\mu y_1, \mu y_2, \mu y_3) = (\lambda x_1, \lambda x_2, 0) + (\mu y_1, \mu y_2, 0) = (\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, 0)$.

La imagen de (2, 3, 7) es $f_1(2, 3, 7) = (2, 3, 0)$, y geométricamente la identificamos con la proyección o sombra del vector (2, 3, 7) sobre el plano $x_3 = 0$.

Para $f_2: \mathbb{R}^3 \to \mathbb{R}^3 / f_2(x_1, x_2, x_3) = (x_1, 0, x_3)$ se hace exactamente igual, y es un pequeño ejercicio para que el lector compruebe si ha entendido la definición de aplicación lineal. Esta aplicación es la proyección sobre el plano $y = 0$.

Sabemos que si $V = U_1 \oplus U_2$, cualquier vector de V se puede escribir de forma única como $\overline{v} = \overline{u}_1 + \overline{u}_2$, siendo $\overline{u}_1 \in U_1$, $\overline{u}_2 \in U_2$.

Las aplicaciones $p_1: V \to U_1 / p_1(\overline{v}) = \overline{u}_1$ y $p_2: V \to U_2 / p_2(\overline{v}) = \overline{u}_2$, se llaman proyecciones de V sobre los subespacios U_1 y U_2.

Ejemplo 2.1.3

La aplicación $f: \mathbb{R} \to \mathbb{R} / f(x) = \text{sen}\, x$ no es una aplicación lineal.
No conserva la suma:

$x \to f(x) = \text{sen}\, x$
$y \to f(y) = \text{sen}\, y$
$x + y \to \text{sen}\,(x + y) = \text{sen}\, x \cos y + \text{sen}\, y \cos x \neq \text{sen}\, x + \text{sen}\, y$

Tampoco conserva el producto por escalares:

$$\text{sen}(10.\frac{\pi}{2}) \neq 10.\text{sen}(\frac{\pi}{2});$$

$$\text{sen}(10.\frac{\pi}{2}) = \text{sen}(5\pi) = \text{sen}(\pi) = 0 \qquad 10.\text{sen}(\frac{\pi}{2}) = 10.1 = 10$$

La aplicación $i: V \to V$ que transforma cada vector $\overline{u} \in V$ en el mismo vector $\overline{u}$, se llama aplicación identidad, y es evidente que es una aplicación lineal.

La aplicación $0: V \to W$ que transforma cada vector $\overline{u} \in V$, en el vector $\overline{0} \in W$, se llama aplicación nula, y también es evidente que es una aplicación lineal.

2.1.2. Consecuencias

1. Para cualquier aplicación lineal $f : V \to W$, se verifica que $f(\overline{0}) = \overline{0}$

Demostración:

Sabemos que $\overline{0} = 0\overline{v}$, por ser f una aplicación lineal $f(\overline{0}) = 0f(\overline{v}) = \overline{0}$

2. Para cualquier aplicación lineal $f : V \to W$, se verifica que $f(-\overline{v}) = -f(\overline{v})$.

Demostración:

$f(\overline{v}) + f(-\overline{v}) = f(\overline{v} + (-\overline{v})) = f(\overline{0}) = \overline{0} \Rightarrow f(-\overline{v}) = -f(\overline{v})$

3. Si dos vectores son linealmente dependientes, sus imágenes mediante una aplicación lineal también lo son.

Demostración:

Sea $\{\overline{u}_1, \overline{u}_2\}$ un sistema ligado $\Rightarrow$ Si $\lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 = \overline{0}$ hay algún $\lambda_i \neq 0$
$f(\lambda_1\overline{u}_1 + \lambda_2\overline{u}_2) = f(\overline{0}) = \overline{0}$
$f(\lambda_1\overline{u}_1 + \lambda_2\overline{u}_2) = \lambda_1 f(\overline{u}_1) + \lambda_2 f(\overline{u}_2) = \overline{0}$ con algún $\lambda_i \neq 0$, por tanto $\{f(\overline{u}_1), f(\overline{u}_2)\}$ es un sistema ligado.

4. Si dos vectores son linealmente independientes, sus imágenes mediante una aplicación lineal pueden no serlo.

Demostración:

Sea $f : \mathbb{R}^2 \to \mathbb{R} \,/\, f(x_1, x_2) = x_1$ una aplicación lineal (es muy fácil comprobarlo),
$\{(2, 1), (1, 3)\}$ es un sistema libre en $\mathbb{R}^2$
$f(2, 1) = 2; f(1, 3) = 1$
$\{2, 1\}$ es un sistema ligado en $\mathbb{R}$.

5. La composición de dos aplicaciones lineales es una aplicación lineal.

Demostración:

Sean f y g tales que $V \xrightarrow{f} W \xrightarrow{g} Y$, si ambas son aplicaciones lineales, podemos escribir:

$$\begin{aligned}(g \circ f)(\lambda\overline{u} + \mu\overline{v}) &= g(f(\lambda\overline{u} + \mu\overline{v})) = g(\lambda f(\overline{u}) + \mu f(\overline{v})) = \\ &= \lambda(g(f(\overline{u})) + \mu(g(f(\overline{v})) = \lambda(g \circ f)(\overline{u}) + \mu(g \circ f)(\overline{v})\end{aligned}$$

La igualdad anterior muestra que $(g \circ f)$ es una aplicación lineal.

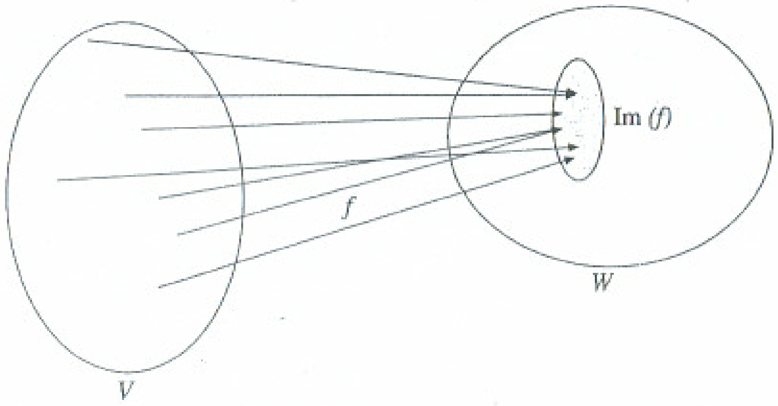
2.2. Subespacios distinguidos: núcleo e imagen

En cualquier aplicación entre espacios vectoriales hay dos subespacios que tienen un interés especial: núcleo e imagen de la aplicación.

La imagen está constituida por todos los vectores de W que tiene un original en V.

2.2.1. Definición de imagen de una aplicación lineal

> La imagen de la aplicación lineal $f: V \to W$ se define como:
> $\mathrm{Im}(f) = \{\overline{w} \in W$, tales que $\exists \overline{v} \in V$ que verifica $\overline{w} = f(\overline{v})\}$



2.2.2. Proposición

> La imagen de una aplicación lineal $f: V \to W$ es un subespacio de W.

Demostración:

Vamos a comprobar que $\mathrm{Im}(f)$ verifica la condición que caracteriza un subespacio.

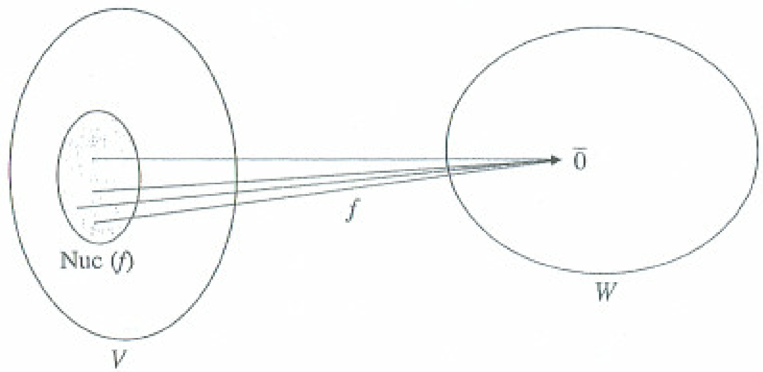
Aplicando la definición de imagen, si $\overline{w}_1, \overline{w}_2 \in \mathrm{Im}(f) \Rightarrow \exists \overline{v}_1, \overline{v}_2 \in V$, tales que $f(\overline{v}_1) = \overline{w}_1, f(\overline{v}_2) = \overline{w}_2$, por tanto, $f(\overline{v}_1 + \overline{v}_2) = f(\overline{v}_1) + (\overline{v}_2) = \overline{w}_1 + \overline{w}_2$, como $\overline{v}_1 + \overline{v}_2 \in V$ por ser V un espacio vectorial $\Rightarrow \overline{w}_1 + \overline{w}_2 \in \mathrm{Im}(f)$.

También se verifica $\forall \lambda \in \mathbb{R}, f(\lambda \overline{v}) = \lambda f(\overline{v}) = \lambda \overline{w}$, como $\lambda \overline{w} \in \mathrm{Im}(f)$ porque $\lambda \overline{v} \in V \Rightarrow \mathrm{Im}(f)$ es un subespacio de W.

El otro subespacio importante es el subconjunto de V cuyos elementos se transforman en el vector neutro de W.

2.2.3. Definición de núcleo de una aplicación lineal

> Núcleo de la aplicación lineal $f: V \to W$ es el conjunto
> $\mathrm{Nuc}(f) = \{\overline{v} \in V$, tales que $f(\overline{v}) = \overline{0} \in W\}$



2.2.4. Proposición

El núcleo de una aplicación lineal $f: V \to W$ es un subespacio de V

Demostración:

Vamos a comprobar que Nuc(f) verifica la condición que caracteriza un subespacio.

Si $\overline{v}_1, \overline{v}_2 \in$ Nuc(f) $\Rightarrow f(\overline{v}_1 + \overline{v}_2) = f(\overline{v}_1) + f(\overline{v}_2)$ por ser f una aplicación lineal; como $f(\overline{v}_1) + f(\overline{v}_2) = \overline{0} + \overline{0} = \overline{0} \in W \Rightarrow \overline{v}_1 + \overline{v}_2 \in$ Nuc(f).

También se verifica: $\forall \lambda \in \mathbb{R}$, $\forall \overline{v} \in$ Nuc(f) $\Rightarrow f(\lambda\overline{v}) = \lambda f(\overline{v}) = \overline{0} \Rightarrow \lambda\overline{v} \in$ Nuc(f), por tanto, Nuc(f) es un subespacio de V.

Ejemplo 2.2.1

"Sea f:(M, +, $\mathbb{R}$) $\to$ ($\mathbb{R}^4$, +, $\mathbb{R}$) la aplicación lineal definida como $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} \to$

$\to (a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0)$ en el ejemplo 2.1.1"

La imagen de la aplicación lineal estará formada por el subconjunto de $\mathbb{R}^4$, cuyos elementos sean de la forma $(x_1, x_2, x_3, x_4) = (a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0)$.

$$\left.\begin{array}{l} x_1 = a_{11} + a_{12} \\ x_2 = a_{12} - a_{11} \\ x_3 = a_{21} \\ x_4 = 0 \end{array}\right\} \Rightarrow \boxed{\mathrm{Im}(f) = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 / x_4 = 0\}}$$

El núcleo de la aplicación lineal estará formado por el subconjunto de M cuya imagen sea $(0, 0, 0, 0) = \overline{0} \in \mathbb{R}^4$, es decir, los elementos $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} \in$ M que cumplan $f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} = (a_{11} + a_{12}, a_{12} - a_{11}, a_{21}, 0) = (0, 0, 0, 0)$

$$\left.\begin{array}{r} a_{11} + a_{12} = 0 \\ a_{12} - a_{11} = 0 \\ a_{21} = 0 \end{array}\right\} \Rightarrow a_{11} = a_{12} = 0 \Rightarrow \boxed{\mathrm{Nuc}(f) = f^{-1}(0, 0, 0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}$$

2.2.5. Teorema

Si V es de dimensión finita (dim $V = n$), entonces, se verifica:
dim Nuc (f) + dim Im (f) = dim V

Demostración:

El núcleo es un subespacio vectorial de V (V es un espacio finito), por tanto, tiene una base finita. Llamemos B a esta base, los vectores de B son linealmente independientes, y se pueden completar con otro sistema de vectores linealmente independientes hasta obtener una base $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_p, \overline{u}_1, ..., \overline{u}_q\}$ de V.

Hemos tomado B como base del núcleo, cuya dimensión es p.

Nos falta ver que $\{f(\overline{u}_1), ..., f(\overline{u}_q)\}$ es una base de la imagen de f.

Veamos primero que es generador.

Por la tercera conclusión de *[2.1.2]*, sabemos que si $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_p, \overline{u}_1, ..., \overline{u}_q\}$ es generador de $V \Rightarrow \{f(\overline{e}_1), f(\overline{e}_2), ..., f(\overline{e}_p), f(\overline{u}_1), ..., f((\overline{u}_q)\}$ es generador de $f(V)$.

Como B $\subset$ Nuc(f), $f(\overline{e}_1) = ... = f(\overline{e}_p) = \overline{0}$, y $\{f((\overline{u}_1), ..., f((\overline{u}_q)\}$ genera $f(V)$, que es Im(f).

Veamos ahora que $\{f((\overline{u}_1), ..., f((\overline{u}_q)\}$ son linealmente independientes:

Si $\lambda_1 f((\overline{u}_1) + ... + \lambda_q f((\overline{u}_q) = \overline{0}$, por ser f una aplicación lineal $f(\lambda_1\overline{u}_1 + ... + \lambda_q\overline{u}_q) = \overline{0}$, es decir $\lambda_1\overline{u}_1 + ... + \lambda_q\overline{u}_q \in$ Nuc(f), y por tanto, se puede expresar como combinación de los vectores de una de sus bases, por ejemplo de B.

$\lambda_1\overline{u}_1 + ... + \lambda_q\overline{u}_q = \alpha_1\overline{e}_1 + ... + \alpha_p\overline{e}_p = \overline{0}$, podemos escribir

$\alpha_1\overline{e}_1 + ... + \alpha_p\overline{e}_p + \lambda_1\overline{u}_1 + ... + \lambda_q\overline{u}_q = \overline{0}$, de donde se deduce necesariamente que $\alpha_1 = ... = \alpha_p = \lambda_1 = ... = \lambda_q = 0$ por ser $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_p, \overline{u}_1, ..., \overline{u}_q\}$ base de V.

Por tanto, $\lambda_1 f(\overline{u}_1) + ... + \lambda_q f(\overline{u}_q) = \overline{0} \Rightarrow \lambda_1 = ... = \lambda_q = 0 \Rightarrow \{f(\overline{u}_1), ..., f(\overline{u}_q)\}$ es un conjunto libre, y la dimensión de la imagen de f es q.

Hemos demostrado que $p + q = n$.

- A la dimensión de la imagen de f se la llama Rango de la aplicación lineal: rg(f).

Ejemplo 2.2.2

En el ejemplo 2.2.1 se verifica dim Nuc(f) + dim Im(f) = dim$V$. En efecto:

La dimensión de $M = \left\{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix}\right\}$ es tres porque $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} = a_{11}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} +$ $+ a_{12}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + a_{21}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ y $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right\}$ es un sistema libre, es decir, es una base de M formada por tres elementos.

dim Nuc (f) = 0, porque Nuc (f) = $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

dim Im (f) = 3, porque:

$\mathrm{Im}(f) = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,/\, x_4 = 0\} = \{(x_1, x_2, x_3, 0) \in \mathbb{R}^4\}$

como $\begin{cases} (x_1, x_2, x_3, 0) = x_1(1,0,0,0) + x_2(0,1,0,0) + x_3(0,0,1,0) \\ \{(1,0,0,0), (0,1,0,0), (0,0,1,0)\} \text{ es un sistema libre} \end{cases}$ $\Rightarrow$ este conjunto es una base de $\mathrm{Im}(f) \subset \mathbb{R}^4$ formada por tres elementos.

Ya conocemos la importancia de las demostraciones en matemáticas, pero además de su valor intrínseco, a lo largo de ellas, se utilizan y demuestran propiedades importantes que tienen valor fuera de dicha demostración.

Como parte de la demostración anterior ha quedado demostrada la propiedad expresada en la proposición siguiente:

2.2.6. Proposición

Una aplicación lineal $f: V \rightarrow W$ conserva la independencia lineal si el núcleo de la aplicación lineal es $\{\overline{0}\}$

La parte de la demostración anterior que demuestra esta proposición es:

"Debemos comprobar que $\{\overline{u}_1, \ldots, \overline{u}_q\}$ linealmente independientes $\Rightarrow$ $\Rightarrow \{f(\overline{u}_1), \ldots, f(\overline{u}_q)\}$ son linealmente independientes *si el núcleo de la aplicación lineal es* $\{\overline{0}\}$; $\lambda_1\overline{u}_1 + \ldots + \lambda_q\overline{u}_q = \overline{0} \Rightarrow f(\lambda_1\overline{u}_1 + \ldots + \lambda_q\overline{u}_q) = \lambda_1 f(\overline{u}_1) + \ldots + \lambda_q f(\overline{u}_q) = f(\overline{0}) = \overline{0}$. Como $\lambda_1 = \ldots = \lambda_q = 0$ porque $\{\overline{u}_1, \ldots \overline{u}_q\}$ son linealmente independientes $\Rightarrow f(\overline{u}_1), \ldots, f(\overline{u}_q)$ son linealmente independientes."

2.3. Aplicaciones lineales y matrices

Determinar una aplicación $f: V \to W$ es dar un procedimiento que permita calcular las imágenes de los elementos de V.

Hay distintos procedimientos para determinar una aplicación:

- Si el conjunto tiene un número finito de elementos (no es lo mismo que de dimensión finita), se pueden dar las imágenes de todos y cada uno de ellos.
- Se puede dar una ley de formación que permita asignar a cada elemento de V un elemento de W, de modo que f verifica las condiciones dadas en *[2.1.1]* como hemos hecho en *2.2*.
- Si la aplicación es lineal, no es necesario dar esa ley de formación ni las imágenes de todos los elementos, es suficiente dar las imágenes de los vectores de una base, porque por ser lineal, con ello quedan determinadas las imágenes de todos los vectores de V, es decir, queda determinada la aplicación.
  Las imágenes de los elementos de V son elementos de W, y vienen dados respecto a una base de W, que hay que fijar previamente.

2.3.1. Teorema

Si V, W son dos espacios vectoriales sobre $\mathbb{R}$; $B = \{\bar{e}_1, \bar{e}_2, ..., \bar{e}_n\}$ es una base de V y $S = \{\bar{w}_1, \bar{w}_2, ..., \bar{w}_n\}$ un subconjunto de W, entonces, existe una aplicación lineal única $f: V \to W$, tal que, $f(\bar{e}_1) = \bar{w}_1, f(\bar{e}_2) = \bar{w}_2, ..., f(\bar{e}_n) = \bar{w}_n$.

Demostración:

Como hemos hecho en otras ocasiones para demostrar que algo existe, lo construimos. Queremos demostrar que existe una aplicación lineal $f: V \to W$, vamos a construirla:

Cuando se fija la base B en V, cualquier vector $\bar{x} \in V$, se puede escribir dando sus coordenadas respecto a la base: $\bar{x} = (x_1, x_2, ..., x_n)$, es decir, $\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + ... + x_n\bar{e}_n$.

Construimos de la siguiente manera la imagen de cada elemento de la aplicación:

$$f(\bar{x}) = f(x_1\bar{e}_1 + x_2\bar{e}_2 + ... + x_n\bar{e}_n) = x_1 f(\bar{e}_1) + x_2 f(\bar{e}_2) + ... + x_n f(\bar{e}_n) =$$
$$= x_1\bar{w}_1 + x_2\bar{w}_2 + ... + x_n\bar{w}_n \Rightarrow \text{La aplicación } f \text{ existe.}$$

Como las coordenadas $(x_1, x_2, ..., x_n)$ de un vector $\bar{x}$ son únicas respecto a una base $B \Rightarrow$ la aplicación f es única.

Para ver que es lineal debemos comprobar que la aplicación construida así, verifica las condiciones dadas en *[2.1.1]*:

$$f(\lambda\bar{x} + \mu\bar{y}) = \lambda f(\bar{x}) + \mu f(\bar{y}); \quad \forall \bar{x}, \bar{y} \in V; \forall \lambda, \mu \in \mathbb{R}$$

En efecto,

$$\begin{aligned}
f(\lambda\bar{x} + \mu\bar{y}) &= f((\lambda x_1 + \mu y_1)\bar{e}_1 + ... + (\lambda x_n + \mu y_n)\bar{e}_n) = \\
&= [(\lambda x_1 + \mu y_1)\bar{w}_1 + ... + (\lambda x_n + \mu y_n)\bar{w}_n] = \\
&= (\lambda x_1\bar{w}_1 + ... + \lambda x_n\bar{w}_n) + (\mu y_1\bar{w}_1 + ... + \mu y_n\bar{w}_n) = \\
&= (\lambda x_1 f(\bar{e}_1) + ... + \lambda x_n f(\bar{e}_n)) + (\mu y_1 f(\bar{e}_1) + ... + \mu y_n f(\bar{e}_n)) = \\
&= \lambda(x_1 f(\bar{e}_1) + ... + x_n f(\bar{e}_n)) + \mu(y_1 f(\bar{e}_1) + ... + y_n f(\bar{e}_n)) = \\
&= \lambda f(x_1\bar{e}_1 + ... + x_n\bar{e}_n) + \mu f(y_1\bar{e}_1 + ... + y_n\bar{e}_n) = \lambda f(\bar{x}) + \mu f(\bar{y})
\end{aligned}$$

$\Rightarrow$ La aplicación f es lineal.

Ejemplo 2.3.1

Sea $f: \mathbb{R}^2 \to \mathbb{R}^3 / f(1, 0) = (1, 2, 1); f(0, 1) = (1, 1, 0)$ una aplicación lineal. Determínense unas ecuaciones de dicha aplicación.

Solución:

$B = \{(1, 0), (0, 1)\}$ es una base de $\mathbb{R}^2$, $\{(1, 2, 1), (1, 1, 0)\} \subset \mathbb{R}^3$, (no es base)
$(x_1, x_2) = x_1(1, 0) + x_2(0, 1)$; por ser f lineal podemos escribir:

$$f(x_1, x_2) = x_1 f(1, 0) + x_2 f(0, 1) = x_1(1, 2, 1) + x_2(1, 1, 0) = (x_1 + x_2, 2x_1 + x_2, x_1) \Rightarrow$$

$$\Rightarrow \begin{cases} y_1 = x_1 + x_1 \\ y_2 = 2x_1 + x_2 \\ y_3 = x_1 \end{cases}$$ son unas ecuaciones de la aplicación lineal dada, que

permiten calcular la imagen (y_1, y_2, y_3) de cualquier elemento (x_1, x_2) de $\mathbb{R}^2$.

Ejemplo 2.3.2

¿Se puede determinar alguna aplicación lineal $f: \mathbb{R}^2 \to \mathbb{R}^2$, tal que, $f(2, 3) = (0, 1)$; $f(-2, -3) = (1, 0)$?

Solución:

No puede haber una aplicación lineal que cumpla las condiciones dadas, ya que, si fuera lineal, debería cumplirse: $f(-2, -3) = f[-(2, 3)] = -[f(2, 3) = = -(1, 0)$, mientras que, en el enunciado aparece $f(-2, -3) = (1, 0)$.

Una aplicación lineal se puede determinar dando sus ecuaciones, como hemos hecho en el ejemplo anterior, pero también se puede hacer dando una matriz que va asociada a cada aplicación lineal cuando hemos fijado una base en el conjunto V y otra en el conjunto W.

2.3.2. Matriz asociada a una aplicación lineal

Si V, W son dos espacios vectoriales sobre $\mathbb{R}$, $B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\}$ es una base de V, $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_m\}$ una base de W y f una aplicación lineal de V en W.
Se llama matriz asociada a la aplicación lineal f respecto a las bases B y S, al conjunto A de elementos ordenados en una tabla rectangular de m filas y n columnas

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{siendo} \begin{cases} f(\overline{e}_1) = a_{11}\overline{u}_1 + a_{21}\overline{u}_2 + \dots + a_{m1}\overline{u}_m \\ f(\overline{e}_2) = a_{12}\overline{u}_1 + a_{22}\overline{u}_2 + \dots + a_{m2}\overline{u}_m \\ \dots \\ f(\overline{e}_n) = a_{1n}\overline{u}_1 + a_{2n}\overline{u}_2 + \dots + a_{mn}\overline{u}_m \end{cases}$$

Observemos que las columnas de la matriz A son las coordenadas de los vectores de $f(B)$ respecto a la base S de W.

El concepto de matriz ha surgido aquí de forma natural asociado a las aplicaciones lineales para cuyo estudio son una herramienta indispensable, y veremos su estructura algebraica a partir de la de las aplicaciones lineales, pero las matrices existen independientemente de las aplicaciones y forman una clase de objetos matemáticos estudiados con anterioridad por el lector en distintos momentos.

En Informática las matrices son una herramienta utilizada con mucha frecuencia desde cosas tan elementales como representar la posición de los pixels en la pantalla como un elemento de una matriz de 800 columnas y 640 fila (o 1024×800, ...) donde el color en cada uno es el valor del elemento en esa posición, hasta los Arrays utilizados para dimensionar variables, donde los subíndices del mismo indican la posición del elemento.

Utilizar los conocimientos anteriores de matrices, nos permite expresar la aplicación de una forma más cómoda, es su expresión analítica:

Si $f(x_1, x_2, ..., x_n) = (y_1, y_2, ..., y_m)$, la aplicación construida en *[2.3.2]* se puede expresar:

$$\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \Leftrightarrow Y = AX$$

Ejemplo 2.3.3

Escríbanse las ecuaciones matriciales de la aplicación dada en el ejemplo 2.3.1.

Solución:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Leftrightarrow \begin{cases} y_1 = x_1 + x_2 \\ y_2 = 2x_1 + x_2 \\ y_3 = x_1 \end{cases}$$ (ecuaciones obtenidas en el ejemplo 2.3.1).

2.3.3. Proposición

Si B es una base de V, su imagen $f(B)$ mediante la aplicación lineal f es un sistema generador de $f(V)$

Demostración:

Cualquier elemento de $f(V)$ tiene un original en $V \Rightarrow$ se puede escribir como combinación lineal de los elementos de $f(B)$.

$$f(\overline{x}) = f(x_1\overline{e}_1 + x_2\overline{e}_2 + \ldots + x_n\overline{e}_n) = x_1 f(\overline{e}_1) + x_2 f(\overline{e}_2) + \ldots + x_n f(\overline{e}_n)$$

2.3.4. Definición

Rango de una matriz es el rango de la aplicación lineal asociada denotada rg (A)

2.3.5. Teorema

El rango de una matriz es el número de vectores columna linealmente independientes que hay en ella

Demostración:

Si $f(B)$ es la imagen de una base de V, es un sistema generador de $f(V)$, por tanto, la dimensión de $f(V)$ será el número de vectores linealmente independientes que haya en $f(B)$.

Las columnas de la matriz A son los vectores de $f(B)$, por tanto, el número de vectores linealmente independientes que hay en $f(B)$ será el número de vectores columna linealmente independientes que hay en A.

El rango de la aplicación lineal se puede definir como la dimensión de su imagen o el rango de la matriz asociada.

Ejemplo 2.3.4

En la aplicación del ejemplo 2.3.1, la dimensión del subespacio imagen es 2, porque

$\begin{pmatrix} y_1 \\ y_1 \\ y_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, es decir $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ y $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ generan la imagen; además, como son linealmente independientes, son base de $\mathrm{Im}(f)$ y su dimensión es 2.

2.4. Los espacios vectoriales ($L(V, W)$, +, $\mathbb{R}$) y ($M_{m\times n}$,+, $\mathbb{R}$)

Llamaremos $L(V, W)$ el conjunto de todas las aplicaciones lineales posibles entre dos espacios vectoriales sobre $\mathbb{R}$: V de dimensión n, y W de dimensión m.

Estamos acostumbrados a sumar aplicaciones, multiplicar aplicaciones por escalares y componer dos aplicaciones. En este apartado vamos a ir viendo las propiedades que tienen las operaciones citadas, y como consecuencia, la estructura algebraica que tiene $L(V, W)$ con las distintas operaciones definidas con sus elementos.

2.4.1. Teorema

> ($L(V, W)$, +, $\mathbb{R}$) es un espacio vectorial sobre $\mathbb{R}$

Demostración:

Si f y g pertenecen a ($L(V, W)$, +), la suma de aplicaciones se define como:

$$(f+g)(\overline{v}) = f(\overline{v}) + g(\overline{v})$$

La suma es una aplicación lineal:

$$(f+g)(\alpha\overline{v}_1+\beta\overline{v}_2) = f(\alpha\overline{v}_1+\beta\overline{v}_2) + g(\alpha\overline{v}_1+\beta\overline{v}_2) = f(\alpha\overline{v}_1) + f(\beta\overline{v}_2) + g(\alpha\overline{v}_1) + g(\beta\overline{v}_2) = (f+g)(\alpha\overline{v}_1) + (f+g)(\beta\overline{v}_2) = \alpha(f+g)(\overline{v}_1) + \beta(f+g)(\overline{v}_2)$$

Es evidente que, tiene las propiedades conmutativas, asociativa, existencia de neutro (la función nula, tal que $0(\overline{v}) = \overline{0}$, $\forall\overline{v}\in V$), y elemento opuesto de f : $-f$, tal que $(-f)(\overline{u}) = -f(\overline{u})$.

Por tanto, ($L(V, W)$, +) es **un grupo conmutativo**.

El producto de una aplicación lineal por un escalar se define como:

$(\lambda f)(\overline{v}) = \lambda f(\overline{v})$, siendo $f\in L(V, W)$ y $\lambda \in \mathbb{R}$.

El producto de escalar por aplicación lineal es aplicación lineal:

$$(\lambda f)(\alpha\overline{v}_1+\beta\overline{v}_2) = \lambda f(\alpha\overline{v}_1+\beta\overline{v}_2) = \lambda[f(\alpha\overline{v}_1)+f(\beta\overline{v}_2)] = \lambda[\alpha f(\overline{v}_1)+\beta f(\overline{v}_2)] = \alpha\lambda f(\overline{v}_1)+\beta\lambda f(\overline{v}_2) = \alpha(\lambda f)(\overline{v}_1)+\alpha(\lambda f)(\overline{v}_2)$$

Es claro que esta operación tiene las propiedades que sirvieron para definir un espacio vectorial en *[1.2.2]*:

1. Distributiva de escalares respecto a aplicaciones: $\lambda(f+g) = \lambda f + \lambda g$.
2. Distributiva de aplicaciones respecto a escalares: $(\lambda+\mu)f = \lambda f + \mu f$.
3. Asociativa: $(\lambda\mu)f = \lambda(\mu f)$.
4. Existe el escalar unidad, 1, que cumple $1f = f$.

Hemos demostrado que ($L(V, W)$, +, $\mathbb{R}$) **es un espacio vectorial sobre $\mathbb{R}$.**

Por ser ($L(V, W)$, +, $\mathbb{R}$) un espacio vectorial, las aplicaciones lineales del espacio vectorial V (cuyos elementos son los vectores de V) en el espacio vectorial W (cuyos elementos son los vectores de W) son vectores.

Sea $f\in L(V, W)$, recordemos que $\dim(V) = n$ y $\dim(W) = m$, y que si fijamos una base en V y otra en W, cada aplicación $f : V \to W$ lleva asociada una matriz de orden $m\times n$.

Representaremos por $M_{m\times n}$ el conjunto de todas las matrices de orden $m\times n$.

Recordemos que, una aplicación tal que su inversa también es aplicación, es una biyección o correspondencia biyectiva.

Este es el caso de la correspondencia que existe entre las aplicaciones lineales y las matrices, cada aplicación lineal de $L(V, W)$ lleva asociada una matriz única de $M_{m\times n}$, y cada matriz de $M_{m\times n}$ determina una aplicación lineal única de $L(V, W)$.

$$L(V, W) \underset{g^{-1}}{\overset{g}{\rightleftarrows}} M_{m\times n}$$

La correspondencia biyectiva existe entre el objeto de estudio $L(V, W)$ y la herramienta utilizada, $M_{m\times n}$, nos induce a analizar la estructura de ambos conjuntos simultáneamente.

2.4.2. Teorema

$(M_{m\times n}, +, \mathbb{R})$ es un espacio vectorial sobre $\mathbb{R}$

Demostración:

$M_{m\times n} \times M_{m\times n} \xrightarrow{+} M_{m\times n}$: **Operación de sumar matrices de orden** $m \times n$.

Si $A = (a_{ij})$ es la matriz asociada a $f_A : V \to W$ y $B = (b_{ij})$ es la matriz asociada a $f_B : V \to W$, llamaremos matriz suma $A + B$ a la matriz asociada a la aplicación $f_A + f_B$: $V \to W$.

Basta ver la definición de suma de aplicaciones para establecer que $A + B =$ $= (a_{ij} + b_{ij})$.

La expresión analítica de $f_A + f_B$ es $Y = (A + B)X$.

Por la correspondencia biyectiva existente entre $L(V, W)$ y $M_{m\times n}$, tal que conserva la operación de sumar, podemos asegurar que $(M_{m\times n}, +)$ tiene las mismas propiedades que $(L(V, W), +)$, es decir, es un **grupo conmutativo.**

$M_{m\times n} \times \mathbb{R} \longrightarrow M_{m\times n}$: **Operación de multiplicar una matriz** $A \in M_{m\times n}$, **por un escalar** $\lambda \in \mathbb{R}$.

Si $A = (a_{ij})$ es la matriz asociada a $f_A : V \to W$, la operación de multiplicar la matriz $A \in M_{m\times n}$, por el escalar $\lambda \in \mathbb{R}$, da como resultado la matriz $\lambda A \in M_{m\times n}$, que es la matriz asociada a λf_A.

Utilizando la definición de aplicación por escalar $\Rightarrow \lambda A = (\lambda a_{ij})$.

Por la correspondencia biyectiva existente entre $L(V, W)$ y $M_{m\times n}$, tal que conserva la operación de multiplicar por escalares, podemos asegurar que ambos conjuntos tienen las mismas propiedades al multiplicar sus elementos por un escalar:

1. Distributiva de escalares respecto a matrices: $\lambda(A + B) = \lambda A + \lambda B$
2. Distributiva de matrices respecto a escalares: $(\lambda + \mu)A = \lambda A + \mu A$
3. Asociativa: $(\lambda\mu)A = \lambda(\mu A)$
4. Existe el escalar unidad, 1, que cumple $1A = A$

La aplicación $f : L(V, W) \to M_{m\times n}$, que asocia a cada aplicación lineal una matriz, es una aplicación biyectiva que transforma la suma de aplicaciones en suma de matrices, y al producto de escalar por aplicación en escalar por matriz, es decir, conserva las operaciones del espacio vectorial.

Una aplicación biyectiva que conserva la estructura de espacio vectorial recibe el nombre de isomorfismo entre los espacios vectoriales.

Hemos demostrado que **el conjunto** $(M_{m\times n}, +, \mathbb{R})$ **es un espacio vectorial isomorfo a** $(L(V, W), +, \mathbb{R})$.

Ejemplo 2.4.1

Dadas las aplicaciones lineales asociadas a las matrices $A = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 1 & -1 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & -5 & 6 & -1 \\ 2 & 0 & -2 & -3 \end{pmatrix}$, se pide establecer la expresión analítica de las aplicaciones lineales asociadas f_A, f_B, la aplicación $f_A + 2f_B$, y la matriz correspondiente a esta aplicación.

Solución:

$\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 1 & -1 \end{pmatrix}$ es la matriz de la aplicación lineal $f_A : \mathbb{R}^4 \to \mathbb{R}^2$, cuya expresión es:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \Leftrightarrow \begin{cases} y_1 = 1x_1 + 2x_2 - 3x_3 + 4x_4 \\ y_2 = 0x_1 - 5x_2 + 1x_3 - 1x_4 \end{cases}$$

$\begin{pmatrix} 3 & -5 & 6 & -1 \\ 2 & 0 & -2 & -3 \end{pmatrix}$ es la matriz de la aplicación lineal $f_B : \mathbb{R}^4 \to \mathbb{R}^2$, cuya expresión es:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3 & -5 & 6 & -1 \\ 2 & 0 & -2 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \Leftrightarrow \begin{cases} y_1 = 3x_1 - 5x_2 + 6x_3 - 1x_4 \\ y_2 = 2x_1 + 0x_2 - 2x_3 - 3x_4 \end{cases}$$

$f_A : \mathbb{R}^4 \to \mathbb{R}^2 \,/\, (x_1, x_2, x_3, x_4) \to (1x_1 + 2x_2 - 3x_3 + 4x_4,\ 0x_1 - 5x_2 + 1x_3 - 1x_4)$

$f_B : \mathbb{R}^4 \to \mathbb{R}^2 \,/\, (x_1, x_2, x_3, x_4) \to (3x_1 - 5x_2 + 6x_3 - 1x_4,\ 2x_1 + 0x_2 - 2x_3 - 3x_4)$

$f_{A+2B} : \mathbb{R}^4 \to \mathbb{R}^2 \,/\, (x_1, x_2, x_3, x_4) \to (7x_1 - 8x_2 + 9x_3 + 2x_4,\ 4x_1 - 5x_2 - 3x_3 - 7x_4)$

La expresión analítica de $f_A + 2f_B$ es $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 7 & -8 & 9 & 2 \\ 4 & -5 & -3 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$

$\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 1 & -1 \end{pmatrix} + 2\begin{pmatrix} 3 & -5 & 6 & -1 \\ 2 & 0 & -2 & -3 \end{pmatrix} = \begin{pmatrix} 7 & -8 & 9 & 2 \\ 4 & -5 & -3 & -7 \end{pmatrix}$ es la matriz de la aplicación lineal f_{A+2B}.

2.4.3. Teorema

$Dim(L(V, W), +, \mathbb{R}) = \dim(M_{m\times n}, +, \mathbb{R}) = m \times n$

Demostración:

Para saber la dimensión de $(M_{m\times n}, +, \mathbb{R})$, es suficiente conocer el número de elementos de una de sus bases.

$\{E_{ij}, i = 1, \dots, m, j = 1, \dots, n\}$ es una base si cada E_{ij} es la matriz de orden $m \times \times n$ que tiene 0 todos los elementos, excepto el que ocupa la fila i, columna j, que vale 1.

La base así construida es la base canónica y tiene $m \times n$ matrices E_{ij}, y por tanto, la dimensión de $(M_{m\times n}, +, \mathbb{R})$ es $m \times n$.

La dimensión de $(L(V, W), +, \mathbb{R})$ es también $m \times n$ por ser isomorfo a $(M_{m\times n}, +, \mathbb{R})$.

Ejemplo 2.4.2

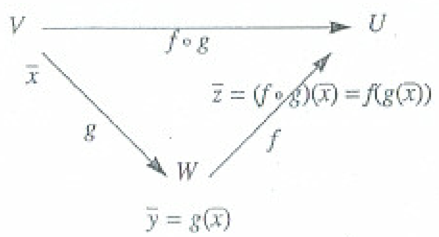
La base canónica de $M_{2\times 3}$ es:

$\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$, tiene 6 matrices linealmente indepedientes.

2.5. Composición de aplicaciones lineales y producto de matrices

A veces hay que realizar una serie sucesiva de transformaciones a un elemento de un conjunto inicial, de manera que el transformado mediante una aplicación es el elemento a transformar mediante la siguiente aplicación; ese proceso es el de componer aplicaciones.

Si V, W, U son tres espacios vectoriales reales, g y f son las siguientes aplicaciones lineales, se puede construir la aplicación lineal $f \circ g$ (se lee g compuesta con f), de modo que, la imagen mediante la aplicación sucesiva de g y f a cualquier vector de V, dé el mismo resultado que la aplicación de $f \circ g$ a dicho vector.



◆ Obsérvese que se escriben en orden inverso a como se aplican.

La forma de multiplicar matrices parecería artificiosa si no fuera la expresión analítica de la composición de aplicaciones; vamos a ver por qué está definida de esa extraña manera.

2.5.1. Teorema

La composición de dos aplicaciones lineales es una aplicación lineal

Demostración:

Si g y f son las aplicaciones del gráfico anterior, por ser lineales podemos escribir:

$f \circ g(\lambda\bar{v}_1 + \mu\bar{v}_2) = f[g(\lambda\bar{v}_1 + \mu\bar{v}_2)] = f(\lambda g(\bar{v}_1) + \mu g(\bar{v}_2)) = \lambda f(g(\bar{v}_1)) + \mu f(g(\bar{v}_2)) =$
$= \lambda(f \circ g)(\bar{v}_1) + \mu(f \circ g)(\bar{v}_2)$, quedando demostrado que $f \circ g$ es lineal.

Si $\dim(V) = n$, $\dim(W) = p$ y B es la matriz asociada a la aplicación lineal respecto a las bases canónicas, $g : V \to W \Rightarrow B \in M_{p\times n}$.

Si $\dim(W) = p$, $\dim(U) = m$ y A es la matriz asociada a la aplicación lineal, respecto a las bases canónicas, $f : W \to U \Rightarrow A \in M_{m\times p}$.

Como $\dim(V) = n$, $\dim(U) = m$ la matriz asociada a la aplicación lineal, respecto a las bases canónicas, $f \circ g : V \to U$ será de orden $m \times n$.

En estas condiciones definimos el producto de matrices como la matriz de la aplicación composición.

2.5.2. Definición de producto de matrices

AB es la matriz asociada a $f \circ g : V \to U$, siendo A la matriz asociada a $f : W \to U$ y B la matriz asociada a $g : V \to W$

2.5.3. Cálculo del producto de matrices

La aplicación lineal $g : V \to W$, cuya matriz asociada es $B = (b_{ij})$ $M_{p\times n}$, tiene por ecuaciones:

$(y_1, y_2, ..., y_p) = g\,(x_1, x_1, ..., x_n)$; Cada $y_h = b_{h1}x_1 + b_{h2}x_2 + ... + b_{hn}x_n = \sum_{j=1}^{n} b_{ij}x_j$; $h = 1, ..., p$.

La aplicación lineal $f : W \to U$, cuya matriz asociada es $A = (a_{ij}) \in M_{m \times p}$, tiene por ecuaciones:

$(z_1, z_1, \dots, z_m) = f(y_1, y_2, \dots, y_p)$; Cada $z_i = a_{i1}y_1 + a_{i2}y_2 + \dots + a_{ip}y_p = \sum_{h=1}^{p} a_{ih}y_h$; $i = 1, \dots, m$

Las ecuaciones de la aplicación lineal $f \circ g : V \to U$, cuya matriz asociada es $AB = (c_{ij}) \in M_{n \times m}$, tendrá por ecuaciones:

$$(z_1, z_1, \dots, z_m) = f(y_1, y_2, \dots, y_p) = f[g(x_1, x_2, \dots, x_n)] = f \circ g(x_1, x_2, \dots, x_n)$$

$$\text{Cada } z_i = \sum_{h=1}^{p} a_{ih}y_h = \sum_{h=1}^{p} a_{ih}\Big(\sum_{j=1}^{n} b_{hj}x_j\Big) = \sum_{j=1}^{n}\Big(\sum_{h=1}^{p} a_{ih}b_{hj}\Big)x_j \text{ para } i = 1, \dots, m$$

En Informática el uso de expresiones de este tipo equivale al uso de bucles (De $i = 1, \dots, m$; For $i = 1, \dots, m$), son válidos también para expresar flujos condicionales (i está entre 1 y m, con $i \neq j$) o para flujos repetitivos (i está entre m y n).

Utilizar el símbolo Σ simplifica la escritura de una suma, aunque puede hacer más difícil la interpretación de la expresión, por ello, vamos aclarar su significado.

Al escribir la matriz asociada a la aplicación lineal que tiene estas ecuaciones, se obtiene una matriz tal que, el elemento que ocupa la fila i, columna j : c_{ij} de $AB = (c_{ij}) \in M_{m \times n}$, se ha obtenido haciendo la siguiente manipulación:

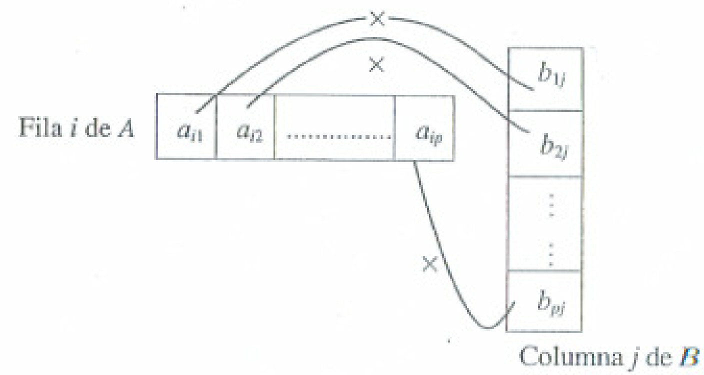
"Primer elemento de la fila i de la matriz A por primer elemento de la columna j de la matriz B, más segundo elemento de la fila i de la matriz A por segundo elemento de la columna j de la matriz B, más..., más elemento p -esimo de la fila i de la matriz A por elemento p -esimo de la columna j de la matriz B".

Todo este proceso se puede escribir con la siguiente expresión:

$C = (c_{ij}) \in M_{m \times n}$, siendo $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$; $i = 1, 2 \dots m$; $j = 1, \dots, n$.

Esta es la forma en que el estudiante ha visto siempre el producto de matrices.

Abusando del lenguaje, se dice que se obtiene el elemento c_{ij} multiplicando la fila i de A por la columna j de B.



♦ Obsérvese que para poder componer aplicaciones, el espacio final de la primera transformación que se aplica, debe ser el inicial de la segunda transformación, que en lenguaje de matrices equivale a decir "sólo se pueden multiplicar dos matrices cuyos órdenes son $m \times p$ y $p \times n$".

2.5.4. Teorema

La composición de dos aplicaciones lineales verifica:

1. No es conmutativa: $f \circ g$ puede ser $\neq g \circ f$
2. Existe la aplicación identidad: $i \circ f = f$
3. Asociativa: $(f \circ g) \circ h = f \circ (g \circ f)$
4. Distributiva respecto a la suma por la izquierda: $h \circ (f + g) = h \circ f + h \circ g$
5. Distributiva respecto a la suma por la derecha: $(f + g) \circ h = f \circ h + g \circ h$
6. Asociativa con escalares $\lambda(f \circ g) = (\lambda f) \circ g = f \circ (\lambda g)$

Demostraciones:

El fundamento de estas sencillas demostraciones es comprobar si la imagen de cualquier vector de V mediante las aplicaciones que hay a ambos lados del signo "=" es la misma o distinta.

1. Cuando se busca demostrar que la igualdad no es cierta, es suficiente en buscar un contraejemplo, es lo que hacemos en este caso:

Definimos las funciones $f: \mathbb{R}^2 \to \mathbb{R}^2$ como $f(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$, y $g: \mathbb{R}^2 \to \mathbb{R}^2$ como $g(x_1, x_2) = (x_1, x_1)$, al componer ambas funciones en orden distinto se obtiene:

$$\mathbb{R}^2 \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$$

$$(x_1, x_2) \to (x_1 + x_2, x_1 - x_2) \to (x_1 + x_2, x_1 + x_2) \Rightarrow (g \circ f)(x_1, x_2) = (x_1 + x_2, x_1 + x_2)$$

$$\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}^2$$

$$(x_1, x_2) \to (x_1, x_1) \to (2x_1, 0) \Rightarrow (f \circ g)(x_1, x_2) = (2x_1, 0)$$

Por tanto, $(g \circ f) \neq (f \circ g)$.

2. i es la aplicación identidad definida en **2.1** como $i(\overline{v}) = \overline{v}\ \forall \overline{v}$.
3. $[(f \circ g) \circ h](x) = (f \circ g)(h(x)) = f(g(h(x)))$
   $[f \circ (g \circ h)](x) = f[(g \circ h)(x)] = g(h(x)))$
4. $[h \circ (f + g)](x) = h((f + g)(x)) = h((f(x) + g(x)) = h(f(x)) + h(g(x))$
   $(h \circ f + h \circ g)(x) = (h \circ f)(x) + (h \circ g)(x) = h(f(x)) + h(g(x))$
5. Es idéntica a la 2.
6. $[\lambda(f \circ g)](x) = \lambda(f \circ g)(x) = \lambda(f(g(x)) = \lambda f(g(x))$
   $[(\lambda f) \circ g](x) = (\lambda f)g(x) = \lambda f(g(x))$

2.5.5. Teorema

El producto de matrices verifica:
1. No es conmutativo: AB puede ser $\neq BA$
2. Existe la matriz identidad: $I \in M_{nxn}$ tal que $I.A = A$ y $I \in M_{pxp}$ tal que $I \cdot B = B$
3. Asociativa: $(AB)C = A(BC)$
4. Distributiva respecto a la suma por la izquierda: $A(B + C) = AB + AC$
5. Distributiva respecto a la suma por la derecha: $(A + B)C = AC + BC$
6. Asociativa con escalares $\lambda(AB) = (\lambda A)B = A(\lambda B)$

Para las matrices no es necesario hacer las demostraciones de las propiedades correspondientes, ya que, el producto de matrices es la matriz asociada a la composición.

2.6. Álgebras de endomorfismos y matrices cuadradas

En **2.4** estudiamos el conjunto de todas las aplicaciones lineales posibles entre dos espacios vectoriales sobre $\mathbb{R}$: V de dimensión n y W de dimensión m, lo llamamos $L(V, W)$, y vimos que tenía estructura de espacio vectorial.

Si nos ceñimos a las aplicaciones lineales de V en sí mismo, estamos ante un caso particular de lo estudiado en **2.4**, las aplicaciones lineales de un espacio vectorial en sí mismo, se llaman endomorfismos

Naturalmente, en el caso particular se verifican todas las propiedades que se verificaban en el general, pero a veces, su singularidad puede conferirle alguna propiedad más.

La primera particularidad que tienen es que la matriz asociada a un endomorfismo es cuadrada.

Si $f : V \to V$, la matriz asociada tendrá las n columnas (cada una formada por la imagen de un vector de una base de V), expresadas en la misma u otra base de V, y, por tanto, cada columna tendrá n coordenadas, que dan lugar a n filas.

$$L(V, V) \underset{f^{-1}}{\overset{f}{\rightleftarrows}} M_{n \times n}$$

Ejemplo 2.6.1

$f : \mathbb{R}^2 \to \mathbb{R}^2 \,/\, f(x_1, x_2) = (2x_1, 0)$ es un endomorfismo de $\mathbb{R}^2$, la matriz asociada a f respecto a la base canónica de $\mathbb{R}^2$, tendrá por columnas las imágenes de la base, $f(0, 1) = (2, 0)$ y $f(1, 0) = (0, 0)$; $(2, 0)$ y $(0, 0)$ tienen 2 coordenadas por ser vectores de $\mathbb{R}^2$.

Respecto a la suma de aplicaciones, los endomorfismos no presentan ninguna particularidad, y como consecuencia, $(L(V, V), +)$ es un grupo conmutativo.

Respecto al producto de aplicaciones lineales por escalares, tampoco ocurre nada especial en los endomorfismos, teniendo la misma estructura de espacio vectorial que tenía $(L(V, W), +, \mathbb{R})$.

2.6.1. Teorema

$(L(V, V), +, \mathbb{R})$ es un subespacio vectorial de $(L(V, W), +, \mathbb{R})$

Demostración:

Es evidente, ya que, $(L(V, V), +, \mathbb{R}) \subset (L(V, W), +, \mathbb{R})$ y $(L(V, V), +, \mathbb{R})$ es espacio vectorial.

La particularidad más significativa del conjunto de endomorfismos de V es que se pueden componer siempre porque el conjunto final de la primera aplicación es el conjunto inicial de la segunda.

2.6.2. Teorema

La composición de dos endomorfimos verifica:

1. Asociativa: $(f \circ g) \circ h = f \circ (g \circ h)$
2. Distributiva respecto a la suma por la izquierda: $h \circ (f + g) = h \circ f + h \circ g$
3. Distributiva respecto a la suma por la derecha: $(f + g) \circ h = f \circ h + g \circ h$
4. Asociativa con escalares: $\lambda(f \circ g) = (\lambda f) \circ g = f \circ (\lambda g)$
5. No conmutativa: $f \circ g \neq g \circ f$ en general
6. Existe un elemento neutro i que es la aplicación identidad en $V \,/\, i \circ f = f \circ i = f$

Demostración:

Verifica las propiedades *1, 2, 3, 4* porque los endomorfismos de V son un caso particular de $L(V, W)$ como hemos dicho anteriormente.

Para poder afirmar la *5* es suficiente con encontrar un contraejemplo como el siguiente:

$$f_A(x_1, x_2) = (x_1 + 2x_2, 3x_1 + 4x_2); f_B(x_1, x_2) = (x_2, x_1)$$
$$f_B(f_A(x_1, x_2)) = (3x_1 + 4x_2, x_1 + 2x_2) \neq f_A(f_B(x_1, x_2)) = (x_2 + 2x_1, 3x_2 + 4x_1)$$

Para ver que la *6* es cierta, es suficiente construir la aplicación i de los gráficos siguientes:

$$\begin{tikzcd} V \arrow[rr, "f \circ i"] \arrow[dr, "i"'] & & V \\ & V \arrow[ur, "f"'] & \end{tikzcd}$$

$\bar{x}$ $\quad (f \circ i)(\bar{x}) = f(i(\bar{x})) = f(\bar{x})$

$i(\bar{x}) = \bar{x}$

$$\begin{tikzcd} V \arrow[rr, "i \circ f"] \arrow[dr, "f"'] & & V \\ & V \arrow[ur, "i"'] & \end{tikzcd}$$

$\bar{x}$ $\quad (i \circ f)(\bar{x}) = i(f(\bar{x})) = f(\bar{x})$

$f(\bar{x})$

Cuando en un conjunto hay definidas dos operaciones, de modo que, respecto a una de ellas es grupo conmutativo, la segunda es asociativa y es distributiva respecto a la primera, tenemos la estructura algebraica llamada anillo.

Podemos recopilar las propiedades que hemos visto en el siguiente teorema:

2.6.3. Teorema

El conjunto de endomorfismos de V es un anillo no conmutativo con elemento neutro
En el conjunto $L(V, V)$ hay definidas dos operaciones: $+$, $\circ$
$L(V, V), +)$ es un grupo conmutativo
$L(V, V), \circ)$ verifica la propiedad asociativa
$\circ$ es distributiva respecto a $+$ por la derecha y por la izquierda
$\circ$ no es conmutativa
$\exists$ elemento neutro

El conjunto de endomorfismos de V tiene estructura de anillo considerando las operaciones $+$, $\circ$, y estructura de espacio vectorial considerando las operaciones $+$, y multiplicación por escalares, además verifica la condición de asociatividad con escalares. Esta estructura tan compleja se llama álgebra.

El hecho de que los endomorfismos de un espacio vectorial se puedan componer siempre, nos permite multiplicar siempre sus matrices asociadas.

Si A es la matriz asociada a f_A y B es la matriz asociada a f_B, siempre es posible hacer AB, que es la matriz asociada a $f_A \circ f_B$.

$$V \xrightarrow{f_B} V \xrightarrow{f_A} V$$
$$\bar{x} \rightarrow f_B(\bar{x}) \rightarrow f_A(f_B(\bar{x}))$$

Ejemplo 2.6.2

En el contraejemplo que pusimos anteriormente para la no conmutatividad, establecimos los siguientes endomorfismos de $\mathbb{R}^2$:

$$f_A(x_1, x_2) = (x_1 + 2x_2, 3x_1 + 4x_2); f_B(x_1, x_2) = (x_2, x_1).$$

La composición $f_B \circ f_A$ es:

$$\mathbb{R}^2 \xrightarrow{f_A} \mathbb{R}^2 \xrightarrow{f_B} \mathbb{R}^2$$

$$(x_1, x_2) \to (x_1 + 2x_2, 3x_1 + 4x_2) \to (3x_1 + 4x_2, x_1 + 2x_2)$$

Las matrices asociadas a f_A y f_B son $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ respecto a las bases canónicas. Su producto es siempre posible, porque siempre es posible componer los endomorfismos en este caso la matriz asociada a $f_B \circ f_A$ es:

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

Hemos estudiado el conjunto de endomorfismos, cada uno de ellos lleva asociada una matriz cuadrada. En el conjunto $M_{n\times n}$ (matrices cuadradas de orden n), están definidas las operaciones de sumar, multiplicar matrices por escalares y multiplicar matrices por matrices con las propiedades que se derivan del isomorfismo $L(V, V) \underset{f^{-1}}{\overset{f}{\rightleftarrows}} M_{n\times n}$

Por ello, podemos asegurar que $M_{n\times n}$ tiene estructura de álgebra, y como consecuencia, $(M_{n\times n}, +, \cdot)$ es un anillo.

2.7. Matriz asociada a un cambio de base en V. Matriz asociada a una aplicación lineal cuando cambian las bases

Recordemos que fijada una base $B = \{\bar{e}_1, \bar{e}_2, ..., \bar{e}_n\}$, cualquier vector $\bar{x} \in V$, se puede expresar en función de los vectores de dicha base $\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + ... + x_n\bar{e}_n$, y llamamos coordenadas del vector $\bar{x}$ respecto a la base dada a los coeficientes $x_1, x_2, ..., x_n$.

Elegir una base diferente $B' = \{\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n\}$ en V, significa que el vector $\bar{x} \in V$ se puede expresar de la forma $\bar{x} = \bar{x}'_1\bar{e}'_1 + \bar{x}'_2\bar{e}'_2 + ... + \bar{x}'_n\bar{e}'_n$, y sus coordenadas respecto a la nueva base son: $\bar{x}'_1, \bar{x}'_2, ..., \bar{x}'_n$.

El número de coordenadas respecto a las dos bases es el mismo, porque todas las bases de un espacio vectorial tienen el mismo número de elementos, su dimensión, pero las coordenadas podrán ser distintas respecto a una base y a otra.

Establezcamos el endomorfismo identidad, de modo que, la imagen de cada vector es el mismo expresado respecto a otra base, entonces:

$$(V, B') \xrightarrow{i} (V, B)$$

$$\bar{x} = \bar{x}'_1\bar{e}'_1 + \bar{x}'_2\bar{e}'_2 + ... + \bar{x}'_n\bar{e}'_n \xrightarrow{i} \bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + ... + x_n\bar{e}_n$$

Las imágenes de los vectores de B' mediante el endomorfismo i vienen dadas por las expresiones siguientes:

$$\begin{cases} i(\bar{e}'_1) = \bar{e}'_1 = q_{11}\bar{e}_1 + q_{21}\bar{e}_2 + ... + q_{n1}\bar{e}_n \\ i(\bar{e}'_2) = \bar{e}'_2 = q_{12}\bar{e}_1 + q_{22}\bar{e}_2 + ... + q_{n2}\bar{e}_n \\ ... \quad ... \quad ... \quad ... \quad ... \quad ... \quad ... \quad ... \\ i(\bar{e}'_n) = \bar{e}'_n = q_{1n}\bar{e}_1 + q_{2n}\bar{e}_2 + ... + q_{nn}\bar{e}_n \end{cases}; \qquad Q = \begin{pmatrix} q_{11} & q_{12} & ... & q_{1n} \\ q_{21} & q_{22} & ... & q_{2n} \\ ... & ... & ... & ... \\ q_{n1} & q_{n2} & ... & q_{nn} \end{pmatrix}$$

La matriz Q es la matriz asociada a i, se llama matriz del cambio de base

Las columnas de Q son las coordenadas de los vectores de la base B' respecto a B.

Las ecuaciones de la aplicación identidad que permiten pasar de las coordenadas de un vector respecto a la base B' a las coordenadas respecto a la base B se llaman ecuaciones de cambio de base, y son:

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \dots & \dots & \dots & \dots \\ q_{n1} & q_{n2} & \dots & q_{nn} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_n \end{pmatrix} \Leftrightarrow \overline{x} = Q\overline{x}'$$

Ejemplo 2.7.1.

Siendo $B' = \{\overline{e}'_1 = (1, 0, 0, -1), \overline{e}'_2 = (0, 1, 2, 0), \overline{e}'_3 = (1, -1, 2, -2), \overline{e}'_4 = (0, 0, 1, -1)\}$ y B la base canónica, determínense la matriz de cambio de base en $\mathbb{R}^4$, las ecuaciones del cambio de base que permiten pasar de las coordenadas respecto a B' a las coordenadas respecto a B, y la imagen respecto a B del vector $(1, -3, 4, -5)_{B'}$.

Solución:

$$(V, B') \xrightarrow{\ i\ } (V, B)$$

vectores de B' respecto a B' vectores de B' respecto a B

$$\begin{aligned}
(1, 0, 0, 0) &\longrightarrow (1, 0, 0, -1) \\
(0, 1, 0, 0) &\longrightarrow (0, 1, 2, 0) \\
(0, 0, 1, 0) &\longrightarrow (1, -1, 2, -2) \\
(0, 0, 0, 1) &\longrightarrow (0, 0, 1, -1)
\end{aligned}$$

Matriz Q del cambio de base: $Q = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 2 & 2 & 1 \\ -1 & 0 & -2 & -1 \end{pmatrix}$

Ecuaciones del cambio de base:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 2 & 2 & 1 \\ -1 & 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 = x'_1 + x'_3 \\ x_2 = x'_2 - x'_3 \\ x_3 = 2x'_2 + 2x'_3 + x'_4 \\ x_4 = -x'_1 - 2x'_3 - x'_4 \end{cases}$$

El vector conocido es $\overline{x}' = (1, -3, 4, -5)_{B'}$, para calcular $\overline{x}$ debemos hacer la sustitución de $(1, -3, 4, -5)_{B'}$ en la expresión anterior:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 2 & 2 & 1 \\ -1 & 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \\ 4 \\ -5 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 = 5 \\ x_2 = -7 \\ x_3 = -3 \\ x_4 = -4 \end{cases}$$

Comprobemos que $(1, -3, 4, -5)_{B'} = (5, -7, -3, -4)_B$:

$$\overline{x} = (1, -3, 4, -5)_{B'} = (1, 0, 0, -1) - 3(0, 1, 2, 0) + 4(1, -1, 2, -2) - 5(0, 0, 1, -1)$$
$$\overline{x} = (5, -7, -3, -4)_B = 5(1, 0, 0, 0) - 7(0, 1, 0, 0) - 3(0, 0, 1, 0) - 4(0, 0, 0, 1)$$

De la misma manera que hemos actuado para expresar en base B los vectores que están en base B', podemos repetir el proceso y obtener en base B' los vectores que están en base B.

$$(V, B) \xrightarrow{\ i\ } (V, B')$$

Cada vector de B pertenece a V, y su imagen mediante el endomorfismo i establecido, es el mismo vector expresado como combinación lineal de los vectores de la base B' de V:

$$\begin{cases} i(\overline{e}_1) = \overline{e}_1 = p_{11}\overline{e}'_1 + p_{21}\overline{e}'_2 + \dots + p_{n1}\overline{e}'_n \\ i(\overline{e}_2) = \overline{e}_2 = p_{12}\overline{e}'_1 + p_{22}\overline{e}'_2 + \dots + p_{n2}\overline{e}'_n; \\ \dots \\ i(\overline{e}_n) \quad \overline{e}_n = p_{1n}\overline{e}'_1 + p_{2n}\overline{e}'_2 + \dots + p_{nn}\overline{e}'_n \end{cases} \qquad P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}$$

La matriz P también es una matriz de cambio de base.
Las ecuaciones de este cambio de base son: $\overline{x}' = P\overline{x}$

Es evidente que, si tenemos un vector expresado respecto a B', pasamos mediante Q a su expresión respecto a B, y la imagen obtenida la expresamos respecto a B' utilizando P, el resultado es el vector original.

Si utilizamos i_Q para representar la identidad cuando es Q la matriz del cambio de base en V, podemos escribir:

$$\begin{array}{ccccc} (V, B') & \xrightarrow{i_Q} & (V, B) & \xrightarrow{i_P} & (V, B') \\ \overline{x}' & \longrightarrow & \overline{x} = Q\overline{x}' & \longrightarrow & \overline{x}' = P\overline{x} = P\,Q\overline{x}' \end{array}$$

La expresión $\overline{x}' = P\overline{x} = P\,Q\overline{x}'$, demuestra que las matrices de cambio de base tienen inversa porque son matrices cuadradas, tales que $PQ = I \Leftrightarrow \begin{cases} P = Q^{-1} \\ Q = P^{-1} \end{cases}$,

y por tanto, $i_Q = i_P^{-1}$ e $i_P = i_Q^{-1}$.

En *[2.3.2]* vimos que "Si V, W son dos espacios vectoriales sobre $\mathbb{R}$, B es una base de V, S una base de W y $f: V \to W$ es una aplicación lineal, la aplicación f lleva asociada una matriz A".

¿Cómo varía esa matriz si en V hacemos un cambio de la base B por la base B', y en W cambiamos la base S por la base S'?

Los cambios producidos se ven en el diagrama siguiente:

$$\begin{array}{ccc} (V, B) & \xrightarrow{f_A} & (W, S) \\ \overline{x} & & f(\overline{x}) \\ \uparrow i_Q & & i_P^{-1}\downarrow\uparrow i_P \\ (V, B') & \xrightarrow{f_{A'}} & (W, S') \\ \overline{x}' & & (f(\overline{x}))' \end{array}$$

Para expresar la imagen de los vectores de (V, B') mediante la aplicación lineal f en (W, S'), tenemos que hacer la composición de aplicaciones $i_P^{-1} \circ f_A \circ i_Q$.

La matriz asociada a la aplicación lineal al cambiar las bases es $P^{-1}AQ$ si Q es la matriz del cambio de base B' por B en V y P es la matriz de cambio de S' por S en W.

La expresión $A' = P^{-1}AQ$ es la **fórmula de cambio de la matriz A asociada a una aplicación lineal cuando cambian ambas bases.**

Ejemplo 2.7.2

$f: \mathbb{R}^2 \to \mathbb{R}^3$ es la aplicación lineal tal que, $f(1, 2) = (1, 0, 1)$ y $f(0, 1) = (2, 1, 0)$. Los vectores $(1, 2)$, $(0, 1)$ están referidos a una base $B = \{\overline{e}_1, \overline{e}_2\} \subset \mathbb{R}^2$, $B' = \{\overline{e}_1 + \overline{e}_2, \overline{e}_1 - \overline{e}_2\}$ es otra base de $\mathbb{R}^2$; los vectores $(1, 0, 1)$, $(2, 1, 0)$ están referidos a una base $S = \{\overline{s}_1, \overline{s}_2, \overline{s}_3\} \subset \mathbb{R}^3$, $S' = \{\overline{s}_1 + \overline{s}_3, \overline{s}_1 - \overline{s}_2, \overline{s}_2\}$ es otra base de $\mathbb{R}^3$. Se pide la matriz asociada a la aplicación lineal respecto a las base $B \subset \mathbb{R}^2$ y $S \subset \mathbb{R}^3$. Se pide la matriz asociada a la aplicación lineal respecto a las bases $B \subset \mathbb{R}^2$ y $S \subset \mathbb{R}^3$, la matriz asociada a la aplicación lineal respecto a las bases $B' \subset \mathbb{R}^2$ y $S' \subset \mathbb{R}^3$, y la imagen respecto a S' de un vector cuyas coordenadas respecto a B' son $(2, 3)$.

Solución:

$f(0, 1) = f(0\overline{e}_1 + \overline{e}_2) = f(\overline{e}_2) = (2, 1, 0)$; $\quad f(1, 2) = f(1\overline{e}_1 + 2\overline{e}_2) = 1f(\overline{e}_1) + 2f(\overline{e}_2) \Rightarrow$
$\Rightarrow f(\overline{e}_1) = f(1, 2) - 2f(\overline{e}_2) = (1, 0, 1) - 2(2, 1, 0) = (-3, -2, 1)$

$f: (\mathbb{R}^2, B) \to (\mathbb{R}^3, S)$ lleva asociada la matriz $A = \begin{pmatrix} -3 & 2 \\ -2 & 1 \\ 1 & 0 \end{pmatrix}$

La matriz del cambio de la base $B' = \{\overline{e}_1 + \overline{e}_2, \overline{e}_1 - \overline{e}_2\}$ a la base $B = \{\overline{e}_1, \overline{e}_2\}$ es $Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

La matriz del cambio de base de S' a S es $P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Necesitamos la matriz del cambio de la base $S = \{\overline{s}_1, \overline{s}_2, \overline{s}_3\}$ a la base $S' = \{\overline{s}_1 + \overline{s}_3, \overline{s}_1 - \overline{s}_2, \overline{s}_2\}$, que es P^{-1}, como aún no hemos dado un procedimiento de cálculo de la matriz inversa, calcularemos directamente las coordenadas de S en función de las de S':

$$S' = \begin{cases} \overline{u}_1 = \overline{s}_1 + \overline{s}_3 \\ \overline{u}_2 = \overline{s}_1 - \overline{s}_2 \\ \overline{u}_3 = \overline{s}_2 \end{cases} \Rightarrow \begin{cases} \overline{s}_1 = \overline{u}_1 + \overline{u}_3 \\ \overline{s}_2 = \overline{u}_3 \\ \overline{s}_3 = \overline{u}_1 - \overline{u}_2 - \overline{u}_3 \end{cases} \Rightarrow P^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{pmatrix}$$

La matriz asociada a la aplicación lineal referida a las bases B' y S' es

$$A' = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -3 & 2 \\ -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & -6 \\ -3 & -9 \end{pmatrix}$$

$$(f(2,3))_{S'} = \begin{pmatrix} 1 & 1 \\ -2 & -6 \\ -3 & -9 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ -22 \\ -33 \end{pmatrix}.$$

Si no hubiéramos utilizado la fórmula encontrada, los pasos seguidos serían:

1. Expresar $\overline{x}' = (2, 3)$ respecto a $B : \overline{x} = Q\overline{x}'$;

$$\overline{x} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

2. Calcular la imagen por f de $(5, 1)$ respecto a $S : f(\overline{x}) = A\overline{x}$;

$$f(5,-1) = \begin{pmatrix} -3 & 2 \\ -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \begin{pmatrix} -17 \\ -11 \\ 5 \end{pmatrix}$$

3. Expresar $f(\overline{x}) = (-17, -11, 5)$ respecto a $S' : P^{-1}(f(\overline{x}))$;

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} -17 \\ -11 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ -22 \\ -33 \end{pmatrix}.$$

2.8. Operaciones elementales en una matriz. Matriz elemental

Sabemos que una matriz $A \in M_{n\times n}$ está asociada a un endomorfismo de V cuando se han fijado unas bases; decimos que son matrices equivalentes todas las matrices que representan el mismo endomorfismo aunque cambiemos las bases en V.

Las sucesivas operaciones que permiten el paso de una matriz a otra equivalente se llaman operaciones elementales en las filas o en las columnas, según donde se efectúen.

Dichas operaciones son:

1. Permutación, o intercambio entre filas.
2. Sustitución de una fila por su suma con una combinación lineal de otras.
3. Multiplicación de una fila por un escalar.

Ejemplo 2.8.1

La matriz $A = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 3 \\ 2 & -2 & 1 \end{pmatrix}$ es equivalente a la matriz $B = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1/2 \\ 0 & 1 & 3 \end{pmatrix}$ porque se puede pasar de una a otra mediante las siguientes operaciones elementales en las filas.

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 3 \\ 2 & -2 & 1 \end{pmatrix} \xrightarrow{F_3 \to F_3 - 2F_1} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & -1 \end{pmatrix} \xrightarrow{F_2 \leftrightarrow F_3} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 2 & -1 \\ 0 & 1 & 3 \end{pmatrix} \xrightarrow{F_2 \to 1/2\, F_2} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1/2 \\ 0 & 1 & 3 \end{pmatrix} = B$$

Obsérvese que todas las matrices que se van obteniendo en pasos intermedios son equivalentes.

De entre las matrices equivalentes a una matriz dada, algunas son especialmente interesantes, porque es más fácil trabajar con ellas que con la original. Esto ocurre con las matrices escalonadas por filas.

$A = (a_{ij})$ es una matriz escalonada por filas (también llamada *forma canónica de Hermite*) cuando:

1. El elemento $a_{11} = 1$ (aunque muchos autores admiten la misma definición imponiendo sólo que $a_{11} \neq 0$).
2. Cada fila, excepto la primera, se inicia con una sucesión de ceros.
3. La sucesión de ceros que inicia cada fila tiene algún cero más que la que inicia la fila anterior.
4. El primer elemento distinto de cero de cada fila, si lo hay, es un uno (o distinto de cero si sólo se pide que $a_{11} \neq 0$) y se llama cabecera de fila.
5. Las filas cuyos elementos son todos ceros, si las hay, son las últimas.
6. En cada columna, los elementos que quedan por debajo del que sirve de cabecera de fila, son ceros.

De forma intuitiva podemos resumir todas estas condiciones dibujando la forma que ha de tener una matriz escalonada:

De orden dos: $\begin{pmatrix} 1 & a_{12} \\ 0 & a_{22} \end{pmatrix}$; a_{22} puede ser 0 ó 1.

De orden tres: $\begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$; $\begin{cases} \text{Si } a_{22} = 1 \rightarrow a_{33} = 0 \text{ ó } a_{33} = 1 \\ \text{Si } a_{22} = 0 \rightarrow a_{23} = 0 \text{ ó } a_{23} = 1 \text{ y } a_{33} = 0 \end{cases}$

2.8.1. Teorema

Toda matriz se puede transformar en otra equivalente escalonada por filas mediante transformaciones elementales.

Una vez más, para ver que algo es posible es suficiente hacerlo; vamos a construir el algoritmo que permite esta transformación.

1. Conseguir que el primer elemento de la primera columna sea distinto de 0. Para ello, si $a_{11} = 0$, se intercambia la fila uno con alguna cuyo primer elemento sea distinto de 0.
2. Convertir el pivote en uno. Se sustituye la fila que contiene el elemento elegido en el paso anterior por ella misma dividida por dicho elemento. Llamaremos pivote al primer elemento distinto de cero (queda situado en la cabecera de la fila).
3. Hacer que sean ceros todos los elementos de la columna del pivote que quedan por debajo de él. Se consigue sustituyendo la fila en cuestión por ella misma menos su coeficiente multiplicado por la fila del pivote.
4. Se repite el procedimiento partiendo de la fila siguiente.

Ejemplo 2.8.2

Dada la matriz: $A = \begin{pmatrix} 2 & -1 & 0 \\ 3 & 2 & 1 \\ 0 & 7 & 2 \\ -3 & 5 & 1 \end{pmatrix}$, para encontrar una matriz equivalente de forma escalonada, procederemos de la siguiente manera:

1. Tomamos $a_{11} = 2$ como elemento a transformar en elemento pivote.
2. Para que el pivote sea 1, hacemos la sustitución $F_1 \rightarrow 1/2\ F_1$.
3. Para que queden ceros en la columna del pivote por debajo de él, hacemos las sustituciones: $F_2 \rightarrow F_2 - 3F_1$ y $F_4 \rightarrow F_4 + 3F_1$.
4. Volvemos a comenzar el ciclo tomando como elemento para convertir en pivote $a_{22} = 7/2$.

Las transformaciones a aplicar son: $F_2 \rightarrow 2/7F_2$; $F_3 \rightarrow F_3 - 7\ F_2$; $F_4 \rightarrow F_4 - 7/2\ F_2$.

$$A=\begin{pmatrix}2&-1&0\\3&2&1\\0&7&2\\-3&5&1\end{pmatrix}\to\begin{pmatrix}1&-1/2&0\\3&2&1\\0&7&2\\-3&5&1\end{pmatrix}\to\begin{pmatrix}1&-1/2&0\\0&7/2&1\\0&7&2\\0&7/2&1\end{pmatrix}\to\begin{pmatrix}1&-1/2&0\\0&1&2/7\\0&7&2\\0&7/2&1\end{pmatrix}\to\begin{pmatrix}1&-1/2&0\\0&1&2/7\\0&0&0\\0&0&0\end{pmatrix}$$

La matriz última es escalonada como queríamos.

El uso de matrices escalonadas permite obtener algunas conclusiones fácilmente, por ejemplo, el cálculo del número de vectores linealmente independientes que forman sus vectores filas y sus vectores columnas.

Si $A = (a_{ij})$ es una matriz de orden $m \times n$, puede ser considerada como un conjunto de m vectores fila o como un conjunto de n vectores columna.

De A se puede obtener una matriz escalonada B con un número r de filas no nulas:

$$A=\begin{pmatrix}a_{11}&a_{12}&\dots&\dots&\dots&\dots&\dots&a_{1n}\\a_{21}&a_{22}&\dots&\dots&\dots&\dots&\dots&a_{2n}\\\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots\\\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots\\\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots\\\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots\\\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots\\a_{m1}&a_{m2}&\dots&\dots&\dots&\dots&\dots&a_{mn}\end{pmatrix}\to\begin{pmatrix}a&\bullet&\bullet&\bullet&\bullet&\dots&\bullet&\bullet\\0&b&\bullet&\bullet&\bullet&\dots&\bullet&\bullet\\0&0&0&0&c&\bullet&\bullet&\bullet\\\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots\\0&0&0&0&0&\dots&r&\bullet\\0&0&0&0&0&\dots&0&0\\\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots\\0&0&0&0&0&\dots&0&0\end{pmatrix}=B$$

El símbolo "$\bullet$" expresa cualquier escalar, y debe entenderse que hay r filas (o escalones) no nulas, con tantos pivotes: $a, b, \dots, r$, como "escalones".

La relación que hay entre las filas linealmente independientes, las columnas linealmente independientes y el número de escalones viene dada por el teorema siguiente:

2.8.2. Teorema

> El número de vectores fila de una matriz A que son linealmente independientes coincide con el de columnas linealmente independientes y es igual al número de filas no nulas de su matriz escalonada

El teorema enunciado se puede desglosar en dos afirmaciones que demostraremos sucesivamente.

Esta técnica también es empleada a menudo en programación, cuando un problema se divide en subproblemas que son resueltos mediante subrutinas.

1. ***El número de vectores fila de una matriz, A, que son linealmente independientes coincide con el número de filas no nulas de su matriz escalonada, r.***

Demostración:

Los vectores fila de B se han obtenido mediante operaciones elementales realizadas en los vectores fila de A, y estas operaciones que consisten en cambiar el orden de los vectores, multiplicar los vectores por escalares y sustituir un vector por él más una combinación lineal de los otros, no modifican el número de vectores linealmente independientes según vimos en el capítulo uno. Por tanto, el número de vectores fila linealmente independientes que hay en B, es igual a r.

2. ***El número de vectores columna de una matriz A que son linealmente independientes coincide con el número de columnas no nulas, r, de su matriz escalonada.***

Demostración:

Para hacer esta demostración suponemos que el lector recuerda la teoría relativa a sistemas de ecuaciones que vio en cursos anteriores. Si no fuera así, puede saltar las demostraciones y volver a ellas cuando haya estudiado el capítulo cuatro.

Sean $\overline{c}_1, \overline{c}_2, \dots, \overline{c}_r$, los vectores columna de la matriz A correspondiente a las r columnas que contienen los pivotes en B; para ver que son linealmente independientes, de la expresión: $\lambda_1\overline{c}_1 + \lambda_2\overline{c}_2 + \dots + \lambda_r\overline{c}_r = \overline{0}$ se debe deducir $\lambda_1 = \lambda_2 = \dots = \lambda_r = 0$.

La expresión $\lambda_1\overline{c}_1 + \lambda_2\overline{c}_2 + \dots + \lambda_r\overline{c}_r = \overline{0}$ expresada componente a componente es un sistema de ecuaciones lineal homogéneo equivalente al planteado con las columnas correspondientes de la matriz escalonada B, ya que las operaciones que se han hecho son las operaciones elementales, y en el sistema B es evidente que la única solución es la trivial: $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 0$.

Además, no puede haber $r + 1$ vectores columnas de A que sean independientes, ya que la matriz escalonada B, tiene solo r filas no nulas, pasaríamos el vector columna $\overline{c}_{r+1}$ detrás del signo igual, λ_{r+1} actuaría de parámetro, y como consecuencia, en el sistema de ecuaciones anterior habría infinitas soluciones para λ_i.

Ejemplo 2.8.3

En el ejemplo 2.8.2 la matriz $A = \begin{pmatrix} 2 & -1 & 0 \\ 3 & 2 & 1 \\ 0 & 7 & 2 \\ -3 & 5 & 1 \end{pmatrix}$ se puede considerar como el conjunto de vectores fila: $F = \{(2, -1, 0), (3, 2, 1), (0, 7, 2), (-3, 5, 1)\}$, o el conjunto de vectores columna $C = \{(2, 3, 0, -3), (-1, 2, 7, 5), (0, 1, 2, 1)\}$.

De la matriz A se pasa a la B mediante transformaciones elementales, y por tanto, podemos afirmar que, el número de vectores linealmente independientes de F es igual al número de vectores linealmente independientes de C, igual al número de escalones de B, igual a dos.

Si la matriz A es cuadrada, puede obtenerse como matriz escalonada por filas una matriz triangular superior, aunque algún elemento de la diagonal puede ser cero. Además, si todas las filas son linealmente independientes, la diagonal principal de la escalonada tendrá todos los elementos distintos de cero.

En este último caso, se puede seguir transformando la matriz obtenida hasta llegar mediante operaciones elementales en sus filas, a una matriz diagonal y posteriormente a la matriz unidad I.

2.8.3. Teorema

Si el rango de los vectores fila de una matriz cuadrada A es igual al orden de la matriz, entonces, se puede transformar A en la matriz unidad mediante operaciones elementales en sus filas

Para demostrar este teorema es suficiente construir el algoritmo que permite el paso de una a otra, de la misma manera que se hizo en el teorema anterior.

Las operaciones elementales que hemos visto, se pueden realizar multiplicando la matriz dada por una matriz de una forma determinada, las matrices que sirven para expresar el cambio se llaman matrices elementales.

2.8.4. Definición de matriz elemental

Se llama matriz elemental a la matriz cuadrada que se obtiene de la matriz unidad al efectuar una sola operación elemental en las filas (o columnas)

2.8.5. Teorema

Si $A \in M_{nxn}$ y $E \in M_{nxn}$ es una matriz elemental y T es una operación elemental cualquiera $\Rightarrow \begin{cases} I \xrightarrow{T} E \\ A \xrightarrow{T} EA \end{cases}$

$EA \in M_{nxn}$ es la matriz que se obtiene de A mediante la misma operación elemental que permite obtener E desde I.

Demostración:

Se puede hacer comprobando que es cierta la afirmación para cada una de las operaciones elementales al construir la matriz elemental correspondiente.

Demostrar por casos, sería equivalente en Informática a la verificación de un programa con condicionantes.

1. Supongamos que T es la operación elemental "intercambio de filas", y E la matriz elemental que se obtiene de I al intercambiar dos filas, si elegimos las filas 1 y 2 no restamos generalidad al resultado (podíamos haber elegido cualquier otro par).

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},\ E=\begin{pmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

$$EA=\begin{pmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{11} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}=\begin{pmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

que es la matriz que se deduce de A al efectuar en ella la misma operación elemental en las filas que habíamos efectuado en E.

2. Supongamos que T es la operación elemental "multiplicar una fila por un escalar λ", si la fila multiplicada es la dos, por ejemplo, la matriz elemental E obtenida desde I es:

$$E=\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}\Rightarrow EA=\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

3. Si T representa la operación elemental sustituir una fila (elijamos la uno) por ella misma más una combinación lineal de otras (elijamos λ veces la dos, por ejemplo), la matriz elemental E obtenida desde I al efectuar esa operación elemental es

$$E=\begin{pmatrix} 1 & \lambda & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

$$\Rightarrow EA=\begin{pmatrix} 1 & \lambda & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}=\begin{pmatrix} a_{11}+\lambda a_{21} & a_{12}+\lambda a_{22} & \dots & a_{1n}+\lambda a_{2n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Ejemplo 2.8.4

La matriz unidad I se convierte en la matriz elemental E mediante la siguiente operación elemental en las filas: $I=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\xrightarrow[F_3 \to F_3+F_1+F_2]{}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}=E.$

Cada operación elemental en las filas de I tiene, por tanto, su matriz elemental correspondiente. En el ejemplo, E es la matriz elemental correspondiente a la operación elemental "Sustituir la fila tres por la fila tres más la fila uno más la fila dos".

Si se realizan las operaciones elementales en las columnas se obtiene un resultado análogo, pero el producto es AE en vez de EA.

2.8.6. Teorema

Si $E \in M_{n\times n}$ es una matriz elemental y $E' \in M_{n\times n}$ es la matriz elemental de la operación inversa entonces, ambas son regulares y $E' = E^{-1}$

Demostración:

Es evidente, sólo hay que observar la forma de los tres tipos de matrices elementales, cada uno de ellos corresponde a una operación elemental realizada en las filas de la matriz unidad.

2.8.7. Teorema

$A \in M_{n \times n}$ tiene inversa $\Leftrightarrow$ $rg(A) = n$

Demostración:

$\Rightarrow$) Si A tiene inversa, es equivalente a la matriz unidad del mismo orden; los vectores fila independientes que hay en I son n (ya que el rango no varía por la aplicación de operaciones elementales) $\Rightarrow$ En A también hay n vectores fila independientes.

$\Leftarrow$) Si el rango de A es n $\Rightarrow$ en A hay n filas independientes $\Rightarrow$ A se puede transformar en la matriz unidad mediante operaciones elementales $\Rightarrow$ A tiene inversa.

Además es muy fácil el cálculo de dicha inversa: Sean $E_1, E_2, ..., E_n$ las matrices elementales correspondientes a cada operación elemental que llevan a la matriz A hasta la unidad I $\Rightarrow$ $E_n ... E_2 \cdot E_1 \cdot A = I$, multiplicando a la derecha por A^{-1} ambos lados de la igualdad $\Rightarrow$ $E_n ... E_2 \cdot E_1 \cdot = A^{-1}$.

EJERCICIOS

Ejercicio 2.1

Determínese si $f: \mathbb{R}^3 \to \mathbb{R}^2 / f(x_1, x_2, x_3) = (x_1 - 2x_2, x_2 + x_3)$ es una transformación lineal.

Solución:

Conserva la suma: $f(\overline{x} + \overline{y}) = f(\overline{x}) + (\overline{y})$:

$(x_1, x_2, x_3) \to (x_1 - 2x_2, x_2 + x_3)$
$(y_1, y_2, y_3) \to (y_1 - 2y_2, y_2 + y_3)$
$(x_1 + y_1, x_2 + y_2, x_3 + y_3) \to (x_1 + y_1 - 2x_2 - 2y_2, x_2 + y_2 + x_3 + y_3) = (x_1 - 2x_2, x_2 + x_3) + (y_1 - 2y_2, y_2 + y_3)$

Conserva el producto por un escalar: $f(\lambda\overline{x}) = \lambda f(\overline{x})$:

$\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3) \to (\lambda x_1 - 2\lambda x_2, \lambda x_2 + \lambda x_3)$
$(x_1, x_2, x_3) \to (x_1 - 2x_2, x_2 + x_3)$; $\lambda(x_1 - 2x_2, x_2 + x_3) = (\lambda x_1 - 2\lambda x_2, \lambda x_2 + \lambda x_3)$

Por conservar la suma y el producto es una transformación lineal.

Ejercicio 2.2

Determínese si $f: \mathbb{R}^3 \to \mathbb{R}^3 / f(x_1, x_2, x_3) = (2x_1, x_1 + x_2, 0)$ es una aplicación lineal, y si lo es, estúdiense su núcleo y su imagen.

Solución:

Es una aplicación o transformación del espacio vectorial $\mathbb{R}^3$ en sí mismo. Las aplicaciones lineales de un espacio vectorial en sí mismo se llaman **endomorfismos**.

Conserva la suma: $f(\overline{x} + \overline{y}) = f(\overline{x}) + f(\overline{y})$

$(x_1, x_2, x_3) \to (2x_1, x_1 + x_2, 0)$
$(y_1, y_2, y_3) \to (2y_1, y_1 + y_2, 0)$
$(x_1 + y_1, x_2 + y_2, x_3 + y_3) \to (2x_1 + 2y_1, x_1 + x_2 + y_1 + y_2, 0) = (2x_1, x_1 + x_2, 0) +$
$+ (2y_1, y_1 + y_2, 0)$

Conserva el producto por un escalar: $f(\lambda\overline{x}) = \lambda f(\overline{x})$

(x_1, x_2, x_3) $(2x_1, x_1 + x_2, 0)$
$\lambda(2x_1, x_1 + x_2, 0) = (2\lambda x_1, \lambda x_1 + \lambda x_2, 0)$
$\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3) \to (2\lambda x_1, \lambda x_1 + \lambda x_2, 0)$

Observación: estas dos condiciones se podían haber comprobado a la vez utilizando la caracterización de aplicaciones lineales.

- Por conservar la suma y el producto, **la transformación es lineal.**

El núcleo es $f^{-1}(0, 0, 0) = \{(x_1, x_2, x_3)/(2x_1, x_1 + x_2, 0) = (0, 0, 0)\}$

$$\left.\begin{array}{r} 2x_1 = 0 \\ x_1 + x_2 = 0 \end{array}\right\} \Rightarrow x_1 = -x_2 = 0$$

Un subespacio vectorial se puede determinar nediante sus ecuaciones cartesianas, paramétricas o mediante una base.

- Las **ecuaciones cartesianas del núcleo** son: $\begin{cases} x_1 + x_2 = 0 \\ x_1 = 0 \end{cases}$

- Las **ecuaciones paramétricas del núcleo** son: $\begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = \lambda \end{cases}$

Ya hemos encontrado unas ecuaciones, determinemos una base:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

- El vector (0, 0, 1) es un sistema generador del núcleo; por ser sólo uno, es independiente $\Rightarrow \{(0, 0, 1)\}$ es **una base del núcleo.**
- El **núcleo tiene dimensión 1** (recordemos que es el número de elementos de sus bases = número de parámetros en las ecuaciones paramétricas).

$\mathrm{Im}(f) = \{f(x_1, x_2, x_3) = (2x_1, x_1 + x_2, 0)\}$

$$\begin{pmatrix} 2x_1 \\ x_1 + x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix};$$

- Los vectores (2, 1, 0) y (0, 1, 0) generan la imagen de f y son linealmente independientes, por tanto $B = \{(2, 1, 0), (0, 1, 0)\}$ **es una base de** $\mathrm{Im}(f)$.
- **La dimensión de la imagen** es 2
- Unas **ecuaciones paramétricas de la imagen** son:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow x_1 = 2\lambda;\ x_2 = \lambda + \mu;\ x_3 = 0$$

- Las **ecuaciones cartesianas de la imagen** son: $x_3 = 0$

Comprobamos que se cumple la relación entre las dimensiones del núcleo, de la imagen y del espacio origen $\mathbb{R}^3$: $\dim(\mathbb{R}^3) = \dim \mathrm{Nuc}(f) + \dim \mathrm{Im}(f)$.

Ejercicio 2.3

Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 / f(3, 1) = (1, 2); f(-1, 0) = (1, 1)$ es una aplicación lineal. Se pide: a) Determinar la matriz asociada a f tomando la base canónica como base de $\mathbb{R}^2$. b) Las ecuaciones de la aplicación. c) La imagen de (2, 4).

Solución:

a) La base canónica de $\mathbb{R}^2$ es $B = \{(1, 0), (0, 1)\}$, para calcular la matriz asociada a f hay que calcular $f(1, 0)$ y $f(0, 1)$:

$$(1, 0) = x(3, 1) + y(-1, 0) \Rightarrow \begin{cases} 1 = 3x - y \\ 0 = x \end{cases} \Rightarrow x = 0;\ y = -1$$

Por ser f lineal, se verifica:

$$f(1, 0) = f[0(3, 1) - 1(-1, 0)] = 0f(3, 1) - 1f(-1, 0) = 0(1, 2) - 1(1, 1) = (-1, -1)$$

$$(0, 1) = x(3, 1) + y(-1, 0) \Rightarrow \begin{cases} 0 = 3x - y \\ 1 = x \end{cases} \Rightarrow x = 1;\ y = 3$$

Por ser f lineal, se verifica:

$f(0, 1) = f[1(3, 1) + 3(-1, 0)] = 1f(3, 1) + 3f(-1, 0) = 1(1, 2) + 3(1, 1) = (4, 5)$

Matriz asociada a f: $\begin{pmatrix} -1 & 4 \\ -1 & 5 \end{pmatrix}$

b) Ecuaciones de la aplicación:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -1 & 4 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{cases} y_1 = -x_1 + 4x_2 \\ y_2 = -x_1 + 5x_2 \end{cases}$$

c) La imagen de (2, 4) se puede calcular utilizando la definición de aplicación lineal y las imágenes de los elementos de la base.

$(2, 4) = 2(1, 0) + 4(0, 1)$.

Por ser f lineal, se verifica:

$f[2(1, 0) + 4(0, 1)] = 2f(1, 0) + 4f(0, 1) = 2(-1, -1) + 4(4, 5) = (14, 18)$

También podíamos haber utilizado las ecuaciones de la aplicación:

$$\begin{matrix} y_1 = -x_1 + 4x_2 \\ y_2 = -x_1 + 5x_2 \end{matrix} \Rightarrow \begin{cases} y_1 = -2 + 4.4 = 14 \\ y_2 = -2 + 5.4 = 18 \end{cases}$$

O podíamos haber determinado $f(2, 4)$ a partir de la matriz:

$$f\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 & 4 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 14 \\ 18 \end{pmatrix}$$

Ejercicio 2.4

Si $f: \mathbb{R}^3 \to \mathbb{R}^2 / f(1, 0, 0) = (0, 1); f(0, 1, 0) = (1, 1); f(0, 0, 1) = (1, 0)$ es una aplicación lineal. Se pide: a) Determinar la matriz asociada a f tomando las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$. b) Determinar las ecuaciones de la imagen del subespacio $x_1 + x_2 - x_3 = 0$.

Solución:

a) Matriz asociada a f:

Es la matriz cuyas columnas son las imágenes de la base dada en $\mathbb{R}^3$:

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

Las ecuaciones de la aplicación son: $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

b) Unas ecuaciones paramétricas del subespacio $x_1 + x_2 - x_3 = 0$ son:

$$\begin{cases} x_1 = \lambda \\ x_2 = \mu \\ x_3 = \lambda + \mu \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$\{(1, 0, 1), (0, 1, 1)\}$ es una base de $x_1 + x_2 - x_3 = 0$, por tanto $\{f(1, 0, 1), f(0, 1, 1)$ es un sistema de generadores de $f(x_1 + x_2 - x_3 = 0)$, (recordemos que las aplicaciones lineales pueden no conservar la independencia lineal ni la dimensión de los subespacios).

$$\Rightarrow f\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; f\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$\Rightarrow f(x_1 + x_2 - x_3 = 0)$ está generado por (1, 1)

Unas ecuaciones paramétricas de $f(x_1 + x_2 - x_3 = 0)$ son: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Leftrightarrow$

$$\Leftrightarrow \begin{cases} x_1 = 2\lambda \\ x_2 = \lambda \end{cases}$$

Unas ecuaciones paramétricas son: $x_1 = 2x_2$.

Ejercicio 2.5

Determínese las condiciones que deben cumplir dos matrices de igual tamaño para que su producto sea una matriz simétrica.

Solución:

$$\left.\begin{array}{l}\text{Tienen igual tamaño}\\ \text{Se pueden multiplicar matrices de orden } m\times p \text{ y } p\times m\end{array}\right\} \Rightarrow m=p=n$$

- La primera condición es que sean cuadradas.

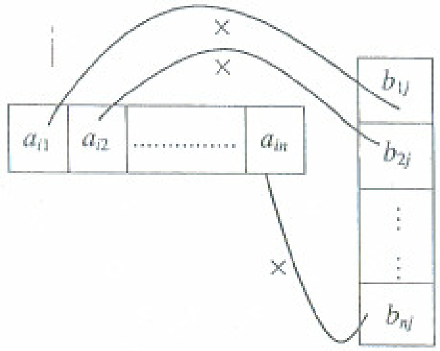
Si $C = A_{nxn}B_{nxn} = (c_{ij}) \in M_{nxn}$, para que el producto sea simétrico debe cumplirse $c_{ij} = c_{ji}$ siendo $c_{ij} = \sum_{r=1}^{n} a_{ir}b_{rj}$ y $c_{ji} = \sum_{r=1}^{n} a_{jr}b_{ri}$ para todos los valores de i, j.

- Luego la segunda condición es que se verifique que, la fila i de A por la columna j de B sea igual a la fila j de A por la columna i de B.

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = a_{j1}b_{1i} + a_{j2}b_{2i} + \ldots + a_{jn}b_{ni} = c_{ji}$$

Recordemos la forma de multiplicar matrices:

Fila i de A



Columna j de B

Ejercicio 2.6

Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ una aplicación lineal tal que $f(x_1, x_2, x_3) = (x_1 - x_3, x_2 + x_3)$, y sea $g : \mathbb{R}^2 \to \mathbb{R}^3$, una aplicación lineal tal que $g(x_1, x_2) = (x_1 + x_2, x_1 - x_2, 2x_1 + 3x_2)$. Se pide determinar las ecuaciones de las aplicaciones $f \circ g$ y $g \circ f$, y comprobar, mediante las matrices asociadas a ambas aplicaciones respecto a las bases canónicas, que el producto de matrices no es conmutativo.

Solución:

En este caso se pueden realizar $f \circ g$ y $g \circ f$:

$$\begin{tikzcd} \mathbb{R}^3 \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & \mathbb{R}^2 \arrow[r, "g"] \arrow[rr, bend right, "f\circ g"'] & \mathbb{R}^3 \arrow[r, "f"] & \mathbb{R}^2 \end{tikzcd}$$

Ecuaciones de $g \circ f$:

$$\begin{aligned} &f(x_1, x_2, x_3) = (x_1 - x_3, x_2 + x_3) = (y_1, y_2)\\ &g(f(x_1, x_2, x_3)) = g(y_1, y_2) = (y_1 + y_2, y_1 - y_2, 2y_1 + 3y_2) =\\ &= (x_1 - x_3 + x_2 + x_3, x_1 - x_3 - (x_2 + x_3), 2(x_1 - x_3) + 3(x_2 + x_3)) =\\ &= (x_1 + x_2, x_1 - x_2 - 2x_3, 2x_1 + 3x_2 + x_3)\end{aligned}$$

Ecuaciones de $f \circ g$:

$$g(x_1, x_2) = (x_1 + x_2, x_1 - x_2, 2x_1 + 3x_2) = (y_1, y_2, y_3)$$
$$f(g(x_1, x_2)) = f(y_1, y_2, y_3) = (y_1 - y_3, y_2 + y_3) = (x_1 + x_2 - (2x_1 + 3x_2), x_1 - x_2 + 2x_1 + 3x_2) = (-x_1 - 2x_2, 3x_1 + 2x_2)$$

Es evidente que las aplicaciones $g \circ f$: $\mathbb{R}^3 \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^3$ y $f \circ g$: $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^3 \xrightarrow{f} \mathbb{R}^2$ no pueden coincidir al tener distintos los conjuntos original e imagen.

Matriz asociada a f:

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 0 \end{pmatrix}, \text{ porque } \begin{cases} (1, 0, 0) \to (1, 0) \\ (0, 1, 0) \to (0, 1) \\ (0, 0, 1) \to (-1, 1) \end{cases}$$

Matriz asociada a g:

$$B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 3 \end{pmatrix}, \text{ porque } \begin{cases} (1, 0) \to (1, 1, 2) \\ (0, 1) \to (1, -1, 3) \end{cases}$$

Matriz asociada a $g \circ f$:

$$BA = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & -2 \\ 2 & 3 & 1 \end{pmatrix},$$

Expresión analítica de $g \circ f$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Matriz asociada a $f \circ g$:

$$AB = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 3 & 2 \end{pmatrix},$$

Expresión analítica de $f \circ g$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Como ya sabíamos, $A\ B \neq B\ A$ porque $f \circ g \neq g \circ f$

Ejercicio 2.7

Consideramos en $\mathbb{R}^3$ las bases: $B = \{(1, 2, 3), (3, 4, 0), (1, 1, 0)\}$ y $B' = \{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$. Se pide encontrar las matrices de cambio de base, comprobar que son inversas, escribir las ecuaciones del cambio de base, y calcular las coordenadas de un vector en la base B sabiendo que en B' sus coordenadas son (2, 1, 1).

Solución:

- Coordenadas de B' respecto a B:

$$\alpha(1, 2, 3) + \beta(3, 4, 0) + \gamma(1, 1, 0) = (1, 1, 0) \Rightarrow \alpha = 0; \quad \beta = 0; \quad \gamma = 1$$
$$\alpha(1, 2, 3) + \beta(3, 4, 0) + \gamma(1, 1, 0) = (0, 1, 1) \Rightarrow \alpha = 1/3; \quad \beta = 2/3; \quad \gamma = -7/3$$
$$\alpha(1, 2, 3) + \beta(3, 4, 0) + \gamma(1, 1, 0) = (1, 0, 1) \Rightarrow \alpha = 1/3; \quad \beta = -4/3; \gamma = 14/3$$

Por tanto:

$(\mathbb{R}^3, B') \to (\mathbb{R}^3, B)$
$(1, 1, 0) \to (0, 0, 1)$
$(0, 1, 1) \to (1/3, 2/3, -7/3)$
$(1, 0, 1) \to (1/3, -4/3, 14/3)$

- Matriz para pasar de las coordenadas de un vector respecto a B' a las coordenadas respecto a B:

$$Q = \begin{pmatrix} 0 & 1/3 & 1/3 \\ 0 & 2/3 & -4/3 \\ 1 & -7/3 & 14/3 \end{pmatrix}$$

- Coordenadas de B respecto a B':

$$\alpha(1, 1, 0) + \beta(0, 1, 1) + \gamma(1, 0, 1) = (1, 2, 3) \Rightarrow \alpha = 0; \quad \beta = 2; \quad \gamma = 1$$
$$\alpha(1, 1, 0) + \beta(0, 1, 1) + \gamma(1, 0, 1) = (3, 4, 0) \Rightarrow \alpha = 7/2; \quad \beta = 1/2; \quad \gamma = -1/2$$
$$\alpha(1, 1, 0) + \beta(0, 1, 1) + \gamma(1, 0, 1) = (1, 1, 0) \Rightarrow \alpha = 1, \quad \beta = 0; \quad \gamma = 0$$

- Matriz para pasar de las coordenadas de un vector respecto a B a las coordenadas respecto a B':

$$P=\begin{pmatrix}0 & 7/2 & 1\\ 2 & 1/2 & 0\\ 1 & -1/2 & 0\end{pmatrix}$$

- Las matrices Q y P son inversas:

$$QP=\begin{pmatrix}0 & 1/3 & 1/3\\ 0 & 2/3 & -4/3\\ 1 & -7/3 & 14/3\end{pmatrix}\begin{pmatrix}0 & 7/2 & 1\\ 2 & 1/2 & 0\\ 1 & -1/2 & 0\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}=$$

$$=\begin{pmatrix}0 & 7/2 & 1\\ 2 & 1/2 & 0\\ 1 & -1/2 & 0\end{pmatrix}\begin{pmatrix}0 & 1/3 & 1/3\\ 0 & 2/3 & -4/3\\ 1 & -7/3 & 14/3\end{pmatrix}=PQ$$

- Ecuaciones del cambio de base:

Las ecuaciones que permiten pasar de las coordenadas de un vector respecto a la base B' a las coordenadas respecto a la base B son:

$$\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0 & 1/3 & 1/3\\ 0 & 2/3 & -4/3\\ 1 & -7/3 & 14/3\end{pmatrix}\begin{pmatrix}x'_1\\ x'_2\\ x'_3\end{pmatrix}\Leftrightarrow \overline{x}=Q\overline{x}'$$

Las ecuaciones que permiten pasar de las coordenadas de un vector respecto a la base B a las coordenadas respecto a la base B' son:

$$\begin{pmatrix}x'_1\\ x'_2\\ x'_3\end{pmatrix}=\begin{pmatrix}0 & 7/2 & 1\\ 2 & 1/2 & 0\\ 1 & -1/2 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}\Leftrightarrow \overline{x}=Q^{-1}\overline{x}$$

Para calcular las coordenadas de un vector en la base B conociéndolas en la otra, aplicaremos las ecuaciones oportunas, que en este caso son:

$$\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0 & 1/3 & 1/3\\ 0 & 2/3 & -4/3\\ 1 & -7/3 & 14/3\end{pmatrix}\begin{pmatrix}2\\ 1\\ 1\end{pmatrix}\Rightarrow\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}2/3\\ -2/3\\ 13/3\end{pmatrix}$$

Para cambiar de coordenadas un solo vector no está justificado el esfuerzo de calcular las ecuaciones del cambio. Lo haríamos directamente:

$\alpha(1,2,3)+\beta(3,4,0)+\gamma(1,1,0)=2(1,1,0)+1(0,1,1)+1(1,0,1)\Rightarrow \alpha=$ $=2/3;\ \beta=-2/3;\ \gamma=13/3$

Ejercicio 2.8

Dada la matriz $A=\begin{pmatrix}0 & 3 & 1\\ 1 & 2 & 0\\ 1 & 1 & 3\end{pmatrix}$, se pide: a) Encontrar una matriz equivalente escalonada por filas. b) Transformar la matriz obtenida en la matriz identidad mediante operaciones elementales.

Solución:

a)

$$A=\begin{pmatrix}1 & 3 & 1\\ 1 & 2 & 0\\ 1 & -1 & 3\end{pmatrix}\xrightarrow[\substack{F_2\to F_2-F_1\\ F_3\to F_3-F_1}]{}\begin{pmatrix}1 & 3 & 1\\ 0 & -1 & -1\\ 0 & -4 & 2\end{pmatrix}\xrightarrow[F_2\to -F_2]{}\begin{pmatrix}1 & 3 & 1\\ 0 & 1 & 1\\ 1 & -4 & 2\end{pmatrix}\xrightarrow[F_3\to F_3+4F_2]{}\begin{pmatrix}1 & 3 & 1\\ 0 & 1 & 1\\ 0 & 0 & 6\end{pmatrix}\xrightarrow[F_3\to 1/6\,F_3]{}\begin{pmatrix}1 & 3 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}$$

b)

$$\begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 6 \end{pmatrix} \xrightarrow[F_3 \to 1/6\, F_3]{} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[\substack{F_1 \to F_1 - F_3 \\ F_2 \to F_2 - F_3}]{} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[F_1 \to F_1 - 3F_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Ejercicio 2.9

Demuéstrese que cuando se verifica $a^2 + b^2 + c^2 = 1$ con $a, b, c \in \mathbb{R}$, todas las potencias de la matriz $\begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}$ son iguales a ella misma.

Solución:

Queremos probar que $\begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}^n = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}, \forall n \in \mathbb{R}$

Las tres etapas del método de inducción son:

- Comprobar que es cierto para $n = 1$.

Es evidente que $\begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix} = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}$

- Suponer que es cierto para $n - 1$.

$$\begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}^{n-1} = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}$$

- Comprobar que también es cierto para n.

$$\begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}^{n-1} \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix} = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix} \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix} =$$

$$= \begin{pmatrix} a^4 + a^2b^2 + a^2c^2 & a^3b + ab^3 + abc^2 & a^3c + ab^2c + ac^3 \\ a^3b + ab^3 + abc^2 & a^2b^2 + b^4 + b^2c^2 & a^2bc + b^3c + bc^3 \\ a^3c + ab^2c + ac^3 & a^2bc + b^3c + bc^3 & a^2c^2 + b^2c^2 + c^4 \end{pmatrix} =$$

$$= \begin{pmatrix} a^2(a^2+b^2+c^2) & ab(a^2+b^2+c^2) & ac(a^2+b^2+c^2) \\ ab(a^2+b^2+c^2) & b^2(a^2+b^2+c^2) & ac(a^2+b^2+c^2) \\ ac(a^2+b^2+c^2) & bc(a^2+b^2+c^2) & c^2(a^2+b^2+c^2) \end{pmatrix} = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}$$

Cuando una matriz A tiene la propiedad $A^n = A$ se llama idempotente.

Ejercicio 2.10

Demuéstrese que si A es una matriz que verifica la relación $A^2 - A - I = 0$, entonces existe A^{-1}.

Solución:

$A^2 - A - I = 0 \Rightarrow I = A^2 - A = A(A - I) \Rightarrow A^{-1} = A - I$

Existe y vale $A - I$.

CAPÍTULO 3

DETERMINANTE DE UNA MATRIZ CUADRADA

INTRODUCCIÓN

Los determinantes aparecen en Matemáticas en relación con los problemas de eliminación y resolución de sistemas lineales.

En 1693, Leibniz usó un conjunto sistemático de índices para designar los coeficientes de un sistema de tres ecuaciones lineales con dos incógnitas. Por un proceso de eliminación, obtuvo una expresión (la resultante del sistema, que no es más que el determinante de los coeficientes), cuya anulación era condición necesaria y suficiente para que el sistema tuviera solución.

La búsqueda de fórmulas explícitas de resolución de sistemas lineales, iniciada por MacLaurin en 1729 parece ser el origen más directo de la teoría de determinantes. Vandermonde, en una Memoria aparecida en 1772, fue el primero en dar una exposición sistemática y detallada de la teoría. Demuestra las propiedades generales de los determinantes, como el hecho de que son funciones multilineales alternadas de sus filas y de sus columnas, la igualdad de un determinante y su transpuesto y también la regla para desarrollar un determinante por una fila o columna. Como la mayor parte de los algebristas de la época, Vandermonde se contenta con verificar las propiedades para valores pequeños de n.

El nombre de determinante se debe a Cauchy, quien también introdujo el uso de dobles subíndices y la disposición de los elementos en un cuadrado de n^2 puntos (las líneas verticales fueron introducidas posteriormente por Cayley). Como en muchos otros temas, es Cauchy quien establece la teoría general, dando demostraciones rigurosas y completas. Establece también la fórmula general que da el producto de dos determinantes como otro determinante así como las relaciones entre los determinantes formados con los menores de distintos órdenes de un determinante dado.

A partir del trabajo de Cauchy, los determinantes van a convertirse en una herramienta básica en todos los problemas de Álgebra Lineal, y su estudio fue objeto preferente de atención de los matemáticos del siglo XIX.

CAUCHY (1789-1879)

Cauchy nació el 21 de agosto de 1789 en un París convulso por la Revolución Francesa y las revueltas populares.

La ejecución de rey en 1793 en la guillotina, los cambios impulsados por Robespierre que también termina en la guillotina y la llegada de Napoleón, propician una época de organización del estado y de estabilidad para los científicos.

En 1814, Napoleón abdica y se instaura en Francia, de nuevo, la monarquía en la persona de Luis XVIII. Las contrarreformas que inicia el rey, producen un gran descontento popular. Napoleón se hace con el poder de nuevo. Las monarquías europeas forman un ejército y derrotan a Napoleón en Waterloo.

En este período, desarrolla Cauchy su actividad, que a grandes rasgos es la siguiente:

En 1805 entró en la Escuela Politécnica de París donde se formó.

A lo largo de su vida fue coherente con sus ideas políticas y religiosas, así, cuando en 1830 fue requerido para que jurase lealtad al nuevo régimen se negó a hacerlo y perdió los puestos académicos que tenía.

Su brillantez como científico no le acompañaba como docente y cuando daba clases, sus alumnos se quejaban de que no podían entenderlas. Entre otros alumnos, tuvo al nieto de Carlos X, pero dadas las limitaciones de uno como alumno y de otro como profesor, el resultado no fue bueno y Cauchy regresó a París en 1838, volvió a la Academia pero no a la enseñanza, porque se negó a jurar lealtad al régimen.

Regresó a su puesto en la Universidad en 1848, cuando Luis Felipe fue derrocado.

Murió: 23 de mayo de 1857 en Sceaux (París-Francia).

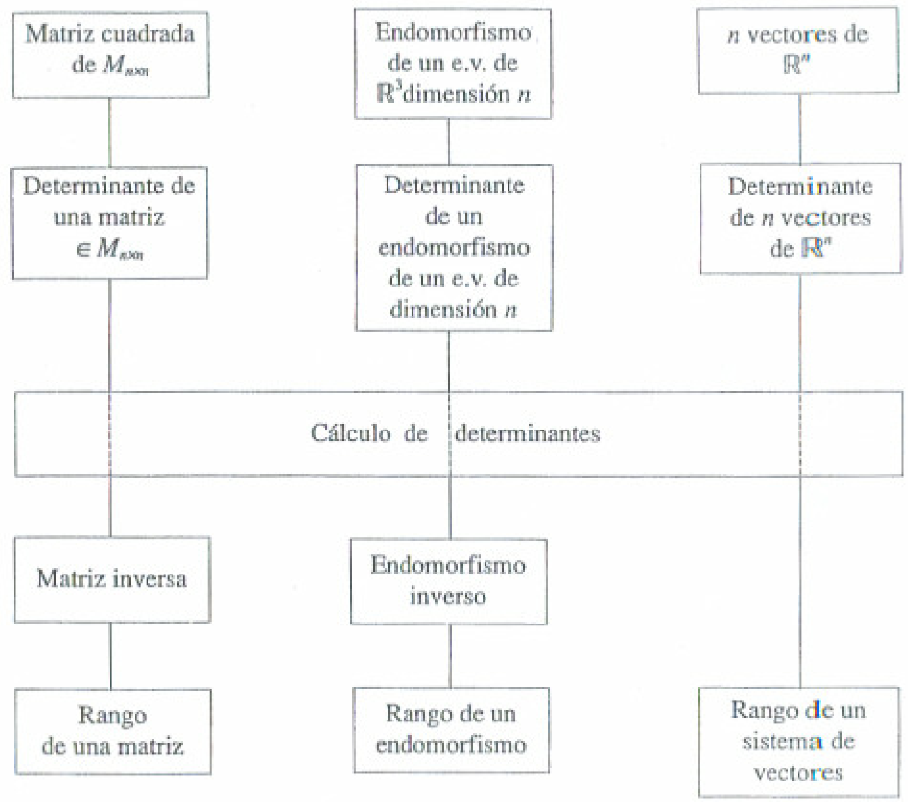
Con 789 trabajos distribuidos en 27 volúmenes, es uno de los matemáticos más prolíficos de la historia.

Su nombre aparece ligado a gran cantidad de campos: teoría de funciones complejas, series, ecuaciones, solución de ecuaciones diferenciales, creó la teoría de las funciones analíticas, desarrolló la teoría de determinantes.

Gracias a él, el análisis infinitesimal adquiere bases sólidas, adoptando métodos rigurosos que el mismo describe:

"He tratado de dar a los métodos todo el rigor que se exige en geometría, sin acudir jamás a los argumentos tomados de la generalidad del álgebra..."

CAPÍTULO 3


Matriz cuadrada de $M_{n\times n}$
Endomorfismo de un e.v. de $\mathbb{R}^3$dimensión n
n vectores de $\mathbb{R}^n$
Determinante de una matriz $\in M_{n\times n}$
Determinante de un endomorfismo de un e.v. de dimensión n
Determinante de n vectores de $\mathbb{R}^n$
Cálculo de determinantes
Matriz inversa
Endomorfismo inverso
Rango de una matriz
Rango de un endomorfismo
Rango de un sistema de vectores

3.1. Determinante de una matriz cuadrada

Hasta aquí hemos estado trabajando con vectores y matrices, y las herramientas de que hemos dispuesto para su estudio han sido poco eficientes. En este capítulo vamos a utilizar los determinantes como herramienta que facilita muchas de las tareas cuyo estudio hasta ahora ha sido poco ágil, como la dependencia e independencia lineal, el rango de una matriz, el cálculo de la matriz inversa, etc.

Vamos a estudiar el determinante de la matriz asociada a un endomorfismo, por ello, es conveniente repasar con detenimiento *[2.6]* antes de seguir adelante.

Dos conjuntos formados por los mismos elementos, pero ordenado de forma diferente, decimos que son dos permutaciones o sustituciones diferentes del mismo conjunto.

3.1.1. Definición de permutación

σ es una permutación de $S = \{1, 2, ..., n\} \Leftrightarrow \sigma$ es una aplicación biyectiva de $S \to S$

La imagen de cada elemento $i \in S$ es $\sigma(i) \in S$.

Es frecuente expresar la permutación de la siguiente manera: $\begin{pmatrix} 1 & 2 & \dots & n \\ \downarrow & \downarrow & \dots & \downarrow \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$,

o simplemente $\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(1) & \dots & \sigma(n) \end{pmatrix}$.

Dos elementos están en inversión cuando el orden en que figuran es distinto del de la permutación principal $(1, 2, ..., n)$

Una permutación se llama par si tiene un número par de inversiones, e impar si tiene un número impar de inversiones.

A las permutaciones pares se les asigna el signo +, y a las permutaciones impares se les asigna el signo –.

Ejemplo 3.1.1

Si σ es una permutación de $S = \{1, 2, 3\}$, en el gráfico siguiente se detallan todas las permutaciones posibles, así, como, el número de inversiones, el tipo de permutación y el signo.

Original $\xrightarrow{\sigma}$ *Imagen* $\rightarrow$	*N.º de inversiones* $\rightarrow$	*Tipo de permutación*	*Signo*
$1, 2, 3 \xrightarrow{\sigma} 1, 2, 3$	0	Par	+
$1, 2, 3 \xrightarrow{\sigma} 1, \underline{3, 2}$	1	Impar	–
$1, 2, 3 \xrightarrow{\sigma} \underline{2, 1}, 3$	1	Impar	–
$1, 2, 3 \xrightarrow{\sigma} 2, \underline{3, 1}$	2	Par	+
$1, 2, 3 \xrightarrow{\sigma} \underline{3, 1}, 2$	2	Par	+
$1, 2, 3 \xrightarrow{\sigma} 3, \underline{2, 1}$	3	Impar	–

El número de permutaciones que se puede hacer con un conjunto de 3 elementos es $3! = 3.2.1$.

Si el conjunto S tiene n elementos, el número de permutaciones es $n! = n.(n-1)... 3.2.1$.

Recibe el nombre de determinante la aplicación que asigna a cada matriz el escalar obtenido de la forma siguiente:

La imagen de una matriz cuadrada $A = (a_{ij}) \in M_{n\times n}$ es el escalar obtenido al sumar algebraicamente $n!$ sumandos, tales que:

Cada sumando es el producto de n factores.

Cada factor es un elemento de la matriz perteneciente a una fila distinta del conjunto $\{1, 2, ..., n\}$, (colocadas en el orden natural), y a una columna distinta perteneciente al conjunto $\{\sigma(1), \sigma(2), ..., \sigma(n)\}$.

Cada sumando estará afectado del signo + o –, según que la permutación de las columnas $\sigma(1), \sigma(2), ..., \sigma(n)$, sea par o impar.

La aplicación definida da lugar a la siguiente definición:

3.1.2. Definición de determinante de una matriz

Determinante de una matriz A es la aplicación:

$$M_{n\times n} \xrightarrow{\text{det}} \mathbb{R}$$

$$A = (a_{ij}) \xrightarrow{\text{det}} \det(A) = |A| = \sum (signo\ \sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

Con el fin de aclarar los pasos dados, vamos a analizar detenidamente el proceso para una matriz de orden 2.

$$M_{2\times 2} \xrightarrow{\text{det}} \mathbb{R}$$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{\text{det}} |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

Como $n = 2$, el número de sumandos es: $2! = 2.1 = 2$.

$a_{11}a_{22}$ y $a_{12}a_{21}$

Cada sumando debe tener 2 factores:

$a_{11}a_{22}$ tiene a_{11} y a_{22}; $a_{12}a_{21}$ tiene a_{12} y a_{21}.

En el sumando $a_{11}a_{22}$, el factor a_{11} es de la fila 1, columna 1; el factor a_{22} es de la fila 2 columna 2.

En el sumando $a_{12}a_{21}$, el factor a_{12} es de la fila 1, columna 2; el factor a_{21} es de la fila 2 columna 1.

El sumando $a_{11}a_{22}$ lleva signo más porque una vez colocados sus factores de modo que las filas estén en el orden natural $(1, 2)$, la permutación que representa las columnas $(1, 2)$ no presentan ninguna inversión.

El sumando $a_{12}a_{21}$ lleva signo menos porque una vez colocados sus factores de modo que las filas estén en el orden natural $(1, 2)$, la permutación que representa las columnas $(2, 1)$ presenta un número impar de inversiones (1).

En los órdenes pequeños estamos acostumbrados a utilizar la definición dada, aunque en general, no decimos que es la definición, sino que utilizamos la "Regla de Sarrus".

$$M_{2\times 2} \xrightarrow{\text{det}} \mathbb{R}$$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{\text{det}} |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \xrightarrow{\text{det}} |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

El determinante de orden 4 tiene $4! = 24$ sumandos, el de orden 5 tiene $5! = 120$ sumandos, y para matrices de órdenes superiores, el número de sumandos $n!$ es muy elevado, de modo que, aunque la definición indica un algoritmo de fácil programación para resolver en ordenador, cuando n sea grande se necesita un ordenador muy potente y mucho tiempo. Por ello, es conveniente recurrir a otros procedimientos de cálculo derivados de las propiedades de los determinantes.

La forma de algunas matrices permite el cálculo sencillo de sus determinantes mediante la utilización directa de la definición.

3.1.3. Consecuencias

1. El determinante de una matriz triangular es el producto de los elementos de la diagonal principal.

Al utilizar la definición encontramos que el único sumando que no contiene un factor cero es el producto de los elementos de la diagonal.

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_{nn} \end{vmatrix} = a_{11}a_{22}\dots a_{nn}.$$

2. ***El determinante de una matriz diagonal es el producto de los elementos de la diagonal principal.***

Sale como consecuencia de lo anterior.

3. ***El determinante de la matriz unidad es 1: $|I| = 1$.***

La matriz I es un caso particular de las matrices diagonales.

Sabemos que cada matriz corresponde a un endomorfismo. Si $A = (a_{ij})$ es la matriz asociada al endomorfismo $f: V \to V$, llamaremos determinante del endormorfismo al determinante de la matriz asociada A.

3.1.4. Definición de determinante de un endomorfismo

Determinante de un endomorfismo $f: V \to V \Leftrightarrow |f| = |A|$, siendo A la matriz asociada a f

También vimos en *[2.3]* que si $f: V \to V$, la matriz asociada A tiene n columnas, cada columna es la imagen de un vector de una base de V, expresada en la misma u otra base de V, y por tanto, cada columna tendrá n coordenadas, que dan lugar a n filas.

Se llama determinante de un conjunto de n vectores de n coordenadas cada uno al determinante de la matriz que forman.

3.1.5. Definición de determinante de un conjunto de vectores de $\mathbb{R}^n$

Determinante de $\{\overline{x}_1, \overline{x}_2, \dots, \overline{x}_n\} = |A|$, siendo A la matriz cuyas columnas son $\overline{x}_1, \overline{x}_2, \dots, \overline{x}_n$

Ejemplo 3.1.2

En el ejemplo 2.6.1, habíamos definido el enfomorfismo de $\mathbb{R}^2, f: \mathbb{R}^2 \to \mathbb{R}^2$ $/ f(x_1, x_2) = (2x_1, 0)$, calculado la matriz A asociada a f respecto a la base canónica de $\mathbb{R}^2$, cuyas columnas son las imágenes de la base canónica $\{(1, 0), (0, 1)\}$, $f(0, 1) = (2, 0)$ y $f(1, 0) = (0, 0)$.

El determinante de f, el determinante de A y el determinante de los vectores columna tienen la misma expresión: $|f| = |A| = \begin{vmatrix} 2 & 0 \\ 0 & 0 \end{vmatrix} = 0.$

3.2. Cálculo de determinantes

Hemos visto que la definición de determinante de una matriz hace muy largo su cálculo si la matriz es de un orden elevado, vamos a dedicar esta sección y la siguiente al estudio de las propiedades que permiten desarrollar procedimientos eficientes, tanto para el cálculo manual, como para su incorporación a programas de ordenador.

Cuando elegimos los elementos que forman la intersección de algunas filas y algunas columnas de una matriz, sin variar la disposición en que estaban, hemos formado una <u>submatriz</u> de la matriz dada.

Si la submatriz es cuadrada, se puede calcular su determinante, y ese determinante se llama <u>menor</u>.

3.2.1. Definición de menor

> Menor de una matriz $A \in M_{nxn}$ es el determinante de cualquier submatriz cuadrada suya B.

3.2.2. Definición de menor complementario

> Menor complementario de la submatriz B de una matriz $A \in M_{nxn}$ es el determinante de la submatriz cuadrada de A que resulta de eliminar en A las filas y columnas a las que pertenece B. Se designa por α_B.

3.2.3. Definición de adjunto

> Adjunto o cofactor de la submatriz B de una matriz $A \in M_{nxn}$ es el menor de B afectado del signo correspondiente, se designa por A_B y su valor es: $A_B = (-1)^{\Sigma i + \Sigma j}\, \alpha_B$.

En la definición anterior $\sum i$ es la suma de los subíndices que indican las filas, $\sum j$ la suma de los índices que indican las columnas, que contiene la submatriz B en la matriz $A \in M_{nxn}$.

Cuando la submatriz B elegida está formada por un solo elemento, estamos ante un caso particular de los conceptos anteriores, y las definiciones de menor, menor complementario y adjunto no varían.

Ejemplo 3.2.1

Sea la matriz $A = \begin{pmatrix} 2 & -1 & 3 & 3 \\ \mathit{\mathbf{1}} & -1 & 0 & \mathit{\mathbf{1}} \\ \mathit{\mathbf{1}} & 2 & 4 & \mathit{\mathbf{2}} \\ 1 & 1 & 1 & 0 \end{pmatrix} \in M_{4x4}$, la submatriz $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ está formada por los elementos que están en negrita y cursiva, es la intersección de las filas 2 y 3 y las columnas 1 y 4 de A.

El menor correspondiente es: $|B| = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1$.

El menor complementario de B es: $\alpha_B = \begin{vmatrix} -1 & 3 \\ 1 & 1 \end{vmatrix} = -4$.

El adjunto de B es: $A_B = (-1)^{(2+3)+(1+4)}\, \alpha_B = (-1)^{10}\,(-4) = -4$.

Si elegimos como submatriz el elemento $a_{21} = 1 \in A$, el menor es el mismo elemento, el menor complementario $\alpha_{21} = \begin{vmatrix} -1 & 3 & 3 \\ 2 & 4 & 2 \\ 1 & 1 & 0 \end{vmatrix} = 2$, y el adjunto $A_{21} =$

$= (-1)^{1+2}\, \alpha_{12} = -2$.

Nuestro propósito es reducir el cálculo de un determinante al cálculo de determinantes de orden inferior al dado; vamos a demostrar que esto es posible.

Antes de hacer la demostración recordemos que la expresión que define un determinante se puede escribir de más de una manera, así, un determinante de orden tres admite las siguientes expresiones:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} -$$

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

Observemos que hemos sacado factor común los elementos de la primera fila, y lo que queda dentro de cada paréntesis es el menor complementario del factor común; el menor complementario afectado del signo correspondiente es el adjunto de dicho factor común. Por tanto, hemos reducido el cálculo de un determinante de orden tres al cálculo de determinantes de orden dos.

No ofrece ninguna dificultad de concepto, aunque sí de notación, reproducir para el determinante de orden n el proceso seguido en el caso de orden tres.

3.2.4. Teorema

El determinante de una matriz es igual a la suma de los productos de los elementos de una fila por sus adjuntos correspondientes

Demostración:

Si $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$, la definición de determinante dada en *[3.1.2]*

nos dice que $|A|$ tiene $n!$ sumandos, cada sumando es el producto de n factores, cada factor es un elemento de la matriz perteneciente a una fila distinta: 1, 2, ..., n, y a una columna distinta $\sigma(1), \sigma(2), \dots, \sigma(n)$.

Si nos fijamos en la primera fila $(a_{11}, a_{12}, \dots, a_{1n})$ de la matriz, podemos sacar factor común a_{11} en los $n-1$ sumandos que contienen a_{11} en el desarrollo del determinante. El factor que multiplica a a_{11}, tendrá $n-1$ sumandos, cada sumando tendrá $n-1$ factores, y cada factor será de las filas 2, 3, ..., n, y de las columnas 2, 3, ..., n, ya que en el desarrollo del determinante primitivo de orden n los factores de cada sumando son de fila y columna distinta, y al quitar el a_{11}, los que le multiplican son de las restantes filas y columnas. Dicho factor es, por tanto, el determinante que resulta de eliminar la primera fila y la primera columna. Es el menor complementario de a_{11} : α_{11}.

Aplicando la definición de adjunto del menor a_{11}, vemos que el menor complementario coincide con su adjunto, porque $A_{11} = (-1)^{1+1}\,\alpha_{11} = \alpha_{11}$.

Repitiendo el proceso con los sumandos que contienen a_{12}, el factor que le multiplica tiene $n-1$ sumandos, cada sumando $n-1$ factores, y cada factor es de las filas 2, 3, ..., n, y de las columnas 1, 3, ..., n. Es decir, es $-\alpha_{12}$ porque es el desarrollo del menor complementario de a_{12}, pero todos los sumandos tienen el signo cambiado. Hemos obtenido, en definitiva el adjunto $A_{12} = -\alpha_{12}$.

Operando de manera análoga con a_{13}, luego con a_{14}, etc., hasta a_{1n}, se tiene el desarrollo del determinante por los elementos de la primera fila:

$$|A| = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n}.$$

La elección de la primera fila no ha restado generalidad al procedimiento, si hubiéramos elegido cualquier otra fila, el resultado sería el mismo, tal como hemos expresado de forma general en el enunciado del teorema.

Podíamos haber optado por seguir el proceso anterior con los elementos de cualquier columna, habríamos obtenido un resultado análogo, pero sustituyendo la palabra fila por la palabra columna. El teorema diría que el desarrollo del determinante por los elementos de la columna i es:

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}$$

Una generalización de este teorema permite desarrollar el determinante por los elementos de varias filas. Esta forma de cálculo de determinantes se llama regla de Laplace, aunque no vamos a hacer su demostración, sí vamos a enunciarla dada su utilidad.

3.2.5. Regla de Laplace

> El determinante de una matriz cuadrada de orden n es igual a la suma de todos los productos posibles de los menores de orden p,(con $p < n$) que se pueden formar con p filas (o columnas), por sus adjuntos correspondientes.

Ejemplo 3.2.2

Calcúlese $|A| = \begin{vmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 4 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & 3 & -1 & 0 \end{vmatrix}$ mediante la regla de Laplace por los adjuntos de la segunda y tercera fila.

Solución:

Hay seis menores de orden 2 de la segunda y tercera fila:

$$\Delta_1 = \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -1; \quad \Delta_2 = \begin{vmatrix} 1 & 4 \\ 1 & 0 \end{vmatrix} = -4; \quad \Delta_3 = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1;$$

$$\Delta_4 = \begin{vmatrix} 2 & 4 \\ 1 & 0 \end{vmatrix} = -4; \quad \Delta_5 = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3; \quad \Delta_6 = \begin{vmatrix} 4 & 1 \\ 0 & 2 \end{vmatrix} = 8;$$

Sus menores complementarios son:

$$\alpha_1 = \begin{vmatrix} 3 & 0 \\ -1 & 0 \end{vmatrix} = 0; \quad \alpha_2 = \begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} = 0; \quad \alpha_3 = \begin{vmatrix} 1 & 3 \\ 3 & -1 \end{vmatrix} = -10;$$

$$\alpha_4 = \begin{vmatrix} 2 & 0 \\ 1 & 0 \end{vmatrix} = 0; \quad \alpha_5 = \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = -5; \quad \alpha_6 = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5;$$

Los adjuntos correspondientes son:

$$A_1 = (-1)^{\Sigma i + \Sigma j} \quad \alpha_1 = (-1)^{(2+3)+(1+2)} \quad \alpha_1 = (-1)^8\, \alpha_1 = 0; \quad A_2 = (-1)^{2+3+1+3} \quad \alpha_2 = 0;$$

$$A_3 = (-1)^{2+3+1+4} \quad \alpha_3 = -10; \qquad A_4 = (-1)^{2+3+2+3} \quad \alpha_4 = 0;$$

$$A_5 = (-1)^{2+3+2+4} \quad \alpha_5 = 5; \qquad A_6 = (-1)^{2+3+3+4} \quad \alpha_6 = 5;$$

Aplicando la Regla de Laplace se obtiene el valor del determinante:

$$|A| = \sum_{i=1}^{6} \Delta_i A_i = -1 \cdot 0 + (-4) \cdot 0 + 1 \cdot (-10) + (-4) \cdot 0 + 3 \cdot 5 + 8 \cdot 5 = 45.$$

3.3. Propiedades de los determinantes

El teorema *[3.2.4]*, además de permitir una forma de cálculo de un determinante reduciéndolo a suma algebraica de otros de orden una unidad menor, permite probar algunas de las propiedades de los determinantes.

Es muy frecuente que en la vida cotidiana saquemos conclusiones de hechos aislados sin suficiente fundamento lógico. Por ejemplo, "como la terminación 5 se ha repetido varios años seguidos en los sorteos de la lotería, inducimos que la terminación 5 traerá suerte en el sorteo siguiente". Es evidente que la conclusión es falsa.

El hecho de que una situación se repita, no nos permite asegurar que va a ser cierta siempre. Por ello, no es suficiente dar ejemplos para demostrar un hecho, hace falta una demostración. El método inductivo bien utilizado permite asegurar que las conclusiones obtenidas son ciertas, y con su uso se articulan demostraciones válidas.

La aplicación del método de inducción consiste en seguir los siguientes pasos:

- Comprobar que una proposición es cierta una vez (la primera).
- Comprobar que si es cierta una vez (la vez n–1), lo es la vez siguiente (la vez n).

Así, como es cierta la vez uno, lo es la dos, como es cierta la dos lo es la tres...

En algunas demostraciones posteriores, utilizaremos este método, que también será ampliamente utilizado en programación para demostrar la corrección formal de ciertos tipos de algoritmos (algoritmos recursivos no finales), donde se cuenta con un parámetro de entrada sobre el que se establece un subproblema, de tamaño menor que el problema. En función de su solución se puede resolver el problema original.

3.3.1. Propiedades de los determinantes

1. ***El determinante de una matriz cuadrada coincide con el de su transpuesta***

Demostración:

Sea n el orden de la matriz A.

1. Si $n = 1$ ó $n = 2$ la propiedad es cierta, como puede comprobarse fácilmente utilizando la definición.
2. Suponiendo que la propiedad es cierta para matrices de orden n–1, vamos a demostrar que es cierta para las de orden n.

En efecto, es desarrollo de $|A|$ por los menores de la primera fila es:

$$|A| = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n}$$

Sea B la matriz transpuesta de $A \Leftrightarrow B = A^t \Leftrightarrow$ Cada $b_{ij} = a_{ji}$.
Desarrollando $|B|$ por los elementos de la primera columna:

$$|B| = b_{11}B_{11} + b_{21}B_{21} + \dots + b_{n1}B_{n1}$$

donde B_{ij} es el adjunto del elemento b_{ij}.

Observemos que la matriz de orden n–1 que resulta al eliminar en la matriz B la fila i y la columna j es la transpuesta de la que resulta al eliminar la fila j y la columna i en la matriz A.

Por lo tanto, utilizando la hipótesis de que en las matrices de orden n–1 que son transpuestas sus determinantes coinciden, cada adjunto B_{ji}, del desarrollo de $|B|$, es igual al correspondiente A_{ij} del desarrollo de $|A|$ en nuestro caso.

Como consecuencia: $|A| = |A^t|$.

Hemos demostrado que la propiedad es cierta para los determinantes de orden n, esta propiedad supone que todas las características que un determinante tenga para las filas, encuentra su análogo para las columnas.

2. ***Si en una matriz A se permutan dos filas entre sí, el determinante de la matriz B que resulta es el determinante de A cambiado de signo***

Demostración:

Haremos la demostración por inducción.

El resultado es cierto para matrices de orden 2 como puede comprobarse utilizando la definición de determinante.

Suponiendo que es cierto para matrices de orden $n-1$, comprobemos que también lo es para las de orden n.

Si A es una matriz de orden n, y B la matriz que resulta de intercambiar las filas i y j de A, queremos demostrar que $|B| = -|A|$.

Sea la fila k otra fila distinta a las anteriores ($k \neq i$; $k \neq j$). Si desarrollamos ambos determinantes por los elementos de la fila k, obtenemos:

$$|A| = a_{k1}A_{k1} + a_{k2}A_{k2} + \dots + a_{kn}A_{kn}; \quad |B| = b_{k1}B_{k1} + b_{k2}B_{k2} + \dots + b_{kn}B_{kn}.$$

Los elementos B_{kh} y A_{kh} para $h = 1, \dots, n$, de cada par de adjuntos son de una matriz de orden $n-1$, en los que están intercambiados la fila i con la fila j, y por la hipótesis de inducción damos por cierto que $B_{kh} = -A_{kh}$.

Por tanto, es $|B| = -b_{k1}A_{k1} - b_{k2}A_{k2} - \dots - b_{kn}A_{kn} \Rightarrow |B| = -|A|$.

3. ***Si una matriz tiene dos filas iguales, su determinante es cero.***

Esta propiedad es consecuencia inmediata de la anterior.

Demostración:

Sea A una matriz con dos filas iguales.

Intercambiando entre sí esas dos filas, la matriz B que resulta es la misma $\Rightarrow |B| = |A|$.

Por otra parte, $|B| = -|A|$ porque la matriz B se ha obtenido al intercambiar dos filas en la matriz A.

La única posibilidad que existe para que un valor real sea igual a su opuesto, es que este valor sea 0.

$$\left.\begin{array}{l} |A| = |B| \\ |A| = -|B| \end{array}\right\} \Rightarrow |A| = |B| = 0.$$

4. ***Si en una matriz se multiplican todos los elementos de una fila por un mismo escalar λ, el determinante de la matriz resultante queda multiplicado por λ.***

Demostración:

Sea A una matriz de orden n y sea B la matriz que resulta de multiplicar en A cada elemento de la fila i por el escalar λ.

Desarrollando $|B|$ por los elementos de esta fila i, es: $|B| = b_{i1}B_{i1} + b_{i2}B_{i2} + \dots + b_{in}B_{in}$, en la fila i, cada elemento $b_{ij} = \lambda a_{ij}$, y cada adjunto $B_{ij} = A_{ij}$ ya que las matrices A y B difieren sólo en la fila i, y los adjuntos anteriores no contienen elementos de ella.

Se verifica, por tanto:

$$|B| = b_{i1}B_{i1} + b_{i2}B_{i2} + \dots + b_{in}B_{in} = \lambda a_{i1}A_{i1} + \lambda a_{i2}A_{i2} + \dots + \lambda a_{in}A_{in} = \lambda(a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}) = \lambda |A|$$

5. ***Si en una matriz A, se sustituye una fila por ella misma más el resultado de multiplicar otra fila distinta por el escalar λ, el determinante de la matriz resultante no varía.***

Demostración:

Sustituyendo en la matriz A la fila i por ella más la fila j ($j \neq i$) multiplicada por λ, se obtiene la matriz B donde cada elemento de la fila i es $b_{ir} = a_{ir} + \lambda a_{jr}$ con $r = 1, \dots, n$.

Para calcular $|B|$ por los elementos de la fila i, se efectúa:

$$|B| = (a_{i1} + \lambda a_{j1})B_{i1} + (a_{i2} + \lambda a_{j2})B_{i2} + \dots + (a_{in} + \lambda a_{jn})B_{in}$$

Como cada adjunto $B_{ik} = A_{ik}$, porque las matrices A y B coinciden excepto en la fila i, podemos escribir:

$$|B| = (a_{i1} + \lambda a_{j1})A_{i1} + (a_{i2} + \lambda a_{j2})A_{i2} + \dots + (a_{in} + \lambda a_{jn})A_{in} =$$
$$= (a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}) + \lambda(a_{j1}A_{i1} + a_{j2}A_{i2} + \dots + a_{jn}A_{in}) = |A| + \lambda|C|$$

donde C es una matriz que tiene iguales la fila i y la fila j, cuyo determinante ha sido desarrollado por los elementos de la fila i.

Por tanto $|C| = 0 \Rightarrow |B| = |A|$.

6. ***Si en una matriz hay una fila combinación lineal de otras filas, su determinante es cero.***

Para demostrar esta propiedad vamos a utilizar las dos anteriores.

Sea A una matriz, tal que, su fila i es combinación lineal de otras filas. Al desarrollar $|A|$ por los elementos de la fila i, obtendremos una suma algebraica de determinantes, en cada uno de ellos podemos sacar una constante fuera (*propiedad 4*), todas las matrices de los determinantes obtenidos tienen dos filas iguales, y su valor es 0.

7. ***Si A y B son dos matrices cuadradas de orden n, se verifica:***
$|AB| = |A| \cdot |B|$

Demostración:

Sean $A = (a_{ij})$ y $B = (b_{ij})$ dos matrices cuadradas de orden n, construyamos la matriz C siguiente

$$C = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} & 0 & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & b_{2n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} & 0 & 0 & .. & 0 \\ -1 & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1n} \\ 0 & -1 & \dots & 0 & a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \left(\begin{array}{c|c} \mathrm{B} & 0 \\ \hline -I & A \end{array}\right)$$

C es una matriz cuadrada de orden $2n$ distribuida en cuatro submatrices según indica el gráfico.

Transformamos C mediante las siguientes sustituciones elementales de filas sin que se altere el valor de su determinante; llamaremos D a la matriz obtenida.

$F_1 \to F_1 + b_{11}F_{n+1} + b_{12}F_{n+2} + \dots + b_{1n}F_{2n} \Rightarrow$ Las primeras n posiciones de la primera fila quedan ocupadas por ceros, y las n restantes por los elementos de la primera fila de la matriz producto BA.

$F_2 \to F_2 + b_{21}F_{n+1} + b_{22}F_{n+2} + \dots + b_{2n}F_{2n} \Rightarrow$ Las primeras n posiciones de la segunda fila quedan ocupadas por ceros, y las n restantes por los elementos de la segunda fila de la matriz BA.

. . . .

$F_n \to F_n + b_{n1}F_{n+1} + b_{n2}F_{n+2} + \dots + b_{nn}F_{2n} \Rightarrow$ Las primeras n posiciones de la fila n quedan ocupadas por ceros, y las n restantes por los elementos de la fila n de la matriz BA.

El resultado es la matriz D:

$$D = \left(\begin{array}{ccc|c} & 0 & & BA \\ \hline -1 & \dots & 0 & \\ \dots & \dots & \dots & A \\ 0 & \dots & -1 & \end{array}\right)$$

Cálculo de $|C|$:

Tomamos la submatriz B de C, y utilizamos la regla de Laplace, el único menor distinto de cero en las n primeras filas y en las n primeras columnas de C es $|B|$. Su menor complementario es $\alpha_A = |A|$, y su adjunto C_B:

$$C_B = (-1)^{\Sigma i + \Sigma j}\,|A| = (-1)^{(1+2+\dots+n)+(1+2+\dots+n)}|A| = |A|$$

Por tanto: $|C| = |B| \cdot |A|$

Cálculo de $|D|$:

Tomamos la submatriz BA de D, y utilizamos la regla de Laplace, el único menor distinto de cero en las n primeras filas y en las n últimas columnas de D es $|BA|$. Su menor complementario es $\alpha_{BA} = \begin{vmatrix} -1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & -1 \end{vmatrix} = (-1)^n$, cuyo adjunto vale:

$$D_{BA} = (-1)^{\Sigma i + \Sigma j}\,\alpha_{AB} = (-1)^{(1+2+\dots+n)+((n+1)+\dots+2n)}\,(-1)^n = (-1)^{S+n} = 1.$$

(Obsérvese que el exponente 1+2+...+n+(n+1)+...+2n es la suma de términos de una progresión aritmética cuyo valor es $S = \frac{1+2n}{2} 2n$, y $S + n =$ $= (1 + 2n)n + n = 2n^2 + 2n$ es un número par).

Por tanto, $|D| = |BA|$.

Como sabemos que $|C| = |D|$, podemos asegurar que $|BA| = |B|.|A|$, ya que $|C| = |B|\,|A|$ y $|D| = |AB|$.

Obsérvese que si bien el producto de matrices no es conmutativo, sí lo es el producto de escalares, y los determinantes lo son; por tanto podemos decir que $|AB| = |A|\cdot|B|$ aunque hayamos hecho los cálculos para $|BA| = |B|.|A|$.

Las propiedades de los determinantes permiten establecer procedimientos para abreviar su cálculo.

Siempre es posible la transformación de la matriz dada en una matriz triangular mediante operaciones elementales en la misma como vimos en *[2.8]*, aplicar el efecto de estas transformaciones al determinante y proceder al cálculo del determinante de una matriz triangular.

Ejemplo 3.3.1

Dada la matriz $A = \begin{pmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \\ 3 & 2 & 0 \end{pmatrix}$, se pide calcular su determinante a través de una matriz escalonada equivalente.

Paso de A a la escalonada equivalente	Efecto de las transformaciones en su determinante
$\begin{pmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \\ 3 & 2 & 0 \end{pmatrix}$	$\begin{vmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \\ 3 & 2 & 0 \end{vmatrix}$
$\downarrow F_1 \leftrightarrow F_2$	Cambio el determinante de signo. (Prop. 2)
$\begin{pmatrix} -1 & 0 & 1 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{pmatrix}$	$-\begin{vmatrix} -1 & 0 & 1 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{vmatrix}$
$\downarrow F_1 \rightarrow F_1$	Multiplica el determinante por (–1). (Prop. 4)
$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{pmatrix}$	$\begin{vmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{vmatrix}$
$\downarrow F_3 \rightarrow F_3 - 3F_1$	Ninguno. (Prop. 5)
$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 0 & 2 & 3 \end{pmatrix}$	$\begin{vmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 0 & 2 & 3 \end{vmatrix}$
$\downarrow F_2 \rightarrow 1/2F_2$	Multiplica el determinante por 1/2. (Prop. 4)
$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 2 & 3 \end{pmatrix}$	$2\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 2 & 3 \end{vmatrix}$
$\downarrow F_3 \rightarrow F_3 - 2F_2$	Ninguno. (Prop. 5)
$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 2 \end{vmatrix}$

Observaciones:

La matriz escalonada ha sido obtenida mediante el algoritmo esbozado en *[2.8]*.

Cuando decimos que multiplicamos el determinante por 1/2, estamos diciendo que el determinante de la nueva matriz es 1/2 del de la matriz A, por tanto, $|A|$ es dos veces el determinante de la matriz transformada.

Lo mismo ocurre en los dos cambios de signo de las primeras transformaciones.

La matriz obtenida es triangular y su determinante vale el producto de los elementos de la diagonal principal *[2.8]*.

$$\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 2 \end{vmatrix} = 2 \Rightarrow |A| = \begin{vmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \\ 3 & 2 & 0 \end{vmatrix} = 2 \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 2 \end{vmatrix} = 2.2 = 4.$$

3.4. Cálculo de la matriz inversa

En el capítulo 2 estudiamos las operaciones en el conjunto de matrices cuadradas, y dijimos que en dicho conjunto existe elemento identidad I, que es la unidad para la multiplicación, pero que no siempre existe el elemento inverso de una matriz dada A.

Si existe la matriz inversa de A es la matriz A^{-1} que verifica: $AA^{-1} = A^{-1}A = I$

3.4.1. Definición

A^{-1} es la matriz inversa de $A \Leftrightarrow AA^{-1} = A^{-1}A = I$

Las matrices que tienen elemento inverso se llaman inversibles o regulares

3.4.2. Definición

$A \in M_{n\times n}$ es una matriz regular $\Leftrightarrow$ Existe su inversa $A^{-1} \in M_{n\times n}$

Las matrices que no tienen inversa se llaman matrices singulares

3.4.3. Definición

$A \in M_{n\times n}$ es una matriz singular $\Leftrightarrow$ No tiene inversa

Las matrices inversibles verifican las mismas propiedades que los elementos de un anillo que tienen inverso y esto nos permite enunciar las consecuencias siguientes.

3.4.4. Consecuencias

1. ***Si una matriz A es regular, su inversa A^{-1} también lo es, y verifica $(A^{-1})^{-1} = A$.***

Demostración:

Es consecuencia directa de la definición.

2. ***Si A es una matriz cuadrada con inversa A^{-1}, ésta es la única.***

Demostración:

Si existieran dos matrices C y D inversas ambas de A, utilizando la asociatividad del producto de matrices, podríamos escribir $C(AD) = (CA)D$.

$$\left.\begin{array}{l}\text{Por ser } D \text{ inversa de } A \text{ es } AD = I\\ \text{Por ser } C \text{ inversa de } A \text{ es } CA = I\end{array}\right\} \Rightarrow \left\{\begin{array}{l} C(AD) = CI = C\\ (CA)D = ID = D\end{array}\right\} \Rightarrow C = D.$$

3. ***Si A y B son matrices regulares, también lo son sus productos AB y BA; y se verifica $(A \cdot B)^{-1} = B^{-1}A^{-1}$.***

Demostración:

Si utilizamos la ley asociativa del producto de matrices podemos escribir las siguientes cadenas de igualdades:

$$\left\{\begin{array}{l}(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = (AI)A^{-1} = AA^{-1} = I\\ (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = (B^{-1}I)B = B^{-1}B = I\end{array}\right\} \Rightarrow (AB)^{-1} = B^{-1}A^{-1}$$

El cálculo de la inversa de una matriz cuadrada se puede hacer utilizando distintos métodos, haremos un ejemplo o daremos sólo una indicación de cada procedimiento, excepto del último que es el más útil y generalizado.

- Utilizando la definición.

Ejemplo 3.4.1

Determínese la matriz inversa de $I - A$ en función de las potencias de A, sabiendo que A es una matriz regular, tal que, $A^3 = 0$.

Solución:

$$\begin{array}{l}(I-A)^2\,(I-A)(I-A)^{-1} = (I-A)^2\\ (I-A)^3 = I - 3A + 3A^2 - A^3 = I - 3A + 3A^2\\ (I - 3A + 3A^2)\,(I-A)^{-1} = (I-A)^2\\ [I - 3A(I-A)]\,(I-A)^{-1} = I(I-A)^{-1} - 3A(I-A)(I-A)^{-1} = (I-A)^2\\ (I-A)^{-1} = (I-A)^2 + 3A = I - 2A + A^2 + 3A = A^2 + A + I\end{array}$$

- Resolviendo un sistema de ecuaciones lineales.

Si $A^{-1} = (x_{ij}) \in M_{n\times n}$ inversa de $A = (a_{ij}) \in M_{n\times n}$ es la matriz a hallar, debe verificarse $AA^{-1} = I$

$$\begin{pmatrix} a_{21} & a_{12} & \dots & a_{1n}\\ a_{11} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n}\\ x_{21} & x_{22} & \dots & x_{2n}\\ \dots & \dots & \dots & \dots\\ x_{n1} & x_{n2} & \dots & x_{nn}\end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & 1\end{pmatrix}$$

Si realizamos el producto de las dos matrices e igualamos los valores de los elementos que ocupan la misma posición en ambos miembros, obtenemos un sistema lineal de n^2 ecuaciones lineales con n^2 incógnitas, que es laborioso de resolver, aunque no tiene ninguna dificultad técnica.

Ejemplo 3.4.2

Determínese la inversa de $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$, utilizando un sistema de ecuaciones.

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{cases} x_{11} + 2x_{21} = 1 \\ 3x_{21} = 0 \\ x_{12} + 2x_{22} = 0 \\ 3x_{22} = 1 \end{cases} \Rightarrow \begin{cases} x_{11} = 1 \\ x_{21} = 0 \\ x_{12} = -2/3 \\ x_{22} = 1/3 \end{cases}$$

Para comprobarlo basta hacer los siguientes productos:

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

♦ Mediante transformaciones elementales.

Es conveniente repasar *[2.8]* antes de hacer el ejemplo siguiente, porque vamos a utilizar los resultados que obtuvimos allí.

Si A tiene inversa, mediante transformaciones elementales de filas en A se obtiene la matriz unidad I, si aplicamos esas mismas transformaciones en I, se obtiene una matriz M, de manera que $M \cdot A = I$, es decir $M = A^{-1}$ (teorema *[2.8.5]*).

Ejemplo 3.4.3

Se desea calcular la inversa de $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 1 & 1 & 1 \end{pmatrix}$ utilizando operaciones elementales.

Solución:

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & -1 & -1 & -2 & 1 & 0 \end{array}\right) \to$$

$$F_2 \to F_2 - 2F_1 \qquad F_2 \leftrightarrow F_1 \qquad F_3 \to F_3 + F_2$$
$$F_3 \to F_3 - F_1$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 0 & -2 & -3 & 1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 3/2 & -1/2 & -1/2 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 1 & 1 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 & 3/2 & -1/2 & -1/2 \end{array}\right)$$

$$F_3 \to -1/2\, F_3 \qquad F_1 \to F_1 - 2F_3$$
$$F_2 \to F_2 + F_3$$

La matriz inversa es $A^{-1} = \begin{pmatrix} -2 & 1 & 1 \\ 1/2 & -1/2 & 1/2 \\ 3/2 & -1/2 & -1/2 \end{pmatrix}$; para comprobarlo hay que ver:

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} -2 & 1 & 1 \\ 1/2 & -1/2 & 1/2 \\ 3/2 & -1/2 & -1/2 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 1 \\ 1/2 & -1/2 & 1/2 \\ 3/2 & -1/2 & -1/2 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

♦ Mediante determinantes.

Para poder hacerlo vamos a introducir algunos conceptos más.

3.4.5. Definición

$\overline{A} \in M_{nxn}$ es la matriz adjunta de $A \Leftrightarrow \overline{A}$ es la matriz que resulta de sustituir cada elemento por su adjunto correspondiente.

Ejemplo 3.4.4

Determínese la matriz adjunta $\overline{A}$ de $A = \begin{pmatrix} 1 & 2 & 1 \\ -2 & 0 & 3 \\ 1 & 1 & 3 \end{pmatrix}$

Solución:

Los adjuntos de los distintos elementos son:

$$A_{11}=+\begin{vmatrix}0&3\\1&3\end{vmatrix}=-3;\quad A_{12}=-\begin{vmatrix}-2&3\\1&3\end{vmatrix}=9;\quad A_{13}=\begin{vmatrix}-2&0\\1&1\end{vmatrix}=-2;\quad A_{21}=-\begin{vmatrix}2&1\\1&3\end{vmatrix}=-5;$$

$$A_{22}=\begin{vmatrix}1&1\\1&3\end{vmatrix}=2;\quad A_{23}=-\begin{vmatrix}1&2\\1&1\end{vmatrix}=1;\quad A_{31}=\begin{vmatrix}2&1\\0&3\end{vmatrix}=6;\quad A_{32}=-\begin{vmatrix}1&1\\-2&3\end{vmatrix}=-5;$$

$$A_{33}=\begin{vmatrix}1&2\\-2&0\end{vmatrix}=4$$

$$\overline{A}=\begin{pmatrix}-3&9&-2\\-5&2&1\\6&-5&4\end{pmatrix}.$$

3.4.6. Teorema

El producto de una matriz cuadrada A por la transpuesta de su adjunta es el valor de su determinante por la matriz unidad. $A(\overline{A})^t=|A|\,I$

Demostración

Sea la matriz $A=\begin{pmatrix}a_{11}&a_{12}&\dots&a_{1n}\\a_{21}&a_{22}&\dots&a_{2n}\\\dots&\dots&\dots&\dots\\a_{n1}&a_{n1}&\dots&a_{nn}\end{pmatrix}$ y $(\overline{A})^t=\begin{pmatrix}A_{11}&A_{21}&\dots&A_{n1}\\A_{12}&A_{22}&\dots&A_{n2}\\\dots&\dots&\dots&\dots\\A_{1n}&A_{2n}&\dots&A_{nn}\end{pmatrix}$ la traspuesta de su adjunta.

Al hacer el producto de ambas se ve que $A(\overline{A})^t=(b_{ij})$ siendo $b_{ij}=\begin{cases}|A|\,\text{si } i=j\\0 \text{ si } i\neq j\end{cases}$,

aplicando la propiedad *3* de *[3.3.1]*.

$$A(\overline{A})^t=\begin{pmatrix}|A|&0&\dots&0\\0&|A|&\dots&0\\\dots&\dots&\dots&\dots\\0&0&\dots&|A|\end{pmatrix}=|A|\,I.$$

De la misma manera se deduce que, $(\overline{A})^t A=|A|I$, por tanto, $A(\overline{A})^t=(\overline{A})^t A=|A|I$.

3.4.7. Teorema

Una matriz cuadrada A tiene inversa (es regular) $\Leftrightarrow$ $|A|\neq 0$

Demostración:

$\Rightarrow$) Si A regular $\Rightarrow |A|\neq 0$
Si A regular, existe matriz inversa A^{-1}, tal que, $A\,A^{-1}=A^{-1}\,A=I$
Como el determinante de un producto es el producto de los determinantes $\Rightarrow |A|\,|A^{-1}|=|A^{-1}|\,|A|=1$, necesariamente ha de ser $|A|\neq 0$.

$\Leftarrow$) Si $|A|\neq 0 \Rightarrow A$ es regular
Si $|A|\neq 0$ existe una matriz $B=\dfrac{1}{|A|}\cdot(\overline{A})^t$ (basta construirlo como hemos hecho en el teorema anterior).
Al aplicar el teorema *[3.4.5]*, se obtienen los resultados siguientes:

$$AB=A\left(\frac{1}{|A|}(\overline{A})^t\right)=\frac{1}{|A|}\left(A\cdot(\overline{A})^t\right)=\frac{1}{|A|}\cdot|A|\cdot I=I$$

$$BA=\left(\frac{1}{|A|}(\overline{A})^t\right)A=\frac{1}{|A|}\left((\overline{A})^t\cdot A\right)=\frac{1}{|A|}\cdot|A|\cdot I=I$$

esta expresión nos proporciona otra forma de calcular la inversa de una matriz cuando existe, es decir, de una matriz regular.

3.4.8. Cálculo de la matriz inversa de la matriz A

B es la inversa de $A \Leftrightarrow B = A^{-1} = \dfrac{1}{|A|} \cdot (\bar{A})^t$

Ejemplo 3.4.5

Determínese la matriz inversa A^{-1} de la matriz $A = \begin{pmatrix} 1 & 2 & 1 \\ -2 & 0 & 3 \\ 1 & 1 & 3 \end{pmatrix}$ del ejemplo 3.4.4.

Solución:

Habíamos determinado la matriz adjunta $\bar{A} = \begin{pmatrix} -3 & 9 & -2 \\ -5 & 2 & 1 \\ 6 & -5 & 4 \end{pmatrix} \Rightarrow (\bar{A})^t =$

$= \begin{pmatrix} -3 & -5 & 6 \\ 9 & 2 & -5 \\ -2 & 1 & 4 \end{pmatrix}$; $|A| = \begin{vmatrix} 1 & 2 & 1 \\ -2 & 0 & 3 \\ 1 & 1 & 3 \end{vmatrix} = 13$; $A^{-1} = 1/13 \begin{pmatrix} -3 & -5 & 6 \\ 9 & 2 & -5 \\ -2 & 1 & 4 \end{pmatrix}$.

Podemos comprobar que $1/13 \begin{pmatrix} 1 & 2 & 1 \\ -2 & 0 & 3 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} -3 & -5 & 6 \\ 9 & 2 & -5 \\ -2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

3.4.9. Consecuencias

1) *Si A es una matriz regular de orden n, entonces,* $|A^{-1}| = \dfrac{1}{|A|}$

Demostración:

Si A tiene inversa A^{-1}, se verifica $A\,A^{-1} = I$, y $|A|\,|A^{-1}| = |I| = 1 \Rightarrow |A^{-1}| =$

$= \dfrac{1}{|A|}$.

2) ***Para toda matriz A de orden n se verifica*** $|A|\,|(\bar{A})^t| = |A|^n$.

Demostración:

La matriz producto $|A| \cdot I$ es diagonal, su determinante es $|A|^n$.

3.5. Rango de una matriz; rango de un endomorfismo; rango de un sistema de vectores

En *[1.3.6]* vimos la definición de vectores linealmente independientes y en *[1.3.7.]* su caracterización, así podíamos seleccionar los vectores que dependían de otros, o saber que el sistema de vectores dado es libre.

En *[2.8.1.]* utilizamos las matrices escalonadas para conseguir el mismo fin, y después de estudiar los determinantes disponemos de una herramienta poderosa que facilita otros criterios para hacer lo mismo.

El teorema siguiente contiene una condición de independencia lineal de vectores.

3.5.1. Teorema

Un sistema S de n vectores de un espacio vectorial de dimensión n es libre $\Leftrightarrow$ El determinante de la matriz cuyas filas o columnas son los vectores de S es $\neq 0$

Demostración

$\Rightarrow$) Supongamos que un sistema S de n vectores de un espacio vectorial de dimensión n es libre, si utilizamos *[2.8.7]* $\Rightarrow$ Si A es la matriz cuyas filas son los vectores de S, A tiene inversa al tener rango máximo, utilizando *[3.4.7]* $\Rightarrow$ $\Rightarrow |A| \neq 0$.

$\Leftarrow$) Para demostrar la implicación "$\Rightarrow$", hemos utilizado sólo la condición necesaria de cada teorema, pero como también es cierta la suficiente, es válida toda la cadena de implicaciones tomadas desde el final hasta el principio.

Ejemplo 3.5.1

El sistema de vectores $F = \{(1, -1, 3), (2, 1, 1), (1, 0, 0)\} \subset \mathbb{R}^3$, es linealmente independiente, ya que el determinante $\begin{vmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \\ 3 & 1 & 0 \end{vmatrix} = -4 \neq 0.$

Como ya hemos anunciado, el cálculo del rango de una matriz se puede hacer utilizando determinantes.

3.5.2. Teorema

El rango de una matriz es el mayor de los órdenes de sus menores no nulos

Demostración

Sea $A \in M_{m \times n}$ una matriz de rango r, y sea A_1 una submatriz de A regular de orden p, tal que el determinante de cualquier otra submatriz cuadrada de orden mayor que p es cero.

Vamos a demostrar que $r = p$ demostrando que $r \geq p$ y $r \leq p$.

1. $r \geq p$

Sean $i_1, i_2, ..., i_p$ las filas que ocupa A_1 en la matriz A, y sean $j_1, j_2, ..., j_p$ las columnas.

Considérese la matriz B con las mismas filas que A_1, pero con las columnas de A. Es una matriz con p filas y n columnas.

La matriz B tiene sus p filas linealmente independiente porque son de A_1 y esta es regular. Por tanto, $\text{rg}(B) = p$, y la matriz A tiene por lo menos esas mismas filas independientes. Por tanto $\text{rg}(A) = r \geq p$.

2. $r \leq p$

Supongamos que existe una submatriz B de A, tal que $\text{rg}(B) = r > p$, en la matriz A hay r filas independientes. Sean $i_1, i_2, ..., i_r$ dichas filas y consideremos la matriz B formada por esas r filas y todas las columnas de A. La matriz B será de orden $r \times n$ y tendrá rango r porque las r filas son independientes, luego hay un menor en B, y por tanto, en A de orden r distinto de cero.

Pero si esto ocurre, estamos en contradicción con la hipótesis de que no había ninguna submatriz cuadrada de orden mayor que p de determinante distinto de cero $\Rightarrow r \leq p$.

Como $\left.\begin{matrix} r \geq p \\ r \leq p \end{matrix}\right\} \Rightarrow r = p.$

Este teorema se puede utilizar para calcular el rango de una matriz. Aunque el procedimiento es un poco largo, se puede agilizar si se hace de una forma ordenada y se tienen presentes las propiedades de los determinantes.

El algoritmo a seguir es el siguiente:

1. Detectar si hay líneas iguales o alguna es combinación lineal de otras, en cuyo caso se eliminan, porque el rango de la matriz resultante tras la eliminación es el mismo que el de la matriz primitiva.
2. Buscar una submatriz cuyo determinante sea distinto de cero.
3. Ampliar dicha submatriz con una fila y una columna de la matriz de todas las formas posibles. Si los determinantes de todas las ampliaciones son cero, el rango de la matriz es el orden del menor elegido. Si alguno es distinto de cero, estamos en el paso 2, y desde él se continúa el algoritmo hasta que todos los menores ampliados sean cero.

Ejemplo 3.5.2.

Para calcular el rango de $A = \begin{pmatrix} 5 & -1 & 3 & -2 & 4 \\ 17 & -2 & 5 & 1 & 7 \\ 19 & -1 & 1 & 8 & 2 \end{pmatrix}$, comenzamos por detectar mediante una simple observación que la 5ª columna es la 3ª menos la 2ª $\Rightarrow$ se puede eliminar cualquiera de las tres, que es combinación de las otras dos. La matriz que se obtiene es del mismo rango.

$$\text{rg}(A) = rg\begin{pmatrix} 5 & -1 & 3 & -2 & 4 \\ 17 & -2 & 5 & 1 & 7 \\ 19 & -1 & 1 & 8 & 2 \end{pmatrix} = \text{rg}\begin{pmatrix} 5 & -1 & 3 & -2 \\ 17 & -2 & 5 & 1 \\ 19 & -1 & 1 & 8 \end{pmatrix}$$

En este paso, es posible que no hubiéramos detectado nada, y hubiéramos pasado al siguiente:

El menor $\begin{vmatrix} 5 & -1 \\ 17 & -2 \end{vmatrix} = 7 \neq 0 \Rightarrow$ El rango es, al menos, 2.

Las posibles submatrices ampliadas de $\begin{pmatrix} 5 & -1 \\ 17 & -2 \end{pmatrix}$ son:

$$\begin{pmatrix} 5 & -1 & -3 \\ 17 & -2 & 5 \\ 19 & -1 & 1 \end{pmatrix} \text{y} \begin{pmatrix} 5 & -1 & -2 \\ 17 & -2 & 1 \\ 19 & -1 & 8 \end{pmatrix}$$

Sus determinantes son:

$$\begin{vmatrix} 5 & -1 & -3 \\ 17 & -2 & 5 \\ 19 & -1 & 1 \end{vmatrix} = 0 \quad ; \quad \begin{vmatrix} 5 & -1 & -2 \\ 17 & -2 & 1 \\ 19 & -1 & 8 \end{vmatrix} = 0$$

Como ninguno es $\neq 0 \Rightarrow \text{rg}(A) = 2$.

En *[2.3.4]* definimos el rango de un matriz como el de la aplicación asociada, los endomorfismos son un caso particular de las aplicaciones lineales y como es natural la definición es válida.

En el teorema *[2.3.5]* demostramos que, "el rango de una matriz es el número de vectores columna linealmente independientes que hay en ella", y a este número de vectores independientes le llamamos rango del sistema de vectores; como consecuencia, podemos concluir con el resultado:

Rango de un endomorfismo = Rango de la matriz asociada = Rango del sistema de vectores que forma la matriz.

EJERCICIOS

Ejercicio 3.1.

a) Calcúlese mediante determinantes la inversa de la matriz

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

b) Determínense las relaciones existentes entre A y A^{-1}; A^2 e I.

Solución:

a) Matriz inversa:

La matriz tiene inversa porque su determinante es distinto de cero.

♦ Cálculo de $|A|$.

Como es un determinante de orden cuatro, el cálculo mediante la aplicación directa de la definición es muy largo, utilizaremos las propiedades necesarias para reducirlo a otros de menor orden.

Los pasos a seguir son:

1. Hacer ceros los elementos de la primera columna. Mediante la sustitución de cada fila por ella misma menos la primera fila.
2. Desarrollar el determinante por los dos adjuntos de los elementos de la 1.ª columna.
3. Hacer cero el elemento a_{33} de la nueva matriz sustituyendo la 3.ª fila por sí misma menos la 2.ª fila.
4. Desarrollar el determinante por los adjuntos de los elementos de la 1.ª columna.

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

$$= -2 \cdot (-1) \cdot \begin{vmatrix} -2 & -2 \\ -2 & 2 \end{vmatrix} = -16$$

♦ Cálculo de la matriz adjunta de A.

Los adjuntos de cada elemento de la matriz son:

$$A_{11} = \begin{vmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{vmatrix} = -4; \qquad A_{22} = A_{33} = A_{44} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{vmatrix} = -4;$$

$$A_{12} = \begin{vmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{vmatrix} = -4; \qquad A_{13} = \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \end{vmatrix} = -4;$$

$$A_{14} = \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{vmatrix} = -4; \qquad A_{23} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \end{vmatrix} = -(-4) = 4;$$

$$A_{24} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{vmatrix} = -4; \qquad A_{34} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{vmatrix} = -(-4) = 4;$$

Observemos que la matriz es simétrica, y por tanto, el adjunto de un elemento coincide con el del simétrico respecto a la diagonal principal $A_{ij} = A_{ji}$.

La matriz adjunta $\overline{A}$ es: $\overline{A} = \begin{pmatrix} -4 & -4 & -4 & -4 \\ -4 & -4 & 4 & 4 \\ -4 & 4 & -4 & 4 \\ -4 & 4 & 4 & -4 \end{pmatrix}$

La matriz inversa $A^{-1} = \dfrac{1}{|A|}(\overline{A})' = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & -1/4 & -1/4 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ 1/4 & -1/4 & -1/4 & 1/4 \end{pmatrix} = \dfrac{1}{4}A.$

b, La relación que existe entre A y A^{-1} es: $A^{-1} = \dfrac{A}{4}$.

La relación que existe entre A^2 e I es: $4I = A^2$, porque $I = AA^{-1} = A\,\dfrac{A}{4} =$
$= \dfrac{A^2}{4}$.

Ejercicio 3.2

Descompóngase en producto de factores el siguiente determinante:

$$\begin{vmatrix} x^2 & xy & y^2 \\ 2x & x+y & 2y \\ 1 & 1 & 1 \end{vmatrix}.$$

Solución:

$$\begin{vmatrix} x^2 & xy & y^2 \\ 2x & x+y & 2y \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} x^2-y^2 & xy-y^2 & y^2 \\ 2x-2y & x-y & 2y \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} x^2-y^2 & xy-y^2 \\ 2x-2y & x-y \end{vmatrix} =$$

$$= (x-y)^2 \begin{vmatrix} x+y & y \\ 2 & 1 \end{vmatrix} = (x-y)^3$$

Hemos hecho las transformaciones $C_1 \longrightarrow C_1 - C_3$ y $C_2 \longrightarrow C_2 - C_3$ en la primera matriz, y posteriormente hemos sacado factor común $(x - y)$ en cada una de las filas (o columna).

Ejercicio 3.3

Demuéstrese, aplicando las propiedades (sin resolver) que el siguiente determinante es múltiplo de 246: $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 8 & 6 & 1 \end{vmatrix}$.

Solución:

El número $246 = 2 \times 123$

Sustituyendo $C_1 \longrightarrow 100C_1 + 10C_2 + C_3$, se obtiene una matriz cuya primera columna está formada por múltiplos de 123, y la segunda columna por múltiplos de 2.

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 8 & 6 & 1 \end{vmatrix} = \begin{vmatrix} 123 & 2 & 3 \\ 246 & 4 & 6 \\ 861 & 6 & 1 \end{vmatrix} = 123 \cdot 2 \begin{vmatrix} 1 & 1 & 3 \\ 2 & 2 & 6 \\ 7 & 3 & 1 \end{vmatrix}.$$

Ejercicio 3.4

Hállase el rango de la matriz: $A = \begin{pmatrix} 1 & a & -1 & 2 \\ 2 & -1 & a & 5 \\ 1 & 10 & -6 & 1 \end{pmatrix}$ para los distintos valores de a.

Solución:

El rango es al menos 2, ya que el menor: $\begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} = 1 \neq 0$.

Ampliando ese menor con filas y columnas de A se obtienen los dos siguientes:

$$\Delta_1 = \begin{vmatrix} 1 & a & 2 \\ 2 & -1 & 5 \\ 1 & 10 & 1 \end{vmatrix}; \qquad \Delta_2 = \begin{vmatrix} 1 & -1 & 2 \\ 2 & a & 5 \\ 1 & -6 & 1 \end{vmatrix};$$

cuyos valores son:

$$\Delta_1 = \begin{vmatrix} 1 & a & 2 \\ 2 & -1 & 5 \\ 1 & 10 & 1 \end{vmatrix} = \begin{vmatrix} 1 & a & 2 \\ 0 & -1-2a & 1 \\ 0 & 10-a & -1 \end{vmatrix} = \begin{vmatrix} -1-2a & 1 \\ 10-a & -1 \end{vmatrix} = 1 + 2a - 10 + a = 3a - 9 =$$
$$= 3(a-3);$$

$$\Delta_2 = \begin{vmatrix} 1 & -1 & 2 \\ 2 & a & 5 \\ 1 & -6 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 \\ 0 & a+2 & 1 \\ 0 & -5 & -1 \end{vmatrix} = \begin{vmatrix} a+2 & 1 \\ -5 & -1 \end{vmatrix} = -a + 3;$$

Si $a = 3 \Rightarrow \Delta_1 = 0$ y $\Delta_2 = 0 \Rightarrow \mathrm{rg}(A) = 2$.

Si $a \neq 3 \Rightarrow \mathrm{rg}(A) = 3$

Ejercicio 3.5

Demuéstrese que la condición necesaria y suficiente para que una matriz cuadrada regular A de orden n y elementos enteros, tenga por inversa otra matriz A^{-1} con elementos enteros es que $|A| = \pm 1$.

Solución:

Demostrar una condición necesaria y suficiente es demostrar una doble implicación. En este caso, para demostrar que "Los elementos de A^{-1} son enteros $\Leftrightarrow |A| = \pm 1$", hay que ver que:

- Si los elementos de A^{-1} son enteros $\Rightarrow |A| = \pm 1$.

Si los elementos de A y A^{-1} son números enteros, también los son $|A|$ y $|A^{-1}|$.

Como $|A^{-1}| = \dfrac{1}{|A|}$, se deduce que $|A| = \pm 1$, ya que, los únicos enteros que admiten inverso entero también son ± 1.

- Si $|A| = \pm 1$ y los elementos de A son enteros $\Rightarrow$ los elementos de A^{-1} son enteros.

Si la matriz A tiene sus elementos números enteros, también los tendrá su matriz adjunta, $\bar{A}$, y la transpuesta de la adjunta $(\bar{A})'$.

Por tanto, al ser $|A| = \pm 1$, y $A^{-1} = \dfrac{1}{|A|} \cdot (\bar{A})'$ la matriz A^{-1} tiene todos sus elementos enteros.

Ejercicio 3.6

a) Calcúlese el determinante: $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix}$;

b) Basada en la respuesta a), escríbase una solución para $\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix}$

y compruébese después.

Solución:

a)

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ a & b-a & c-a \\ a^2 & b^2-a^2 & c^2-a^2 \end{vmatrix} = (b-a)(c-a)\begin{vmatrix} 1 & 1 \\ b+a & c+a \end{vmatrix} =$$

$$= (b-a)(c-a)[(c+a)-(b+a)] = (b-a)(c-a)(c-b).$$

b) El determinante dado es de la misma forma que el primero, todos los que tienen dicha forma se llaman de Vandermonde.

Podemos aventurar sin miedo a error que:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix} = (b-a)(c-a)(d-a)(c-b)(d-b)(d-c)$$

porque el proceso seguido es válido, sea cual sea el orden del determinante.

Se deja al lector la comprobación del resultado.

Los determinantes de Vandermonde tienen aplicación en interpolación polinómica de curvas, muy empleada en diseño gráfico por ordenador.

Ejercicio 3.7

Demuéstrese utilizando sus propiedades que los determinantes de los siguientes cuadrados mágicos tienen el mismo valor absoluto:

$$\Delta_1 = \begin{vmatrix} 6 & 1 & 8 \\ 7 & 5 & 3 \\ 2 & 9 & 3 \end{vmatrix}; \quad \Delta_1 = \begin{vmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{vmatrix}; \quad \Delta_3 = \begin{vmatrix} 2 & 7 & 6 \\ 9 & 5 & 1 \\ 4 & 3 & 8 \end{vmatrix}; \quad \Delta_1 = \begin{vmatrix} 6 & 7 & 2 \\ 1 & 5 & 9 \\ 8 & 3 & 4 \end{vmatrix};$$

$$\Delta_5 = \begin{vmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{vmatrix}; \quad \Delta_6 = \begin{vmatrix} 2 & 9 & 4 \\ 7 & 5 & 3 \\ 6 & 1 & 8 \end{vmatrix}; \quad \Delta_7 = \begin{vmatrix} 8 & 3 & 4 \\ 1 & 5 & 9 \\ 6 & 7 & 2 \end{vmatrix}; \quad \Delta_8 = \begin{vmatrix} 4 & 3 & 8 \\ 9 & 5 & 1 \\ 2 & 7 & 6 \end{vmatrix};$$

Solución:

$\Delta_1 \Longleftrightarrow \Delta_2$ se pasa intercambiando C_1 con $C_3 \Rightarrow$ tienen el mismo valor absoluto.

$\Delta_3 \Longleftrightarrow \Delta_4$ se pasa intercambiando C_1 con $C_3 \Rightarrow$ tienen el mismo valor absoluto.

$\Delta_5 \Longleftrightarrow \Delta_6$ se pasa intercambiando C_1 con $C_3 \Rightarrow$ tienen el mismo valor absoluto.

$\Delta_7 \Longleftrightarrow \Delta_8$ se pasa intercambiando C_1 con $C_3 \Rightarrow$ tienen el mismo valor absoluto.

$\Delta_3 \Longleftrightarrow \Delta_6$ son traspuestos $\Rightarrow$ tienen el mismo valor absoluto.

$\Delta_5 \Longleftrightarrow \Delta_8$ son traspuestos $\Rightarrow$ tienen el mismo valor absoluto.

$\Delta_1 \Longleftrightarrow \Delta_6$ se pasa intercambiando F_1 con F_3 tienen el mismo valor absoluto.

- Estos cuadrados considerados mágicos desde la antigüedad porque su suma da el mismo resultado en todas las direcciones son los únicos que se pueden formar con los nueve primeros números enteros y positivos.

Ejercicio 3.8

Calcúlense las raíces reales de la ecuación: $\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+x & 1 \\ 1 & 1 & 1 & 1-x \end{vmatrix} = 81.$

Solución

$$\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+x & 1 \\ 1 & 1 & 1 & 1-x \end{vmatrix} = \begin{vmatrix} 1+x & 1 & 1 & 1 \\ -x & -x & 0 & 0 \\ -x & 0 & x & 0 \\ -x & 0 & 0 & -x \end{vmatrix} = -\begin{vmatrix} -x & 0 & 0 \\ -x & x & 0 \\ -x & 0 & -x \end{vmatrix} - x\begin{vmatrix} 1+x & 1 & 1 \\ -x & x & 0 \\ -x & 0 & -x \end{vmatrix} =$$

$$= -x^3 - xxx\begin{vmatrix} 1+x & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & -1 \end{vmatrix} = -x^3 - x^3\begin{vmatrix} x & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & -1 \end{vmatrix} = -x^3 - x^3\begin{vmatrix} x & 1 \\ -1 & 1 \end{vmatrix} =$$

$$-x^3 + x^4 - x^3 = x^4.\ x^4 = 81 \Rightarrow x = \pm 3.$$

Las operaciones elementales que hemos ido haciendo para llegar a este resultado son:

1. Sustituir las filas 2.ª y 3.ª y 4.ª por ellas mismas menos la 1.ª fila.
2. Desarrollar por los adjuntos de la segunda columna.
3. Sacar factor común una x en la 2.ª fila y otra x en la 3.ª fila del segundo determinante de orden tres.
4. Sustituir la 1.ª fila por la 1.ª fila más la 3.ª fila en dicho determinante de orden tres.
5. Desarrollar ese determinante por los elementos de la última columna.

Ejercicio 3.9

Dado el sistema de vectores $\{(1, 2, 3), (2, 1, 5), (2, 3, a) \in \mathbb{R}^3\}$, se pide el valor que debe tener a para que el sistema sea una base de $\mathbb{R}^3$.

Solución:

Para que formen base de $\mathbb{R}^3$ deben ser tres vectores linealmente independientes, y por tanto, el rango del sistema de vectores deben ser tres, es decir,

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 5 \\ 2 & 3 & a \end{vmatrix} \neq 0.$$

Si $a = 17/3$, los tres vectores forman base.

Si $a = 17/3$, el tercer vector es combinación lineal de los otros y no forman base.

Ejercicio 3.10

Determínese para qué valores reales de x pueden existir matrices cuadradas no nulas B tales que $AB = 0$, siendo $A = \begin{pmatrix} 1 & 3 \\ 12 & x^2 \end{pmatrix}$.

Solución:

La condición necesaria es que $|A| = 0$, ya que si no, existe A^{-1} y se verifica: $A^{-1}(AB) = IB = 0 \Rightarrow B = 0$.

$$Si\ |A| = \begin{vmatrix} 1 & 3 \\ 12 & x^2 \end{vmatrix} = 0 \Rightarrow x = \pm 6.$$

CAPÍTULO

4 SISTEMAS DE ECUACIONES LINEALES

INTRODUCCIÓN

Las ecuaciones lineales aparecen en los primeros testimonios escritos de que tenemos noticia. Ya los babilonios habían desarrollado métodos concretos de resolución de sistemas lineales sencillos basados en la "eliminación" sucesiva de incógnitas hasta reducir el problema a una sucesión de "reglas de tres", es decir, ecuaciones de una sola incógnita.

Distintas reglas prácticas se establecen a lo largo de la Historia para mejorar el proceso y aligerar los cálculos, pero hasta el siglo XVIII nadie parece preocuparse de caracterizar cuándo es posible este proceso de reducción. En todo caso, si no se obtiene la solución, se dice que el problema es "imposible" o "indeterminado". Solamente cuando, al estudiar ciertos problemas de Geometría Algebraica y Mecánica, aparecen sistemas de ecuaciones lineales cuyos coeficientes son, a su vez, funciones de parámetros variables, empieza a desarrollarse a partir de 1750 una teoría general de tales sistemas.

La primera fórmula de cálculo explícito de la solución de un sistema de n ecuaciones lineales con n incógnitas, con coeficientes indeterminados, se debe a MacLaurin en 1729. MacLaurin obtuvo las fórmulas para $n = 2$ y $n = 3$ y trató de encontrar una regla general. En 1750, aparentemente sin conocer el trabajo de MacLaurin, G. Cramer describe explícitamente la fórmula para n cualquiera, dando la solución como cociente de 2 expresiones polinomiales en los coeficientes. Con el desarrollo de la teoría de determinantes, estos resultados toman la forma definitiva que conocemos en la actualidad. La introducción por parte de Frobenius de la noción de rango de una matriz en 1879, le permitió formular el criterio general de resolución de un sistema lineal de m ecuaciones con n incógnitas, que hoy conocemos como teorema de Rouché-Frobenius.

Salvo la determinación de métodos más efectivos y directos de cálculo (método de Gauss, o métodos iterativos: Gauss, Seidel, Jacobi, etc.), la teoría general de resolución de sistemas lineales en dimensión finita se encuentra ya, alrededor de 1870, en el mismo estado que en la actualidad.

ADA AUGUSTA LOVELACE (1815-1852)

Hija del poeta romántico Lord Byron fue una niña enfermiza pero, con una gran fuerza de voluntad, superó la enfermedad y la adicción a la droga que le produjeron las medicinas que le suministraron para curarla.

Su madre había estudiado Matemáticas y tenía gran fe en el papel redentor del estudio, por ello procuró a Ada una buena educación científica, que la niña asimiló rápidamente.

A los 17 años se interesó e impresionó por la máquina de diferencias finitas de Babbage, comenzó a estudiar y quiso ser matemática, pero como Babbage estaba muy ocupado en recaudar fondos para su máquina, y no prestaba atención a sus trabajos, abandonó la idea.

Ayudada por su esposo, el conde de Lovelace, consiguió que Babbage "la entrenara".

El ingenio analítico o máquina de Babbage tenía un dispositivo de entrada (parecido a las tarjetas perforadas del telar de Jacquard), un almacén (Memoria), un molino (Procesador) y un dispositivo de salida en forma de tarjetas.

En unas notas sobre el ingenio, Ada tuvo la idea de reutilizar las tarjetas encargadas de un procedimiento todas las veces que hiciera falta dentro del mismo programa, incorporando con ello la herramienta de las subrutinas al programa principal.

La ingeniería de la época no era suficientemente avanzada para construir la máquina, ella creyó que sus cálculos le harían ganar en las carreras de caballos, y con las garantías, tener dinero para construir la máquina, pero los continuos fallos llevarían a la familia a la banca rota, por ello, quemó todos sus trabajos.

. Está considerada precursora de la programación de ordenadores porque ideó varios programas para hacer cálculos matemáticos avanzados en la máquina analítica.

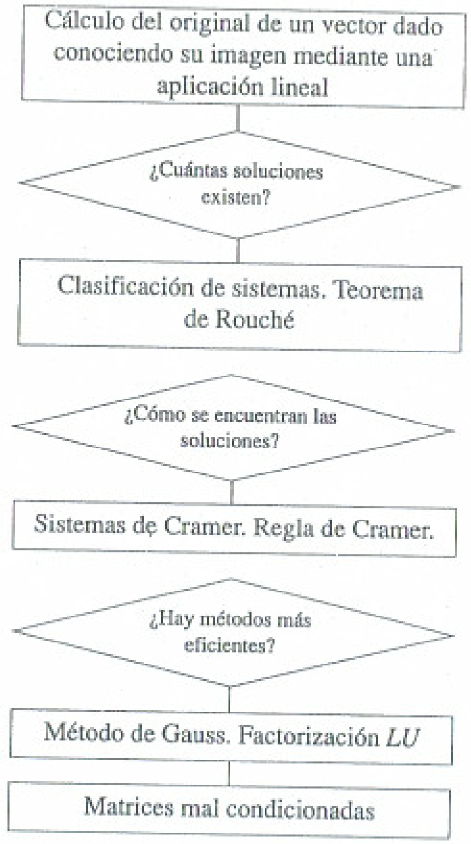
Ha dado nombre a un lenguaje de alto nivel basado en el Pascal que fue desarrollado a partir de 1977 en el Departamento de Defensa de EE.UU.

En 1843 enferma gravemente y sufre dolores horribles, nuevamente es tratada con opiáceos alternando con morfina, que contrarrestaba con grandes cantidades de brandi.

En 1844 está convencida de que un exceso de matemáticas ha perjudicado su salud, y abandona definitivamente las matemáticas, pero no el juego; nada le ayudaba a superar el dolor que le producía un cáncer terminal.

La máquina de Babbage y sus teorías fueron olvidadas, y con ellas Ada Byron, hasta que 101 años después son rescatadas del olvido por Bowden, un pionero inglés de la reinvención de los ordenadores.

CAPÍTULO 4



4.1. Sistemas de ecuaciones lineales

En este capítulo queremos alcanzar los siguientes objetivos:

- Determinar cuándo el sistema tiene solución y cuando no.
- En caso de tener solución, determinar cuántas soluciones tiene.
- Dar un procedimiento para encontrar todas las soluciones.

Antes de abordar la forma de alcanzar los objetivos propuestos con la herramienta que nos proporcionan los conocimientos adquiridos en capítulos anteriores, es conveniente recordar, aunque sea brevemente, algunos conceptos y definiciones básicas para poder unificar las distintas notaciones utilizadas anteriormente por el lector.

4.1.1. Definición de sistema de ecuaciones lineales

Un sistema de m ecuaciones lineales con n incógnitas es todo conjunto de ecuaciones de la forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases},$$

$a_{ij}, b_i \in \mathbb{R}$ y x_j son las incógnitas ($i = 1, \ldots, m$ y $j = 1, \ldots, n$)

Cuando todos los términos independientes $b_1, b_2, \ldots, b_m$ son cero, el sistema se llama homogéneo, y en caso contrario no homogéneo.

Para todo sistema existe un sistema homogéneo asociado que resulta de hacer cero los términos independientes del mismo.

Se llama solución del sistema a toda n-upla $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ de elementos de R, tales que sustituidos en $(x_1, x_2, \ldots, x_n)$, todas las ecuaciones del sistema se convierten en identidades. Resolver el sistema es encontrar todas las $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ que lo verifican.

Dos sistemas de ecuaciones con el mismo conjunto de soluciones, se llaman equivalentes. Convertir un sistema lineal en otro equivalente, pero más sencillo de resolver es una técnica habitual para buscar sus soluciones. Como veremos más adelante también aquí utilizaremos esa técnica.

El sistema lineal de m ecuaciones con n incógnitas definido en *[4.1.1]* se puede expresar también en forma matricial.

4.1.2. Expresión matricial de un sistema de ecuaciones lineales

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \ldots \ldots \ldots \ldots \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \Leftrightarrow \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix} \Leftrightarrow AX = B$$

donde A es la matriz de los coeficientes de las incógnitas, y X, B son los vectores columna de las incógnitas y de los términos independientes, respectivamente.

Esta expresión matricial de un sistema de ecuaciones lineal es la expresión analítica de una aplicación lineal, como se vio en el capítulo 2, el vector imagen Y de un vector original X es ahora el vector de los términos independientes B.

4.1.3. Un sistema de ecuaciones lineales es la expresión analítica de una aplicación lineal

El sistema $AX = B$ de m ecuaciones lineales con n incógnitas es la expresión analítica de una aplicación lineal: $f: V \to W$, donde:
V es un espacio vectorial real de dimensión n (n.° de incógnitas)
W es un espacio vectorial real de dimensión m (n.° de ecuaciones)
y cada columna i de su matriz asociada A es el vector formado por los coeficientes de la variable x_i en las m ecuaciones.

Ejemplo 4.1.1

El sistema de ecuaciones lineales $\begin{cases} 3x_1 + 2x_2 - x_3 = -1 \\ 2x_1 + x_3 = 5 \end{cases}$, se puede expresar

matricialmente: $\begin{pmatrix} 3 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$

Esta expresión indica que la imagen del vector (x_1, x_2, x_3) mediante la aplicación cuya matriz asociada es $\begin{pmatrix} 3 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}$ es el vector $(-1, 5)$.

Si en la expresión matricial cambiamos el vector $\begin{pmatrix} -1 \\ 5 \end{pmatrix}$ por el vector $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, obtenemos: $\begin{pmatrix} 3 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, que es la expresión analítica de la aplicación lineal asociada a la misma matriz:

$$(x_1, x_2, x_3) \to \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Recordemos que las columnas de la matriz son las imágenes de los vectores de una base de $\mathbb{R}^3$ respecto a una base de $\mathbb{R}^2$ (las canónicas, si no se especifica lo contrario).

$f: \mathbb{R}^3 \to \mathbb{R}^2$
$f(1, 0, 0) = (3, 2); f(0, 1, 0) = (2, 0)\ ; f(0, 0, 1) = (-1, 1)$

en las bases $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ de $\mathbb{R}^3$, y $S = \{(1, 0), (0, 1)\}$ de $\mathbb{R}^2$.

Una solución del sistema $\begin{pmatrix} 3 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ es $(1, 0, 0)$.

Una solución del sistema $\begin{pmatrix} 3 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ es $(0, 1, 0)$.

Una solución del sistema $\begin{pmatrix} 3 & 2 & -1 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ es $(0, 0, 1)$.

La comprobación de los resultados anteriores es muy fácil.
El hecho de que un sistema de ecuaciones lineales sea la expresión analítica de una aplicación lineal, nos permite decir que una solución de un sistema de ecuaciones es un original de un vector dado, B. Por tanto, la existencia o no, de soluciones depende de que dicho vector pertenezca, o no, respectivamente a la imagen de la aplicación.

4.1.4. Cálculo del original de un vector en una aplicación lineal

El conjunto de soluciones del sistema $AX = B$ es el conjunto original del vector $B \in W$ en la aplicación lineal $f: V \to W$ de expresión analítica $AX = Y$.

Ejemplo 4.1.2

En el ejemplo anterior, hallar el conjunto de soluciones del sistema, equivale a encontrar el conjunto original del vector $B = (-1, 5)$ en la aplicación lineal correspondiente.

La solución del sistema es

$$\begin{cases} 3x_1 + 2x_2 - x_3 = -1 \\ 2x_1 \qquad + x_3 = 5 \end{cases} \Leftrightarrow \begin{cases} 3x_1 + 2x_2 - x_3 = -1 \\ x_3 = -2x_1 + 5 \end{cases} \Rightarrow 5x_1 + 2x_2 = 4 \Rightarrow \begin{cases} x_1 = \lambda \\ x_2 = \dfrac{4-5\lambda}{2} \\ x_3 = -2\lambda + 5 \end{cases}$$

El conjunto S de soluciones $S = \left\{(x_1, x_2, x_3) = \left(\lambda, \dfrac{4-5\lambda}{2}, -2\lambda + 5\right), \lambda \in \mathbb{R}\right\}$

es el conjunto original del vector $B = (-1, 5)$ en la aplicación lineal f dada.

Recordemos que el núcleo de una aplicación lineal $f: V \to W$ es $\mathrm{Nuc}(f) = \{\overline{v} \in V$, tales que $f(\overline{v}) = \overline{0}\}$, es decir, el conjunto de vectores que forman el núcleo de una aplicación lineal es el conjunto de soluciones del sistema homogéneo asociado al sistema que la define.

4.1.5. Cálculo del núcleo de una aplicación lineal

> El núcleo $\mathrm{Nuc}(f)$ de la aplicación lineal $f: V \to W$ de expresión analítica $AX = Y$ es el conjunto de soluciones del sistema homogéneo $AX = 0$

Ejemplo 4.1.3

El núcleo de la aplicación de los ejemplos anteriores es la solución del sistema homogéneo $\begin{cases} 3x_1 + 2x_2 - x_3 = 0 \\ 2x_1 \qquad + x_3 = 0 \end{cases}$

Es decir:

$$\begin{cases} 3x_1 + 2x_2 - x_3 = 0 \\ 2x_1 \qquad + x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} 3x_1 + 2x_2 - x_3 = 0 \\ x_3 = -2x_1 \end{cases} \Rightarrow 5x_1 + 2x_2 = 0 \Rightarrow \begin{cases} x_1 = \lambda \\ x_2 = \dfrac{-5\lambda}{2} \\ x_3 = -2\lambda \end{cases}$$

$$\mathrm{Nuc}(f) = \left\{\left(\lambda, \frac{-5}{2}\lambda, -2\lambda\right), \lambda \in \mathbb{R}\right\}$$

4.1.6. Proposición

> El conjunto de soluciones de un sistema lineal, si el sistema tiene solución, es igual al conjunto de soluciones del sistema homogéneo asociado más una solución particular del no homogéneo

Demostración:

Sea $AX = B$ un sistema lineal no nomogéneo y f la aplicación lineal correspondiente, $f: V \to W$ de matriz asociada A.

Si $S = \{X \in V \,/\, f(X) = B\}$ es el conjunto de soluciones del sistema, y G una solución particular del mismo, se verifica $f(G) = B$.

Sabemos que el conjunto de soluciones del sistema homogéneo $AX = \overline{0}$ asociado es el núcleo $\mathrm{Nuc}(f)$. Por tanto, $S = \{X \in V \,/\, f(X) = B\} = \{X \in V \,/\, f(X) = f(G)\} = \{X \in V \,/\, f(X - G) = \overline{0}\} = \{X \in V \,/\, (X - G) \in \mathrm{Nuc}(f)\}$, como consecuencia $S = \mathrm{Nuc}(f) + G$.

Ejemplo 4.1.4

En el sistema analizado en ejemplos anteriores, $\mathrm{Nuc}(f)=\left\{\left(\lambda, \frac{-5}{2}\lambda, -2\lambda\right),\ \lambda\in\mathbb{R}\right\}$

Una solución particular del sistema $\begin{cases}3x_1+2x_2-x_3=-1\\ 2x_1 \qquad\quad +x_3=5\end{cases}$ es $G=(2,-3,1)$,

ya que $\begin{pmatrix}3&2&-1\\2&0&1\end{pmatrix}\begin{pmatrix}2\\-3\\1\end{pmatrix}=\begin{pmatrix}-1\\5\end{pmatrix}$.

Por tanto, el conjunto de soluciones del sistema es:

$$S=\left\{(2,-3,1)+\left(\lambda,\frac{-5}{2}\lambda,-2\lambda\right),\ \lambda\in\mathbb{R}\right\}.$$

4.2. Clasificación de sistemas de ecuaciones lineales

En cualquier algoritmo para buscar soluciones de un sistema, parece natural comprobar en primer lugar si existen, porque si no es así, se evitan todos los pasos intermedios que conducen a encontrarlas.

Los sistemas se pueden clasificar en dos grupos: compatibles e incompatibles, según tengan, o no, solución; en el primer caso puede ser determinado o indeterminado dependiendo de que tenga una solución o más.

El criterio para resolver el problema de existencia de soluciones viene dado por el Teorema de Rouché-Frobenius, con el que el lector está muy familiarizado.

4.2.1. Teorema de Rouché-Frobenius

El sistema $AX=B$ tiene solución $\Leftrightarrow$ El rango de A = rango de la matriz A ampliada con la columna B

Demostración:

Dado el sistema de m ecuaciones lineales con n incógnitas:

$$\begin{cases}a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2\\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m\end{cases}$$

de expresión matricial $AX=B$, es posible la siguiente expresión vectorial:

$$x_1\begin{pmatrix}a_{11}\\a_{21}\\\ldots\\a_{m1}\end{pmatrix}+x_2\begin{pmatrix}a_{12}\\a_{22}\\\ldots\\a_{m2}\end{pmatrix}+\ldots+x_n\begin{pmatrix}a_{21}\\a_{2n}\\\ldots\\a_{mn}\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\\ldots\\b_m\end{pmatrix}$$

Si llamamos $A_1, A_2, ..., A_n$, a los vectores columna de la matriz A (cada vector columna A_i tiene como elementos los coeficientes de la incógnita x_i en todas las ecuaciones del sistema) y B es el vector de los términos independientes, podemos escribir el sistema mediante la siguiente expresión: $x_1A_1 + x_2A_2 + ... + x_nA_n = B$.

Evidentemente, existirán valores de $x_1, x_2, ..., x_n$ que hagan cierta esa igualdad si, y sólo si el vector B es combinación lineal de los vectores $A_1, A_2, ..., A_n$;

es decir, si las matrices: $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$ y $(A \mid B) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{pmatrix}$ tienen el mismo rango.

4.2.2. Proposición

> Un sistema lineal compatible tiene solución única (es determinado) $\Leftrightarrow$ El rango de A = Número de incógnitas n

Demostración:

Sea $AX = B$ un sistema compatible y sea G una solución del mismo.

Sabemos que $AX = Y$ es la expresión matricial de una aplicación lineal $f: V \to W$, tal que, A es la matriz asociada a f, V es un espacio vectorial de dimensión n (n.º de incógnitas), y que el conjunto de soluciones de un sistema lineal, si el sistema tiene solución, es igual al conjunto de soluciones del sistema homogéneo asociado más una solución particular del no homogéneo. *[4.1.6]*, es decir, el conjunto de soluciones del sistema es Nuc(f) + G.

Como dim Nuc (f) + dim Im(f) = dim $V = n$ *[2.2.5]* y rg (A) = dim Im(f) *[2.3.4]* $\Rightarrow$ dim Nuc(f) + rg (A) = n.

El cumplimiento de esta igualdad abre dos posibilidades:

1. rg (A) = n, entonces dim Nuc(f) = 0, el conjunto de soluciones es sólo la solución particular G, y el sistema es determinado.
2. rg (A) < n, entonces dim Nuc(f) > 0 y el sistema tiene infinitas soluciones; es un sistema indeterminado.

El siguiente cuadro resume los criterios dados en el teorema de Rouché y la proposición anterior:

Sistema Compatible (tiene solución)
Rango de A = Rango de la matriz A ampliada con la columna B de términos independientes

- Determinado (solución única)
Rango de A = Número de incógnitas n
- Indeterminado (infinitas soluciones)
Rango de A < Número de incógnitas n

Ejemplo 4.2.1

Dado el sistema de ecuaciones lineales $\begin{cases} ax + y + z = 1 \\ x + ay + z = 1 \\ x + y + az = 1 \end{cases}$, se pide clasificarlo en función del valor del parámetro real a.

Solución:

La matriz de los coeficientes de las incógnitas es $A = \begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$, y la matriz ampliada con el vector de los términos independientes es

$$(A|B)=\begin{pmatrix} a & 1 & 1 & 1\\ 1 & a & 1 & 1\\ 1 & 1 & a & 1\end{pmatrix}.$$

$|A| = (a+2)(a-1)^2 \Rightarrow |A| = 0$ si $a=-2$ o $a=1$.

Vamos a analizar qué ocurre con los rangos de ambas matrices en cada caso:

Si $a=-2 \Rightarrow A=\begin{pmatrix} -2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{pmatrix}$ y $\text{rg}(A)=2$; $(A|B)=\begin{pmatrix} -2 & 1 & 1 & 1\\ 1 & -2 & 1 & 1\\ 1 & 1 & -2 & 1\end{pmatrix}$ y $\text{rg}(A|B)=3$.

Como los rangos de A y $A|B$ son distintos, el sistema es incompatible.

Si $a=1 \Rightarrow \text{rg}(A)=\text{rg}(A|B)=1$

Como los rasgos de A y $A|B$ son iguales, el sistema es compatible.

En este caso, el rango es menor que el número de incógnitas $n=3 \Rightarrow$ es un sistema compatible e indeterminado.

4.2.3. Sistemas lineales homogéneos

Un sistema homogéneo es siempre compatible.
Un sistema homogéneo es indeterminado (tiene soluciones distintas de la trivial) $\Leftrightarrow$
El rango de $A <$ Número de incógnitas n.

Los sistemas homogéneos son un caso particular de los sistemas lineales. En ellos el vector B de los términos independientes es el vector $\overline{0}$, por tanto, el rango de la matriz A de las incógnitas coincide con el rango de la matriz ampliada $A|\overline{0}$, es decir, el sistema es siempre compatible; este hecho es evidente porque

un sistema homogéneo tiene siempre la solución trivial: $x_1 = x_2 = \dots = x_n = 0$.

El interés es el caso de este tipo de sistemas radica en el conocimiento de soluciones distintas de la trivial, que existirán cuando el rango de la matriz de coeficientes sea menor que el número de incógnitas.

Para los sistemas homogéneos una solución particular es $G=(0,0,\dots,0) \Rightarrow$ $\Rightarrow S = G + \text{Nuc}(f) = \text{Nuc}(f)$.

Ejemplo 4.2.2

El sistema $\begin{cases} x \quad\quad -2z = 0\\ -x+y+\ z = 0\\ -x+y-3z=0\end{cases}$ es compatible por ser homogéneo, y es indeterminado, ya que $\text{rg}A = \text{rg}\begin{pmatrix} 1 & 0 & -2\\ -1 & 1 & 1\\ 1 & 1 & -3\end{pmatrix} = 2 <$ n.º incógnitas $= 3$.

4.3. Cálculo de soluciones

Cuando se tiene la certeza de la existencia de soluciones de un sistema lineal (sistema compatible), el paso siguiente en cualquier algoritmo eficiente corresponde al cálculo de las mismas.

Vamos a exponer dos métodos concretos de resolución de sistemas; consideraremos en primer lugar la regla de Cramer, el segundo procedimiento a que nos referimos es el Método de Gauss.

Como vamos a ver, teóricamente con el primer procedimiento queda zanjado de forma definitiva el problema de la resolución de sistemas de ecuaciones, pero sólo de forma teórica, porque cuando el orden es grande, el método de Cramer (o, lo que es lo mismo, la obtención de la inversa de la matriz de los coeficientes) no es práctico por el elevado número de operaciones que necesita.

Valga como ejemplo decir que, la resolución de un sistema de Cramer de orden n por la regla de Cramer, conlleva un total del orden de $(n + 1)!$ sumas, $(n + 2)!$ productos y n divisiones. Esto supone, por ejemplo, que para $n = 10$ se requieran aproximadamente 400.000.000 operaciones.

4.3.1. Definición de sistema de Cramer

Un sistema de ecuaciones lineales se llama de Cramer si tiene el mismo número de ecuaciones que de incógnitas y la matriz de los coeficientes es regular

Estamos analizando el caso particular en el que f es un isomorfismo, o lo que es lo mismo, A es una matriz cuadrada inversible.

Ejemplo 4.3.1.

El sistema $\begin{cases} x-2y+z=1 \\ 3x+y+z=-2 \\ x+y+z=1 \end{cases}$ es de Cramer, porque el número de ecuaciones = número de incógnitas = 3.

Además, la matriz A es regular

$$|A| = \begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -6 \neq 0 \quad A \text{ es regular.}$$

Para los sistemas de Cramer es obvio que existe solución, y es única: $X = A^{-1}B$.

Vamos a utilizar los desarrollos matriciales apropiados para obtener la conocida regla que enunciamos a continuación.

4.3.2. Regla de Cramer

Todo sistema de Cramer tiene solución única que viene dada por la expresión $x_k = \frac{c_k}{|A|}$ (c_k es el valor del determinante de la matriz que resulta al reemplazar en A su columna k-ésima por el vector columna B de los términos independientes)

Demostración:

Considérese el sistema de Cramer, cuya matriz A es de orden n.

$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$, su expresión matricial es $AX = B$, con A regular.

Como consecuencia del teorema de Rouché, se deduce que el sistema es compatible y determinado (tiene solución única).

Por ser A regular, existe A^{-1}, y se pueden multiplicar por ella por la izquierda ambos miembros de la igualdad $AX = B$, obteniendo como resultado:

$$X = A^{-1}B = \frac{1}{|A|}(\overline{A})^t B.$$

(Recordemos que $\overline{A}$ es la matriz adjunta de A, obtenida al sustituir en A cada elemento a_{ij} por su adjunto A_{ij}).

Ahora bien, el producto de matrices $(\overline{A})^t B$ es una matriz columna. Si $c_1, c_2, ..., c_n$ son sus elementos, podemos escribir:

$$(\overline{A})^t B = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{pmatrix}, \text{ siendo } c_K = b_1A_{1K} + b_2A_{2K} + \dots +$$

$$+ b_nA_{nK}.$$

Esta expresión coincide con el valor del determinante de la matriz que resulta al sustituir en A su coluna k por la columna de términos independientes B,

como es evidente al desarrollar el determinate $c_k = \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{n2} \\ a_{12} & \dots & b_2 & \dots & a_{n1} \\ \dots & \dots & \dots & \dots & \dots \\ a_{1n} & \dots & b_n & \dots & a_{nn} \end{vmatrix}$ por los

elementos de dicha columna

$$\text{Por tanto: } X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \frac{1}{|A|}(\overline{A})^t B = \frac{1}{|A|}\begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{pmatrix}.$$

Sabemos que para que dos vectores sean iguales, las coordenadas correspondientes han de ser iguales $\Rightarrow x_k = \dfrac{c_k}{|A|}$.

Ejemplo 4.3.2

Resuélvase el sistema $\begin{cases} 3x_1 - 4x_2 = 1 \\ x_1 + 5x_2 = -1 \end{cases}$

Solución:

Es un sistema de Cramer, tal que $|A| = \begin{vmatrix} 3 & -4 \\ 1 & 5 \end{vmatrix} = 19$

$$x_1 = \frac{\begin{vmatrix} 1 & -4 \\ -1 & 5 \end{vmatrix}}{\begin{vmatrix} 3 & -4 \\ 1 & 5 \end{vmatrix}} = \frac{1}{19}; \quad x_2 = \frac{\begin{vmatrix} 3 & 1 \\ 1 & -1 \end{vmatrix}}{\begin{vmatrix} 3 & -4 \\ 1 & 5 \end{vmatrix}} = \frac{-4}{19}$$

El caso general de un sistema compatible se reduce a uno de Cramer, eligiendo un subsistema principal de rango máximo en el que aparecen como parámetros en el término independiente las variables no principales.

La reducción de un sistema compatible a uno de Cramer es siempre posible, sin más que eliminar las ecuaciones que sean combinación lineal de otras y sustituir por parámetros las incógnitas que no sean necesarias para que la submatriz de A que define el nuevo sistema sea regular (la forma de elegir esta sumatriz no es única, pero las soluciones obtenidas sí lo son por ser sistemas equivalentes).

En el ejemplo siguiente indicaremos paso a paso el proceso indicado.

Ejemplo 4.3.3

Resuélvase el sistema $\begin{cases} 2x - 3y + z = 2 \\ x - y - z = -1 \\ 3x - 5y + 3z = 5 \end{cases}$ utilizando la regla de Cramer.

Solución:

$$A=\begin{pmatrix}2&-3&1\\1&-1&-1\\3&-5&3\end{pmatrix};\ (A\,|\,B)=\begin{pmatrix}2&-3&1&2\\1&-1&-1&-1\\3&-5&3&5\end{pmatrix};\ \mathrm{rg}\,(A)=\mathrm{rg}(A\,|\,B)=2.$$

Un menor de A de orden 2, no nulo es, por ejemplo $\begin{vmatrix}2&-3\\1&-1\end{vmatrix}=1\neq 0$

Por tanto, la tercera fila es combinación lineal de las otras dos. Eliminando esta tercera fila obtenemos el sistema equivalente $\begin{cases}2x-3y+z=2\\ x-\ y-z=-1\end{cases}$

Haciendo $z=\lambda$ (incógnita que no intervien en el menor no nulo) y pasándola como término independiente, resulta el sistema de Cramer $\begin{cases}2x-3y=2-\lambda\\ x-\ y=-1+\lambda\end{cases}$

En dicho sistema se puede aplicar la regla con los siguientes resultados:

$$x=\frac{\begin{vmatrix}2-\lambda&-3\\-1+\lambda&-1\end{vmatrix}}{\begin{vmatrix}2&-3\\1&-1\end{vmatrix}}=4\lambda-5\,;\ y=\frac{\begin{vmatrix}2&2-\lambda\\1&-1+\lambda\end{vmatrix}}{\begin{vmatrix}2&-3\\1&-1\end{vmatrix}}=3\lambda-4;\ z=\lambda.$$

Si tuviéramos un sistema cuya matriz de coeficientes fuera trangular superior el sistema $\begin{cases}a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n=b_1\\ \qquad a_{22}x_2+\dots+a_{2n}x_n=b_2\\ \qquad\dots\dots\dots\dots\\ \qquad\qquad a_{nn}x_n=b_n\end{cases}$ se resolvería fácilmente, despejando x_n de la última ecuación, después sustituyendo este valor en la penúltima y calculando x_{n-1}, etc.

Este método de cálculo, por razones obvias, lo designaremos como retro-sustitución. Un sencillo cómputo muestra que el número de operaciones necesarias en este caso es del orden de $n(n-1)/2$ sumas algebraicas, el mismo número de productos y n divisiones.

El fundamento del método de Gauss para la resolución práctica de un sistema de Cramer $AX=B$, con A inversible, consiste en la determinación de un sistema equivalente (es decir, con las mismas soluciones) en el que la matriz de coeficientes sea triangular.

4.3.3. Método de Gauss

Este método se utiliza para la resolución de sistema de Cramer (o de sistemas que pueden transformarse en un sistema de ese tipo); consta de las siguientes fases:

a) Eliminación gaussiana. Se basa en la determinación de una matriz no singular M tal que $U=MA$ sea triangular superior. La matriz M se obtiene como composición de matrices asociadas a transformaciones elementales.
b) Cálculo simultáneo de $(U\,|\,C)=(MA\,|\,MB)$ mediante la aplicación de las transformaciones elementales a las filas de la matriz ampliada.
c) Solución del sistema equivalente $MAX=MB$ por retro-sustitución.

Ejemplo 4.3.4

Resuélvase por el método de Gauss el sistema $\begin{cases}2x_1+x_2-x_3=-1\\ \qquad x_2-3x_3=-7\\ x_1-x_2+2x_3=10\end{cases}$

$$(A\,|\,B)=\left(\begin{array}{ccc|c}2&1&-1&-1\\0&1&-3&-7\\1&-1&2&10\end{array}\right)$$

a) Reducción de $(A\,|\,B)$ a $(U\,|\,C)$ mediante transformaciones elementales en las filas.

Utilizando como pivote el elemento $a_{11} = 2$ y haciendo cero los situados debajo en la misma columna, se obtiene:

$$(A\,|\,B) = \left(\begin{array}{ccc|c} 2 & 1 & -1 & -1 \\ 0 & 1 & -3 & -7 \\ 1 & -1 & 2 & 10 \end{array}\right) \xrightarrow{1} \left(\begin{array}{ccc|c} 2 & 1 & -1 & -1 \\ 0 & 1 & -3 & -7 \\ 0 & -3/2 & 5/2 & 21/2 \end{array}\right)$$

De la misma forma, con $a_{22} = 1$ como pivote se hacen cero los elementos situados debajo de él en la misma columna.

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & -1 \\ 0 & 1 & -3 & -7 \\ 0 & -3/2 & 5/2 & 21/2 \end{array}\right) \xrightarrow{2} \left(\begin{array}{ccc|c} 2 & 1 & -1 & -1 \\ 0 & 1 & -3 & -7 \\ 1 & 0 & -2 & 0 \end{array}\right) = (U\,|\,C)$$

Las transformaciones elementales utilizadas son:

1. Sustitución de la tercera fila por la suma de ella misma más la primera fila multiplicada por $-1/2$.
2. Sustitución de la tercera fila por ella más la segunda fila multiplicada por $3/2$.

Esta última matriz ya tiene la matriz de coeficientes triangular superior.

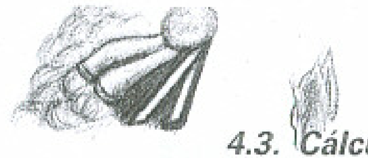
Corresponde a un sistema equivalente al primero. (Hemos aplicado las transformaciones elementales simultáneamente al vector columna de términos independientes.)

$$\left\{\begin{array}{r} 2x_1 + x_2 - x_3 = -1 \\ x_2 - 3x_3 = -7 \\ 2x_3 = 0 \end{array}\right. \Leftrightarrow MAX = MB$$

La última fase a aplicar es la de retro-sustitución:

De la 3.ª ecuación $x_3 = 0$, sustituyendo este valor en la 2.ª ecuación $\Rightarrow x_2 = -7$ y sustituyendo ambos en la primera $\Rightarrow x_1 = 3$.
Por tanto, la solución única es $(3, -7, 0)$.

Si recorremos el camino seguido para resolver por este procedimiento el mismo sistema $AX = B$ de orden n, y vamos contando el número de operacio-



nes ejecutadas, podemos comprobar que son del orden de $3n^2/2$ sumas, $3n^2/2$ multiplicaciones y n divisiones. Para $n = 10$ da un total de unas 300 operaciones, lo que supone una tremenda mejora sobre las 400.000.000 necesarias para la resolución del mismo sistema por la regla de Cramer.

Además presenta la ventaja de ser fácilmente programable y la ejecución del programa es posible en cualquier P.C, recomendamos al lector que esboce el programa del algoritmo en cualquier lenguaje que conozca.

Quizá el lector se pregunte la razón para elegir Ada Byron como personaje paradigmático del capítulo cuatro, la respuesta la encontrará al saber que es autora del primer programa para resolver sistemas de ecuaciones en una máquina.

Cuadro resumen del conjunto de instrucciones dadas por las tarjetas y resultados intermedios obtenidos en el programa de Ada Byron:

Dado el sistema $\left\{\begin{array}{l} mx + ny = d \\ m'x + n'y = d' \end{array}\right.$, llamaremos V_i a las diferentes columnas de ruedas cifradas y supondremos los coeficientes m, n, d, m', n', d' en las ocho primeras columnas de la manera siguiente:

$$V_0 = m,\ V_1 = n,\ V_1 = d,\ V_3 = m',\ V_4 = n',\ V_5 = d',\ V_6 = n,\ V_7 = n'$$

Número de la operación	*Naturaleza de la operación*	*Columnas sobre las que se efectúa la operación*	*Columnas que reciben los resultados*	*Resultados de la operación*
1	*	$V_2 * V_4$	V_8	dn'
2	*	$V_5 * V_1$	V_9	$d'n$
3	*	$V_4 * V_0$	V_{10}	$n'm$
4	*	$V_1 * V_3$	V_{11}	nm'
5	–	$V_8 - V_9$	V_{12}	$dn'-d'n$
6	–	$V_{10} - V_{11}$	V_{13}	$n'm - nm'$
7	/	$V_{12} - V_{13}$	V_{14}	$(dn' - d'n)/(n'm - nm')$

El valor de x está dado por la columna V_{14}. Un cálculo suplementario permite calcular y de la misma manera.

Como en los ordenadores actuales, el molino va a buscar en la columna apropiada el dato que necesita para efectuar la operación en curso, y la reinscribe para que esté disponible para una operación posterior.

4.4. Otro método de resolución. La factorización *LU*

Aunque como hemos visto el método de Gauss representa una mejora considerable frente a la regla de Cramer, cuando el número de variables de un sistema es muy grande, tampoco resulta operativo.

Parece natural la búsqueda de otros métodos de resolución basados en algoritmos programables que reduzcan considerablemente el tiempo invertido por el ordenador.

Uno de estos métodos es la factorización *LU*, que consiste en la creación de una matriz, llamada matriz L, que contiene información sobre las transformaciones que es necesario realizar en A para convertirla en una matriz triangular superior U, para poder, mediante esa información, resolver el sistema por el método de Gauss.

4.4.1. Obtención de la matriz L

Sea el sistema $AX = B$ (A de orden n) que se va a resolver por el método de Gauss. Se prepara la matriz A de modo que tenga el elemento $a_{11} \neq 0$, para que de esta manera las únicas operaciones elementales que haya que realizar para transformar A en la matriz triangular superior U, sean sustituciones de filas por su suma con otra multiplicada por un determinado número, es decir, que no sea necesario efectuar intercambio de filas.

La información sobre esas operaciones elementales necesarias para transformar A en la matriz triangular superior U, se va a guardar en una matriz llamada "matriz L" para que el ordenador lo pueda utilizar siempre que sea necesario, por ejemplo para resolver al mismo tiempo distintos sistemas con la misma matriz A y términos independientes distintos.

La matriz L se va a crear a partir de la matriz unidad I del mismo orden que A, de la siguiente forma:

En el proceso de obtención de la matriz triangular superior U a partir de la matriz de coeficientes $A = (a_{ij})$*, por cada transformación elemental que suponga la sustitución de la fila i por su suma con la j multiplicada por* α*, en la matriz unidad I se cambia el cero que ocupa la posición (i, j), por* $-\alpha$*.*

Por ejemplo, si en A se sustituye la segunda fila por su suma con la primera multiplicada por 2, entonces, en la matriz unidad I el cero que ocupa la posición (fila 1, columna 2), se cambia por -2.

Ejemplo 4.4.1.

Constrúyase la matriz L de la matriz $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 2 & -3 \\ 3 & -1 & -2 \end{pmatrix}$.

Solución:

Sustituciones necesarias para transformar A en la triangular superior U

$$A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 2 & -3 \\ 3 & -1 & -2 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & -4 \\ 3 & -1 & -2 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & -4 \\ 0 & -7 & 1 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & -4 \\ 0 & 0 & -6 \end{pmatrix} = U$$

Se parte ahora de la matriz unidad $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ en la que se van a hacer los siguientes cambios, uno por cada una de las transformaciones anteriores:

Transformación 1: Sustitución de la fila 2.ª por su suma con la 1.ª multiplicada por 1.
Cambio inducido en I: El cero de la posición fila 2, columna 1 se sustituye por -1.
Transformación 2: Sustitución de la fila 3.ª por su suma con la 1.ª multiplicada por -3.
Cambio inducido en I: El cero de la fila 3, columna 1 se sustituye por 3.
Transformación 3: Sustitución de la fila 3.ª por su suma con la 2.ª multiplicada por 7/4.
Cambio inducido en I. El cero de la fila 3, columna 2 se sustituye por $-7/4$.

Matriz L obtenida:

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & -7/4 & 1 \end{pmatrix} = L$$

En el ejemplo 4.4.2 resolveremos el sistema.

Hay que observar que la matriz L es triangular inferior, con todos los elementos de la diagonal unos y, obviamente, sólo podrá utilizarse este procedimiento cuando la matriz A sea regular.

FACTORIZACIÓN LU DE LA MATRIZ A.

La matriz L, construida de esta forma, tiene la siguiente propiedad:

4.4.2. Proposición

> Para toda matriz regular A, se verifica $A = LU$, donde U es la matriz triangular superior deducida de ella mediante transformaciones elementales en sus filas que supongan sólo sustitución entre ellas, y L es la matriz L correspondiente a A según la definición anterior

Demostración:

Si para pasar de la matriz A a la matriz triangular superior U se han tenido que hacer n transformaciones elementales, se puede expresar $E_nE_{n-1} \dots E_2E_1A = U$, siendo E_i la matriz correspondiente a la transformación elemental i-ésima.
Sabemos que existe E_i^{-1}, por tanto, $A = E_1^{-1} E_2^{-1} \dots E_{n-1}^{-1} E_n^{-1} U$.
Sólo queda comprobar que la matriz L de A es, precisamente $L = E_1^{-1} E_2^{-1} \dots E_{n-1}^{-1} E_n^{-1}$, que equivale a $L = E_1^{-1} E_2^{-1} \dots E_{n-1}^{-1} E_n^{-1} I$.
Esta comprobación es sencilla recordando lo que son, tanto la matriz L como cada una de las inversas de las matrices elementales.
Para cada operación elemental que suponga sustituir en A la fila i por su suma con la fila j multiplicada por α, la matriz elemental correspondiente E es la que se obtiene de I al hacer la misma operación, y la matriz inversa E^{-1} es la que se deduce de I sustituyendo la fila i por su suma con la fila j multi-

plicada por $-\alpha$ (transformación elemental que deshace la transformación realizada por E).

Si se repasan todos los productos, primero $E_n^{-1}I$, luego $E_{n-1}^{-1}E_n^{-1}I$, y así sucesivamente se observa que se van construyendo los elementos que tiene la matriz L.

Se recomienda al lector que constate esta observación.

Ejemplo 4.4.2

Compruébese que para la matriz $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 2 & -3 \\ 3 & -1 & -2 \end{pmatrix}$ se verifica que $A = LU$.

Es la matriz del ejemplo anterior 4.4.1, donde ya se hallaron las matrices U y L correspondientes.

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 2 & -3 \\ 3 & -1 & -2 \end{pmatrix}; \qquad U = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & -4 \\ 0 & 0 & -6 \end{pmatrix}; \qquad L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & -7/4 & 1 \end{pmatrix}$$

Se verifica $A = LU$.

Hay que observar que aunque, efectivamente, una matriz regular A se factoriza en LU de la forma anterior, es decir, como producto de una matriz triangular inferior por otra triangular superior, sin embargo esta factorización no es única. Por ejemplo, la matriz anterior también es el producto de estas dos del mismo modo triangulares.

$$A = (kL) \cdot \left(\frac{1}{k} U\right) \text{ si } k \neq 0$$

4.4.3. Resolución de sistemas mediante factorización LU

Para de resolver el sistema $AX = B$ utilizando la matriz L, sólo hay que efectuar las transformaciones que se indican en la misma para transformar la matriz de términos independientes B, y después mediante sustituciones resolver el sistema equivalente obtenido.

Ejemplo 4.4.3

Resuélvase mediante factorización LU el sistema $\begin{cases} x_1 + 2x_2 - x_3 = 3 \\ -x_1 + 2x_2 - 3x_3 = 13 \\ 3x_1 - x_2 - 2x_3 = -5 \end{cases}$

Solución:

Es el sistema $AX = B$, tal que $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 2 & -3 \\ 3 & -1 & -2 \end{pmatrix}$; $B = \begin{pmatrix} 3 \\ 13 \\ -5 \end{pmatrix}$; A es la matriz de los ejercicios anteriores para la que ya calculamos $U = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & -4 \\ 0 & 0 & -6 \end{pmatrix}$ y $L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & -7/4 & 1 \end{pmatrix}$

Los elementos de L situados debajo de la diagonal indican las transformaciones a realizar en B, que recordemos son las que se efectúan en A para obtener la triangular U. Realizándolas en B se obtiene:

$$B = \begin{pmatrix} 3 \\ 13 \\ -5 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 3 \\ 16 \\ -5 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 3 \\ 16 \\ -14 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 3 \\ 16 \\ 14 \end{pmatrix} = C$$

El sistema transformado es $UX = C$: $\begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & -4 \\ 0 & 0 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 16 \\ 14 \end{pmatrix} \Leftrightarrow$

$$\Leftrightarrow \begin{cases} x_1 + 2x_2 - x_3 = 3 \\ 4x_2 - 4x_3 = 16 \\ -6x_3 = 14 \end{cases}$$

En este sistema, por retro-sustitución se halla la solución:

$$x_1 = -8/3\ , \quad x_2 = 5/3\ , \quad x_3 = -7/3$$

La utilidad de la factorización LU es especialmente significativa cuando la información contenida en la matriz L debe utilizarse varias veces, por ejemplo cuando han de resolverse varios sistemas con la misma matriz de coeficientes y distintos términos independientes.

4.4.4. Posibles errores en el cálculo de sistemas por ordenador. Matrices mal condicionadas

Al realizar cálculos algebraicos con ayuda de ordenador, puede suceder que se cometan errores debido a la pérdida de cifras que no siendo significativas en un principio, al ser arrastradas en cálculos posteriores, pueden llegar a originar resultados muy diferentes a los correctos.

Por ejemplo, si a un ordenador se le pide calcular 1/3, el resultado que ofrecerá no se corresponde con el valor exacto, sino que dará 0,33 o bien 0,34 y en ambos casos comete un error de redondeo que en el arrastre para posteriores operaciones pueden dar lugar a resultados muy extraños.

En ciertos sistemas de ecuaciones lineales, cuando se resuelven con ayuda informática ese problema puede reflejarse de una manera muy significativa.

Por ejemplo, el sistema $\begin{cases} x_1 + x_2 = 10 \\ 0{,}001x_1 + x_1 = 10 \end{cases}$ tiene por solución $x_1 = 0$, $x_2 = 10$.

Ahora bien, puede suceder que teniendo en cuenta los errores de redondeo se obtenga un sistema "casi igual" como es el siguiente: $\begin{cases} x_1 + x_2 = 10 \\ x_1 + 0{,}001x_2 + 10 \end{cases}$, donde la variación en los coeficientes de la segunda ecuación es insignificante, de 0,001 y sin embargo la solución del nuevo sistema es $x_1 = 10$, $x_2 = 0$, que es muy diferente de la real.

Los sistemas de este tipo, es decir, aquellos en los que una variación pequeña en los coeficientes produce una gran variación en los resultados, se llaman sistemas inestables o mal condicionados. La matriz A de los sistemas de este tipo es una matriz mal condicionada.

Es necesario tener ciertas cautelas con los programas que se diseñan para resolver sistemas de esas características, y establecer condiciones que corrijan las desviaciones que se puedan originar.

Una forma de corrección es lo que se llama cambio de escala, que consiste en multiplicar por un número grande la ecuación del sistema con alguno de sus coeficientes muy próximo a cero, para que así el posible redondeo no influya significativamente en el resultado.

Por ejemplo, en el sistema anterior se puede multiplicar la ecuación que tiene un coeficiente muy cercano al valor cero y viene arrastrada por redondeos anteriores, por un factor que los incremente de forma significativa, por ejemplo 1000, resultando así el sistema $\begin{cases} x_1 + x_2 = 10 \\ 1000x_1 + x_2 = 10000 \end{cases}$, que es equivalente al primero.

Otro problema relativo a errores que puede plantear el ordenador en un momento determinado es el contrario al anterior, es decir, la obtención de un número distinto de cero de un valor que sí debe serlo. Una forma de corregirlo es la asignación de cero a esa magnitud en cuanto esta sea menor que un determinado número muy pequeño (fijar la cota de error).

En general, la corrección de errores por problemas de redondeo son aplicables y pueden ser originales para cada situación que se plantee. Los descritos aquí son únicamente dos casos frecuentes al resolver sistemas de ecuaciones lineales, sin embargo pueden surgir otros de características diferentes para los que hay que aplicar medidas de corrección oportunas en cada caso.

EJERCICIOS

Ejercicio 4.1

Hállase el valor de a que hace indeterminado el sistema $\begin{cases} 2x_1 + ax_2 - x_3 = 0 \\ x_1 - x_2 + ax_3 = 0 \\ 11x_1 + 7x_2 + 17x_3 = 0 \end{cases}$ y, resuélvase para los valores enteros de a, si hay alguno.

Solución:

Es un sistema homogéneo con el mismo número de ecuaciones que incógnitas, por tanto, siempre es compatible.

Es indeterminado si el rango de la matriz de coeficientes A es menor que el número de incógnitas. Es decir, si rg $A < 3 \Leftrightarrow$ Si $|A| = 0$.

$$|A| = \begin{vmatrix} 2 & a & -1 \\ 1 & -1 & a \\ 11 & 7 & 17 \end{vmatrix} = 11a^2 - 31a - 52 = 0, \text{ para } a = 4,\ a = -13/11.$$

Para $a = 4$ el sistema es $\begin{cases} 2x_1 + 4x_2 - x_3 = 0 \\ x_1 - x_2 + 4x_3 = 0 \\ 11x_1 + 7x_2 + 17x_3 = 0 \end{cases}$

El rango de la matriz de coeficientes de este sistema es:

$$\text{rg }(A) = \text{rg}\begin{pmatrix} 2 & 4 & -1 \\ 1 & -1 & 4 \\ 11 & 7 & 17 \end{pmatrix} = 2 \Rightarrow \text{es un sistema compatible indeterminado.}$$

Para resolver el sistema se escoge un menor de orden 2 no nulo, por ejemplo

$$\begin{vmatrix} 2 & 4 \\ 1 & -1 \end{vmatrix} = -6 \neq 0$$

El sistema que resulta de eliminar de él la ecuación no contenida en el menor no nulo anterior, es equivalente al primero:

$$\begin{cases} 2x_1 + 4x_2 - x_3 = 0 \\ x_1 - x_2 + 4x_3 = 0 \end{cases}$$

Haciendo $x_3 = \lambda$ y pasando al segundo miembro queda el sistema

$$\begin{cases} 2x_1 + 4x_2 = \lambda \\ x_1 - x_2 = -4\lambda \end{cases}$$

de solución $x_1 = -5/2\lambda$; $x_2 = 3/2\lambda$, que junto con $x_3 = \lambda$ es la solución del sistema pedido.

Ejercicio 4.2

Se considera el sistema $\begin{cases} x_1 + x_2 - 2x_3 = 1 \\ 2x_1 - x_2 - x_3 = -1 \\ x_1 - 3x_2 + 2x_3 = -3 \end{cases}$, se pide:

a) Comprobar que $(x_1, x_2, x_3) = (0, 1, 0)$ es solución del mismo.
b) Hallar el conjunto de soluciones del sistema homogéneo asociado.
c) Hallar el conjunto de soluciones del sistema no homogéneo.

Solución:

a) Para ver que la terna $(x_1, x_2, x_3) = (0, 1, 0)$ es solución del sistema, es suficiente ver que al sustituir dicha solución en el sistema se obtienen tres identidades:

$$\begin{cases} 0 + 1 - 0 = 1 \\ 2.0 - 1 - 0 = -1 \\ 0 - 3.1 + 2.0 = -3 \end{cases}$$

b) Sistema homogéneo asociado

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

Comprobamos previamente si es indeterminado mediante el teorema de Rouché:

$(A) = \begin{pmatrix} 1 & 1 & -2 \\ 2 & -1 & -1 \\ 1 & -3 & 2 \end{pmatrix}$; $|A| = 0$; rg $(A) = 2$, algún menor como $\begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -3 \neq 0$

Al ser $\text{rg}(A) = 2 <$ nº de incógnitas $= 3$, el sistema es indeterminado (tiene soluciones distintas de la trivial).

Para resolverlo se elimina la ecuación que no interviene en el menor distinto de cero, y se obtiene el sistema: $\begin{cases} x_1 + x_2 - 2x_3 = 0 \\ 2x_1 - x_2 - x_3 = 0 \end{cases}$, que es equivalente al primero.

Haciendo $x_3 = \lambda$ y cambiando de miembro, se obtiene $\begin{cases} x_1 + x_2 = 2\lambda \\ 2x_1 - x_2 = \lambda \end{cases}$, cuya solución es:

$x_1 = \lambda$, $x_2 = \lambda$.

La solución del sistema homogéneo es el conjunto de soluciones: $(x_1, x_2, x_3) =$

$= (\lambda, \lambda, \lambda)$, que también podemos escribir así: $\begin{cases} x_1 = \lambda \\ x_2 = \lambda \\ x_3 = \lambda \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

c) En el sistema no homogéneo $\text{rg}(A) = \text{rg}(A\,|\,B) = 2$. Por tanto, es compatible e indeterminado.

Recordando que el conjunto de soluciones de un sistema compatible es el conjunto de soluciones de su sistema homogéneo asociado más una solución particular del mismo, podemos escribir:

- Conjunto de soluciones del sistema no homogéneo:

$(x_1, x_2, x_3) = (0, 1, 0) + (\lambda, \lambda, \lambda) = (\lambda, 1 + \lambda, \lambda)$, que también se puede expresar de la siguiente manera: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Ejercicio 4.3

Calcúlense los valores de a, b para los que las soluciones del sistema $\begin{cases} x_1 + ax_2 = 1 \\ 2x_1 - bx_3 = 2 \end{cases}$, son tales que el valor de x_1 más el de x_2 es 1.

Solución:

El problema es equivalente a resolver el sistema que resulta al añadir al sistema dado la ecuación $x_1 + x_2 = 1$ que indica la condición adicional.

El sistema obtenido es $\begin{cases} x_1 + ax_2 = 1 \\ 2x_1 - bx_3 = 2 \\ x_1 + x_2 = 1 \end{cases}$

Las matrices de coeficientes (A) y ampliada $(A\,|\,B)$ de este sistema son:

$$(A) = \begin{pmatrix} 1 & a & 0 \\ 2 & 0 & -b \\ 1 & 1 & 0 \end{pmatrix} \text{ y } (A\,|\,B) = \left(\begin{array}{ccc|c} 1 & a & 0 & 1 \\ 2 & 0 & -b & 2 \\ 1 & 1 & 0 & 1 \end{array}\right)$$

El determinante de A vale:

$$|A| = \begin{vmatrix} 1 & a & 0 \\ 2 & 0 & -b \\ 1 & 1 & 0 \end{vmatrix} = b(1-a) \Rightarrow |A| = 0 \text{ para } b = 0 \text{ ó } a = 1.$$

Para que existan soluciones el sistema ha de ser compatible, es decir, $\text{rg}(A) = \text{rg}(A\,|\,B)$; analicemos por separado qué ocurre con los rangos de las matrices en función de los valores de a y b:

1. Si $b = 0 \Rightarrow (A\,|\,B) = \left(\begin{array}{ccc|c} 1 & a & 0 & 1 \\ 2 & 0 & 0 & 2 \\ 1 & 1 & 0 & 1 \end{array}\right)$

$\text{rg}(A) = \text{rg}(A\,|\,B) = 2$ para todo valor de a.

1. Si $a = 1 \Rightarrow (A\,|\,B) = \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 2 & 0 & -b & 2 \\ 1 & 1 & 0 & 1 \end{array}\right)$

$\text{rg}(A) = \text{rg}(A\,|\,B) = 2$ para todo valor de b.

Por tanto, el sistema cumple la condición adicional indicada, si $a = 1$ ó $b = 0$.

Ejercicio 4.4

Sea A una matriz cuadrada de orden 2. Demuéstrese que la condición necesaria y suficiente para que A sea singular es que exista otra matriz cuadrada M de orden 2 no nula y tal que $AM = 0$.

Solución:

Vamos a ver primero la implicación: A es singular $\Rightarrow$ Existe $M \neq 0$ con $AM = 0$.

En efecto, sea la matriz $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ singular.

Si A es singular, el sistema homogéneo $AX = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ tiene soluciones distintas de la trivial porque $|A| = 0 \Rightarrow \text{rg}(A) <$ número de incógnitas $= 2$.

Sea $C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ una de las soluciones, es decir, se verifica $\begin{cases} a_{11}c_1 + a_{12}c_2 = 0 \\ a_{21}c_1 + a_{22}c_2 = 0 \end{cases}$, y sea M la matriz $M = \begin{pmatrix} c_1 & 0 \\ c_2 & 0 \end{pmatrix}$ que evidentemente es una matriz no nula.

Por ser C solución del sistema, la matriz producto

$$AM = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} c_1 & 0 \\ c_2 & 0 \end{pmatrix} = \begin{pmatrix} a_{11}c_1 + a_{12}c_2 & 0 \\ a_{21}c_1 + a_{22}c_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Veamos ahora que si A no es singular $\Rightarrow$ No existe $M \neq 0$ con $AM = 0$. Utilicemos la reducción al absurdo:

Supongamos que A es regular, entonces existe la matriz inversa A^{-1}.

Si $AM = 0 \Rightarrow A^{-1}AM = 0$. Es decir, $M = 0$.

Ejercicio 4.5

Aplicando el teorema de Rouché, analícese la compatibilidad del sistema

$$\begin{cases} ax + y + z = a^2 \\ x + y + az = a \\ x + y + 2az = 2 \end{cases}$$ para los distintos valores de a.

Solución:

Cálculo del rango de las matrices de coeficientes (A) y ampliada $(A|B)$:

$$(A)=\begin{pmatrix} a & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2a \end{pmatrix};\ (A|B)=\left(\begin{array}{ccc|c} a & 1 & 1 & a^2 \\ 1 & 1 & a & a \\ 1 & 1 & 2a & 2 \end{array}\right)$$

$$|A|=\begin{vmatrix} a & 1 & 1 \\ 1 & 1 & a \\ 1 & 1 & 2a \end{vmatrix}=a(a-1);\ |A|=0 \text{ para } a=0, a=1.$$

Analicemos cada caso por separado:

$$\text{Para } a=0,\ (A|B)=\left(\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 \end{array}\right)$$

$\text{rg}(A)=2$; $\text{rg}(A|B)=3$, el sistema es incompatible.

$$\text{Para } a=1,\ (A|B)=\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{array}\right)$$

$\text{rg}(A)=2$; $\text{rg}(A|B)=2$, el sistema es compatible e indeterminado.

Para $a\neq 0$ y $a\neq 1$, al ser $|A|\neq 0 \Rightarrow \text{rg}(A)=\text{rg}(A|B)=3$, el sistema es compatible y determinado.

Ejercicio 4.6

Sea el espacio vectorial real V y $E=\{\bar{e}_1, \bar{e}_2, \bar{e}_3\}$ una base del mismo. Se considera la familia de endomorfismos $\{f\}$ determinados de la siguiente forma:

$$\begin{aligned} f(\bar{e}_1)&=a\bar{e}_1+b\bar{e}_2+b\bar{e}_3 \\ f(\bar{e}_2)&=b\bar{e}_1+a\bar{e}_2+b\bar{e}_3 \\ f(\bar{e}_3)&=b\bar{e}_1+b\bar{e}_2+a\bar{e}_3 \end{aligned}$$

Se pide determinar:

a) Valores de a, b para los que f es biyectiva.
b) Para los endomorfismos no biyectivos, analizar si el vector $\bar{u}=3\bar{e}_1-3\bar{e}_2$ pertenece al conjunto $\text{Im}(f)$, siendo f cualquier endomorfismo de la familia dada.

Solución:

Expresión analítica de la familia de endomorfismos, en la base E:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

que corresponde al sistema de ecuaciones lineales

$$\begin{cases} y_1=ax_1+bx_2+bx_3 \\ y_2=bx_1+ax_2+bx_3 \\ y_3=bx_1+bx_2+ax_3 \end{cases}$$

donde $X=\begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix}$ e $Y=\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ están expresados en la base E, y $f(X)=Y$.

Determinante de la matriz A de la familia: $|A|=\begin{vmatrix} a & b & b \\ b & a & b \\ b & b & a \end{vmatrix}=(a+2b)(a-b)^2$

$|A|=0$ para $a=-2b$, $a=b$.

a) Los endomorfismos biyectivos son aquellos cuya su matriz es regular. Es decir, para $a\neq -2b$ y también $a\neq b$.
b) Analicemos los no biyectivos. Valores de $a=-2b$ ó $a=b$.

Se supone que $b\neq 0$, en caso contrario no existe el sistema.

El vector $\bar{u}$ pertenece a $\text{Im}(f)$ si el sistema que define la aplicación es compatible, es decir si:

$\operatorname{rg}(A) = \operatorname{rg}(A \mid B)$, siendo $B = \begin{pmatrix} 3 \\ -3 \\ 0 \end{pmatrix}$.

Analicemos cada caso.

Si $a = -2b$, $\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} -2b & b & b \\ b & -2b & b \\ b & b & -2b \end{pmatrix} = 2$ porque $b \neq 0$, y $\operatorname{rg}(A \mid B) =$

$$= \operatorname{rg}\left(\begin{array}{ccc|c} -2b & b & b & 3 \\ b & -2b & b & -3 \\ b & b & -2b & 0 \end{array}\right) = 2$$

El sistema es compatible, existen vectores que son antiimagen de $\overline{u}$, es decir: $\overline{u} \in \operatorname{Im}(f)$ para cualquier f perteneciente a la familia.

Si $a = b$, $\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} b & b & b \\ b & b & b \\ b & b & b \end{pmatrix} = 1$; $\operatorname{rg}(A \mid B) = \operatorname{rg}\left(\begin{array}{ccc|c} b & b & b & 3 \\ b & b & b & -3 \\ b & b & b & 0 \end{array}\right) = 2$

El sistema es incompatible, $\overline{u} \notin \operatorname{Im}(f)$ para ningún f perteneciente a la familia.

Ejercicio 4.7

Compruébese la compatibilidad del sistema $\begin{cases} x + 3y + z - t = 2 \\ 2x + 7y + 3z - 4t = 1 \\ x + y + 2z + t = 1 \end{cases}$, y resuélvase por el método de Gauss cuando sea compatible.

Solución:

Matriz ampliada: $(A \mid B) = \left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 2 \\ 2 & 7 & 3 & -4 & 1 \\ 1 & 1 & 2 & 1 & 1 \end{array}\right)$

$\operatorname{rg}A = \operatorname{rg}(A \mid B) = 3$. Por ejemplo, es $\begin{vmatrix} 1 & 3 & 1 \\ 2 & 7 & 3 \\ 1 & 1 & 2 \end{vmatrix} = 3 \neq 0$

El sistema es compatible e indeterminado al ser 4 el nº de incógnitas.

Haciendo $t = \lambda$ y pasando en cada ecuación al segundo miembro el sumando que lo contiene, se obtiene el sistema con tres incógnitas

$\begin{cases} x + 3y + z = 2 + \lambda \\ 2x + 7y + 3z = 1 + 4\lambda \\ x + y + 2z = 1 - \lambda \end{cases}$, que es de Cramer, con matriz ampliada $(A \mid B) =$

$$= \left(\begin{array}{ccc|c} 1 & 3 & 1 & 2 + \lambda \\ 2 & 7 & 3 & 1 + 4\lambda \\ 1 & 1 & 2 & 1 - \lambda \end{array}\right).$$

Mediante transformaciones elementales en las filas se obtiene:

$$(A \mid B) = \left(\begin{array}{ccc|c} 1 & 3 & 1 & 2 + \lambda \\ 2 & 7 & 3 & 1 + 4\lambda \\ 1 & 1 & 2 & 1 - \lambda \end{array}\right) \xrightarrow{1} \left(\begin{array}{ccc|c} 1 & 3 & 1 & 2 + \lambda \\ 0 & 1 & 1 & -3 + 2\lambda \\ 1 & 1 & 2 & 1 - \lambda \end{array}\right) \xrightarrow{2}$$

$$\xrightarrow{2} \left(\begin{array}{ccc|c} 1 & 3 & 1 & 2 + \lambda \\ 0 & 1 & 1 & -3 + 2\lambda \\ 0 & -2 & 1 & -1 - 2\lambda \end{array}\right) \xrightarrow{3} \left(\begin{array}{ccc|c} 1 & 3 & 1 & 2 + \lambda \\ 0 & 1 & 1 & -3 + 2\lambda \\ 0 & 0 & 3 & -1 + 2\lambda \end{array}\right) = (U \mid C)$$

que corresponde al sistema $\begin{cases} x + 3y + z = 2 + \lambda \\ y + z = -3 + 2\lambda \\ 3z = -7 + 2\lambda \end{cases}$

Utilizando la retrosustitución, de la tercera ecuación $\Rightarrow z = \dfrac{-7 + 2\lambda}{3}$

sustituyendo en la segunda $y = -3 + 2\lambda - z = \dfrac{4\lambda - 2}{3}$

y por último, en la primera ecuación $x = 2 + \lambda - 3y - z = \dfrac{19 - 11\lambda}{3}$

La solución es: $x = \dfrac{19 - 11\lambda}{3}$; $y = \dfrac{4\lambda - 2}{3}$; $z = \dfrac{-7 + 2\lambda}{3}$; $t = \lambda$

Ejercicio 4.8

Calcúlense las ecuaciones paramétricas y cartesianas del subespacio de $\mathbb{R}^3$ generado por los vectores $\{(1, 0, -1), (2, 1, 1)\}$.

Solución:

El sistema $\{(1, 0, -1), (2, 1, 1)\}$ es libre porque $\text{rg}\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 1 \end{pmatrix} = 2$.

Los vectores del subespacio pedido son los vectores (x_1, x_2, x_3) que dependen linealmente de los vectores del sistema generador. Por tanto, el subespacio vectorial generado es: $\{(x_1, x_2, x_3) = \alpha(1, 0, -1) + \beta(2, 1, 1)\ \forall\ \alpha, \beta \in \mathbb{R}\}$.

Es decir, los vectores que verifican: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$

Para que los vectores que hay a ambos lados del signo "=" sean iguales, han de ser iguales coordenada o coordenada $\Leftrightarrow \begin{cases} x_1 = \alpha + 2\beta \\ x_2 = \beta \\ x_3 = -\alpha + \beta \end{cases}$, que son las ecuaciones paramétricas.

Las ecuaciones no paramétricas o cartesianas son las que resultan de eliminar los parámetros en el sistema anterior. (Recordemos que el número de ecuaciones cartesianas de un subespacio es el número de incógnitas menos la dimensión del subespacio, es decir: $3 - 2 = 1$ ecuación.)

El sistema $\begin{cases} x_1 = \alpha + 2\beta \\ x_2 = \quad\ \ \beta \\ x_3 = -\alpha + \beta \end{cases}$ es equivalente al sistema lineal homogéneo

$$\begin{cases} \alpha + 2\beta - x_1 = 0 \\ \beta - x_2 = 0 \\ -\alpha + \beta - x_3 = 0 \end{cases}$$

Para que este sistema homogéneo tenga solución distinta de la trivial, es necesario que $\text{rg}\begin{pmatrix} 1 & 2 & -x_1 \\ 0 & 1 & -x_2 \\ -1 & 1 & -x_3 \end{pmatrix} = 2 \Leftrightarrow \begin{vmatrix} 1 & 2 & -x_1 \\ 0 & 1 & -x_2 \\ -1 & 1 & -x_3 \end{vmatrix} = 0 \Leftrightarrow x_1 - 3x_2 + x_3 = 0$, que es la ecuación cartesiana pedida.

Ejercicio 4.9

Hállese la factorización LU de la matriz $A = \begin{pmatrix} 1 & -1 \\ 3 & -4 \end{pmatrix}$.

Solución:

- Obtención de la matriz L

Sustituciones necesarias para transformar A en U triangular superior.

$$A = \begin{pmatrix} 1 & -1 \\ 3 & -4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} = U$$

La única sustitución es: 2ª fila por su suma con la 1ª fila multiplicada por -3. La matriz L es por lo tanto la que resulta de cambiar en la matriz unidad I el cero de la fila 2 columna 1 por $+3$ (el multiplicador cambiado de signo): $L = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$.

- La factorización LU de A es, por tanto:

$$A = LU = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}.$$

Ejercicio 4.10

Resuélvese el sistema $A^3X = B$ mediante el cálculo de la matriz L de A, siendo $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 6 \\ 18 \\ 0 \end{pmatrix}$.

Solución:

- Cálculo de la matriz L:

Sustituciones necesarias para transformar A en una matriz triangular superior.

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & -1 & 1 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix} = U$$

1. Sustitución de la 2.ª fila por su suma con la primera multiplicada por -1.
2. Sustitución de la 3.ª fila por su suma con la primera multiplicada por -1.
3. Sustitución de la 3.ª fila por su suma con la segunda multiplicada por $+1$.

que corresponden a los siguientes cambios en la matriz identidad I:

1. Cambio del 0 de la posición (2, 1) por 1 (opuesto del factor que multiplica).
2. Cambio del 0 de la posición (3, 1) por 1.
3. Cambio del 0 de la posición (3, 2) por -1.

A partir de I se obtiene la matriz L:

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = L$$

- Para resolver el sistema se utiliza reiteradamente la información existente en la matriz L:

1. $A^3X = B$ se expresa como $A(A^2X) = B$; haciendo $A^2X = Y$, queda el sistema $AY = B$ que se resuelve por el método de Gauss. (Hay que transformar primero B.)

Sistema $AY = B$. Mediante la información contenida en L el vector B se transforma en C.

$$B = \begin{pmatrix} 6 \\ 18 \\ 0 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 6 \\ 12 \\ 0 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 6 \\ 12 \\ -6 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 6 \\ 12 \\ 6 \end{pmatrix} = C$$

El sistema equivalente a $AY = B$ es $UY = C$.

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 12 \\ 6 \end{pmatrix}, \text{ de solución } Y = \begin{pmatrix} 8 \\ 8 \\ 2 \end{pmatrix}$$

2. Una vez hallado Y, $A^2X = Y$, se puede expresar de la forma $A(AX) = Y$, y haciendo $AX = Z$ se resuelve $AZ = Y$.

Sistema $AZ = Y$; con la información de L la matriz Y se transforma en D

$$Y = \begin{pmatrix} 8 \\ 8 \\ 2 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 8 \\ 0 \\ 2 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 8 \\ 0 \\ -6 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 8 \\ 0 \\ -6 \end{pmatrix} = D$$

El sistema equivalente a $AZ = Y$ es $UZ = D$.

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \\ -6 \end{pmatrix}, \text{ de solución } Z = \begin{pmatrix} 6 \\ 4 \\ -2 \end{pmatrix}$$

3. Por último, conociendo Z puede hallarse X a partir de $AX = Z$.

El vector Z se transforma en E.

$$Z = \begin{pmatrix} 6 \\ 4 \\ -2 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 6 \\ -2 \\ -2 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 6 \\ -2 \\ -8 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 6 \\ -2 \\ -10 \end{pmatrix} = E$$

El sistema equivalente a $AX = Z$ es $UX = E$.

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \\ -10 \end{pmatrix}, \text{ de solución } X = \begin{pmatrix} 8/3 \\ 14/3 \\ -10/3 \end{pmatrix}$$ que es la solución

del sistema.

CAPÍTULO 5

PRODUCTO ESCALAR DE VECTORES Y ESPACIO EUCLÍDEO

INTRODUCCIÓN

Aunque los griegos habían reconocido ya que el espacio matemático o abstracto es distinto del espacio percibido a través de los sentidos, la mayoría de los matemáticos hasta alrededor de 1800 estaban convencidos de que la Geometría Euclídea era la única y correcta idealización del espacio físico sensible.

La razón es que un vector geométrico como segmento orientado sirve para modelizar muchas aplicaciones en Física, como así también los términos lontigud y ángulo.

Esta profunda convicción en el valor absoluto de la Geometría Euclídea queda claramente de manifiesto a través de los muchos intentos de basar la Aritmética, Álgebra y Análisis, cuyos fundamentos lógicos eran bastante oscuros en la Geometría Euclídea, y garantizar así la verdad y consistencia de estas áreas de la Matemática.

Por otro lado, la introducción de las coordenadas por Descartes y Fermat y el desarrollo de la Geometría Analítica, hicieron que los métodos algebraicos y analíticos pasaran a dominar la Geometría de modo casi exclusivo. Sin embargo, poco a poco, empezaron a surgir voces para reivindicar los métodos geométricos puros (la Geometría Sintética), cuya elegancia y claridad habían cautivado a tantos matemáticos en el pasado. Muchas de las objeciones a los métodos analíticos en Geometría se basaban en la pérdida de significado geométrico que conllevaba su uso, diluido en los cada vez más engorrosos y largos cálculos.

Esta reacción contra los métodos analíticos se hace cada vez más intensa. Así, Carnot proclama su deseo de "liberar la Geometría de los jeroglíficos del análisis".

El movimiento hacia la Geometría Sintética, iniciado a finales del XVIII y encabezado por figuras como Monge y Carnot, va a tener como consecuencia la clarificación de las distintas nociones geométricas y, en particular, la creación de la Geometría Proyectiva como una nueva rama de las Matemáticas. El principal innovador en este campo fue Poncelet, discípulo de Monge y Carnot quien, sirviendo como oficial de Napoleón en la campaña de Rusia, fue capturado y pasó un año en prisión. Allí reelaboró sus estudios geométricos sin ayuda de libros y obtuvo nuevos resultados. Este hecho significó la creación de una nueva disciplina: La Geometría Proyectiva.

El impulso dado a la Geometría por este nuevo punto de vista fue considerable, y la Geometría Proyectiva se convirtió pronto en una de las ramas más activas y más populares de las Matemáticas. Una de sus principales ventajas era evitar casi completamente los cálculos, lo que la hacía principalmente atractiva como tema de los primeros cursos de los estudios de Matemáticas.

Hacia 1850, los conceptos generales y los objetivos de la Geometría Proyectiva y su diferencia con los de la Geometría Euclídea, estaban claros; sin embargo, no se había clarificado lo suficiente la relación lógica entre ambas geometrías. Fue Von Staudt, quien desarrolló las ideas para liberar la Geometría Proyectiva de las nociones de longitud y congruencia. A través de sucesivos trabajos en 1856, 1857 y 1860, Von Staudt consiguió desarrollar la Geometría Proyectiva independientemente de la noción de distancia. De esta manera, se puso de manifiesto que la Geometría Proyectiva es previa y más general que la Geometría Euclídea, al tratar propiedades cualitativas y descriptivas de las figuras geométricas.

Por otro lado, los geómetras "algebraicos" pronto se apercibieron de que también era posible utilizar los métodos de la Geometría Analítica en la nueva Geometría. Así aparecen las que hoy conocemos como coordenadas homogéneas, que alcanzan su pleno desarrollo con Plücker.

Los estudios relativos a otras geometrías han continuado hasta nuestros días sin que Geometría Euclídea haya perdido su vigencia.

En esta línea, hay que destacar los trabajos de Cayley; en ellos trata de definir un producto escalar que generalice el usual, y emplear las fórmulas habituales para definir ángulos y distancias. Este proceso es también el utilizado al definir una métrica y es el que se sigue en la actualidad para definir axiomáticamente el Espacio Euclídeo.

La extensión a espacios de dimensión infinita fue llevada a cabo por Hilbert y su escuela, introducen formalmente el lenguaje geométrico habitual del espacio de Hilbert: ortogonalidad, teorema de la proyección ortogonal, bases ortonormales, etc. Sin embargo, la formulación axiomática del espacio de Hilbert general se debe a J. Von Neuman (1929-30), en sus trabajos decisivos sobre la fundamentación matemática de la Mecánica Cuántica.

EUCLIDES

A pesar de ser el matemático más famoso de la Grecia antigua, la vida de Euclides fue tan oscura que no hay asociados a su nombre datos fidedignos de su origen; lo poco que sabemos de su vida nos ha llegado a través de los comentarios de un historiador griego llamado Proclo.

Según estos datos, parece que vivió en Alejandría (Egipto), en torno al año 300 a. C. Allí fundó una escuela de estudios matemáticos. También parece que estudió en la escuela fundada por Platón.

Aunque de la vida de Euclides se conoce poco, sin embargo, su obra sí es ampliamente conocida.

Hasta nuestros días han sobrevivido cinco de sus obras: Los elementos, Los datos, La división de figuras, Los fenómenos y La óptica.

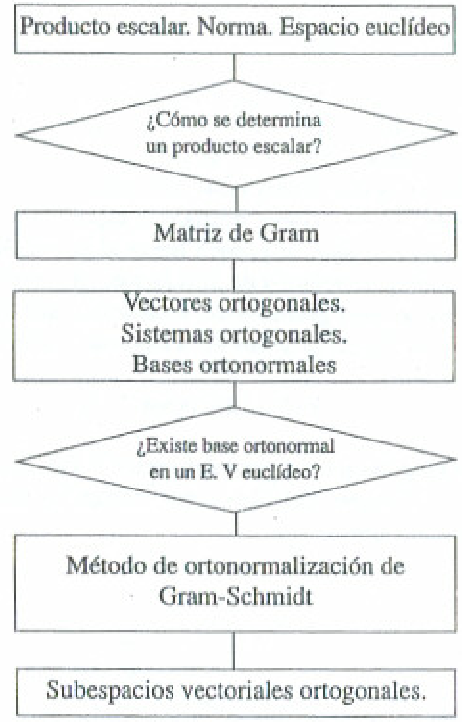
Sin duda, la más conocida es "Los elementos".

Normalmente asociamos "Elementos" a Euclides, pero la palabra "Elementos" significa fundamentos, lo que está ordenado. En general, se entendía por "Elementos" las obras construidas con el método deductivo que se pueden considerar libros de texto. Del año 300 a. C. datan los de Euclides, pero hay una buena lista de precursores que quedaron eclipsados por él, parece que el primero es Hipócrates de Quíos y el último Teudio de Magnesia.

Los Elementos de Euclides son un conjunto de 13 libros escritos hacia el año 300 a. C. que ha sido copiado sin cesar. Nos han llegado copias en distintas versiones, las hay en versión árabe, hay copias latinas de la traducción de la versión árabe y la primera versión impresa en lengua vernácula, en Venecia data de 1482.

Los elementos de Euclides se utilizaron como texto durante 2.000 años (hasta el siglo XVIII) e incluso hoy, una versión modificada de sus primeros libros constituye la base de la enseñanza de la geometría plana en las escuelas secundarias.

CAPÍTULO 5


Producto escalar. Norma. Espacio euclídeo
¿Cómo se determina un producto escalar?
Matriz de Gram
Vectores ortogonales. Sistemas ortogonales. Bases ortonormales
¿Existe base ortonormal en un E. V euclídeo?
Método de ortonormalización de Gram-Schmidt
Subespacios vectoriales ortogonales.

5.1. Producto escalar de vectores y espacio euclídeo

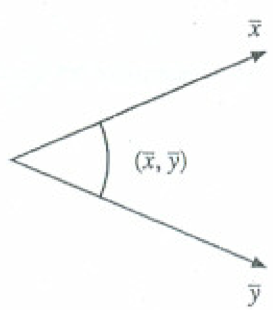
El concepto de vector geométrico como segmento orientado, se relaciona generalmente con las aplicaciones en física, igual que los términos *longitud* y *ángulo*. Sin embargo, en el lenguaje de los espacios vectoriales abstractos estas palabras no figuran en lo que hemos visto hasta ahora.

En este capítulo veremos cómo, dotando al espacio vectorial de la operación llamada producto escalar, se pueden llevar al terreno de lo abstracto las nociones de *longitud* de un vector y de *ángulo* entre dos vectores.

Volviendo a los *vectores geométricos*, recordemos que el producto escalar, tal como se estudia en física, se define de la siguiente manera:

$$\bar{x} \cdot \bar{y} = |\bar{x}|\,|\bar{y}| \cos(\bar{x}, \bar{y})$$

donde $|\bar{x}|$ es el módulo o longitud del vector $\bar{x}$, $|\bar{y}|$ es el módulo del vector $\bar{y}$, y $\cos(\bar{x}, \bar{y})$ es el coseno del ángulo que forman los vectores $\bar{x}$ e $\bar{y}$, por lo que el resultado es un número real.



Estos entes matemáticos son sólo un caso particular de la situación general que estudiaremos a continuación, cuando generalizamos el concepto de producto escalar para un espacio vectorial cualquiera.

5.1.1. Definición de producto escalar de dos vectores

Sea V un espacio vectorial real en el que se considera una aplicación, tal que, a cada par de vectores $\bar{u}$, $\bar{v}$ de V le hace corresponder un número real.
Dicha operación es un producto escalar, que se denota como $\bar{u} \cdot \bar{v}$, si $\forall\, \bar{u}, \bar{v} \in V$ y $\forall \lambda \in \mathbb{R}$ se verifican los siguientes axiomas:

1. $\bar{u} \cdot \bar{v} = \bar{v} \cdot \bar{u}$.
2. $(\lambda \bar{u}) \cdot \bar{v} = \lambda(\bar{u} \cdot \bar{v})$.
3. $(\bar{u}_1 + \bar{u}_2) \cdot \bar{v} = \bar{u}_1 \cdot \bar{v} + \bar{u}_2 \cdot \bar{v}$.
4. $\bar{u} \cdot \bar{u} > 0$ para $\bar{u} \neq \bar{0}$, donde $\bar{0}$ es el vector nulo y además $\bar{0} \cdot \bar{0} = 0$.

5.1.2. Definición de espacio vectorial euclídeo

Se llama espacio vectorial euclídeo a todo espacio vectorial real provisto de un producto escalar.

Ejemplo 5.1.1.

Sean dos vectores de $\mathbb{R}^2$: $\bar{u} = (x_1, x_2)$; $\bar{v} = (y_1, y_2)$. Estúdiese si $\bar{u} \cdot \bar{v}$, definido como $\bar{u} \cdot \bar{v} = x_1y_1 + x_2y_2$, es el producto escalar usual.

Para comprobar si es espacio vectorial euclídeo, o no, será necesario verificar si se cumplen los axiomas lo que definen.

1. $\bar{u} \cdot \bar{v} = x_1y_1 + x_2y_2 = x_2y_2 + x_1y_1 = \bar{v} \cdot \bar{u}$
2. $(\lambda \bar{u}) \cdot \bar{v} = (\lambda x_1, \lambda x_2) \cdot (y_1, y_2) = \lambda x_1y_1 + \lambda x_2y_2 = \lambda(x_1y_1 + x_2y_2) = \lambda(\bar{u} \cdot \bar{v})$
3. $[\bar{u}_1 + \bar{u}_2] \cdot \bar{v} = [(x_1, x_2) + (x'_1, x'_2)] \cdot (y_1, y_2) = (x_1 + x'_1, x_2 + x'_2) \cdot (y_1, y_2) =$
$= (x_1 + x'_1)y_1 + (x_2 + x'_2)y_2 = (x_1y_1 + x_2y_2) + (x'_1y_1 + x'_2y_2) = \bar{u}_1 \cdot \bar{v} + \bar{u}_2 \cdot \bar{v}$

4. $\overline{u} \cdot \overline{u} = x_1x_1 + x_2x_2 = x_1^2 + x_2^2 > 0$ excepto si $x_1 = x_2 = 0$, en cuyo caso sería $\overline{u} \cdot \overline{u} = 0$

Luego la aplicación dada es un producto escalar.

Ejemplo 5.1.2

En el espacio vectorial V de los binomios de primer grado:
$V = \{ax + b \,/\, a, b \in \mathbb{R}\}$, se define la operación: $p(x) \cdot q(x) = \int_0^1 p(x)\, q(x) dx$. Verifíquese que la operación dada es un producto escalar.

Solución:

1. $p(x) \cdot q(x) = \int_0^1 p(x) \cdot q(x) dx = \int_0^1 q(x) \cdot p(x) dx = q(x) \cdot p(x)$
2. $[\lambda p(x)] \cdot q(x) = \int_0^1 \lambda p(x) \cdot q(x) dx = \lambda \int_0^1 p(x) \cdot q(x) dx = \lambda [p(x) \cdot q(x)]$
3. $[p_1(x) + p_2(x)] \cdot q(x) = \int_0^1 [p_1(x) + p_2(x)] \,.\, q(x) dx =$
$$= \int_0^1 p_1(x) \cdot q(x) dx + \int_0^1 p_2(x) \cdot q(x) dx = p_1(x) \cdot q(x) + p_2(x) \cdot q(x)$$
4. $p(x) \cdot p(x) = \int_0^1 p(x) \cdot p(x) dx = \int_0^1 [p(x)]^2\, dx$

Como $p(x) = ax + b$, resulta:

$$p(x) \cdot p(x) = \int_0^1 (ax + b)^2\, dx =$$

$$= \left[\frac{1}{3a} (ax + b)^3 \right]_0^1 = \frac{a^2 + 3ab + 3b^2}{3} = \frac{1}{3} \left(a + \frac{3b}{2} \right)^2 + \left(\frac{\sqrt{3}b}{2} \right)^2 \geq 0$$

Ejemplo 5.1.3

En el espacio euclídeo definido en el ejemplo anterior, hállese el producto escalar, si $p(x) = x$; $q(x) = 1 - x$.

$$p(x) \cdot q(x) = \int_0^1 x(1 - x) dx = \int_0^1 (x - x^2) dx = \left[\frac{1}{2} x^2 - \frac{1}{3} x^3 \right]_0^1 = 1/6.$$

5.2. Norma de un vector y ángulo entre dos vectores

Si recordamos la definición dada para el producto escalar de vectores geométricos:

$\bar{x} \cdot \bar{y} = |\bar{x}|\,|\bar{y}| \cos(\bar{x}, \bar{y})$

podemos obtener el módulo (longitud) de un vector, de la siguiente forma:

$\bar{x} \cdot \bar{x} = |\bar{x}|\,|\bar{x}| \cos(0) \Rightarrow |\bar{x}| = \sqrt{\bar{x} \cdot \bar{x}}$, pues $\cos(0) = 1$.

Generalizando la definición anterior para el caso de un espacio vectorial euclídeo, se puede dar la siguiente definición:

5.2.1. Definición de norma de un vector

> Se llama norma de un vector $\bar{x}$, perteneciente a un espacio vectorial euclídeo V, a la raíz cuadrada positiva del producto escalar $\bar{x} \cdot \bar{x}$ (se denota como $\|\bar{x}\|$) $\Leftrightarrow \|\bar{x}\| = \sqrt{\bar{x} \cdot \bar{x}}$

Observemos, que como $\bar{x} \cdot \bar{x} \geq 0$, siempre existe $\|\bar{x}\|$.

Ejemplo 5.2.1

Determínese la norma del vector $\bar{u} = (x_1, x_2)$, que figura en el ejemplo 5.1.1, correspondiente al producto escalar que allí se define.

Solución:

$\|\bar{u}\| = \sqrt{\bar{u} \cdot \bar{u}} = \sqrt{x_1^2 + x_2^2}$

Ejemplo 5.2.2.

Determínese la norma de los vectores $p(x) = x$; $q(x) = 1 - x$, de acuerdo con la definición de producto escalar dada en el ejemplo 5.1.2.

Solución:

$$\|p(x)\| = \sqrt{\int_0^1 x^2\, dx} = \sqrt{\left[\frac{x^3}{3}\right]_0^1} = \frac{1}{\sqrt{3}}$$

$$\|q(x)\| = \sqrt{\int_0^1 (1-x)^2\, dx} = \sqrt{\left[-\frac{1}{3}(1-x)^3\right]_0^1} = \frac{1}{\sqrt{3}}$$

5.2.2. Definición de vector unitario

> Se dice que un vector es unitario o normalizado cuando su norma es igual a uno $\Leftrightarrow$ $\Leftrightarrow \|\bar{u}\| = 1$

Podemos ver que dado un vector $\bar{u}$, no nulo, el vector $\bar{u}' = \dfrac{\bar{u}}{\|\bar{u}\|}$ es unitario,

ya que: $\|\bar{u}'\|^2 = \bar{u}' \cdot \bar{u}' = \left(\dfrac{\bar{u}}{\|\bar{u}\|}\right) \cdot \left(\dfrac{\bar{u}}{\|\bar{u}\|}\right) = \dfrac{\bar{u} \cdot \bar{u}}{\|\bar{u}\|^2} = \dfrac{\|\bar{u}\|^2}{\|\bar{u}\|^2} = 1$

5.2.3. Propiedades de la norma

Dados dos sectores $\bar{x}$ e $\bar{y}$ del espacio vectorial euclídeo V y $\forall\lambda \in \mathbb{R}$ se cumplen las siguientes propiedades de la norma:

1. $\| \lambda\bar{x} \| = \lambda \| \bar{x} \|$
2. $| \bar{x} \cdot \bar{y} | \le \| \bar{x} \| \| \bar{y} \|$ Desigualdad de Schwartz.
3. $\| \bar{x} + \bar{y} \| \le \| \bar{x} \| + \| \bar{y} \|$ Desigualdad de Minkowski. (Desigualdad triangular)

Demostraciones:

1. $\| \lambda\bar{x} \| = \sqrt{(\lambda\bar{x}) \cdot (\lambda\bar{y})} = \sqrt{\lambda^2(\bar{x} \cdot \bar{x})} = \lambda \| \bar{x} \|$

2. Desigualdad de Schwartz.

$$|\bar{x} \cdot \bar{y}| \le \| \bar{x} \| \, \| \bar{y} \|$$

Si $\bar{x} = \bar{0}$ ó $\bar{y} = \bar{0}$, no hace falta demostrar nada, lo único que se escribe es $0 \le 0$.

Si $\bar{x} \ne \bar{0}$ y $\bar{y} \ne \bar{0}$, sean $\bar{v}_1 = \dfrac{\bar{x}}{\| \bar{x} \|} + \dfrac{\bar{y}}{\| \bar{y} \|}$ y $\bar{v}_2 = \dfrac{\bar{x}}{\| \bar{x} \|} - \dfrac{\bar{y}}{\| \bar{y} \|}$, es evidente que $\bar{v}_1, \bar{v}_2 \in V$, por tanto, $\bar{v}_1 \cdot \bar{v}_1 \ge 0$ y $\bar{v}_2 \cdot \bar{v}_2 \ge 0 \Leftrightarrow$

$$\Leftrightarrow \left(\frac{\bar{x}}{\| \bar{x} \|} \pm \frac{\bar{y}}{\| \bar{y} \|}\right) \cdot \left(\frac{\bar{x}}{\| \bar{x} \|} \pm \frac{\bar{y}}{\| \bar{y} \|}\right) \ge 0 \Leftrightarrow \frac{\bar{x} \cdot \bar{x}}{\| \bar{x} \|^2} + \frac{\bar{y} \cdot \bar{y}}{\| \bar{y} \|^2} \pm 2\frac{\bar{x} \cdot \bar{y}}{\| \bar{x} \| \| \bar{y} \|} \ge 0,$$

como $\dfrac{\bar{x} \cdot \bar{x}}{\| \bar{x} \|^2} = 1$ y $\dfrac{\bar{y} \cdot \bar{y}}{\| \bar{y} \|^2} = 1$, podemos escribir: $2 \pm 2\dfrac{\bar{x} \cdot \bar{y}}{\| \bar{x} \| \| \bar{y} \|} \ge 0 \Leftrightarrow$

$$\Leftrightarrow \pm 2\frac{\bar{x} \cdot \bar{y}}{\| \bar{x} \| \| \bar{y} \|} \ge -2 \Leftrightarrow \pm \bar{x} \cdot \bar{y} \le \| \bar{x} \| \| \bar{y} \|.$$

3. Desigualdad de Minkowski.

$$\| \bar{x} + \bar{y} \|^2 = (\bar{x} + \bar{y}) \cdot (\bar{x} + \bar{y}) = \| \bar{x} \|^2 + \| \bar{y} \|^2 + 2(\bar{x} \cdot \bar{y})$$

Aplicando la desigualdad de Schwartz resulta: $\| \bar{x} + \bar{y} \|^2 \le \| \bar{x} \|^2 + \| \bar{y} \|^2 + 2\| \bar{x} \| \| \bar{y} \| = (\| \bar{x} \| + \| \bar{y} \|)^2$, y extrayendo la raíz cuadrada, obtenemos la desigualdad de Minkowski: $\| \bar{x} + \bar{y} \| \le \| \bar{x} \| + \| \bar{y} \|$.

Si ahora expresamos la desigualdad de Schwartz: $|\bar{x} \cdot \bar{y}| \le \| \bar{x} \| \| \bar{y} \|$ en la forma equivalente:

$-\| \bar{x} \| \| \bar{y} \| \le \bar{x} \cdot \bar{y} \le \| \bar{x} \| \| \bar{y} \|$, y suponiendo que $\| \bar{x} \|, \| \bar{y} \| \ne 0$, dividimos por $\| \bar{x} \| \| \bar{y} \|$ obtenemos:

$$-1 \le \frac{\bar{x} \cdot \bar{y}}{\| \bar{x} \| \| \bar{y} \|} \le 1.$$

Podemos ver que el cociente anterior toma los mismos valores que el coseno de un ángulo. Además para $\bar{x} = \bar{y}$ se tendría un ángulo igual a cero y el valor del cociente anterior resulta igual a uno.

Aunque, en general los vectores del espacio vectorial euclídeo V no serán vectores geométricos, las consideraciones anteriores permiten llegar a la siguiente definición:

5.2.4. Definición de coseno del ángulo entre dos vectores

Dados dos vectores no nulos $\bar{x}, \bar{y}$ pertenecientes a un espacio vectorial euclídeo V se define como coseno del ángulo entre dichos vectores a: $\cos(\bar{x}, \bar{y}) = \dfrac{\bar{x} \cdot \bar{y}}{\| \bar{x} \| \| \bar{y} \|}$

Ejemplo 5.2.3.

Calcúlese el coseno del ángulo que forman los vectores $p(x) = x$; $q(x) = 1 - x$, para el producto escalar definido en el ejemplo 5.1.2, teniendo en cuenta los valores obtenidos en los ejemplos 5.1.3 y 5.2.2.

Solución:

$$\cos(p, q) = \frac{p \cdot q}{\|p\| \, \|q\|} = \frac{1/6}{(1/\sqrt{3})(1/\sqrt{3})} = \frac{1}{2}$$

5.2.5. Definición de ortogonalidad

> Dos vectores $\overline{x}$, $\overline{y}$ son <u>ortogonales</u> $\Leftrightarrow$ El producto escalar $\overline{x} \cdot \overline{y}$ es igual a cero, siendo $\overline{x}$, $\overline{y}$ distintos del vector nulo.

De acuerdo con la definición anterior, para que dos vectores sean ortogonales debe cumplirse que $\cos(\overline{x}, \overline{y}) = 0$. Esta definición coincide con la dada para perpendicularidad de vectores geométricos, en la que el ángulo que determinan ambos vectores debe ser recto.

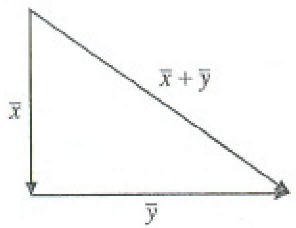
Observemos que se ha definido la ortogonalidad para vectores no nulos. También podemos ver que el vector nulo es ortogonal a cualquier vector $\overline{x}$, ya que $\overline{0} \cdot \overline{x} = 0$.

Dados dos vectores de un espacio euclídeo $\overline{x}$, $\overline{y}$, se verifica:

$$\|\overline{x} + \overline{y}\|^2 = (\overline{x} + \overline{y}) \cdot (\overline{x} + \overline{y}) = \overline{x} \cdot \overline{x} + 2(\overline{x} \cdot \overline{y}) + \overline{y} \cdot \overline{y}$$

pero si dichos vectores son ortogonales, es: $\overline{x} \cdot \overline{y} = 0$, se obtiene la conocida expresión del teorema de Pitágoras:

$$\|\overline{x} + \overline{y}\|^2 = \|\overline{x}\|^2 + \|\overline{y}\|^2$$



Ejemplo 5.2.4.

En el espacio vectorial euclídeo $\mathbb{R}^3$, con el producto escalar usual, determínese un vector que sea ortogonal a los vectores: $\overline{u} = (1, 2, 1)$; $\overline{v} = (0, -1, 1)$; $\overline{w} = (1, 1, 2)$.

Solución:

Sea $\overline{x} = (a, b, c)$ el vector a determinar.

Como $\overline{x}$ debe ser ortogonal a u entonces: $\overline{x} \cdot \overline{u} = 0 \Rightarrow a + 2b + c = 0$

También debe ser ortogonal a $\overline{v}$: $\overline{x} \cdot \overline{v} = 0 \Rightarrow -b + c = 0$.

Por último, debe ser ortogonal a $\overline{w}$: $\overline{x} \cdot \overline{w} = 0 \Rightarrow a + b + 2c = 0$

Por lo tanto, el vector buscado debe ser la solución del sistema:

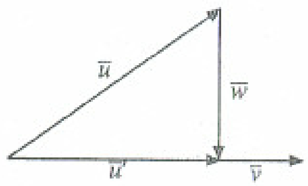
$$\begin{cases} a + 2b + c = 0 \\ -b + c = 0 \\ a + b + 2c = 0 \end{cases}$$

Sistema que tiene por solución: $a = -3\alpha$; $b = \alpha$, $c = \alpha \Rightarrow \overline{x} = (-3\alpha, \alpha, \alpha)$, es decir, el espacio vectorial generado por $(-3, 1, 1)$.

Volviendo al caso de los vectores geométricos, de gran utilidad en física, cuando se habla de la proyección de un vector sobre otro, en general se hace referencia al módulo del vector proyección, pero si lo que se requiere es el vector que se obtiene como proyección ortogonal de un vector $\overline{u}$ sobre otro vector $\overline{v}$, tendremos que determinarlo como la composición de $\overline{u}$ con otro vector $\overline{w}$ ortogonal a $\overline{v}$.

Sea $\overline{u}'$ el vector proyección de $\overline{u}$ sobre $\overline{v}$.

El módulo de $\overline{u}'$ resulta: $\|\overline{u}'\| = \|\overline{u}\| \cos(\overline{u}, \overline{v}) = \dfrac{(\overline{u} \cdot \overline{v})}{\|\overline{v}\|}$



Multiplicando por el vector unitario de la dirección de $\overline{v}$: $\dfrac{\overline{v}}{\|\overline{v}\|}$, se obtiene el vector $\overline{u}' = \dfrac{(\overline{u}\cdot\overline{v})}{\|\overline{v}\|^2}\,\overline{v}$

Es posible extender este concepto para cualquier par de vectores de un espacio vectorial euclídeo y decir que la proyección ortogonal de un vector $\overline{u}$ sobre otro vector $\overline{v}$ de dicho espacio es:

$$\text{Proy } \overline{u}_{\overline{v}} = \frac{(\overline{u}\,.\,\overline{v})}{\|\overline{v}\|^2}\,\overline{v}$$

Ejemplo 5.2.5.

Determínese el vector que se obtiene como proyección ortogonal del vector $p(x) = x$ sobre el vector $q(x) = 1 - x$, con la definición de producto escalar dada en el ejemplo 5.1.2 y teniendo en cuenta los resultados obtenidos en los ejemplos 5.1.3 y 5.2.2.

Solución:

$$\text{Proy } p_q = \frac{(p\cdot q)}{\|q\|^2}\,q = \frac{1/6}{(1/\sqrt{3})^2}\,(1-x) = \frac{1}{2} - \frac{1}{2}\,x$$

5.3. Expresión del producto escalar en una base dada

Si V es un espacio vectorial euclídeo de dimensión n, y $B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\}$ una base cualquiera de V, entonces, cualquier pareja de vectores $\overline{x}$, $\overline{y} \in V$ se pueden expresar como:

$$\overline{x} = x_1\overline{e}_1 + x_2\overline{e}_2 + ... + x_n\overline{e}_n;\ \overline{y} = y_1\overline{e}_1 + y_2\overline{e}_2 + ... + y_n\overline{e}_n.$$

Si conocemos los n^2 productos escalares $\overline{e}_i\,.\,\overline{e}_j$ (para $i, j = 1, 2, ..., n$), el producto escalar $\overline{x}\cdot\overline{y}$ se puede expresar de la siguiente forma:

$$\overline{x}\cdot\overline{y} = (x_1\overline{e}_1 + x_2\overline{e}_2 + ... + x_n\overline{e}_n)\cdot(y_1\overline{e}_1 + y_2\overline{e}_2 + ... + y_n\overline{e}_n)$$

Luego, utilizando las propiedades del producto escalar:

$$\overline{x}\cdot\overline{y} = x_1y_1(\overline{e}_1\cdot\overline{e}_1) + x_1y_2(\overline{e}_1\cdot\overline{e}_2) + ... + x_1y_n(\overline{e}_1\cdot\overline{e}_n) + x_2y_1(\overline{e}_2\cdot\overline{e}_1) + \\ + x_2y_2(\overline{e}_2\cdot\overline{e}_2) + ... + x_2y_n(\overline{e}_2\cdot\overline{e}_n) + ... + x_ny_1(\overline{e}_n\cdot\overline{e}_1) + x_ny_2(\overline{e}_n\cdot\overline{e}_2) + ... + x_ny_n(\overline{e}_n\cdot\overline{e}_n)$$

Si para todo i, j se denota el producto $\overline{e}_i\cdot\overline{e}_j = g_{ij}$, es posible escribir el producto escalar $\overline{x}\cdot\overline{y}$ de forma matricial:

$$\overline{x}\cdot\overline{y} = (x_1\ x_2\ ...\ x_n)\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

De manera abreviada se puede expresar como $\overline{x}\cdot\overline{y} = XGY'$.
donde X e Y son las matrices fila de las coordenadas de los vectores $\overline{x}$ e $\overline{y}$, respectivamente y G es la llamada matriz métrica del producto escalar o matriz de Gram.

Es un sencillo ejercicio demostrar que, para que la expresión XGY' represente el producto escalar de los vectores $\overline{x}$ e $\overline{y}$, respecto de una base dada, es condición necesaria y suficiente que se verifiquen las siguientes condiciones:

1. $g_{ij} = g_{ji}$, $\forall i \neq j$ (La matriz G debe ser simétrica).
2. $g_{ij} > 0$, $\forall i = j$ (Todos los elementos de la diagonal principal deben ser positivos).
3. $g_{ii}g_{jj} - g_{ij}^2 \geq 0$, $\forall_{i,j}$.

Ejemplo 5.3.1.

El producto escalar de los vectores $\overline{x} = (x_1, x_2)$ e $\overline{y} = (y_1, y_2)$ de un espacio euclídeo bidimensional, referidos a la base $B = \{\overline{e}_1, \overline{e}_2\}$ está definido por:

$\cdot y = 4x_1y_1(e_1 \cdot e_1) \cdot$
$x_1y_2(e_1 \cdot e_2) \cdot$
$x_2y_1(e_2 \cdot e_1)$;
$9x_2y_2(e_1 \cdot e_1)$

$$\overline{x} \cdot \overline{y} = 4x_1y_1 + x_1y_2 + x_2y_1 + 9x_2y_2$$

$u \cdot u : 4.2.2 + 2 + 2 + 9 = 13 + 16 = 29$

1. Hállese la matriz que caracteriza a ese producto escalar y verifíquese que cumple con las condiciones anteriormente expresadas.
2. Calcúlense las normas de los vectores de la base.
3. Determínese el coseno del ángulo que forman los vectores de la base.
4. Determínese el producto escalar $\overline{u} \cdot \overline{u}$, siendo $\overline{u} = (2, 1)$.

a) Si la base primitiva B es la formada por los vectores $\overline{e}_1 = (1, 0)$; $\overline{e}_2 = (0, 1)$.

b) En otra base B' dada por los vectores $\overline{e}'_1 = \left(\frac{1}{2}, 0\right)$; $\overline{e}'_2 = \left(0, \frac{1}{3}\right)$.

Solución:

1. $\overline{x} \cdot \overline{y} = (x_1\ x_2)\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (x_1g_{11} + x_2g_{21}\quad x_1g_{12} + x_2g_{22})\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$

$= g_{11}x_1y_1 + g_{21}x_2y_1 + g_{12}x_1y_2 + g_{22}x_2y_2$

Por tanto, $g_{11} = 4$; $g_{12} = g_{21} = 1$; $g_{22} = 9 \Rightarrow G = \begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}$.

G cumple las condiciones requeridas:

1. $g_{12} = g_{21} = 1$.
2. $g_{11} = 4 > 0$; $g_{22} = 9 > 0$.
3. $g_{11} \cdot g_{22} - g_{12}^2 = 4.9 - 1 \geq 0$.

2. Para calcular la norma de los vectores de la base basta hacer:

$$\|\overline{e}_1\|^2 = \overline{e}_1 \cdot \overline{e}_1 = g_{11} = 4 \Rightarrow \|\overline{e}_1\| = \sqrt{g_{11}} = 2$$

$$\|\overline{e}_2\|^2 = \overline{e}_2 \cdot \overline{e}_2 = g_{22} = 9 \Rightarrow \|\overline{e}_2\| = \sqrt{g_{22}} = 3$$

3. $\cos(\overline{e}_1, \overline{e}_2) = \dfrac{\overline{e}_1 \cdot \overline{e}_2}{\|\overline{e}_1\| \|\overline{e}_2\|} = \dfrac{g_{12}}{\|\overline{e}_1\| \|\overline{e}_2\|} = \dfrac{1}{6}$.

4. a) En la base canónica B el vector $\overline{u}$ queda expresado como $\overline{u} = 2\overline{e}_1 + \overline{e}_2$, luego: $\overline{u} \cdot \overline{u} = UGU' = (2\ 1)\begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = (9\ 11)\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 29$.

También podríamos haber calculado el producto escalar a partir de la expresión original.

b) Como $\overline{x} \cdot \overline{y} = 4x_1y_1 + x_1y_2 + x_2y_1 + 9x_2y_2$, si los vectores de la base B' son:

$$\overline{e}'_1 = \left(\frac{1}{2}, 0\right) \text{ y } \overline{e}'_2 = \left(0, \frac{1}{3}\right) \begin{cases} \overline{e}'_1 \cdot \overline{e}'_1 = 4 \cdot \left(\frac{1}{2}\right)^2 + 2 \cdot \frac{1}{2} \cdot 0 + 9 \cdot 0^2 = 1 \\ \overline{e}'_2 \cdot \overline{e}'_2 = 4 \cdot 0^2 + 2 \cdot 0 \cdot \frac{1}{3} + 9 \cdot \left(\frac{1}{3}\right)^2 = 1 \\ \overline{e}'_1 \cdot \overline{e}'_2 = \overline{e}'_2 \cdot \overline{e}'_1 = 4 \cdot \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot 0 + \\ + 9 \cdot 0 \cdot \frac{1}{3} = \frac{1}{6} \end{cases}$$

Por tanto, la matriz métrica que caracteriza el producto escalar en la base B' es $G' = \begin{pmatrix} 1 & 1/6 \\ 1/6 & 1 \end{pmatrix}$

En la base B' el vector $\overline{u}$ queda expresado como $\overline{u} = 4\overline{e}'_1 + 3\overline{e}'_2 \Rightarrow$

$$\Rightarrow \overline{u} \cdot \overline{u} = (4\ 3)\begin{pmatrix} 1 & 1/6 \\ 1/6 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ 3 \end{pmatrix} = (9/2\ \ 11/3)\begin{pmatrix} 4 \\ 3 \end{pmatrix} = 29.$$

Como podemos ver, el valor del producto escalar es el mismo, independientemente de la base elegida.

5.4. Sistemas ortogonales y ortonormales de vectores

Sea V un espacio vectorial euclídeo y S un subconjunto de V, diremos que S es un sistema ortogonal si cualquier vector de S es ortogonal a todos los demás vectores de S.

5.4.1. Definición de sistema ortogonal

El sistema $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_n\} \subset V$ es ortogonal $\Leftrightarrow$ $\overline{u}_i \cdot \overline{u}_j = 0, \ \forall\, i \neq j$

5.4.2. Proposición

$S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_n\}$ es un sistema ortogonal de vectores no nulos $\Rightarrow S$ es linealmente independiente.

Demostración:

Sea $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_n\}$ un sistema ortogonal de vectores no nulos, tal que, $\lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + ... + \lambda_n\overline{u}_n = 0$ con $\lambda_i \in \mathbb{R}$.

Por las propiedades del producto escalar:

$$\overline{u}_1 \cdot (\lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + ... + \lambda_n\overline{u}_n) = \lambda_1\|\overline{u}_1\|^2 + 0 + ... + 0 = 0 \Rightarrow \lambda_1 = 0$$

De manera similar, comprobamos que $\lambda_2 = \lambda_3 = ... = \lambda_n = 0$; luego los vectores del sistema ortogonal S son linealmente independientes.

Si S es un subconjunto de un espacio vectorial euclídeo con todos sus vectores unitarios se dice que es un sistema normado.

5.4.3. Definición de sistema normado

El sistema $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_n\}$ es normado $\Leftrightarrow \|\overline{u}_i\| = 1,\ 1 \le i \le n$

Si un sistema de vectores S, contenido en un espacio vectorial euclídeo, es ortogonal y además normado, entonces se dice que es un sistema ortonormado.

5.4.4. Definición de sistema ortonormado

El sistema $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_n\}$ es ortonormado $\Leftrightarrow \overline{u}_i \cdot \overline{u}_j = \begin{cases} 0 & \forall i \neq j \\ 1 & \forall i = j \end{cases}$

Ejemplo 5.4.1.

En el sistema $S = \{1, \text{sen}x, \cos x\}$ cuyos vectores son las funciones: 1, sen x, cos x, se define el producto escalar de la siguiente forma:

$$f(x) \cdot g(x) = \int_{-\pi}^{\pi} f(x) \,.\, g(x)dx.$$

a) Verifíquese que S es un sistema ortogonal.
b) Encuéntrese un sistema ortonormado a partir de S.

Solución:

a) Si $h(x) = 1$; $k(x) = \text{sen}\ x$; $m(x) = \cos x$

$$h(x) \cdot k(x) = \int_{-\pi}^{\pi} 1 \cdot \text{sen}\ x\, dx = \left[-\cos x\right]_{-\pi}^{\pi} = -1 - (-1) = 0$$

$$h(x) \cdot m(x) = \int_{-\pi}^{\pi} 1 \cdot \cos x\, dx = \left[\text{sen}\ x\right]_{-\pi}^{\pi} = 0$$

$$k(x) \cdot m(x) = \int_{-\pi}^{\pi} \text{sen}\ x \cdot \cos x dx = \int_{-\pi}^{\pi} \text{sen}\ x\, d(\text{sen}\ x) = \left[\frac{1}{2}\text{sen}^2 x\right]_{-\pi}^{\pi} = 0$$

Por tanto, el sistema es ortogonal.

b) Para encontrar un sistema que, además de ortogonal, sea normalizado es necesario calcular primero la norma de cada uno de los vectores dados:

$$\|h(x)\| = \sqrt{\int_{-\pi}^{\pi} 1.1\, dx} = \sqrt{[x]_{-\pi}^{\pi}} = \sqrt{2\pi}$$

$$\|k(x)\| = \sqrt{\int_{-\pi}^{\pi} \text{sen}^2 x\, dx} = \sqrt{\left[\frac{1}{2}(x - \text{sen}\ x \cdot \cos x)\right]_{-\pi}^{\pi}} = \sqrt{\pi}$$

$$\|m(x)\| = \sqrt{\int_{-\pi}^{\pi} \cos^2 x\, dx} = \sqrt{\left[\frac{1}{2}(x + \text{sen}\ x \cdot \cos x)\right]_{-\pi}^{\pi}} = \sqrt{\pi}$$

El sistema ortonormado (ortogonal y normalizado) es:

$$S' = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\text{sen}\ x}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}.\right\}$$

5.4.5. Método de ortonormalización de Gram-Schmidt

Como el valor del producto escalar es invariante cualquiera que sea la base considerada, es conveniente elegir una base en la que dicho producto tenga la expresión más sencilla posible. Esta expresión la podremos obtener cuando la matriz métrica G sea la matriz unitaria I, porque entonces: $\overline{x} \cdot \overline{y} = XIY^t = x_1y_1 + x_2y_2 + ... + x_ny_n$. (Es decir, el producto escalar usual.)

Esto es lo que ocurre cuando la base es ortonormada.

Bases ortonormadas hay infinitas, pero el problema que se plantea es, a partir de una base cualquiera del espacio vectorial euclídeo, encontrar otra base que sea ortonormada.

Este problema es fácil de resolver con el método de Gram-Schmidt, que nos permite obtener una única base formada por vectores ortogonales, que después se normalizan.

- Sea $B = \{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\}$ la base primitiva y $B' = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_n\}$ la base ortogonal a determinar.
- Primero formamos las siguientes relaciones:

$$\overline{u}_1 = \overline{e}_1$$
$$\overline{u}_2 = \overline{e}_2 + \alpha_1\overline{u}_1$$
$$\overline{u}_3 = \overline{e}_3 + \beta_1\overline{u}_1 + \beta_2\overline{u}_2$$
$$\dots\dots\dots\dots$$
$$\overline{u}_n = \overline{e}_n + \lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + \dots + \lambda_{n-1}\overline{u}_{n-1}$$

- A continuación hallamos los escalares $\alpha_1, \beta_1, \lambda_{n-1}$, teniendo en cuenta que debe ser $\overline{u}_i \cdot \overline{u}_j = 0$, $\forall i, j$, tales que $i \neq j$.

$$\overline{u}_1 \cdot \overline{u}_2 = \overline{u}_1 \cdot (\overline{e}_2 + \alpha_1\overline{u}_1) = \overline{u}_1 \cdot \overline{e}_2 + \alpha_1(\overline{u}_1 \cdot \overline{u}_1) = 0 \Rightarrow \alpha_1 = -\frac{(\overline{u}_1 \cdot \overline{e}_2)}{\|\overline{u}_1\|^2}.$$

$$\overline{u}_1 \cdot \overline{u}_3 = \overline{u}_1 \cdot (\overline{e}_3 + \beta_1\overline{u}_1 + \beta_2\overline{u}_2) = \overline{u}_1 \cdot \overline{e}_3 + \beta_1(\overline{u}_1 \cdot \overline{u}_1) + \beta_2(\overline{u}_1 \cdot \overline{u}_2) = 0 \Rightarrow$$

$$\Rightarrow \beta_1 = -\frac{(\overline{u}_1 \cdot \overline{e}_2)}{\|\overline{u}_1\|^2}$$

$$\overline{u}_2 \cdot \overline{u}_3 = \overline{u}_2 \cdot (\overline{e}_3 + \beta_1\overline{u}_1 + \beta_2\overline{u}_2) = \overline{u}_2 \cdot \overline{e}_3 + \beta_1(\overline{u}_2 \cdot \overline{u}_1) + \beta_2(\overline{u}_2 \cdot \overline{u}_2) = 0 \Rightarrow$$

$$\Rightarrow \beta_2 = -\frac{(\overline{u}_2 \cdot \overline{e}_3)}{\|\overline{u}_2\|^2}$$

- A partir de estos resultados obtenemos:

$$\overline{u}_1 = \overline{e}_1$$

$$\overline{u}_2 = \overline{e}_2 - \frac{(\overline{u}_1 \cdot \overline{e}_2)}{\|\overline{u}_1\|^2}\overline{u}_1$$

$$\overline{u}_3 = \overline{e}_3 - \frac{(\overline{u}_1 \cdot \overline{e}_3)}{\|\overline{u}_1\|^2}\overline{u}_1 - \frac{(\overline{u}_2 \cdot \overline{e}_3)}{\|\overline{u}_2\|^2}\overline{u}_2$$

- Reiterando el proceso, logramos encontrar los restantes vectores de la base ortogonal, que de forma general resultarán:

$$\overline{u}_k = \overline{e}_k - \frac{(\overline{u}_1 \cdot \overline{e}_k)}{\|\overline{u}_1\|^2}\overline{u}_1 - \frac{(\overline{u}_2 \cdot \overline{e}_k)}{\|\overline{u}_2\|^2}\overline{u}_2 - \dots - \frac{(\overline{u}_{k-1} \cdot \overline{e}_k)}{\|\overline{u}_{k-1}\|}\overline{u}_{k-1}$$

- Y para obtener la base ortonormada bastará dividir cada vector por su norma:

$$B' = \left\{\frac{\overline{u}_1}{\|\overline{u}_1\|}, \frac{\overline{u}_2}{\|\overline{u}_2\|}, \dots, \frac{\overline{u}_n}{\|\overline{u}_n\|}\right\}$$

Como las expresiones obtenidas anteriormente no son fáciles de recordar, es preferible aplicar en cada ejercicio en particular el algoritmo estudiado de forma general.

Ejemplo 5.4.2

En el espacio vectorial euclídeo del ejemplo 5.3.1, donde el producto escalar de dos vectores $\overline{x} = (x_1, x_2)$, $\overline{y} = (y_1, y_2)$ estaba definido por:

$$\overline{x} \cdot \overline{y} = 4x_1y_1 + x_1y_2 + x_2y_1 + 9x_2y_2,$$

vimos que la base B formada por los vectores $\overline{e}_1 = (1, 0)$; $\overline{e}_2 = (0, 1)$ no es normada ni ortogonal. Se pide:

a) Mediante el método de Gram-Schmidt y a partir de los vectores de la base B, obtener una base ortonormada.

b) Verificar que en dicha base la matriz métrica es la unitaria.
c) Calcular en la base ortonormal obtenida el producto escalar $\overline{u} \cdot \overline{u}$, siendo $\overline{u} = (2, 1)$

Solución:

a) Aplicando el método de Gram-Schmidt, se elige el primer vector de la base ortogonal buscada como: $\overline{u}_1 = \overline{e}_1 = (1, 0)$.
Después se hace: $\overline{u}_2 = \overline{e}_2 + \alpha\overline{u}_1 = (0, 1) + \alpha(1, 0) = (\alpha, 1)$.
Pero debe ser: $\overline{u}_1 \cdot \overline{u}_2 = 0$.
Por tanto, teniendo en cuenta la expresión del producto escalar dada, es:

$$\overline{u}_1 \cdot \overline{u}_2 = (1, 0) \cdot (\alpha, 1) = 4.1.\alpha + 1.1 + \alpha.0 + 9.0.1 = 4\alpha + 1 = 0 \Rightarrow \alpha = -\frac{1}{4}.$$

Luego: $\overline{u}_1 = (1, 0)$; $\overline{u}_2 = \left(-\frac{1}{4}, 1\right)$.

Aunque los vectores $\overline{u}_1, \overline{u}_2$ son ortogonales, para obtener los vectores paralelos a ellos, y que además estén normalizados, habrá que dividir cada uno por la norma correspondiente.

$$\|\overline{u}_1\| = \sqrt{\overline{u}_1 \cdot \overline{u}_1} = \sqrt{4 \,.\, 1^2 + 2 \,.\, 1 \,.\, 0 + 9 \,.\, 0^2} = 2.$$

$$\|\overline{u}_2\| = \sqrt{\overline{u}_2 \cdot \overline{u}_2} = \sqrt{4 \,.\, (-1/4)^2 + 2 \,.\, (-1/4) \,.\, 1 + 9 \,.\, 1^2} = \frac{\sqrt{35}}{2}.$$

Por tanto, la base ortonormada está formada por los vectores:

$$\overline{v}_1 = \frac{\overline{u}_1}{\|\overline{u}_1\|} = \left(\frac{1}{2}, 0\right), \overline{v}_2 = \frac{\overline{u}_2}{\|\overline{u}_2\|} = \left(-\frac{1}{2\sqrt{35}}, \frac{2}{\sqrt{35}}\right).$$

b) Para obtener la matriz métrica es necesario calcular:

$$g_{11} = \overline{v}_1 \cdot \overline{v}_1 = 4 \,.\, (1/2)^2 + 2.(1/2) \,.\, 0 + 9 \,.\, 0^2 = 1$$

$$g_{22} = \overline{v}_2 \cdot \overline{v}_2 = 4 \,.\left(-\frac{1}{2\sqrt{35}}\right)^2 + 2 \,.\left(-\frac{1}{2\sqrt{35}}\right) . \frac{2}{\sqrt{35}} + 9 \,.\left(\frac{2}{\sqrt{35}}\right)^2 = 1$$

$$g_{12} = g_{21} = 4 \,.\, (1/2) \,.\left(-\frac{1}{2\sqrt{35}}\right) + (1/2) \,.\, \frac{2}{\sqrt{35}} + 0 \,.\left(-\frac{1}{2\sqrt{35}}\right) + 9 \,.\, 0 \,.\, \frac{2}{\sqrt{35}} = 0$$

Como era de esperar, la matriz métrica resulta ser la matriz unitaria:

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

c) En la base ortonormada que se ha determinado, el vector $\overline{u}$ se puede expresar como: $\overline{u} = a\overline{v}_1 + b\overline{v}_2 = a\left(\frac{1}{2}, 0\right) + b\left(-\frac{1}{2\sqrt{35}}, \frac{2}{\sqrt{35}}\right) = (2, 1) \Rightarrow$

$$\Rightarrow \begin{cases} \frac{1}{2}\, a - \frac{1}{2\sqrt{35}}\, b = 2 \\ \\ 0 \,.\, a + \frac{2}{\sqrt{35}}\, b = 1 \end{cases} \Rightarrow b = \frac{\sqrt{35}}{2};\ a = \frac{9}{2}.$$

Por tanto, el vector $\overline{u}$, en la base B', resulta: $\overline{u} = \frac{9}{2}\overline{v}_1 + \frac{\sqrt{35}}{2}\overline{v}_2$.

Luego, el producto escalar $\overline{u} \cdot \overline{u}$ es: $\overline{u} \cdot \overline{u} = \begin{pmatrix} \frac{9}{2} & \frac{\sqrt{35}}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{9}{2} \\ \frac{\sqrt{35}}{2} \end{pmatrix} = 29,$

valor que, como era de esperar, coincide con el obtenido en el ejemplo 5.3.1.

5.4.6. Definición de matriz ortogonal

Una matriz cuadrada A se dice que es una matriz ortogonal si: $AA' = A'A = I$ (I matriz unitaria) $\Leftrightarrow A' = A^{-1}$ (esta condición evidencia que A es una matriz regular)

Ejemplo 5.4.3.

a) Vamos a verificar que la siguiente matriz A es ortogonal:

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

En efecto:

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

$$= \begin{pmatrix} \frac{1}{3}+\frac{1}{3}+\frac{1}{3} & \frac{1}{3\sqrt{2}}-\frac{2}{3\sqrt{2}}+\frac{1}{3\sqrt{2}} & \frac{-1}{\sqrt{6}}+0+\frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}}-\frac{2}{3\sqrt{2}}+\frac{1}{3\sqrt{2}} & \frac{1}{6}+\frac{4}{6}+\frac{1}{6} & \frac{-1}{2\sqrt{3}}+0+\frac{1}{2\sqrt{3}} \\ -\frac{1}{\sqrt{6}}+0+\frac{1}{\sqrt{6}} & \frac{-1}{2\sqrt{3}}+0+\frac{1}{2\sqrt{3}} & \frac{1}{2}+0+\frac{1}{2} \end{pmatrix} =$$

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

b) Probaremos que los vectores fila forman una base ortonormal del espacio euclídeo $\mathbb{R}^3$.

Si consideramos los vectores fila:

$$\overline{u} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right); \overline{v} = \left(\frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right); \overline{w} = \left(\frac{-1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

podemos comprobar que:

i) Los tres vectores son normalizados:

$$\|\overline{u}\| = \sqrt{\left(\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2} = \sqrt{\frac{1}{3}+\frac{1}{3}+\frac{1}{3}} = 1$$

$$\|\overline{v}\| = \sqrt{\left(\frac{1}{\sqrt{6}}\right)^2 + \left(\frac{-2}{\sqrt{6}}\right)^2 + \left(\frac{1}{\sqrt{6}}\right)^2} = \sqrt{\frac{1}{6}+\frac{4}{6}+\frac{1}{6}} = 1$$

$$\|\overline{w}\| = \sqrt{\left(\frac{-1}{\sqrt{2}}\right)^2 + 0 + \left(\frac{1}{\sqrt{2}}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1$$

ii) Los tres vectores dados son ortogonales entre sí:

$$\overline{u} \cdot \overline{v} = \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{3}} \cdot \left(\frac{-2}{\sqrt{6}}\right) + \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{6}} = \frac{1}{3\sqrt{2}} - \frac{2}{3\sqrt{2}} + \frac{1}{3\sqrt{2}} = 0$$

$$\overline{u} \cdot \overline{w} = \frac{1}{\sqrt{3}} \cdot \left(\frac{-1}{\sqrt{2}}\right) + \frac{1}{\sqrt{3}} \cdot 0 + \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{2}} = \frac{-1}{\sqrt{6}} + \frac{1}{\sqrt{6}} = 0$$

$$\overline{v} \cdot \overline{w} = \frac{1}{\sqrt{6}} \cdot \left(\frac{-1}{\sqrt{2}}\right) + \left(\frac{-2}{\sqrt{6}}\right) \cdot 0 + \frac{1}{\sqrt{6}} \cdot \frac{1}{\sqrt{2}} = -\frac{1}{3\sqrt{2}} + \frac{1}{3\sqrt{2}} = 0$$

$\Rightarrow$ Los vectores fila de la matriz A forman una base ortonormal del espacio euclídeo $\mathbb{R}^3$.

5.4.7. Producto escalar en un espacio vectorial con una base ortonormal

Si tenemos un espacio vectorial euclídeo V y dos vectores cualquiera del mismo, $\overline{x} = (x_1, x_2, ..., x_n)$, $\overline{y} = (y_1, y_2, ..., y_n)$, dados respecto a una base ortonormal B, el producto escalar $x \,.\, y$ se puede expresar como: $\overline{x} \cdot \overline{y} = x_1y_1 + x_2y_2 + ... + x_ny_n$, es decir, se reduce al producto escalar usual, ya que en este caso, la matriz métrica es la matriz unitaria.

También es inmediato verificar que la norma de un vector $\overline{x}$ se puede determinar como: $\|\overline{x}\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$.

5.4.8. Base ortonormal incompleta

Si V es un espacio euclídeo de dimensión n y $S = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_k\}$ es un sistema ortonormal con $k < n$, es posible completarlo hasta obtener una base ortonormal.

Demostración:

Como $k < n$, existe algún otro vector $\overline{v} \in V$, que es independiente del sistema dado S. Por tanto, el vector $\overline{e} = \overline{v} - (\lambda_1\overline{u}_1 + \lambda_2\overline{u}_2 + ... + \lambda_k\overline{u}_k)$ es no nulo, cualesquiera sean los escalares λ_i; pero para ciertos valores de λ_i resulta $\overline{e}$ ortogonal a $\overline{u}_1, \overline{u}_2, ..., \overline{u}_k$. En efecto, ya que:

$$\overline{e} \cdot \overline{u}_1 = \overline{v} \cdot \overline{u}_1 - (\lambda_1\overline{u}_1 \cdot \overline{u}_1 + \lambda_2\overline{u}_2 \cdot \overline{u}_1 + ... + \lambda_k\overline{u}_k \cdot \overline{u}_1) = 0 \text{ si } \lambda_1 = \overline{v} \cdot \overline{u}_1$$
$$\overline{e} \cdot \overline{u}_2 = \overline{v} \cdot \overline{u}_2 - (\lambda_1\overline{u}_1 \cdot \overline{u}_2 + \lambda_2\overline{u}_2 \cdot \overline{u}_2 + ... + \lambda_k\overline{u}_k \cdot \overline{u}_2) = 0 \text{ si } \lambda_2 = \overline{v} \cdot \overline{u}_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\overline{e} \cdot \overline{u}_k = \overline{v} \cdot \overline{u}_k - (\lambda_1\overline{u}_1 \cdot \overline{u}_k + \lambda_2\overline{u}_2 \cdot \overline{u}_k + ... + \lambda_k\overline{u}_k \cdot \overline{u}_k) = 0 \text{ si } \lambda_k = \overline{v} \cdot \overline{u}_k$$

Entonces, si agregamos al sistema el vector:

$\dfrac{\overline{e}}{\|\overline{e}\|} = \overline{u}_{k+1}$, obtenemos otro sistema:

$S_1 = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_k, \overline{u}_{k+1}\}$, que también es ortonormado pero con un vector más.

Si $k + 1 < n$, reiterando el proceso se obtiene el sistema:

$S' = \{\overline{u}_1, \overline{u}_2, ..., \overline{u}_k, \overline{u}_{k+1}, ..., \overline{u}_n\}$, que es base ortonormal de V.

Ejemplo 5.4.4

Sean $\overline{u}_1 = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}\right)$; $\overline{u}_2 = \left(\frac{2}{\sqrt{6}}, 0, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)$, vectores del espacio vectorial euclídeo canónico $\mathbb{R}^4$.

Determínese un vector $\overline{u}_3$ que pueda formar parte de una base ortonormal de $\mathbb{R}^4$ siguiendo el procedimiento visto en el apartado *[5.4.8]*.

Es fácil comprobar que el sistema dado por los vectores $\overline{u}_1$ y $\overline{u}_2$ es ortonormalizado.

Elegimos un vector cualquiera del espacio dado que sea independiente de $\overline{u}_1$ y $\overline{u}_2$.

Por ejemplo, podemos verificar que un vector, arbitrariamente elegido, $\overline{v} = (0, 0, 3, -1)$ no se puede obtener como combinación lineal de $\overline{u}_1$ y $\overline{u}_2$, pues para que se cumpla que $C_1\overline{u}_1 + C_2\overline{u}_2 + C_3\overline{v} = \overline{0}$ es necesario que sea: $C_1 = C_2 = C_3 = 0$.

Para determinar $\overline{e} = \overline{v} - \lambda_1\overline{u}_1 - \lambda_2\overline{u}_2$ que sea ortogonal a $\overline{u}_1$ y $\overline{u}_2$, debe ser:

$$\lambda_1 = \overline{v} \cdot \overline{u}_1 = -\frac{2}{\sqrt{6}} \text{ y } \lambda_2 = \overline{v} \cdot \overline{u}_2 = \frac{4}{\sqrt{6}}.$$

Por tanto:

$$\overline{e} = \overline{v} - [\lambda_1\overline{u}_1 + \lambda_2\overline{u}_2] = (0, 0, 3, -1)\left[\left(-\frac{1}{3}, \frac{1}{3}, 0, -\frac{2}{3}\right) + \left(\frac{4}{3}, 0, \frac{2}{3}, -\frac{2}{3}\right)\right] =$$

$$= (0, 0, 3, -1) - \left(1, \frac{1}{3}, \frac{2}{3}, -\frac{4}{3}\right) = \left(-1, -\frac{1}{3}, \frac{7}{3}, \frac{1}{3}\right)$$

Es fácil verificar que el vector $\overline{e}$ es ortogonal a los vectores $\overline{u}_1$ y $\overline{u}_2$.

$\overline{u}_3 = \frac{\overline{e}}{\|\overline{e}\|}$ es normalizado, luego $S = \{\overline{u}_1, \overline{u}_2, \overline{u}_3\}$ es un sistema ortonormalizado.

Para poder completar la base tendríamos que obtener un cuarto vector de manera similar, ya que la dimensión de $\mathbb{R}^4$ es cuatro.

5.4.9. Subespacios ortogonales

Dos subespacios de un espacio euclídeo V son ortogonales $\Leftrightarrow$ Todo vector de cualquiera de ellos es ortogonal a todo vector del otro.

Dos subespacios serán ortogonales si todo vector de la base del primero es ortogonal a todo vector de la base del segundo, condición evidentemente necesaria. Que además es suficiente, se puede demostrar descomponiendo los vectores en función de ambas bases y utilizando los axiomas de producto escalar.

Como dos subespacios ortogonales sólo tienen en común el vector nulo, ya que en caso contrario habría un vector ortogonal a sí mismo, en contra del axioma 4 de producto escalar, y dado que dos vectores ortogonales son linealmente independientes, la suma de las dimensiones de los subespacios será n como máximo.

Sea un espacio vectorial euclídeo V, de dimensión n y U un subespacio del mismo. Vamos a llamar complementario ortogonal de U, que se denota como $U^\perp$, al subespacio ortogonal a U, tal que, la suma de las dimensiones de U y $U^\perp$ es n.

Si U tiene dimensión k existe una base ortonormal $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_k\}$. Pero hemos visto que es posible completar la base hasta obtener una base ortonormal de V $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_k, \overline{e}_{k+1}, ..., \overline{e}_n\}$, donde los vectores $\overline{e}_{k+1}, ..., \overline{e}_n$ forman una base de $U^\perp$, ya que resultan ortogonales a todo vector de U.

El conjunto $\{\overline{e}_{k+1}, ..., \overline{e}_n\}$ es una base ortonormal de $U^\perp$, pues en caso contrario existiría algún vector no nulo $\overline{e}_i \in U^\perp$, ortogonal a los vectores que forman dicha base, para lo que debería ser $\overline{e}_i \in U$, pero esto no es posible ya que $U \cap U^\perp = \{\overline{0}\}$ y el vector $\overline{0}$ no puede formar parte de una base.

Además, podemos expresar cualquier vector $\overline{v} \in V$ como $\overline{v} = \overline{u} + \overline{u}'$, donde $\overline{u} \in U$ y $\overline{u}' \in U^\perp$, luego V es suma directa de U y $U^\perp \Leftrightarrow U \oplus U^\perp = V$.

EJERCICIOS

Ejercicio 5.1.

Determínese si la siguiente matriz $G = \begin{pmatrix} 35 & 14 & -8 \\ 14 & 20 & 22 \\ -8 & 22 & 17 \end{pmatrix}$ define un producto escalar en $\mathbb{R}$, respecto de la base canónica:

Solución:

a) Se verifica que $g_{ij} = g_{ji}, \forall i \neq j$.
b) Se verifica que $g_{ij} > 0, \forall i = j$.
c) $g_{11} \cdot g_{22} - g_{12}^2 = 35.20 - 14^2 \geq 0$
$g_{11} \cdot g_{33} - g_{13}^2 = 35.17 - (-8)^2 \geq 0$
$g_{22} \cdot g_{33} - g_{23}^2 = 20.17 - 22^2 \geq 0.$

$\Rightarrow$ **La matriz G define un producto escalar en $\mathbb{R}$.**

Ejercicio 5.2.

Dado dos vectores pertenecientes a un espacio vectorial euclídeo, tales que, $\|\overline{u}_1\| = 1$, $\|\overline{u}_2\| = 3$ y $\cos(\overline{u}_1, \overline{u}_2) = \dfrac{1}{2}$, determínese la matriz que define el producto escalar en el subespacio generado por $\{\overline{u}_1, \overline{u}_2\}$.

Solución:

$$G = \begin{pmatrix} \overline{u}_1 \cdot \overline{u}_1 & \overline{u}_1 \cdot \overline{u}_2 \\ \overline{u}_2 \cdot \overline{u}_1 & \overline{u}_2 \cdot \overline{u}_2 \end{pmatrix}$$

$$\|\overline{u}_1\| = \sqrt{\overline{u}_1 \cdot \overline{u}_1} \Rightarrow \overline{u}_1 \cdot \overline{u}_1 = 1$$

$$\|\overline{u}_2\| = \sqrt{\overline{u}_2 \cdot \overline{u}_2} \Rightarrow \overline{u}_2 \cdot \overline{u}_2 = 9$$

$$\overline{u}_1 \cdot \overline{u}_2 = \|\overline{u}_1\|\|\overline{u}_2\|\cos(\overline{u}_1, \overline{u}_2) = 1 . 3 . \frac{1}{2} = \frac{3}{2}$$

Luego, la matriz pedida es:

$$G = \begin{pmatrix} 1 & 3/2 \\ 3/2 & 9 \end{pmatrix}.$$

Ejercicio 5.3.

Hállese la norma del vector $(\overline{u}_1 + \overline{u}_2)$, siendo $\overline{u}_1$ y $\overline{u}_2$ los vectores del ejercicio anterior.

Solución:

$$\|\overline{u}_1 + \overline{u}_2\| = \sqrt{(\overline{u}_1 + \overline{u}_2) \cdot (\overline{u}_1 + \overline{u}_2)} = \sqrt{(\overline{u}_1 \cdot \overline{u}_1) + 2(\overline{u}_1 \cdot \overline{u}_2) + (\overline{u}_2 \cdot \overline{u}_2)} =$$

$$= \sqrt{1 + 2 . \frac{3}{2} + 9} = \sqrt{13}$$

Ejercicio 5.4.

Compruébese que si V es un espacio vectorial euclídeo (con el producto escalar usual), $\{\overline{e}_1, \overline{e}_2, ..., \overline{e}_n\}$ una de sus bases, y $\lambda_1, \lambda_2, ..., \lambda_n$ números reales, existe un vector único $\overline{x} \in V$, tal que: $\overline{x} \cdot \overline{e}_k = \lambda_k$, para $k = 1, 2, ..., n$.

Solución:

En efecto, $\overline{x} = x_1\overline{e}_1 + x_2\overline{e}_2 + ... + x_n\overline{e}_n$

Entonces:

$$\begin{aligned}
\bar{x}\cdot\bar{e}_1 &= x_1(\bar{e}_1\cdot\bar{e}_1)+x_2(\bar{e}_2\,.\,\bar{e}_1)+\ldots+x_n(\bar{e}_n\cdot\bar{e}_1) = \lambda_1\\
\bar{x}\cdot\bar{e}_2 &= x_1(\bar{e}_1\cdot\bar{e}_2)+x_2(\bar{e}_2\,.\,\bar{e}_2)+\ldots+x_n(\bar{e}_n\cdot\bar{e}_2) = \lambda_2\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\bar{x}\cdot\bar{e}_n &= x_1(\bar{e}_1\cdot\bar{e}_n)+x_2(\bar{e}_2\,.\,\bar{e}_n)+\ldots+x_n(\bar{e}_n\cdot\bar{e}_n) = \lambda_n
\end{aligned}$$

que es un sistema lineal con solución única $(x_1, x_2, \ldots, x_n)$, ya que la matriz métrica del producto escalar es regular.

Ejercicio 5.5.

Determínese un vector $\bar{v}\in\mathbb{R}^4$ de norma uno (con el producto escalar usual), ortogonal a los vectores $\bar{u}_1 = (1, 2, 3, 4)$, $\bar{u}_2 = (4, 3, 2, 1)$, $\bar{u}_3 = (1, 0, 1, 0)$, referido a la base canónica ortonormal.

Solución:

El vector $\bar{v} = (x, y, z, t)$ deberá verificar que sea:

$$\|\bar{v}\| = 1$$
$$\bar{v}\cdot\bar{u}_1 = \bar{v}\cdot\bar{u}_2 = \bar{v}\cdot\bar{u}_3 = 0$$

Lo que es equivalente a:

1. $x^2+y^2+z^2+t^2 = 1$
2. $x+2y+3z+4t = 0$
3. $4x+3y+2z+t = 0$
4. $x+\qquad+z\qquad = 0$

De la ecuación 4 se obtiene: $z = -x$

Si se reemplaza en las ecuaciones 2 y 3 y luego se suman ambas ecuaciones se obtiene: $y = -t$.

Además resulta $y = -x$, entonces $t = x$.

Reemplazando en la ecuación 1 se obtiene: $x^2+(-x)^2+(-x)^2+x^2 = 1 \Rightarrow$ $\Rightarrow x = \pm\sqrt{1/4}$.

Por tanto, la solución del problema son los vectores:

$$\bar{v}_1 = (1/2, -1/2, -1/2, 1/2)$$
$$\bar{v}_2 = (-1/2, 1/2, 1/2, -1/2)$$

Ejercicio 5.6.

Dados dos vectores, $\bar{u}$ y $\bar{v}$, pertenecientes a un espacio euclídeo, encuéntrese un tercer vector $\bar{w}$ ortogonal a $\bar{v}$, y tal que, $\bar{u}$ se descomponga en la suma de $\bar{w}$ con un vector de la dirección de $\bar{v}$.

Solución:

Se tiene que cumplir: $\bar{u} = \lambda\bar{v}+\bar{w} \Rightarrow \bar{w} = \bar{u}-\lambda\bar{v}$.

Pero como además debe ser:

$$\bar{v}\cdot\bar{w} = 0$$
$$\bar{v}\cdot\bar{w} = \bar{v}\cdot(\bar{u}-\lambda\bar{v}) = \bar{v}\cdot\bar{u}-\lambda\|\bar{v}\|^2 = 0$$

Por tanto, resulta:

$$\lambda = \frac{\bar{u}\,.\,\bar{v}}{\|\bar{v}\|^2} \Rightarrow \bar{w} = \bar{u}-\frac{\bar{u}\,.\,\bar{v}}{\|\bar{v}\|^2}\bar{v}$$

Ejercicio 5.7.

Con el producto escalar habitual en $\mathbb{R}^4$, hállese la base $\{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n\}$ aplicando el método de Gram-Schmidt a la familia de vectores $\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$, sabiendo que cada vector $\bar{e}_k$ tiene los primeros k elementos iguales a uno y los demás iguales a cero:

$\bar{e}_1 = (1, 0, 0, ..., 0), \bar{e}_2 = (1, 1, 0, ..., 0), ..., \bar{e}_n = (1, 1, 1, ..., 1).$

Solución:

$\bar{u}_1 = \bar{e}_1 = (1, 0, 0, ..., 0)$

$$\bar{u}_2 = \bar{e}_2 - \frac{\bar{u}_1 \cdot \bar{e}_2}{\|\bar{u}_1\|^2}\bar{u}_1 = (1, 1, 0, ..., 0) - \frac{(1, 0, 0, ..., 0)}{1} = (0, 1, 0, ..., 0)$$

$$\bar{u}_3 = \bar{e}_3 - \frac{\bar{u}_1 \cdot \bar{e}_3}{\|\bar{u}_1\|^2}\bar{u}_1 - \frac{\bar{u}_2 \cdot \bar{e}_3}{\|\bar{u}_2\|^2}\bar{u}_2 = (0, 0, 1, 0, ..., 0)$$

....

$\bar{u}_n = (0, 0, 0, ..., 1)$

Además, los vectores obtenidos son unitarios.

Ejercicio 5.8

Dados dos vectores de un espacio vectorial euclídeo, $\bar{u}$, $\bar{v}$, demuéstrese que: $(\bar{u} + \bar{v}) \cdot (\bar{u} - \bar{v}) = 0$ si y sólo si $\|\bar{u}\| = \|\bar{v}\|$.

Solución:

$$\Rightarrow (\bar{u} + v) \cdot (\bar{u} - \bar{v}) = (\bar{u} \cdot \bar{u}) - (\bar{u} \cdot \bar{v}) + (\bar{v} \cdot \bar{u}) - (\bar{v} \cdot \bar{v}) = \|\bar{u}\|^2 - \|\bar{v}\|^2$$

Entonces:

$$(\bar{u} + \bar{v}) \cdot (\bar{u} - \bar{v}) = \|\bar{u}\|^2 - \|\bar{v}\|^2 = 0 \Rightarrow \|\bar{u}\|^2 = \|\bar{v}\|^2 \Rightarrow \|\bar{u}\| = \|\bar{v}\|$$

$\Leftarrow$ El recíproco también es cierto:

$$\|\bar{u}\| = \|\bar{v}\| \Rightarrow \|\bar{u}\|^2 = \|\bar{v}\|^2 \Rightarrow (\bar{u} + \bar{v}) \cdot (\bar{u} - \bar{v}) = 0$$

Ejercicio 5.9

En el espacio vectorial de los polinomios de grado menor o igual a dos se define el producto escalar como:

$p(x) \cdot q(x) = p(0)q(0) + \int_0^1 p'(x)q'(x)dx$, siendo $p'(x)$ y $q'(x)$ las derivadas de $p(x)$ y $q(x)$ respectivamente.

Determínese el subespacio ortogonal a $V = \{1 - x, 1 - x^2\}$.

Solución:

Sean: $r(x) = ax^2 + bx + c$; $p(x) = 1 - x$; $q(x) = 1 - x^2$, entonces:

$$p(x) \cdot r(x) = c + \int_0^1 -(2ax + b)dx = c - [ax^2 + bx]_0^1 = c - a - b$$

$$q(x) \cdot r(x) = c + \int_0^1 -2x(2ax + b)dx = c - \left[\frac{4ax^3}{3} + bx^2\right]_0^1 = c - \frac{4a}{3} - b$$

Pero debe ser:

$$p(x) \cdot r(x) = 0 \Rightarrow c - a - b = 0$$

$$q(x) \cdot r(x) = 0 \Rightarrow c - \frac{4a}{3} - b = 0$$

De donde se obtiene que: $a = 0$, $b = c \Rightarrow V^\perp = \langle 1 + x \rangle$.

Ejercicio 5.10

Con el producto escalar usual, se quiere encontrar el complemento ortogonal en $\mathbb{R}^3$ (una de cuyas bases es $B = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\}$) del subespacio $V = \{(x_1, x_2, x_3)/x_1 = x_2\}$ de un espacio euclídeo a partir de la base formada por los vectores $\bar{u}_1 = (1, 1, 0)$, $\bar{u}_2 = (0, 0, 1)$.

Solución:

Un elemento $\overline{y}$ pertenece a $V^{\perp}$ (complemento ortogonal de V), si y sólo si $\overline{y}$ es perpendicular a $\overline{u}_1$ y a $\overline{u}_2$.

Como $\overline{y} = y_1\overline{e}_1 + y_2\overline{e}_2 + y_3\overline{e}_3$, debe ser:

$$\begin{cases} \overline{y} \cdot \overline{u}_1 = (y_1\overline{e}_1 + y_2\overline{e}_2 + y_3\overline{e}_3) \cdot (\overline{e}_1 + \overline{e}_2) = y_1 + y_2 = 0 \\ \overline{y} \cdot \overline{u}_2 = (y_1\overline{e}_1 + y_2\overline{e}_2 + y_3\overline{e}_3) \cdot \overline{e}_3 = y_3 = 0 \end{cases}$$

obteniéndose un sistema homogéneo con tres incógnitas que determina un subespacio cuyas ecuaciones son:

$$\begin{cases} y_1 = -y_2 \\ y_3 = 0 \end{cases} \Leftrightarrow \begin{cases} y_1 = -\lambda \\ y_2 = \lambda \\ y_3 = 0 \end{cases} \Leftrightarrow \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \lambda \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right.$$

por tanto, está generado por el vector: $\overline{v} = (-1, 1, 0)$ y es de dimensión 1.

CAPÍTULO 6

MATRICES SEMEJANTES

INTRODUCCIÓN

La noción de valor propio y de vector propio de una matriz (o de la aplicación lineal que representa) había surgido en el siglo XVIII en la teoría de sistemas de ecuaciones diferenciales lineales homogéneas con coeficientes constantes, de la mano de Lagrange y Laplace, y su identificación con los ejes de una cuádrica cuando A es real y simétrica era bien conocida por Cauchy, quien también conocía su coincidencia con las raíces del polinomio característico. Cauchy obtiene los primeros invariantes, probando, por ejemplo, que si A y B son matrices semejantes, tienen los mismos valores propios.

Sylvester, en sus trabajos sobre los haces de cuádricas, y poco después, en 1868, Weierstrass que extendió los resultados de Sylvester para matrices no necesariamente simétricas, dieron un importante impulso a la teoría de los invariantes.

Utilizando la noción de matrices semejantes, Jordan en su monumental *Traité des substitutions* (1870) demostró que toda matriz cuadrada era semejante a otra más sencilla.

El papel sistematizador en la teoría de clasificación y reducción de matrices lo lleva a cabo Frobenius en varias memorias publicadas entre 1877 y 1880. En particular, da la demostración general del teorema de Cayley-Hamilton y plantea el problema de determinar el polinomio de menor grado que satisface la matriz (el polinomio mínimo). Afirma que está formado por factores del polinomio característico, y que es único, lo que no fue demostrado hasta 1904 por Hensel.

Señalemos finalmente que Metzler introdujo en 1892 la exponencial y otras funciones transcendentes de matrices, por medio de series de potencias, que son de gran aplicación en el estudio de sistemas de ecuaciones diferenciales lineales. Para su cálculo efectivo tienen gran importancia la diagonalización de matrices y la descomposición de Jordan, ya que las potencias sucesivas de dichas matrices se calculan muy fácilmente.

ARTHUR CAYLEY (1821-1895)

Arthur Cayley nació en Inglaterra de padre inglés y madre rusa y desde su adolescencia mostró su facilidad para las matemáticas.

Estudió en el King's College, en el Trinity College de Dublín, donde ingresó con la calificación "por encima del primera", y en la Universidad de Cambridge.

Junto con su coetáneo Hamilton, encabezó la prestigiosa escuela de matemáticos ingleses del siglo XIX.

Después de realizar serias investigaciones en matemáticas, se olvidó de ella y se dedicó al turismo por Europa.

En 1846 comenzó a estudiar Derecho, mientras ejercía la abogacía entabló amistad con Sylvester, y las matemáticas salieron beneficiadas de esta relación. Simultáneamente declinaba el interés de ambos por la jurisprudencia.

Por entonces se creó una cátedra de matemáticas en Cambridge, de ella se encargó Cayley y a partir de ese momento no abandonó su interés por esa disciplina.

Sus contribuciones más importantes a las matemáticas son:

- Importantes descubrimientos en la teoría de los determinantes.
- Creó la teoría de los invariantes.
- Es autor de la primera definición abstracta de grupo finito.
- Desarrolló el álgebra de matrices, que sirvió como fundamento para la mecánica cuántica.
- Hizo importantes contribuciones a la teoría de curvas y superficies en geometría analítica.

Sus trabajos en geometría cuatridimensional, proporcionaron a los físicos del siglo XX, especialmente a Albert Einstein, la estructura para desarrollar la teoría de la relatividad.

Su desarrollo de la geometría no dimensional ha sido aplicada en la física al estudio del espacio-tiempo continuo.

Está considerado como el tercer escritor más prolífico de matemáticas, siendo sólo superado por Euler y Cauchy. Sus trabajos matemáticos se publicaron en Cambridge (1889-98).

CAPÍTULO 6

La clase de equivalencia de matrices semejantes

¿Es diagonal alguna matriz de la clase?

Autovalores y Autovectores

¿Cuándo se puede diagonalizar una matriz?

Diagonalización de endomorfismos

Endomorfismos simétricos

Diagonalización de endomorfismos simétricos

6.1. Matrices semejantes

En este capítulo vamos a considerar el estudio de las matrices cuadradas como expresión analítica de los endomorfismos entre espacios vectoriales.

Sabemos que un endomorfismo tiene diferentes expresiones matriciales según consideremos una u otra base en el espacio vectorial dado.

Como veremos más adelante, todas las matrices que representan el mismo endomorfismo pertenecen a la misma clase de equivalencia.

R es una relación de equivalencia en un conjunto X, si es una relación definida entre sus elementos, que verifica las propiedades:

1. Reflexiva: $aRa, \forall\, a \in X$

2. Simétrica: $\left\{\begin{array}{l} aRb \\ bRa \end{array}\right. \Rightarrow a = b$ siendo $a, b \in X$

3. Transitiva: $\left\{\begin{array}{l} aRb \\ bRc \end{array}\right. \Rightarrow aRc$ siendo $a, b, c \in X$

Todos los elementos que están relacionados entre sí mediante una relación de equivalencia forman una clase de equivalencia.

En el conjunto de las matrices cuadradas podemos definir diferentes relaciones de equivalencia, pero dos de ellas serán las que fijarán nuestra atención en este capítulo: la relación de semejanza y la relación de congruencia.

La relación de semejanza en el conjunto de las matrices cuadradas, nos permitirá preguntar, si dada una de ellas, existirá en su clase de equivalencia una matriz diagonal semejante a la dada, y establecer las condiciones necesarias y suficientes para que esto sea posible. La búsqueda de matrices diagonales y de las condiciones de su existencia es el objetivo de este capítulo por la importancia que tienen, tanto en el estudio de disciplinas matemáticas más especializadas, valga el ejemplo de las ecuaciones diferenciales, como por los problemas que resuelven en diferentes ramas de la ingeniería.

Planteareremos el mismo problema para el conjunto de las matrices simétricas de elementos reales, y mediante la relación de congruencia, nos llevará a hacernos idéntica pregunta a la anterior y a establecer los teoremas necesarios para probar la posible existencia de una matriz diagonal en la clase correspondiente a una matriz de aquel conjunto.

6.1.1. Definición de matrices equivalentes

$A, B \in M_{m\times n}$ son matrices equivalentes $\Leftrightarrow$ Existen dos matrices regulares M, N, tales que, $A = MBN$

Es evidente que, la relación entre matrices así definida en $M_{m\times n}$ constituye una relación de equivalencia. La demostración se deja al lector como ejercicio.

Atendiendo a esta definición de matrices equivalentes, se verifica la siguiente propiedad.

6.1.2. Proposición:

Si $A, B \in M_{m\times n}$ con $A = CB$ y C es una matriz regular, entonces, A y B tienen el mismo rango.

Demostración:

Al hacer la multiplicación de matrices $A = CB$:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mm} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{cases} a_{11} = c_{11}b_{11} + c_{12}b_{21} + \dots + c_{1m}b_{m1} \\ a_{12} = c_{11}b_{12} + c_{12}b_{22} + \dots + c_{1m}b_{m2} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{1n} = c_{11}b_{1n} + c_{12}b_{2n} + \dots + c_{1m}b_{mn} \end{cases}$$

Expresión que podemos escribir como:

$$\begin{pmatrix} a_{11} \\ a_{12} \\ \dots \\ a_{1n} \end{pmatrix} = c_{11}\begin{pmatrix} b_{11} \\ b_{12} \\ \dots \\ b_{1n} \end{pmatrix} + c_{12}\begin{pmatrix} b_{21} \\ b_{22} \\ \dots \\ b_{2n} \end{pmatrix} + \dots + c_{1m}\begin{pmatrix} b_{m1} \\ b_{m2} \\ \dots \\ b_{mn} \end{pmatrix}$$

Las filas de la matriz A son combinaciones lineales de las filas de la matriz B, por tanto, el rango de A es menor o igual que el rango de B.

Recíprocamente, como C es regular existe C^{-1} y es regular, por tanto, $B = C^{-1}A$, razonando de la misma forma, podemos afirmar que el rango de B será menor o igual que el rango de A.

Por lo que concluimos que el rango de A es igual que el rango de B.

6.1.3. Proposición

Los rangos de las matrices A y B coinciden $\Leftrightarrow$ A y B son matrices equivalentes en $M_{m\times n}$

Demostración:

Si A y B son equivalentes, entonces $A = MBN$, donde M y N son regulares, y por la proposición *[6.1.2]*, $\mathrm{rg}(A) = \mathrm{rg}(MBN) = \mathrm{rg}(BN) = \mathrm{rg}(B)$, que es lo que queríamos demostrar.

6.1.4. Definición de matrices congruentes

$A, B \in M_{n\times n}$ son matrices congruentes $\Leftrightarrow$ Existe una matriz regular $P \in M_{n\times n}$, tal que, $A = P^tBP$

De esta definición se deducen las siguientes consecuencias:

6.1.5. Consecuencias

1. *Si A y B son congruentes, entonces A y B son equivalentes.*

Demostración:

Evidentemente es trivial, sin más que considerar $M = P^t$ y $N = P$.

2. *Si A y B son congruentes, entonces* $\mathrm{rg}(A) = \mathrm{rg}(B)$.

Demostración:

Sale inmediatamente de *[6.1.3]* y de la consecuencia anterior.

3. *La congruencia de matrices es una relación de equivalencia en* $M_{n\times n}$.

Demostración:

En efecto, se cumplen las tres propiedades que caracterizan a las relaciones de equivalencia:

1. $A = I^t AI$ *(Reflexiva)*
2. $A = P^t BP$, $B = (P^t)^{-1} AP^{-1} \Leftrightarrow B = (P^{-1})^t AP^{-1}$, pues como sabemos, $(P^t)^{-1} = (P^{-1})^t$. *(Simétrica).*

3. $A = P' BP$ y $B = Q'CQ \Rightarrow A = P'Q'CQP = (QP)' CQP$, en virtud de que $(QP)' = P'Q'$. *(Transitiva).*

6.1.6. Definición de matrices semejantes

> $A, B \in M_{n\times n}$ son matrices semejantes $\Leftrightarrow$ Existe una matriz regular $P \in M_{n\times n}$, tal que, $A = P^{-1}BP$, a la matriz P, se le llama matriz de paso.

De esta definición se deducen las siguientes consecuencias:

6.1.7. Consecuencias

1. Si **A** *y* **B** *son semejantes, entonces* **A** *y* **B** *son equivalentes.*

Demostración:

Es trivial, pues basta hacer $M = P^{-1}$ y $N = P$ para tener la definición de equivalencia de matrices.

2. Si **A** y **B** *son semejantes, entonces* $\mathrm{rg}(A) = \mathrm{rg}(B)$.

Demostración:

Es consecuencia inmediata de *[6.1.3.]* y de la consecuencia anterior.

3. La semejanza de matrices es una relación de equivalencia.

Demostración:

La dejamos como ejercicio para el lector.

4. Si $A = P^{-1}BP$, ***entonces*** $A^h = P^{-1}B^hP$.

Demostración:

Vamos a hacer esta demostración por inducción.

Si $A = P^{-1}BP$, se verifica que $A^2 = AA = (P^{-1}BP)(P^{-1}BP) = P^{-1}B^2P$, por la propiedad asociativa.

Supongamos que se verifica para el valor h–1: $A^{h-1} = P^{-1}B^{h-1}P$.

Demostremos que se verifica también para el valor h. En efecto, postmultiplicando por A en la anterior igualdad resulta: $A^h = (P^{-1}B^{h-1}P)(P^{-1}BP) = P^{-1}B^hP$, que es lo que queríamos demostrar.

De todo lo anterior se concluye que las matrices semejantes a una matriz dada de orden n constituyen una clase de equivalencia en el conjunto de las matrices cuadradas $M_{n\times n}$, y que dichas matrices conforman las diversas representaciones analíticas de un mismo endomorfismo f de un espacio vectorial V de dimensión n (cada una de las cuales corresponde a la base elegida en V).

Todo ello nos lleva a plantear el problema de buscar una base de V, en la cual f se represente en la forma más sencilla posible. Dada la simplicidad de la matriz diagonal, intentaremos ver si es posible encontrar una base de V en la que f esté representado por una matriz diagonal semejante a la matriz dada.

Iguales consideraciones haremos con las matrices simétricas de elementos reales, buscando una matriz diagonal que sea congruente con la matriz simétrica dada.

6.1.8. Definición de matriz diagonalizable

> $A \in M_{n\times n}$ es diagonalizable $\Leftrightarrow$ A es semejante a una matriz diagonal $\Lambda \in M_{n\times n}$.

La definición dada se puede expresar también diciendo:

Una matriz A es diagonalizable si existe una matriz regular P de paso, tal que $A = P^{-1}\Lambda P$, siendo Λ una matriz diagonal.

6.1.9. Diagonalización

Diagonalización de $A \Leftrightarrow$ Proceso de búsqueda de las matrices Λ y P dada una matriz A, tales que $A = P^{-1}\Lambda P$.

Como conclusión, diremos que la tarea que nos planteamos es buscar condiciones necesarias y suficientes para que una matriz sea diagonalizable.

6.2. Valores propios y vectores propios

Al principio de este texto, indicamos que nos moveríamos en el cuerpo de los números reales $\mathbb{R}$ siempre que fuera posible, pero para el desarrollo de alguna parte de los capítulos seis y siete es muy útil, especialmente para hacer algunas demostraciones, considerar el cuerpo de los complejos $\mathbb{C}$. Cuando utilicemos $\mathbb{C}$, lo mencionaremos expresamente.

Queremos hacer notar que aunque para las demostraciones se utilice $\mathbb{C}$, los resultados son válidos en $\mathbb{R}$, ya que $\mathbb{R} \subset \mathbb{C}$.

Para alcanzar el objetivo que indicamos en *[6.1]* utilizaremos unos vectores con características especiales a los que llamaremos vectores propios o autovectores de un endomorfismo.

Dado un espacio vectorial V de dimensión n definido sobre el cuerpo $\mathbb{R}$ de los números reales, y una aplicación lineal f de V en sí mismo, se dice que $\overline{x} \neq \overline{0}$ es un vector propio de f cuando se verifica que su imagen por f es un escalar por dicho vector.

6.2.1. Definición de vector propio de un endomorfismo de V

Un vector $\overline{x} \neq \overline{0}$, $\overline{x} \in V$ es un vector propio del endomorfismo f de $V \Leftrightarrow f(\overline{x}) = \lambda\overline{x}$ para algún $\lambda \in \mathbb{R}$

Ejemplo 6.2.1

Sea $V = \{$funciones reales que admiten derivadas de todos los órdenes$\}$ y f el endomorfismo de V, que asigna a cada función de V su derivada.

e^x es un vector propio de f; 1 es valor propio del vector e^x porque $f(e^x) = (e^x)' = 1e^x$.

Sin embargo, sen $x \in V$, no es un vector propio de f, porque f (sen x) = = cos x, y no hay ningún $\lambda \in \mathbb{R}$, tal que, cos $x = \lambda$sen x.

- Obsérvese que así definidos V es un espacio vectorial y f es una aplicación lineal de V en V.

6.2.2. Consecuencia

Si $\exists\lambda$ correspondiente a un vector propio $\overline{x} \Rightarrow \lambda$ es único.

Demostración:

Utilizaremos el método de reducción al absurdo.

Supongamos que hubiera dos valores que lo cumplen: λ, λ', entonces, $f(\overline{x}) = \lambda\overline{x}$ y $f(\overline{x}) = \lambda'\overline{x}$, y por tanto, $\lambda\overline{x} = \lambda'\overline{x}$, $(\lambda - \lambda')\overline{x} = \overline{0} \Rightarrow \lambda = \lambda'$, es decir, ambos valores coincidirían.

A este escalar λ asociado de modo único al vector propio $\overline{x}$, se le llama valor propio asociado.

6.2.3. Definición de valor propio asociado a un vector propio

Valor propio asociado al vector propio $\overline{x} \Leftrightarrow$ Valor de λ que verifica $f(\overline{x}) = \lambda\overline{x}$.

Sea $A = (a_{ij})$ la matriz asociada al endomorfismo f en una cierta base B del espacio vectorial V. Consideremos la matriz unidad I cuadrada de orden n, e i la aplicación identidad en V. Llamaremos X, a la matriz columna cuyas filas son las componentes del vector $\overline{x}$ en la base B dada: $X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$

Podemos escribir: $f(\overline{x}) = \lambda i(\overline{x})$ siendo $\overline{x}$ un vector propio de V, o bien, en forma matricial $AX = \lambda IX$, y por tanto $(A - \lambda I)X = \overline{0}$. Esta ecuación representa un sistema de ecuaciones lineales homogéneas con n incógnitas, que admite solución distinta de la trivial si: $|A - \lambda I| = 0$.

El desarrollo de esta ecuación tiene la forma:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0$$

6.2.4. Definición de ecuación característica de una matriz

Ecuación característica de la matriz $A \Leftrightarrow$ Ecuación $|A - \lambda I| = 0$

Dicha ecuación admite n raíces a las cuales llamaremos raíces características, que son los valores propios del endomorfismo f.

El primer miembro de la igualdad representa un polinomio en λ de grado n al que se denomina polinomio característico de la matriz.

En el ejemplo siguiente seguiremos, paso a paso, para un caso concreto los contenidos anteriores.

Ejemplo 6.2.2.

Consideremos la matriz $A = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}$, si queremos calcular los valores propios y los vectores propios correspondientes, debemos proceder del siguiente modo:

- La ecuación característica es: $|A-\lambda I| = \begin{vmatrix} 3-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 3-\lambda \end{vmatrix} =$

$= (3-\lambda)(\lambda^2 - 5\lambda + 4) = 0$

- Las soluciones de la ecuación son los valores propios: $\lambda_1 = 1$, $\lambda_2 = 3$ y $\lambda_3 = 4$.

Veamos cómo calculamos los vectores propios correspondientes a cada uno de los valores propios anteriores.

En la ecuación matricial $(A-\lambda I)X = \overline{0}$, y considerando la matriz columna:

$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, operaremos de la siguiente manera:

$x_2 = 2x_1$

- Para $\lambda_1 = 1$, se verifica:

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left.\begin{array}{r} 2x_1 - x_2 = 0 \\ -x_1 + x_2 - x_3 = 0 \\ -x_2 + 2x_3 = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{r} 2x_1 - x_2 = 0 \\ -x_2 + 2x_3 = 0 \end{array}\right\} \Rightarrow$$

$$\Rightarrow \begin{cases} x_1 = \delta \\ x_2 = 2\delta, \ \forall\, \delta \in \mathbb{R} \\ x_3 = \delta \end{cases}$$

Hay muchas formas de expresar el subespacio propio asociado a $\lambda_1 = 1$.

Podemos hacerlo mediante las ecuaciones cartesianas $\left.\begin{array}{r} 2x_1 - x_2 = 0 \\ -x_2 + 2x_3 = 0 \end{array}\right\}$,

o las paramétricas $\begin{cases} x_1 = \delta \\ x_2 = 2\delta, \\ x_3 = \delta \end{cases}$ o el subespacio generado por (1, 2, 1), ya

que, $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \delta \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, o bien $L_1 = \{(\delta, 2\delta, \delta), \forall \delta \in \mathbb{R}\}$.

- Para $\lambda_2 = 3$, se verifica:

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left.\begin{array}{r} x_2 = 0 \\ x_1 + x_2 + x_3 = 0 \end{array}\right\} \Rightarrow \begin{cases} x_1 = \beta \\ x_2 = 0 \\ x_3 = -\beta \end{cases} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$= \beta \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ y el conjunto de los vectores propios correspondientes a este valor propio lo podemos escribir como $L_2 = \{(\beta, 0, -\beta), \forall \beta \in \mathbb{R}\}$.

- Para $\lambda_3 = 4$, se verifica:

$$\begin{pmatrix} -1 & -1 & 0 \\ -1 & -2 & -1 \\ 0 & -1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left.\begin{array}{r} x_1 + x_2 = 0 \\ x_1 + 2x_2 + x_3 = 0 \\ x_2 + x_3 = 0 \end{array}\right\} \Rightarrow \begin{cases} x_1 = \alpha \\ x_2 = -\alpha \\ x_3 = \alpha \end{cases} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$= \alpha \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, y su conjunto de vectores propios es $L_3 = \{(\alpha, -\alpha, \alpha), \forall \alpha \in \mathbb{R}\}$.

Ejemplo 6.2.3

La ecuación característica de la matriz $A = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$ es $|A - \lambda I| =$

$= -(\lambda - 2)^3 = 0$, y sólo tiene el valor propio $\lambda = 2$, con multiplicidad algebraica tres.

De las definiciones anteriores se deducen las siguientes proposiciones:

6.2.5. Proposición

El conjunto de los vectores propios asociados al mismo valor propio λ_i, constituye un subespacio vectorial de V, que escribiremos $L_i = \{\overline{x} \in V \,/\, A\overline{x} = \lambda_i \overline{x}\}$

Demostración:

Si trasladamos al caso general el procedimiento de obtención de vectores asociados a un valor propio utilizado en el ejemplo 6.2.2, observamos que, al sustituir cada λ_i en la ecuación matricial $(A - \lambda I)X = \overline{0}$, se obtiene un sistema lineal homogéneo, cuyas soluciones son el conjunto de vectores propios asociados a λ_i, que forman un subespacio vectorial, ya que es el núcleo de la aplicación lineal cuya matriz asociada es $A - \lambda I$.

También es muy fácil hacer la demostración utilizando la caracterización de los subespacios dada en *[1.3.2.]* o *[1.3.3]*.

6.2.6. Proposición

El polinomio característico de la matriz A asociada a un endomorfismo f, es invariante respecto a un cambio de base.

Demostración:

Si llamamos a A' la matriz asociada a f cuando estamos utilizando la base B' en V, el siguiente diagrama conmutativo muestra que: $f_{A'} = i_P^{-1} \circ f_A \circ i_P$.

$$\begin{tikzcd}
(V,B) \arrow[r, "f_A"] \arrow[d, shift right, "i_P^{-1}"'] & (V,B) \arrow[d, shift right, "i_P^{-1}"'] \\
(V,B') \arrow[r, "f_{A'}"] \arrow[u, shift right, "i_P"'] & (V,B') \arrow[u, shift right, "i_P"']
\end{tikzcd}$$

La matriz $A' = P^{-1}AP$ es la matriz transformada de la matriz A al cambiar la base B de V por la base B'.

Su polinomio característico es $|A' - \lambda I|$, que también se puede escribir como: $|P^{-1}AP - \lambda I|$, y como λ es un escalar, se verifica: $\lambda I = P^{-1}\lambda IP$, expresión que llevada a la anterior da lugar a la siguiente:

$$|P^{-1}AP - \lambda I| = |P^{-1}(A - \lambda I)P| = |P^{-1}|\,|A - \lambda I|\,|P| = |P^{-1}P|\,|A - \lambda I| =$$
$$= |A - \lambda I| \Rightarrow |A - \lambda I| = |A' - \lambda I|.$$

Obsérvese que *[6.2.6]* se podría haber enunciado de la siguiente manera: ***Si A y A' son matrices semejantes, entonces tienen la misma ecuación característica.***

No obstante, puede ocurrir que dos matrices tengan la misma ecuación característica y no sean semejantes.

Las matrices: $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ tienen la misma ecuación característica:

$$|A - \lambda I| = (1-\lambda)^3 = 0; \qquad |I - \lambda I| = (1-\lambda)^3 = 0.$$

A e I no son semejantes, pues si así fuera, sería $A = P^{-1}IP \Rightarrow A = I$, que es evidentemente falso.

6.2.7. Proposición

El conjunto de k vectores propios correspondientes a k valores propios distintos de un endomorfismo f, forman un sistema libre.

Demostración:

Utilizaremos el método de reducción al absurdo.

Supongamos que los vectores $\overline{x}_1, \overline{x}_2, ..., \overline{x}_k$ fueran un sistema ligado, sin pérdida de generalidad podemos suponer que los vectores linealmente independientes son los r primeros $\overline{x}_1, \overline{x}_2, ..., \overline{x}_r$, con $(r < k)$. Entonces cualquier $\overline{x}_h$ con $h \geq r$ perteneciente al sistema ligado, se podría expresar de forma única respecto a los que son libres $\Rightarrow \overline{x}_h = \mu_1\overline{x}_1 + \mu_2\overline{x}_2 + ... + \mu_r\overline{x}_r$.

Por ser f una aplicación lineal resultará: $f(\overline{x}_h) = \mu_1 f(\overline{x}_1) + \mu_2 f(\overline{x}_2) + \ldots + \mu_r f(\overline{x}_r) \Rightarrow \lambda_h \overline{x}_h = \mu_1 \lambda_1 \overline{x}_1 + \mu_2 \lambda_2 \overline{x}_2 + \ldots + \mu_r \lambda_r \overline{x}_r$, donde $\lambda_1, \lambda_2, \ldots, \lambda_r, \lambda_h$ son los valores propios asociados a $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_r, \overline{x}_h$. Además $\lambda_h \neq 0$, pues de lo contrario, el sistema $\{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_r\}$ sería ligado.

Pero esto nos lleva a que el vector $\overline{x}_h$ tiene dos expresiones diferentes en el sistema indicado, ya que, todos los valores propios son distintos, lo que es una contradicción.

Por tanto, la hipótesis "si los vectores $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k$ fueran un sistema ligado", no puede ser cierta.

Consecuencia inmediata de *[6.2.7]* es que ***"si la matriz A tiene n valores propios distintos, el conjunto de vectores propios asociados a los n valores propios constituye una base del espacio vectorial V"***.

6.2.8. Proposición

Si λ_i es una raíz característica de la matriz A con multiplicidad algebraica α_i, la dimensión d_i del subespacio de vectores propios asociados a λ_i, verifica la condición: $1 \leq d_i \leq \alpha_i$

La primera parte de la proposición *[6.2.8]*: $1 \leq d_i$ es inmediata, ya que, todo valor propio λ_i, tiene asociado algún vector propio.

Veamos que $d_i \leq \alpha_i$:

Consideremos el valor propio λ_i, y hallemos los valores propios de $M = A - \lambda_i I$ que se obtendrán de $|M - \mu I| = 0$, a lo que es lo mismo, $|A - (\lambda_i + \mu)I| = 0$ y si llamamos $\lambda_i + \mu = \lambda$ quedará $|A - \lambda I| = 0$, y cada raíz λ_j con orden de multiplicidad α_j le corresponde $\mu_j = \lambda_j - \lambda_i$ de $|M - \mu I| = 0$ con el mismo orden de multiplicidad. A la raíz λ_i le corresponde $\mu_i = 0$ con multiplicidad de orden α_i.

Como d_i es la dimensión del subespacio asociado al valor propio λ_i de A existen d_i vectores $\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_{d_i}$ que constituye una base de este subespacio que podrá completarse con $n - d_i$ vectores $\overline{u}_{d_{i+1}}, \ldots, \overline{u}_n$ hasta formar una base de V, con lo cual el endomorfismo asociado a la matriz M se representa en esta base por una matriz de la forma:

$$\begin{pmatrix} 0 & \ldots & 0 & m_{1d_{i+1}} & \ldots & m_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & m_{nd_{i+1}} & \ldots & m_{nn} \end{pmatrix}$$

cuyo polinomio característico contendrá el factor μ al menos con el exponente d_i, luego $\mu = 0$ es una raíz característica de M con un orden de multiplicidad por lo menos igual a d_i, como vimos este orden es precisamente α_i, por tanto $d_i \leq \alpha_i$.

Ejemplo 6.2.4

En el ejemplo 6.2.2, los espacios vectoriales asociados a los valores propios λ_i tienen dimensión uno como es fácil comprobar con solo observar la expresión de L_i:

Los vectores de $L_1 = \{(\delta, 2\delta, \delta), \forall \delta \in \mathbb{R}\}$ dependen sólo de un parámetro δ.
Los vectores de $L_2 = \{(\beta, 0, -\beta), \forall \beta \in \mathbb{R}\}$ dependen sólo de un parámetro β.
Los vectores de $L_3 = \{(\alpha, -\alpha, \alpha), \forall \alpha \in \mathbb{R}\}$ dependen sólo de un parámetro α.

Ejemplo 6.2.5

En el ejemplo 6.2.3. $(A - \lambda I) = \begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{pmatrix}$, cuando $\lambda = 2$ (raíz triple).

La expresión del subespacio vectorial asociado a dicho valor propio es:

$$\begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow -2x_1 + x_2 = 0 \Rightarrow \begin{cases} x_1 = \alpha \\ x_2 = 2\alpha \Rightarrow \\ x_3 = \beta \end{cases} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$L = \{\alpha(1, 2, 0) + \beta(0, 0, 1), \forall \alpha, \beta \in \mathbb{R}\}$, que evidentemente tiene dimensión 2 y verifica $2 = d_i \leq \alpha_i = 3$, como se indica en *[6.2.8]*.

6.3. Diagonalización de matrices

Recordemos que el objetivo perseguido en este capítulo es encontrar, si existe, la matriz más sencilla posible que pertenezca a la misma clase de equivalencia que otra matriz $A \in M_{n\times n}$ mediante la relación de semejanza dada. La matriz buscada es una matriz diagonal Λ.

Por ser matrices semejantes A y Λ están asociadas al mismo endomorfismo f.

Sea un endomorfismo f, sobre un espacio vectorial V de dimensión n, que referido a una base B de V tiene A por matriz asociada. Vamos a establecer un criterio que permita saber si una matriz es diagonalizable.

6.3.1. Condición suficiente de diagonalización

Si $\lambda_1, \lambda_2, ..., \lambda_n$, son n raíces distintas de la ecuación característica de la matriz A,

entonces, existe una matriz Λ diagonal asociada a f, tal que, $\Lambda = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$

Demostración:

En *[6.2.7]* vimos que, *"El conjunto de k vectores propios correspondientes a k valores propios distintos, de un endomorfismo f, forman un sistema libre"*, por tanto, el conjunto de los n vectores propios $\bar{x}_1, \bar{x}_2, ..., x_n$ correspondientes a los valores $\lambda_1, \lambda_2, ..., \lambda_n$ constituye una base de V.

Las columnas de la matriz Λ asociada a f en esa base, están formadas por las componentes de las imágenes de los vectores que forman la base:

$f(\bar{x}_1) = \lambda_1\bar{x}_1$; $f(\bar{x}_2) = \lambda_2\bar{x}_2$; ...; $f(\bar{x}_n) = \lambda_n\bar{x}_n \Rightarrow \Lambda = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$, que evidentemente es una matriz diagonal.

Ahora tenemos dos bases en V: la B, y en ella la matriz asociada a f es A, y la B' formada por vectores propios de f, entonces la matriz asociada a f es Λ.

$$\begin{array}{ccc} (V, B) & \xrightarrow{f_A} & (V, B) \\ {\scriptstyle i_P^{-1}}\downarrow\uparrow{\scriptstyle i_P} & & {\scriptstyle i_P^{-1}}\downarrow\uparrow{\scriptstyle i_P} \\ (V, B') & \xrightarrow{f_\Lambda} & (V, B') \end{array}$$

Si llamamos P a la matriz de paso de la base B' a la B y componemos las aplicaciones correspondientes en el diagrama conmutativo anterior, resultará

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{pmatrix} = \Lambda.$$

Como consecuencia, otra forma de expresar *[6.3.1]* es:

"Si las raíces características $\lambda_1, \lambda_2, ..., \lambda_n$ de una matriz A son todas distintas, existe una matriz regular P que verifica la igualdad $P^{-1}AP = \Lambda$".

Ejemplo 6.3.1

Seguimos con la matriz del ejemplo 6.2.2, y recordamos que la matriz A tiene tres valores propios de multiplicidad uno: $\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = 4$.

El conjunto $B' = \{(1, 2, 1), (1, 0, -1), (1, -1, 1)\}$ formado eligiendo una base en cada uno de los subespacios vectoriales asociados L_1, L_2 y L_3, es una base de V.

El orden de A es tres, como los tres valores propios son distintos *[6.3.1]* nos permite afirmar que A es diagonalizable.

P es la matriz del cambio de las coordenadas de un vector respecto a la base B' a las coordenadas respecto a $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

Para construir la matriz de paso P, elegimos como vector columna cada uno de los vectores de B'.

La matriz de paso será $P = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 1 \end{pmatrix}$.

Su matriz inversa P^{-1}, es la matriz del cambio de las coordenadas de un vector respecto a la base B a las coordenadas respecto a B':

$$P^{-1} = \begin{pmatrix} 1/6 & 1/3 & 1/6 \\ 1/2 & 0 & -1/2 \\ 1/3 & -1/3 & 1/3 \end{pmatrix}.$$

La matriz diagonal semejante es $\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

Haciendo los productos de matrices correspondientes, se comprueba que se verifica la igualdad $A = P\Lambda P^{-1}$.

En *[6.3.1]* vimos que *"tener todos los valores propios diferentes es una condición suficiente para que el endomorfismo sea diagonalizable"*, pero hay veces que el endomorfismo es diagonalizable aunque no cumpla esa condición, por tanto, no es una condición necesaria.

Vamos a establecer una condición no sólo necesaria, sino también suficiente para saber cuando una matriz es diagonalizable.

6.3.2. Condición necesaria y suficiente de diagonalización

> Un endomorfismo f es diagonalizable $\Leftrightarrow$ Existe una base de V formada por vectores propios

Demostración:

$\Rightarrow$ Si f es un endomorfismo de un espacio vectorial, tal que en una base $B = \{\bar{e}_1, \bar{e}_2, ..., \bar{e}_n\}$ de V, la matriz $A = \begin{pmatrix} l_1 & 0 & ... & 0 \\ 0 & l_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & l_n \end{pmatrix}$ asociada a f es diagonal, se verificará: $|A - \lambda I| = (l_1 - \lambda)(l_2 - \lambda) ... (l_n - \lambda) = 0 \Rightarrow l_1, l_2, ..., l_n$ son valores propios de A, y se debe cumplir: $f(\bar{e}_1) = l_1\bar{e}_1, f(\bar{e}_2) = l_2\bar{e}_2, ..., f(\bar{e}_n) = l_n\bar{e}_n$ con $(i = 1, 2, ..., n)$. Entonces existe una base de V formada por los vectores propios.

$\Leftarrow$) Como ya hemos visto en *[6.3.1]* *"si existe una base B de V formada por los vectores propios asociados a valores propios distintos de A, con respecto a esa base B la matriz asociada a f es diagonal"*.

Consideremos ahora el caso en que entre las raíces características haya alguna múltiple, entonces de *[6.2.8]*, si $d_i = \dim L_i$ deducimos que $\sum d_i \leq n$ vectores propios linealmente independientes, y además $\sum d_i$ será igual a n cuando cada d_i, coincida con la multiplicidad α_i del valor propio λ_i.

Como sabemos que d_i es la dimensión del núcleo de $A - \lambda_i I$ que es igual a $n - \text{rg}(A - \lambda_i I)$, el teorema anterior se puede expresar en los siguientes términos:

6.3.3. Teorema

> A es diagonalizable $\Leftrightarrow$ $\text{rg}(A - \lambda_i I)$ es igual a $n - \alpha_i$ para todo valor de i.

Obsérvese que este teorema es equivalente al que tiene por enunciado:

6.3.4. Teorema

Un endomorfismo f es diagonalizable $\Leftrightarrow$ La dimensión del subespacio vectorial asociado a cada valor propio es igual a la multiplicidad algebraica de dicho valor propio.

Cada autovalor tiene asociadas dos multiplicidades, la multiplicidad algebraica (multiplicidad como raíz de la ecuación característica) y la multiplicidad geométrica, que es la dimensión del subespacio asociado a dicho valor propio. Que hemos venido llamando x_i y d_i respectivamente.

Ejemplo 6.3.2

En el ejemplo 6.2.3, el único valor propio de A es $\lambda_1 = 2$, que es triple y verifica: $\text{rg}(A - \lambda_1 I) = 1 \neq 3 - 3 = 0$. Luego la matriz A no es diagonalizable.

Si hubiéramos utilizado el segundo criterio habríamos argumentado: La multiplicidad algebraica de $\lambda_1 = 2$ es tres, la dimensión del subespacio asociado al autovalor $\lambda_1 = 2$ (multiplicidad geométrica) es dos, como consecuencia la matriz A no es diagonalizable.

6.4. Diagonalización de endomorfismos simétricos

De entre los endomorfismos f de un espacio vectorial V tiene interés el estudio de aquellos cuya matriz A, es una matriz simétrica de elementos reales, dado que los teoremas relativos a la existencia de la matriz diagonal constituyen un caso particular de gran utilidad, y se extraen consecuencias diferentes a las ya estudiadas.

Veamos algunas proposiciones sobre propiedades de las matrices simétricas relacionadas con los valores y vectores propios.

6.4.1. Proposición

Todas las raíces de la ecuación característica de una matriz simétrica de elementos reales son reales.

La técnica utilizada en la demostración siguiente incluye como escalares a los complejos, por ello puede ser omitido su estudio por el lector no familiarizado con este cuerpo.

Demostración:

Sean A, una matriz simétrica de elementos reales, X la matriz columna de las coordenadas de un vector $\overline{x} \in V$, y $\lambda \in \mathbb{C}$ un valor propio de la matriz $A \Rightarrow$ $\Rightarrow AX = \lambda X$.

Consideremos una matriz columna $\overline{X}$, que tiene por elementos los conjugados de la matriz X, entonces premultiplicando los miembros de la igualdad anterior por $\overline{X}^t$

$$\overline{X}^t AX = \lambda \overline{X}^t X \qquad (1)$$

tomando la conjugada de la anterior igualdad quedará:

$$X'\,\overline{AX} = \overline{\lambda}X'\,\overline{X}$$

como los elementos de A son reales, la matriz A es igual a su conjugada $\overline{A}$ $\Rightarrow$

$$X'\,A\overline{X} = \overline{\lambda}X'\overline{X}$$

si trasponemos ambos miembros de esta igualdad, se obtiene:

$$\overline{X}'\,AX = \overline{\lambda}\overline{X}'X \quad (2)$$

ya que $A' = A$ en las matrices simétricas.

De igualar (1) y (2) resulta:

$$(\lambda - \overline{\lambda})\,\overline{X}'\,X = \overline{0}$$

como $\overline{X}'\,X = \overline{x}_1x_1 + \overline{x}_2x_2 + \ldots + \overline{x}_nx_n > 0$ por ser X una matriz compleja $\Rightarrow \lambda = \overline{\lambda}$.

6.4.2. Proposición

Si A es una matriz simétrica de elementos reales, los vectores propios asociados a las raíces características distintas son ortogonales con el producto escalar usual.

Como consecuencia de *[6.4.1]*, podemos afirmar que las coordenadas de los vectores propios serán también reales en la base considerada, pues, son las soluciones del sistema de ecuaciones homogéneo $(A - \lambda I\,)X = \overline{0}$.

Consideremos una base ortonormal B_o, y sean $\overline{x}_1$ y $\overline{x}_2$ dos vectores propios asociados a los valores propios λ_1 y λ_2, respectivamente. Se verificará: $AX_1 = \lambda_1X_1$ y $AX_2 = \lambda_2X_2$. (Escrito en forma matricial.)

Multiplicando por la izquierda respectivamente por X_2' y X_1' cada una de las igualdades anteriores, se obtiene: $X_2'AX_1 = \lambda_1X_2'X_1$ y $X_1'AX_2 = \lambda_2X_1'X_2$ (1) y trasponiendo la segunda de ellas, quedará $X_2'AX_1 = \lambda_2X_2'X_1$, ya que, $A = A'$ e igualándolo con la primera de (1), será:

$$(\lambda_1 - \lambda_2)X_2'X_1 = \overline{0}$$

y por tanto se deduce que, como $\lambda_1 \neq \lambda_2$ por hipótesis, entonces $X_2'X_1 = \overline{0}$, es decir, que los vectores propios son ortogonales, ya que el producto escalar de dos vectores se expresa de la forma anterior cuando en el espacio vectorial V está referido a una base ortonormal.

Si A es una matriz simétrica real sus autovalores son reales.

Para demostrar este enunciado es necesario usar el conjunto de los números complejos $\mathbb{C}$, por ello el lector puede omitir su estudio.

Si A es una matriz simétrica real, su polinomio característico tiene todos sus coeficientes reales, por lo tanto si admite una raíz $\beta \in \mathbb{C}$ también debe admitir su conjugada $\overline{\beta}$ que por ser raíz del polinomio característico será también un autovalor. Consideremos dos autovectores $\overline{u}$, $\overline{v}$ correspondientes a β y $\overline{\beta}$ respectivamente es decir:

$$A\overline{u} = \beta\overline{u} \text{ y } A\overline{v} = \overline{\beta}\overline{v}$$

multiplicando por $\overline{u}'$ en la segunda igualdad: $\overline{u}'\,A\overline{v} = \overline{\beta}\overline{u}'\overline{v}$ y tomando traspuesta resulta $(\overline{u}'\,A\overline{v})' = \overline{v}'\,A'\overline{u} = \overline{v}'\beta\overline{u} = \beta\overline{v}'\overline{u}$ y en el otro término de la igualdad $(\overline{\beta}\overline{u}'\overline{v})' = \overline{\beta}\overline{v}'\overline{u}$, por lo tanto $\beta = \overline{\beta}$, es decir $\beta \in \mathbb{R}$.

Este resultado será utilizado en la demostración del teorema siguiente.

6.4.3. Teorema

Dada una matriz simétrica A de elementos reales, existe una matriz ortogonal P, tal que, $P'AP$ es una matriz diagonal formada por los valores propios de A, algunos de los cuales pueden ser iguales.

Este resultado es muy importante, pero dada la dificultad de la demostración, el lector puede omitir su estudio.

Demostración:

Basta con demostrar que en estas condiciones, existe una base ortonormal constituida por vectores propios de A. Puesto que la matriz P sería la matriz que tiene por vectores columnas los vectores de dicha base ortonormal.

Utilizaremos el método de inducción:

Si $A = (a_{11})$ y $P = I$, el teorema se verifica.

Supondremos que se verifica para matrices de orden $n - 1$.

Lo demostraremos para una matriz A de orden n.

Consideremos el endomorfismo f de $\mathbb{R}^n$ asociado a la matriz A que por hipótesis es simétrica real y por el resultado anterior sus autovalores son reales (tiene al menos uno) y hay al menos un autovector $\overline{x} \neq \overline{0}$ para el que $f(\overline{x}) = \lambda\overline{x}$, siendo λ el valor propio real asociado al autovector $\overline{x}$. En consecuencia $\overline{y} = \dfrac{\overline{x}}{\|\overline{x}\|}$ es vector propio de f. Considérese el espacio W ortogonal a y, es decir el constituido por todos los vectores de $\mathbb{R}^n$ ortogonales a $\overline{y}$, el cual es de dimensión n–1. Por lo tanto para todo $\overline{z} \in W$ se tiene que $\overline{z} \cdot \overline{y} = 0$ y en consecuencia.

$$f(\overline{z}) \cdot \overline{y} = (AZ)'\,Y = Z'A'Y = Z'AY = Z'\,(AY) = z \cdot f(\overline{y}) = z \cdot (\lambda\overline{y}) = \lambda(z \cdot y) = 0$$

por consiguiente, si $z \in W$, entonces $f(z)$ es ortogonal a $\overline{y}$. Por hipótesis de inducción el subespacio W posee una base ortonormal de vectores propios de f, si añadimos a esta base el vector $\overline{y}$ obtenemos una base ortonormal formada por vectores propios de f y por tanto se concluye.

CONSECUENCIA

Cuando la matriz A es simétrica de elementos reales, A es congruente con una matriz diagonal.

6.4.4. Proposición

Existen n vectores propios de A (simétrica real) que son ortogonales dos a dos, incluidos los que pertenecen a un mismo valor propio.

Demostración:

Se deduce de la anterior:

Ejemplo 6.4.1

Sea A la matriz simétrica con elementos reales $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$; se trata de de diagonalizar A mediante una matriz ortogonal C aplicando el teorema *[6.4.3]*.

La ecuación característica es: $|A - \lambda I| = \begin{vmatrix} 2-\lambda & 0 & 1 \\ 0 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} = -(\lambda^3 - 7\lambda^2 + {}$ $+ 15\lambda - 9) = 0.$

Como $\lambda^3 - 7\lambda^2 + 15\lambda - 9 = (\lambda - 1)(\lambda^2 - 6\lambda + 9) = 0$ Los valores propios de la ecuación anterior son $\lambda_1 = 3$ con multiplicidad dos, y $\lambda_2 = 1$ con multiplicidad uno.

Consideremos el primer valor propio, y determinemos su subespacio vectorial asociado:

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow -x_1 + x_3 = 0 \Rightarrow \begin{cases} x_1 = \alpha \\ x_2 = \mu \\ x_3 = \alpha \end{cases} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} +$$

$$+\mu\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow L_1 = \{(x_1, x_2, x_3) = \alpha(1, 0, 1) + \mu(0, 1, 0), \forall \alpha, \mu \in \mathbb{R}\}$$

Operando de la misma forma para el segundo valor propio, el subespacio vectorial asociado es: $L_2 = \{(x_1, x_2, x_3) = \beta(1, 0, -1), \forall\beta \in \mathbb{R}\}$

Elegimos en L_1 dos vectores, por ser dos la multiplicidad del valor propio λ_1, si tomamos los más sencillos: $\overline{v}_1 = (1, 0, 1)$ y $\overline{v}_2 = (0, 1, 0)$, que son los que se obtienen para los valores $\alpha = 1, \mu = 0$ y $\alpha = 0, \mu = 1$ respectivamente, podemos observar que ambos vectores son ortogonales.

Del mismo modo elegimos en L_2 el vector $\overline{v}_3 = (1, 0, -1)$, que es ortogonal a los dos anteriores pertenecientes a L_1, como debe ser por la proposición *[6.4.4]*.

Los módulos de estos vectores son $|\overline{v}_1| = \sqrt{2}$; $|\overline{v}_2| = 1$; $|\overline{v}_3| = \sqrt{2}$, por tanto, los tres vectores ortonormales serán:

$$\overline{u}_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right); \overline{u}_2 = (0, 1, 0); \overline{u}_3 = \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$$

que conforman como vectores columnas, la matriz C, por tanto:

$$C = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{pmatrix} = C^t$$

Se verifica, como puede comprobar el lector, que $C^t AC = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \Lambda$,

que en efecto, es una matriz diagonal.

EJERCICIOS

Ejercicio 6.1

Demuéstrese que toda matriz cuadrada, tal que los elementos de cada fila suman la unidad, tiene como valor propio $\lambda = 1$ y como vector propio correspondiente, el vector cuyos únicos elementos son unos.

Solución:

El objetivo de este ejercicio es comprobar que además de los conceptos, se ha asimilado correctamente el lenguaje de matrices, vectores, índices, subíndices, etc., que el estudiante de informática se verá en la necesidad de utilizar.

En la matriz A se verifica: $\sum_{j=1}^{n} a_{ij} = 1$ para cada i, con $1 \leq i \leq n$

i representa cada una de las n filas de la matriz que estamos estudiando, por eso puede variar entre 1 y n.

Para cada i, es decir, en cada fila, los subíndices j que representan las columnas, pueden variar entre 1 y n, y la suma de todos los elementos a_{ij} de esa fila i, vale 1.

La expresión: $A\begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}$, es cierta, ya que:

$$A\begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} \\ \sum_{j=1}^{n} a_{2j} \\ \dots \\ \sum_{j=1}^{n} a_{nj} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \text{ es}$$

un vector propio correspondiente al valor propio $\lambda = 1$ de la matriz A.

Ejercicio 6.2

Dada la matriz $A = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 4 & -1 \\ -2 & 2 & 1 \end{pmatrix}$, se pide calcular los valores propios y los vectores propios.

Solución:

La ecuación característica es:

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 & -1 \\ -2 & 4-\lambda & -1 \\ -2 & 2 & 1-\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 - 11\lambda + 6 = 0.$$

Las raíces de $-\lambda^3 + 6\lambda^2 - 11\lambda + 6 = 0$, se calculan utilizando la regla de Ruffini.

Los valores encontrados son: $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, que son los tres valores propios de la matriz.

Cada valor propio lleva asociado un subespacio de vectores propios.

Consideremos el primer valor propio $\lambda_1 = 1$, se debe verificar:

$$\begin{pmatrix} 1-1 & 1 & -1 \\ -2 & 4-1 & -1 \\ -2 & 2 & 1-1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left\{ \begin{aligned} x_2 - x_3 &= 0 \\ -2x_1 + 3x_2 - x_3 &= 0 \\ -2x_1 + 2x_2 &= 0 \end{aligned} \right. \Rightarrow \left\{ \begin{aligned} x_2 - x_3 &= 0 \\ -x_1 + x_2 &= 0 \end{aligned} \right. \Rightarrow$$

$$\Rightarrow \left\{ \begin{aligned} x_1 &= \alpha \\ x_2 &= \alpha \\ x_3 &= \alpha \end{aligned} \right. \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Para definir el espacio vectorial asociado, L_1, podemos dar:

Sus ecuaciones cartesianas: $\left\{ \begin{aligned} x_2 - x_3 &= 0 \\ -x_1 + x_2 &= 0 \end{aligned} \right.$, o sus ecuaciones paramétricas:

$\left\{ \begin{aligned} x_1 &= \alpha \\ x_2 &= \alpha \\ x_3 &= \alpha \end{aligned} \right.$, o una base: $B = \{(1, 1, 1)\}$, o todos los vectores que lo forman:

$$L_1 = \{(x_1, x_2, x_3) = \alpha(1, 1, 1), \forall \alpha \in \mathbb{R}\}.$$

Como puede comprobar el lector, se verifica:

$\begin{pmatrix} 1 & 1 & -1 \\ -2 & 4 & -1 \\ -2 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, es decir, $(1, 1, 1)$ es un vector propio asociado al valor propio 1.

Para el valor propio $\lambda_2 = 2$, deberemos seguir el mismo camino, y obtendremos:

$$\begin{pmatrix} 1-2 & 1 & -1 \\ -2 & 4-2 & -1 \\ -2 & 2 & 1-2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left\{ \begin{aligned} -x_1 + x_2 - x_3 &= 0 \\ -2x_1 + 2x_2 - x_3 &= 0 \\ -2x_1 + 2x_2 - x_3 &= 0 \end{aligned} \right. \Rightarrow \left\{ \begin{aligned} x_3 &= 0 \\ x_1 - x_2 &= 0 \end{aligned} \right. \Rightarrow$$

$$\Rightarrow \left\{ \begin{aligned} x_1 &= \alpha \\ x_2 &= \alpha \\ x_3 &= 0 \end{aligned} \right. \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \Rightarrow L_2 = \{(x_1, x_2, x_3) = \alpha(1, 1, 0), \forall \alpha \in \mathbb{R}\}.$$

Podemos decir que $(1, 1, 0)$ es un vector propio asociado al valor propio 2.

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

Para el valor propio $\lambda_2 = 3$ obtendremos:

$$\begin{pmatrix} 1-3 & 1 & -1 \\ -2 & 4-3 & -1 \\ -2 & 2 & 1-3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left\{ \begin{aligned} -2x_1 + x_2 - x_3 &= 0 \\ -2x_1 + x_2 - x_3 &= 0 \\ -2x_1 + 2x_2 - 2x_3 &= 0 \end{aligned} \right. \Rightarrow \left\{ \begin{aligned} x_1 &= 0 \\ x_3 - x_2 &= 0 \end{aligned} \right. \Rightarrow$$

$$\Rightarrow \left\{ \begin{aligned} x_1 &= 0 \\ x_2 &= \alpha \\ x_3 &= \alpha \end{aligned} \right. \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \Rightarrow L_3 = \{(x_1, x_2, x_3) = \alpha(0, 1, 1), \forall \alpha \in \mathbb{R}\}.$$

Podemos decir que (0, 1, 1) es un vector propio asociado al valor propio 3.

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

Ejercicio 6.3

Dada una matriz triangular superior $A = (a_{ij})$, y λ_i sus valores propios, demuéstrese que se verifica: $\lambda_i = a_{ii}$, $\forall i = 1, 2, ..., n$

Solución:

En efecto, si A es triangular superior se verifica:

$$|A - \lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda)\dots(a_{nn}-\lambda) = 0.$$

por tanto, $\lambda_i = a_{ii}$, $1 \le i \le n$.

Ejercicio 6.4

Demuéstrese que las matrices A y A' tienen los mismos valores propios. Asimismo, demuéstrese que los vectores propios de A y A' correspondientes a valores propios distintos son ortogonales.

Solución:

A y A' tienen la misma ecuación característica, ya que $(A' - \lambda I) = (A - \lambda I)'$, y $\forall M \in M_{n\times n}$ se verifica que $|M| = |M'|$ además de que $(A + B)' = (A' + B')$.

Veamos la segunda parte.

Sea $\overline{x}$ un vector propio asociado al valor propio de A, $\gamma \Rightarrow A\overline{x} = \gamma\overline{x}$.
Sea $\overline{y}$ un vector propio asociado al valor propio de A', $\beta \Rightarrow A'\overline{y} = \beta\overline{y}$ con $\beta \neq \gamma$.
Tomando traspuestas en la igualdad $A'\overline{y} = \beta\overline{y} \Rightarrow \overline{y}'A = \beta\overline{y}'$.
Multiplicando por la derecha esta igualdad por $\overline{x}$, resulta $\overline{y}'A\overline{x} = \beta\overline{y}'\overline{x} \Rightarrow$ $\Rightarrow \overline{y}'\gamma\overline{x} = \beta\overline{y}'\overline{x} \Rightarrow (\gamma - \beta)\overline{y}'\overline{x} = 0$, como $\gamma \neq \beta$, entonces $\overline{y}'\overline{x} = 0$, que era lo que queríamos demostrar.

Ejercicio 6.5

Estúdiese si es, o no, diagonalizable la matriz $A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$. En caso afirmativo, constrúyase una matriz de paso P.

Solución:

Ecuación característica de la matriz A:

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 1 & -1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix} = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = 0.$$

Valores propios:
Son las raíces de la ecuación característica:
$\lambda_1 = 1$ con multiplicidad dos, es decir, $\alpha_1 = 2$, y $\lambda_2 = 2$ con multiplicidad uno, es decir, $\alpha_2 = 1$.

Subespacios asociados a cada uno de los valores propios:
son $L_1 = \{(x, y, z) = \beta(1, 0, -1) + \delta(0, 1, 1), \forall\beta, \delta \in \mathbb{R}\}$ y $L_2 = \{(x, y, z) = \eta\,(1, 1, -1), \forall\eta \in \mathbb{R}\}$.
Como $\mathrm{rg}(A - \lambda_1 I) = 1 = n - \alpha_1 = 3 - 2 = 1$ y $\mathrm{rg}(A - \lambda_2 I) = 2 = n - \alpha_2 = 3 - 1 = 2$, entonces, la matriz A es diagonalizable.

Habríamos llegado al mismo resultado si hubiéramos argumentado:

$$\left.\begin{array}{l}\text{Multiplicidad algebraica de } \lambda_1 = \text{Multiplicidad geométrica de } L_1 = \alpha_1 = d_1 = 2\\ \text{Multiplicidad algebraica de } \lambda_2 = \text{Multiplicidad geométrica de } L_2 = \alpha_2 = d_2 = 1\end{array}\right\}\Rightarrow$$

$\Rightarrow$ la matriz A es diagonalizable.

Matriz de paso:

La matriz de paso P se forma con dos vectores linealmente independientes de L_1 y uno de L_2, tomados como columnas de dicha matriz, pues sus dimensiones son dos y uno respectivamente, por tanto, $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{pmatrix}$.

Matriz diagonal semejante:

Es la matriz diagonal que tiene los autovalores en la diagonal principal.

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

Comprobación:

Se verifica que $A = P\Lambda P^{-1}$. En efecto:

$$P^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix};$$

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

Las operaciones se pueden realizar utilizando Derive o cualquier otro programa de cálculo simbólico.

Ejercicio 6.6

Dadas las matrices $A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} -1 & 4 \\ -1 & 2 \end{pmatrix}$, se pide demostrar si son, o no, matrices de un mismo endomorfismo.

Solución:

Las matrices A y B son de un mismo endomorfismo respecto a distintas bases, si y sólo si, existe una matriz real (de cambio de base) $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ regular y se verifica $\begin{pmatrix} -1 & 4 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \Rightarrow \begin{pmatrix} -1 & 4 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} =$
$= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \Rightarrow$ De esta expresión se obtiene el sistema de ecuaciones lineal homogéneo $\left\{\begin{array}{r} 2a + 2b - 4c = 0 \\ a - c + 2d = 0 \\ a - b + 4d = 0 \\ 2d - b + c = 0 \end{array}\right.$, con el rango de la matriz de coeficientes tres, por tanto, tiene solución distinta de la trivial, una ellas es, por ejemplo, la que da lugar a la matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

$\Rightarrow$ Ambas matrices representan el mismo endomorfismo pero están referidas a distintas bases.

Ejercicio 6.7

Demuéstrese que dos matrices semejantes A y B, tienen la misma ecuación característica.

Solución:

Las ecuaciones características de las dos matrices son $|A - \lambda I| = 0$ y $|B - \lambda I| = 0$. Si A y B son semejantes, existe una matriz regular P, tal que, $B = P^{-1}AP$. Entonces se verifica $|B - \lambda I| = |P^{-1}AP - \lambda I| = |P^{-1}AP - \lambda P^{-1}P| = |P^{-1}(A-\lambda I)P|P^{-1}||A-\lambda|P|$, y como, $|P^{-1}| = \dfrac{1}{|P|} \Rightarrow |B-\lambda I| = |A - \lambda I|$, que es lo que queríamos demostrar.

Ejercicio 6.8

Estúdiese la existencia de una matriz diagonal semejante a $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & \gamma \\ -2 & 0 & \delta \end{pmatrix}$ para los distintos valores de γ y δ.

Solución:

La ecuación característica es $|A - \lambda I| = \begin{vmatrix} 3-\lambda & 0 & 0 \\ 0 & 1-\lambda & \gamma \\ -2 & 0 & \delta-\lambda \end{vmatrix} =$

$= (3-\lambda)(1-\lambda)(\delta-\lambda) = 0$, cuyos valores propios son $\lambda = 3, \lambda = 1, \lambda = \delta$.

Si $\delta \neq 3$ y $\delta \neq 1$, los tres valores propios son distintos y A es diagonalizable.
Veamos qué ocurre cuando $\delta = 3$. (En cada caso podemos utilizar el criterio de la columna derecha o el de la columna izquierda indistintamente.)
Entonces $\lambda = 3$ y se trata de un valor propio doble.

$(A - 3I) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & \gamma \\ -2 & 0 & 0 \end{pmatrix} \Rightarrow \text{rg}(A - 3I) = 2 \neq 3 - 2 = 1 \Rightarrow A$ no es diagonalizable.

En este caso la multiplicidad algebraica de $\lambda = 3$ es 2, y la multiplicidad geométrica es 1 independientemente del valor de $\gamma \Rightarrow A$ no es diagonalizable.

Si $\delta = 1$, entonces $\lambda = 1$ es un valor propio de multiplicidad algebraica 2, de matriz la A, y $A - I = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & \gamma \\ -2 & 0 & 0 \end{pmatrix}$.

Con $\delta = 1$, si $\gamma \neq 0 \Rightarrow \text{rg}(A - I) = 2 \neq 3 - 2 = 1 \Rightarrow A$ no es diagonalizable.

En este caso las dos multiplicidades no coinciden $\Rightarrow A$ no es diagonalizable.

Con $\delta = 1$, si $\gamma = 0$, entonces $\text{rg}(A - I) = 1 = 3 - 2 = 1$, y la matriz es diagonalizable.

En este caso las dos multiplicidades coinciden $\Rightarrow A$ es diagonalizable.

Ejercicio 6.9

La matriz $A = \begin{pmatrix} a & -1 & b \\ c & 2 & d \\ e & 1 & f \end{pmatrix}$ tiene por vectores propios: (1, 0, 1), (–1, 1, 0), (0, 1, –1). Calcúlense:

a) La matriz A, y los valores propios correspondientes a los tres vectores anteriores. b) Compruébese si es diagonalizable la matriz A. c) Escríbese la matriz de paso P.

Solución:

Como los vectores (1, 0, 1), (–1, 1, 0), (0, 1, –1) son vectores propios de la matriz $\begin{pmatrix} a & -1 & b \\ c & 2 & d \\ e & 1 & f \end{pmatrix}$, se verifica:

$$\begin{pmatrix} a & -1 & b \\ c & 2 & d \\ e & 1 & f \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \lambda\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \begin{cases} a+b=\lambda \\ c+d=0 \\ e+f=\lambda \end{cases} \Rightarrow \begin{cases} c=-d \\ a+b=e+f \end{cases}$$

$$\begin{pmatrix} a & -1 & b \\ c & 2 & d \\ e & 1 & f \end{pmatrix}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \nu\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} -a-1=-\nu \\ -c+2=\nu \\ -e+1=0 \end{cases} \Rightarrow \begin{cases} a+c=1 \\ e=1 \end{cases}$$

$$\begin{pmatrix} a & -1 & b \\ c & 2 & d \\ e & 1 & f \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \varepsilon\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \Rightarrow \begin{cases} -1-b=0 \\ 2-d=\varepsilon \\ 1-f=-\varepsilon \end{cases} \Rightarrow \begin{cases} b=-1 \\ d+f=3 \end{cases}$$

La solución de las ecuaciones anteriores es:

$a = 3$, $b = -1$, $c = -2$, $d = 2$, $e = 1$, $f = 1$, y los valores propios $\lambda = 2$, $\nu = 4$, $\varepsilon = 0$.

b) La matriz A es diagonalizable por tener 3 valores propios diferentes.

c) La matriz de paso P será $P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$.

La matriz diagonal es $\Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, y se verifica que $A = P\Lambda P^{-1}$, como es fácilmente comprobable.

Ejercicio 6.10

Dada la matriz simétrica $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, se pide:

1. Diagonalizar A.
2. Calcular la matriz ortogonal de paso.
3. Comprobar la ortogonalidad de los vectores propios.

Solución:

1. La ecuación característica de la matriz A es

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2(2-\lambda) - (2-\lambda)\,[(1-\lambda)^2 - 1] =$$

$$= (2-\lambda)\left[(1-\lambda)^2 - 1\right] = (2-\lambda)(\lambda^2 - 2\lambda) = (2-\lambda)\,\lambda(\lambda-2) = 0.$$

Los valores propios son: $\lambda_1 = 2$ con multiplicidad algebraica 2 y $\lambda_2 = 0$ con multiplicidad algebraica 1.

El subespacio asociado a $\lambda_1 = 2$ es:

$$\begin{pmatrix} 1-2 & 0 & 1 \\ 0 & 2-2 & 0 \\ 1 & 0 & 1-2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1 - x_3 = 0 \Leftrightarrow \begin{cases} x_1 = \alpha \\ x_2 = \beta \\ x_3 = \alpha \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + $$

$$+\beta\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

El subespacio asociado a $\lambda_2 = 0$ es $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 + x_3 = 0 \\ x_2 = 0 \end{cases} \Leftrightarrow$

$$\Leftrightarrow \begin{cases} x_1 = \alpha \\ x_2 = 0 \\ x_3 = -\alpha \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$\Rightarrow$ La matriz diagonal equivalente a A es $\Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

2. Para calcular la matriz ortogonal de paso, hay que normalizar los autovectores que la van a formar:

Vector normalizado equivalente a $(1, 0, 1)$: $\left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right)$.

Vector normalizado equivalente a (0, 1, 0): (0, 1, 0).

Vector normalizado equivalente a (1, 0, −1): $\left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$.

Matriz de paso: $\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{pmatrix}$

3. Los vectores (1, 0, 1) y (0, 1, 0) son ortogonales entre sí y ortogonales a (1, 0, −1).

Ya que se verifica: $(1, 0, 1) \cdot (0, 1, 0) = (1, 0, 1) \cdot (1, 0, -1) =$
$= (0, 1, 0) \cdot (1, 0, -1) = 0$.

CAPÍTULO 7

FORMAS BILINEALES

INTRODUCCIÓN

Los resultados más profundos acerca de la clasificación y reducción de matrices (y por tanto de sistemas lineales), tuvieron su origen no en el estudio de tales sistemas, sino en el de formas bilineales y cuadráticas.

El problema de la reducción de las formas bilineales y cuadráticas a formas canónicas, lo más simple posible, fue el principal impulsor de las investigaciones sobre reducción y clasificación de matrices.

Fermat ya había identificado las cónicas con las curvas planas definidas por ecuaciones de segundo grado, y procedido a su clasificación. El problema análogo para las cuádricas fue abordado por Euler, quien trató de reducir las ecuaciones de la cuádrica a la forma $ax^2 + by^2 + cz^2 + d = 0$; $(a, b, c \neq 0)$ por medio de un cambio de ejes rectangulares. Se dio cuenta de que ello es posible si, y sólo si, el determinante de la matriz de la forma cuadrática es no nulo, en cuyo caso los números a, b, c son los valores propios de esta matriz. Euler refiere la cuádrica respecto a un nuevo sistema de coordenadas rectangulares de manera que se anulen los términos "rectangulares" en la nueva ecuación de la cuádrica. Poco después, en 1765, Euler trata un problema equivalente (aunque aparentemente sin advertirlo), al estudiar los ejes de inercia de un sólido.

Lagrange, Cauchy y otros abordaron el problema de reducir una forma cuadrática en n variables a suma de cuadrados.

Puede decirse, pues, que a partir de la segunda mitad del siglo XVIII aparece con asiduidad en Álgebra el problema de la determinación de transformaciones lineales (con coeficientes reales o simplemente enteros, según los casos) que permiten reducir las formas cuadráticas de dos o tres variables a unos cuantos tipos simples, y tratar entonces de clasificar éstos en términos de invariantes bajo las transformaciones consideradas, lo que permitirá clasificar las formas generales.

GAUSS (1777 - 1855)

Gauss es uno de los matemáticos más grandes de la historia y estaba convencido de ello, se llamaba a sí mismo **matemático total,** le llamaban **príncipe de las matemáticas,** ambos calificativos se quedaban cortos.

Una anécdota muy conocida, describe cómo a la edad de 7 años asombró a su profesor al entregar en tiempo récord el resultado de sumar desde el número 1 hasta el número 100 con la respuesta correcta.

El profesor le preguntó cómo lo había hecho. Gauss le dijo "$1 + 100 = 101$, $2 + 99 = 101$, $3 + 98 = 101$, siempre suman 101. Como son 50 sumas de 101, el total es 5050".

Esto fue sólo el principio; durante sus estudios de Bachillerato y universitarios redescubrió varios teoremas y aunque algunos eran conocidos, él los ignoraba.

Su pasión era la Aritmética, pero, a lo largo de su vida, trabajó en muchos campos, y en todos hizo aportaciones inestimables.

Siendo muy joven resolvió el problema de la construcción de un polígono regular de 17 lados con regla y compás.

Obsesionado por el rigor y las demostraciones, aficionado al estudio de idiomas y a la literatura inglesa, seguidor de la política mundial, acomplejado hasta el extremo de no querer publicar ni dejar huellas de los pasos intermedios que le llevaban a sus conclusiones por miedo al ridículo, tenía su sello y su lema: un árbol con pocos frutos (pocos pero maduros) y estaba patrocinado por su mecenas: el Duque de Brunsvic.

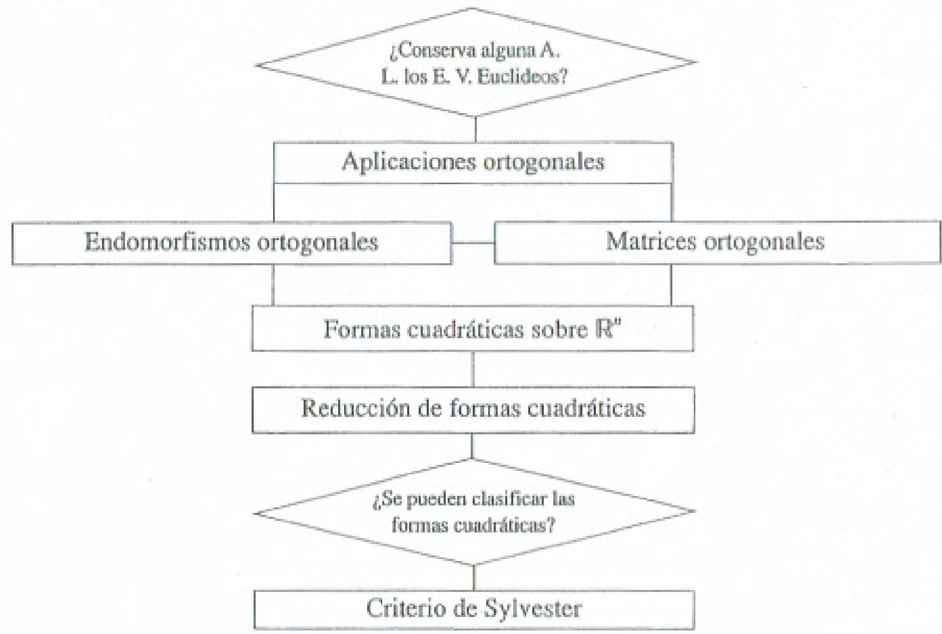
De su trabajo científico destacaremos aquí la primera demostración del teorema fundamental del Álgebra (1799).

Su pasión por las matemáticas le llevó a intercalar otro tipo de actividades, como sus estudios sobre las geometrías no euclídeas por las que se interesó desde muy joven. Discutió el tema con Farkas Bolyai y otros, sin embargo no publicó nada porque creía que su reputación se pondría en entredicho. Más tarde, cuando Lobachevsky publicó su trabajo sobre el tema, dijo en una carta a Schumacher que él estaba convencido desde hacía 54 años.

Nunca abandonó los trabajos de Astronomía y desde 1845 hasta 1851 se encargó, con gran éxito de las finanzas de la Universidad de Göttingen.

Murió el 23 de febrero de 1855 en Göttingen (Alemania).

CAPÍTULO 7


¿Conserva alguna A. L. los E. V. Euclideos?
Aplicaciones ortogonales
Endomorfismos ortogonales
Matrices ortogonales
Formas cuadráticas sobre $\mathbb{R}^n$
Reducción de formas cuadráticas
¿Se pueden clasificar las formas cuadráticas?
Criterio de Sylvester

7.1. Formas bilineales

En el capítulo dos estudiamos las aplicaciones lineales, que son las aplicaciones que conservan la estructura de espacio vectorial. Cuando el conjunto inicial es un producto de dos espacios vectoriales hay unas aplicaciones, las aplicaciones bilineales, que son de gran interés, porque a través de ellas abordaremos el estudio de las formas bilineales y las formas cuadráticas.

El hecho de haber tratado el lector los conceptos de producto escalar, matrices simétricas de elementos reales, vectores ortogonales y bases ortonormales nos permitirá entrar de forma sencilla y natural en este estudio de tanto interés en aplicaciones matemáticas, físicas y tecnológicas.

Si f es una aplicación de un producto de espacios vectoriales en otro espacio vectorial, tal que la aplicación $f: E \times F \to V$ cuando mantenemos fija la primera variable es lineal y la aplicación $f: E \times F \to V$ cuando mantenemos fija la segunda variable es lineal, a f se la llama aplicación bilineal.

7.1.1. Definición de aplicación bilineal

Sean E, F, V espacios vectoriales reales, una aplicación $f: E \times F \to V$ es una aplicación bilineal $\Leftrightarrow \forall\ \overline{x} \in E$, $\forall\ \overline{y} \in F$, $\forall \lambda \in \mathbb{R}$ la aplicación f verifica:

$$\begin{cases} f(\overline{x}+\overline{x}', \overline{y}) = f(\overline{x}, \overline{y}) + f(\overline{x}', \overline{y}) \\ f(\lambda\overline{x}, \overline{y}) = \lambda f(\overline{x}, \overline{y}) \end{cases} \quad \text{y} \quad \begin{cases} f(\overline{x}, \overline{y}+\overline{y}') = f(\overline{x}, \overline{y}) + (\overline{x}, \overline{y}') \\ f(\overline{x}, \lambda\overline{y}) = \lambda f(\overline{x}, \overline{y}) \end{cases}$$

En el caso particular de las aplicaciones bilineales de un espacio vectorial real por sí mismo en $\mathbb{R}$, a la aplicación f se la llama forma bilineal.

7.1.2. Definición de forma bilineal

Sea E un espacio vectorial real, una aplicación f: $E \times E \to \mathbb{R}$ es una forma bilineal $\Leftrightarrow f$: $E \times E \to \mathbb{R}$ es una aplicación bilineal.

Ejemplo 7.1.1

Dada la aplicación $f: \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}$ definida mediante: $f(\overline{x}, \overline{y}) = x_1y_2 + x_2y_3$, siendo $\overline{x} = (x_1, x_2)$ e $\overline{y} = (y_1, y_2, y_3)$ en las bases: $B_1 = \{\overline{e}_1, \overline{e}_2\}$ y $B_2 = \{\overline{e}'_1, \overline{e}'_2, \overline{e}'_3\}$, vamos a comprobar que es forma bilineal.

Aplicando la definición dada anteriormente se tiene:

1. $f(\overline{x} + \overline{x}', \overline{y}) = (x_1 + x'_1)y_2 + (x_2 + \overline{x}'_2)y_3 = x_1y_2 + x'_1y_2 + x_2y_3 + x'_2y_3 =$
   $= x_1y_2 + x_2y_3 + x'_1y_2 + x'_2y_3 = f(\overline{x}, \overline{y}) + f(\overline{x}', \overline{y})$
2. $f(\lambda\overline{x}, \overline{y}) = \lambda x_1y_2 + \lambda x_2y_3 = x_1\lambda y_2 + x_2\lambda y_3 = \lambda(x_1y_2 + x_2y_3) = \lambda f(\overline{x}, \overline{y})$.

De manera similar se comprueba que se verifican las otras dos condiciones.

f es una aplicación bilineal porque:

- Es una aplicación de un producto de espacios vectoriales reales en el espacio vectorial $\mathbb{R}$.
- La aplicación $f: \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}$ cuando mantenemos fija la prim ra variable es lineal.
- $f: \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}$ es lineal al mantener fija la segunda variable.

Además f es una forma bilineal porque:

- Es una aplicación bilineal de $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$.

Si $f: E \times E \to \mathbb{R}$ es una forma bilineal y $B = \{\bar{e}_1, ..., \bar{e}_n\}$ es una base de E, podemos enunciar las siguientes proposiciones:

7.1.3. Proposición

$\forall(\bar{x}, \bar{y}) \in E \times E$, el valor de $f(\bar{x}, \bar{y})$ se puede expresar en función de los valores $f(\bar{e}_i, \bar{e}_j) = a_{ij}$, con $i, j = 1, ..., n$.

Demostración:

En efecto, si E es un espacio vectorial de dimensión n y está referido a una base $B = \{\bar{e}_1, ..., \bar{e}_n\}$, en el cual los vectores $\bar{x}$ e $\bar{y}$ tienen por expresiones respectivas:

$$\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + ... + x_n\bar{e}_n = \sum_{i=1}^{n} x_i\bar{e}_i \text{ e } \bar{y} = y_1\bar{e}_1 + y_2\bar{e}_2 + ... + y_n\bar{e}_n = \sum_{j=1}^{n} y_j\bar{e}_j.$$

Entonces:

$$f(\bar{x}, \bar{y}) = f\left(\sum_{i=1}^{n} x_i\bar{e}_i, \sum_{j=1}^{n} y_j\bar{e}_j\right)$$

y aplicando sucesivamente las propiedades vistas en *[7.1.2]* se obtiene la expresión:

$$f(\bar{x}, \bar{y}) = \sum_{i,j=1}^{n} x_i y_j f(\bar{e}_i, \bar{e}_j) = \sum_{i,j=1}^{n} a_{ij} x_i y_j$$

donde $f(\bar{e}_i, \bar{e}_j) = a_{ij}$ $(i, j = 1, 2, ..., n)$, que es el resultado de la proposición que queríamos demostrar.

A esta altura del texto, el lector debe estar acostumbrado a manejar con soltura las expresiones del tipo anterior, pero si no fuera así, es conveniente que desarrolle algunos sumandos de la expresión, e intente después hacerlo en la forma en que aparece en el libro.

7.1.4. Proposición

Fijada una base, toda forma bilineal se puede representar por una matriz cuadrada, y recíprocamente, a toda matriz cuadrada corresponde una forma bilineal única.

La proposición *[7.1.3]* se puede escribir en forma matricial de la siguiente manera:

$$f(\bar{x}, \bar{y}) = \sum_{i,j} a_{ij}x_i y_j = \sum_i x_i \left(\sum_j a_{ij} y_j\right) = X^t AY.$$

expresión que en forma desarrollada resulta:

$$f(\bar{x}, \bar{y}) = (x_1, x_2, ..., x_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

donde x_i e y_i representan las coordenadas de $\bar{x}$ e $\bar{y}$, siendo X e Y las matrices columna de las coordenadas de $\bar{x}$ e $\bar{y}$ respectivamente y $A = (a_{ij})$, es la llamada matriz de la forma bilineal f en la base $B = \{\bar{e}_1, ..., \bar{e}_n\}$, siendo $a_{ij} = f(\bar{e}_i, \bar{e}_j)$.

De esta forma, fijada una base $B = \{\bar{e}_1, ..., \bar{e}_n\}$, a una forma bilineal f dada, le corresponde una matriz cuadrada A, y sólo una, con lo cual hemos verificado la proposición *[7.1.4]*.

Obsérvese que todos los pasos dados para hacer esta demostración son reversibles, por ello quedan demostradas simultáneamente la implicación directa y su recíproca.

7.1.5. Proposición

La matriz asociada a una forma bilineal $f: E \times E \to \mathbb{R}$ con la base $B = \{\bar{e}_1, ..., \bar{e}_n\}$ es congruente con la matriz asociada a f cuando se efectúa un cambio de base.

Demostración:

Si consideramos una nueva base de E, $B' = \{\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n\}$, siendo X' e Y' las matrices columna de las coordenadas de $\bar{x}$ e $\bar{y}$ respectivamente en la nueva base, se tendrá: $f(\bar{x}, \bar{y}) = (X')^t A' Y'$, donde $A' = (a'_{ij})$ y $a'_{ij} = f(\bar{e}'_i, \bar{e}'_j)$.

Como f se ha definido con independencia del sistema de referencia, se verificará: $(X')^t A' Y' = X^t AY$.

Si P es la matriz que permite pasar de las coordenadas respecto a la antigua base B a las coordenadas respecto a la nueva B', será: $X = PX'$ e $Y = PY'$ y llevando estas igualdades a la expresión anterior resulta:

$$(X')^t A'(Y') = (X')^t P^t APY'$$

de donde se deduce que: $A' = P^t AP$

es decir que la matriz A' asociada a f en la nueva base B' es congruente con la matriz A en la base B.

Ejemplo 7.1.2

Sea el espacio vectorial $\mathbb{R}^3$ y su base canónica $B = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\}$. Dada la forma bilineal $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, expresada por: $f(\bar{x}, \bar{y}) = x_1y_1 - 4x_1y_2 + 6x_2y_3 - - x_3y_1 - 3x_3y_3$. Se pide determinar la matriz asociada y su expresión analítica.

Solución:

$$f(\bar{e}_1, \bar{e}) = f((1, 0, 0), (1, 0, 0)) = 1 \cdot 1 - 4 \cdot 1 \cdot 0 + 6 \cdot 0 \cdot 0 - 0 \cdot 1 - 3 \cdot 0 \cdot 0 = 1$$

De manera similar y teniendo en cuenta que $\bar{e}_2 = (0, 1, 0)$, $\bar{e}_3 = (0, 0, 1)$ se obtiene:

$$f(\bar{e}_1, \bar{e}_2) = -4,\quad f(\bar{e}_1, \bar{e}_3) = 0,\quad f(\bar{e}_2, \bar{e}_1) = 0,\quad f(\bar{e}_2, \bar{e}_2) = 0$$
$$f(\bar{e}_2, \bar{e}_3) = 6,\quad f(\bar{e}_3, \bar{e}_1) = -1,\quad f(\bar{e}_3, \bar{e}_2) = 0,\quad f(\bar{e}_3, \bar{e}_3) = -3$$

luego la matriz asociada A tendrá la forma: $\begin{pmatrix} 1 & -4 & 0 \\ 0 & 0 & 6 \\ -1 & 0 & -3 \end{pmatrix}$

y la expresión analítica de la forma es: $f(\bar{x}, \bar{y}) = (x_1\ x_2\ x_3)\begin{pmatrix} 1 & -4 & 0 \\ 0 & 0 & 6 \\ -1 & 0 & -3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$.

Ejemplo 7.1.3

Sea $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ la matriz que caracteriza a una forma bilineal en la base $B = \{\bar{e}_1, \bar{e}_2\}$ con $\bar{e}_1 = (1, 0)$, $\bar{e}_2 = (1, 1)$. Determínese la matriz A' que caracteriza a dicha forma bilineal en la base $B' = \{\bar{e}'_1, \bar{e}'_2\}$, con $\bar{e}'_1 = (0, 1)$, $\bar{e}'_2 = (1, 2)$, obteniendo previamente la matriz de cambio de base P.

Solución:

Para obtener la matriz de cambio de base, P, hay que encontrar las coordenadas de una base respecto a la otra:

$$(0, 1) = \alpha(1, 0) + \beta(1, 1) \Rightarrow (\alpha, \beta) = (-1, 1)$$
$$(1, 2) = \delta(1, 0) + \gamma(1, 1) \Rightarrow (\delta, \gamma) = (-1, 2)$$

Por tanto: $P = \begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix}$ y $A' = P^t AP = \begin{pmatrix} -1 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$.

Es fácil verificar este resultado calculando la matriz A' directamente, es decir, hallando las imágenes de $\bar{e}'_1$ y $\bar{e}'_2$.

- Recordemos que las coordenadas de $\bar{e}_1 = (1, 0)$ respecto a la base canónica son $(1, 0)_B$ (respecto a la base B) y las coordenadas respecto a B' son:

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

Las coordenadas de $\bar{e}_2 = (1, 1)$ respecto a la base canónica son $(0, 1)_B$ (respecto a B) y las coordenadas respecto a B' son:

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

7.1.6. Definición de formas bilineales simétricas

$f: E \times E \to \mathbb{R}$ es forma bilineal simétrica $\Leftrightarrow f(\bar{x}, \bar{y}) = f(\bar{y}, \bar{x}), \forall (\bar{x}, \bar{y}) \in E \times E.$

7.1.7. Definición de formas bilineales antisimétricas

$f: E \times E \to \mathbb{R}$ es forma bilineal antisimétrica $\Leftrightarrow f(\bar{x}, \bar{y}) = -f(\bar{y}, \bar{x}), \forall (\bar{x}, \bar{y}) \in E \times E.$

Si el espacio vectorial E es $\mathbb{R}^n$ y la matriz asociada A, es simétrica, la forma bilineal es simétrica real.

Un caso particular de aplicaciones bilineales simétricas es el producto escalar.

7.1.8. Teorema

La condición necesaria y suficiente para que una forma bilineal sea simétrica es que la matriz asociada correspondiente sea simétrica.

Demostración:

$\Rightarrow$) En efecto, si f es simétrica, de la definición se deduce que:

$f(\bar{e}_i, \bar{e}_j) = f(\bar{e}_j, \bar{e}_i)$ $(i, j = 1, 2, \ldots, n)$, es decir, $a_{ij} = a_{ji}$

pero esta es la condición para que la matriz $A = (a_{ij})$ sea simétrica.

$\Leftarrow$) Recíprocamente, toda forma bilineal asociada a una matriz simétrica A es simétrica:

La expresión $X^t AY$ es un elemento de $\mathbb{R}$, y por tanto, la matriz que lo tiene como único elemento es igual a su traspuesta: $X^t AY = (X^t AY)^t = Y^t A^t X = Y^t AX \Rightarrow f(\bar{x}, \bar{y}) = f(\bar{y}, \bar{x})$. Obsérvese que este resultado es independiente de la base elegida.

7.1.9. Teorema

La condición necesaria y suficiente para que una forma bilineal sea antisimétrica es que la matriz asociada correspondiente sea antisimétrica.

La demostración de este teorema es similar a la anterior y se deja como ejercicio para el lector.

7.2. Formas cuadráticas

Dada la forma bilineal $f: E \times E \to \mathbb{R}$, si es simétrica, genera un tipo especial de formas, las formas cuadráticas.

7.2.1. Definición de forma cuadrática

$Q: E \to \mathbb{R}$ es la forma cuadrática, generada por la forma bilineal simétrica $f: E \times E \to \mathbb{R} \Leftrightarrow$ es la aplicación $Q: E \to \mathbb{R}$, que verifica $Q(\overline{x}) = f(\overline{x}, \overline{x})$, $\forall \overline{x} \in E$.

Consecuencia inmediata de la definición dada y de la de las formas bilineales, es que dada una base del espacio vectorial E, $B = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_n\}$ y considerando la matriz A simétrica, cuyos elementos son $a_{ij} = f(\overline{e}_i, \overline{e}_j)$, se tendrá: $Q(\overline{x}) = X^t AX$, donde, como es habitual, X es la matriz columna formada por las coordenadas del vector $\overline{x} \in E$ en la base B.

Desarrollando la expresión anterior obtenemos:

$$Q(\overline{x}) = \sum_{i,j}^{n} a_{ij}x_i x_j = \sum_{i}^{n} a_{ii}x_i^2 + 2\sum_{i \neq j}^{n} a_{ij}x_i x_j.$$

Se trata de un polinomio homogéneo de segundo grado respecto de las variables $x_1, x_2, \ldots, x_n$, con coeficientes reales.

Como se puede observar fácilmente, las matrices asociadas a una forma bilineal simétrica y a su forma cuadrática son iguales.

Ejemplo 7.2.1

Dada la forma bilineal simétrica real $f: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, tal que,

$$f(\overline{x}, \overline{y}) = x_1y_1 - x_1y_2 - x_2y_1 + 2x_1y_3 + 2x_3y_1 + x_2y_2 - 3x_2y_3 - 3x_3y_2 + 2x_3y_3.$$

Su forma cuadrática asociada es: $Q: \mathbb{R}^3 \to \mathbb{R}$

$$Q(\overline{x}) = x_1^2 - 2x_1x_2 + 4x_1x_3 + x_2^2 - 6x_2x_3 + 2x_3^2$$

Su matriz asociada es: $A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 2 \end{pmatrix}$.

Obsérvese que Q se puede obtener tanto a partir de f, como A a partir de $Q(\overline{x})$.

Los pasos seguidos son:

1. Sustituir en $f(\overline{x}, \overline{y})$ las coordenadas (y_1, y_2, y_3) por las coordenadas (x_1, x_2, x_3).
2. Para obtener A se desdoblan los términos rectangulares y se colocan en el lugar correspondiente de la matriz. Es decir, como la matriz es simétrica, el término $-6x_2x_3 = -3x_2x_3 - 3x_3x_2 \Rightarrow$ en la matriz A será $a_{32} = a_{23} = -3$.

7.2.2. Rango de una forma cuadrática

Rango de una forma cuadrática $Q(\overline{x})$ = Rango de la matriz A asociada a dicha forma.

La expresión de una forma bilineal depende de la base elegida en el espacio vectorial, lo mismo ocurre con la forma cuadrática generada por ella.

7.2.3. Definición de forma cuadrática canónica

Q es forma cuadrática canónica si es de la forma $Q(\overline{x}) = \sum_{i=1}^{n} k_i x_i^2$

Es evidente que la matriz correspondiente a una forma cuadrática canónica es una matriz diagonal. Según lo visto en el capítulo 6, por ser A una matriz simétrica existe un cambio de base definido por una matriz ortogonal que transforma la matriz simétrica A en una matriz diagonal A'. Es decir, cualquiera que sea una forma cuadrática, se puede reducir a una suma de cuadrados (forma canónica) mediante una transformación ortogonal.

Dado que existen diferentes transformaciones ortogonales que reducen una forma cuadrática a una forma canónica nos podemos preguntar si la transformación elegida influye en dicha forma reducida.

El **teorema de Sylvester** (también denominado ley de inercia) da solución a esta pregunta.

7.2.4. Teorema de Sylvester

Al transformar una forma cuadrática $X'AX$ en suma de cuadrados, el número de estos cuadrados, y el de los términos afectados de coeficientes positivos y negativos, son independientes de la base utilizada.

Demostración:

En efecto, los coeficientes k_i son los términos de la diagonal principal de la matriz diagonal $A' = P' AP$, siendo P la matriz de paso del cambio de base; dichos coeficientes coinciden con las raíces características λ_i de A' y con las de A, pues no se modifican con un cambio de base. Por tanto, las diversas reducciones de la forma cuadrática $X' AX$ corresponden a diferentes ordenaciones de los coeficientes λ_i, lo que no altera el número de ellos, ni el signo de los que sean distintos de cero.

7.2.5. Definición de signatura

Signatura de la forma cuadrática Q = Número de coeficientes $k_i > 0$.

7.2.6. Formas cuadráticas definidas y semidefinidas positivas

Una forma cuadrática Q es definida positiva si $Q(\overline{x}) = X'AX > 0$, $\forall\ \overline{x} \neq \overline{0}$.

Una forma cuadrática Q es semidefinida positiva si $Q(\overline{x}) = X'AX \geq 0$, $\forall\ \overline{x} \neq \overline{0}$.

En general, es imposible comprobar la condición dada para todos los vectores, por ello vamos a buscar criterios que permitan clasificar más cómodamente las formas cuadráticas.

7.2.7. Teorema

La condición necesaria y suficiente para que una forma cuadrática sea definida positiva es que todas las raíces características de la matriz A sean positivas.

Demostración:

$\Rightarrow$) Es necesaria, pues si $Q(\overline{x}) = X' AX > 0$, $\forall\ \overline{x} \neq \overline{0}$ (tomamos como hipótesis que Q es definida positiva) se verifica en su expresión canónica: $X'AX = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2 > 0$ $\forall \overline{x} \in E$ cualesquiera que sean las componentes, no todas nulas del vector $\overline{x}$, lo que exige que todas las λ_i sean positivas.

$\Leftarrow$) Es suficiente dado que todos los valores de λ_i son positivos, y $Q(\overline{x}) =$

$= X^t AX = \sum_i^n \lambda_i x_i^2 > 0 \ \forall \overline{x} \in E \Rightarrow$ La forma cuadrática es definida positiva.

Con la misma técnica se puede demostrar el siguiente teorema:

7.2.8. Teorema

Una condición necesaria y suficiente para que una forma cuadrática sea semidefinida positiva es que la matriz A sea regular y que las raíces no nulas de la ecuación características sean positivas.

7.2.9. Formas cuadráticas definidas y semidefinidas negativas

Una forma cuadrática Q es definida negativa si $Q(\overline{x}) = X^tAX < 0, \ \forall \ \overline{x} \neq \overline{0}$.

Una forma cuadrática Q es semidefinida negativa si $Q(\overline{x}) = X^tAX \leq 0, \ \forall \ \overline{x} \neq \overline{0}$.

Análogas conclusiones tendrán los teoremas *[7.2.7]* y *[7.2.8]* para las formas cuadráticas definidas negativas y semidefinidas negativas, cambiando el signo de las raíces características de la matriz.

Ejemplo 7.2.2

La forma cuadrática Q cuya expresión es: $Q(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2$, tiene como matriz asociada $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

Calculando sus autovalores

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 1 = 1 - 2\lambda + \lambda^2 - 1 = \lambda(\lambda - 2) = 0 \Rightarrow$$

$$\Rightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 2 \end{cases}$$

Su matriz diagonal es: $A' = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, por tanto, como los autovalores son 0 y 2, se trata de una forma cuadrática semidefinida positiva.

7.2.10. Reducción a la forma de Jacobi

Toda forma cuadrática Q, cuya matriz es $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$, con

$$A_1 = a_{11} \neq 0; A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} \neq 0; \dots; A_{n-1} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1(n-1)} \\ a_{12} & a_{22} & \dots & a_{2(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1(n-1)} & a_{2(n-1)} & \dots & a_{(n-1)(n-1)} \end{vmatrix} \neq 0$$

se puede escribir en la forma de Jacobi: $A_1 y_1^2 + \frac{A_2}{A_1} y_2^2 + \dots + \frac{|A|}{A_{n-1}} y_n^2$, mediante el cambio de base $X = TY$, donde $T = \begin{pmatrix} 1 & t_{12} & \dots & t_{1n} \\ 0 & 1 & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ e $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

La demostración de esta afirmación es larga y no viene al caso, pero el lector curioso la puede hacer por inducción.

Al observar la reducción de formas cuadráticas a la forma de Jacobi, es fácil establecer un criterio para clasificarlas sin necesidad de calcular sus raíces características. Dicho criterio es el que encierra la siguiente conclusión:

7.2.11. Conclusión: Criterio de Sylvester

1. La condición necesaria y suficiente para que la forma Q sea definida positiva es que se verifique $A_1 > 0, A_2 > 0, ..., A_{n-1} > 0, |A| > 0$.
2. La condición necesaria y suficiente para que la forma Q sea definida negativa es que se verifique: $A_1 < 0, A_2 > 0, A_3 < 0, ...$
3. Si $|A| = 0$, y se cumplen la condición 1 o la condición 2 hasta el último A_p distinto de cero, la forma Q será semidefinida positiva o semidefinida negativa respectivamente.
4. En otros casos la forma Q será indefinida o de signo variable.

También es posible clasificar las formas cuadráticas mediante los invariantes.

7.2.12. Definición de invariantes

Invariantes de la forma cuadrática $Q: E \to \mathbb{R}$ son las propiedades de cualquier matriz asociada a la forma cuadrática, que no dependen de la base considerada en el espacio vectorial E.

El rango de una forma cuadrática Q es un invariante (el rango de una forma cuadrática Q de E es el rango de la matriz asociada A) tal como se deduce del capítulo 6, ya que, la relación entre las matrices asociadas a E en un cambio de base es una congruencia (caso particular de equivalencia de matrices).

Otro invariante importante es la signatura (recordemos que la signatura de una forma cuadrática es el número de términos positivos que figuran en cualquiera de las matrices asociadas que posee). El teorema de Sylvester demuestra esta importante propiedad.

Con lo dicho hasta aquí, podemos dar un criterio de clasificación de las formas cuadráticas según los valores de sus invariantes:

7.2.13. Clasificación de formas cuadráticas por invariantes

Sea E un espacio vectorial de dimensión n, $\text{rg}(Q)$ el rango de la forma cuadrática $Q: E \to \mathbb{R}$, y $\text{sig}(Q)$ la signatura de dicha forma cuadrática, entonces:

1. Q es definida positiva si: $\text{rg}(Q) = \text{sig}(Q) = n$.
2. Q es definida negativa si: $\text{rg}(Q) = n$ y $\text{sig}(Q) = 0$.
3. Q es semidefinida positiva si: $\text{rg}(Q) = \text{sig}(Q) < n$.
4. Q es semidefinida negativa si: $\text{rg}(Q) < n$ y $\text{sig}(Q) = 0$.
5. Q es indefinida positiva si: $\text{rg}(Q) \neq \text{sig}(Q)$ y $\text{sig}(Q) \neq 0$.

EJERCICIOS

Ejercicio 7.1

Dada una aplicación bilineal $f: V_1 \times V_2 \to \mathbb{R}$, con $\bar{v} \in V_1$, $\bar{u} \in V_2$, tal que: $\bar{v} = \sum_{i=1}^{n} \alpha_i \bar{v}_i$ y $\bar{u} = \sum_{j=1}^{m} \beta_j \bar{u}_j$. Demuéstrese que: $f(\bar{v}, \bar{u}) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j f(\bar{v}_i, \bar{u}_j)$.

Solución:

En efecto,

$f(\bar{v}, \bar{u}) = f\left(\bar{v}, \sum_{j=1}^{m} \beta_j \bar{u}_j\right)$, pero por ser f bilineal, entonces: $f(\bar{v}, \bar{u}) = \sum_{j=1}^{m} \beta_j f(\bar{v}, \bar{u}_j)$,

llevando el valor de $\bar{v}$ a la expresión anterior: $f(\bar{v}, \bar{u}) = \sum_{j=1}^{m} \beta_j f\left(\sum_{i=1}^{m} \alpha_i \bar{v}_i, \bar{u}_j\right)$ aplicando de nuevo el carácter de aplicación bilineal, resulta:

$$f(\bar{v}, \bar{u}) = \sum_{j=1}^{m} \beta_j \sum_{i=1}^{n} \alpha_i f(\bar{v}_i, \bar{u}_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j f(\bar{v}_i, \bar{u}_j).$$

Ejercicio 7.2

Dada la forma bilineal $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, tal que su expresión respecto a la base canónica es: $f((x_1, x_1, x_3), (x_1', x_2', x_3')) = 2x_1x_1' + 3x_1x_2' + x_2x_1' + 6x_1x_3' - 4x_3x_2' - x_3x_3'$.

Obténgase la matriz A' de f en la base $B' = \{(1, 0, 0), (1, 2, 0), (-1, 1, 1)\}$ de $\mathbb{R}^3$.

La matriz A asociada a f respecto a la base canónica es: $A = \begin{pmatrix} 2 & 3 & 6 \\ 1 & 0 & 0 \\ 0 & -4 & -1 \end{pmatrix}$

La expresión analítica de f es $f((x_1, x_2, x_3), (x_1', x_2', x_3')) =$

$$= (x_1 \; x_2 \; x_3) \begin{pmatrix} 2 & 3 & 6 \\ 1 & 0 & 0 \\ 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}.$$

Buscamos la matriz asociada y la expresión analítica de f cuando cambiamos de base en $\mathbb{R}^3$.

Si llamamos $\bar{e}_1' = (1, 0, 0)$, $\bar{e}_2' = (1, 2, 0)$, $\bar{e}_3' = (-1, 1, 1)$ a los vectores de la nueva base B', resulta que:

$$A' = \begin{pmatrix} f(\bar{e}_1', \bar{e}_1') & f(\bar{e}_1', \bar{e}_2) & f(\bar{e}_1, \bar{e}_3') \\ f(\bar{e}_2', \bar{e}_1') & f(\bar{e}_2', \bar{e}_2) & f(\bar{e}_2', \bar{e}_3') \\ f(\bar{e}_3', \bar{e}_1') & f(\bar{e}_3', \bar{e}_2) & f(\bar{e}_3, \bar{e}_3') \end{pmatrix}$$

Si utilizamos la expresión analítica de f para encontrar cada uno de los elementos de la matriz A': $f(\bar{e}_i', \bar{e}_j')$, obtenemos: $A' = \begin{pmatrix} 2 & 8 & 7 \\ 4 & 10 & 5 \\ -1 & -15 & -13 \end{pmatrix}$.

Ejercicio 7.3

Indíquese cuáles de las siguientes aplicaciones corresponden a una forma cuadrática:

(a) $Q: \mathbb{R}^2 \to \mathbb{R}$, tal que, $Q(x_1, x_2) = x_1x_2$
(b) $Q: \mathbb{R}^3 \to \mathbb{R}$, tal que, $Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + x_3$
(c) $Q: \mathbb{R}^3 \to \mathbb{R}$, tal que, $Q(x_1, x_2, x_3) = x_1^2 + x_1x_2 + 2x_2x_3$
(d) $Q: \mathbb{R}^2 \to \mathbb{R}$, tal que, $Q(x_1, x_2) = x_1^2 - x_2^2 + 1$
(e) $Q: \mathbb{R}^3 \to \mathbb{R}$, tal que, $Q(x_1, x_2, x_3) = \operatorname{sen}(x_1^2 + x_2^2 + x_3^2)$

Solución:

Son formas cuadráticas las aplicaciones (a) y (c).

Las aplicaciones (b) y (d), aunque son funciones polinómicas, no son formas cuadráticas porque tienen algún sumando con grado distinto de 2.

La aplicación (e) no es forma cuadrática porque no es una función polinómica.

Y todas las formas cuadráticas son funciones homogéneas de grado dos.

Ejercicio 7.4

Dada la forma cuadrática $Q: \mathbb{R}^3 \to \mathbb{R}$, tal que, $Q(x_1, x_2, x_3) = 2x_1^2 + x_2^2 - x_3^2 - 3x_1x_2 + x_2x_3$, verifíquese que $Q(\overline{x})$ se puede representar matricialmente de la forma: $Q(\overline{x}) = X^t AX$, siendo $A = \begin{pmatrix} 2 & -2 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$.

Solución:

$$X^t AX = (x_1 \; x_2 \; x_3)\begin{pmatrix} 2 & -2 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= \big((2x_1 - x_2)\;(-2x_1 + x_2)\;(x_2 - x_3)\big)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (2x_1 - x_2)x_1 + (-2x_1 + x_2)x_2 +$$

$$+ (x_2 - x_3)x_3 = 2x_1^2 - x_2x_1 - 2x_1x_2 + x_2^2 + x_2x_3 - x_3^2 = 2x_1^2 + x_2^2 - x_3^2 - 3x_1x_2 + x_2x_3$$

Ejercicio 7.5

Determínese el conjunto de matrices A, para las que se verifica $Q(\overline{x}) = X^t AX$, referidas al ejercicio 7.4, para todo $\overline{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$.

Solución:

$$X^t AX = (x_1 \; x_2 \; x_3)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= (x_1a_{11} + x_2a_{21} + x_3a_{31} \quad x_1a_{12} + x_2a_{22} + x_3a_{32} \quad x_1a_{13} + x_2a_{23} + x_3a_{33})\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + (a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + (a_{23} + a_{32})x_2x_3 =$$
$$= 2x_1^2 + x_2^2 - x_3^2 - 3x_1x_2 + x_2x_3$$

Luego debe ser:

$$a_{11} = 2,\ a_{22} = 1,\ a_{33} = -1$$

Además:

$$a_{12} + a_{21} = -3,\ a_{13} + a_{31} = 0,\ a_{23} + a_{32} = 1$$

condiciones que verifica la matriz A dada en el ejercicio anterior, pero existen infinitas matrices que pueden representar la forma cuadrática Q.

Por tanto el conjunto pedido es:

$$\left\{\begin{pmatrix} 2 & a_{12} & a_{13} \\ -3-a_{12} & 1 & 1-a_{32} \\ -a_{13} & a_{32} & -1 \end{pmatrix}\right\} \text{ con } a_{12}, a_{13}, a_{32} \in \mathbb{R}$$

Una matriz de este conjunto es:

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

Ejercicio 7.6

Del conjunto de matrices obtenidas en el ejercicio 7.5.

a) Obténgase la única matriz simétrica B.

b) Compruébese que $B = 1/2\ (A + A^t)$, siendo A la matriz dada en el ejercicio 7.5.

Solución:

a) Para que la matriz sea simétrica se debe verificar:

$$a_{12} = a_{21},\ \ a_{13} = a_{31},\ \ a_{23} = a_{32}$$

además para que pertenezca al conjunto, se debe cumplir que:

$$a_{12}+a_{21}=-3,\quad a_{13}+a_{31}=0,\quad a_{12}+a_{32}=1$$

luego:

$$a_{12}=a_{21}=\frac{-3}{2},\quad a_{13}=a_{31}=0,\quad a_{23}=a_{32}=\frac{1}{2}$$

o sea que la matriz pedida resulta ser:

$$B=\begin{pmatrix}2 & -3/2 & 0\\ -3/2 & 1 & 1/2\\ 0 & 1/2 & -1\end{pmatrix}$$

b)

$$\frac{1}{2}(A+A')=\frac{1}{2}\left(\begin{pmatrix}2 & -2 & 0\\ -1 & 1 & 1\\ 0 & 0 & -1\end{pmatrix}+\begin{pmatrix}2 & -1 & 0\\ -2 & 1 & 0\\ 0 & 1 & -1\end{pmatrix}\right)=\frac{1}{2}\begin{pmatrix}4 & -3 & 0\\ -3 & 2 & 1\\ 0 & 1 & -2\end{pmatrix}=$$

$$=\begin{pmatrix}2 & -3/2 & 0\\ -3/2 & 1 & 1/2\\ 0 & 1/2 & -1\end{pmatrix}=B$$

Se llega al mismo resultado, cualquiera sea la matriz A, que represente la forma cuadrática dada en el ejercicio 7.5.

Ejercicio 7.7.

Hállense el rango y la signatura de las formas cuadráticas cuyas expresiones son:

a) $Q(x_1,x_2)=2x_1^2+3x_1x_2+6x_2^2$
b) $Q(x_1,x_2,x_3)=x_1^2+2x_2^2+4x_2x_3+2x_3^2$
c) $Q(x_1,x_2,x_3)=3x_1^2+2x_1x_2+2x_1x_3+4x_2x_3$

Solución:

a) La matriz asociada a la forma de expresión $Q(x_1,x_2)=2x_1^2+3x_1x_2+6x_2^2$, es:

$$A=\begin{pmatrix}2 & 3/2\\ 3/2 & 6\end{pmatrix};\ |A-\lambda I|=\begin{vmatrix}2-\lambda & 3/2\\ 3/2 & 6-\lambda\end{vmatrix}=\lambda^2-8\lambda+\frac{39}{4}=0$$

Los autovalores son:

$\lambda_1=13/2,\ \lambda_2=3/2\Rightarrow \mathrm{sig}(Q)=2$
$|A|\neq 0\Rightarrow \mathrm{rg}(Q)=2$

b) La matriz asociada a la forma de expresión $Q(x_1,x_2)=x_1^2+2x_2^2+4x_2x_3+2x_3^2$, es:

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

$$|A-\lambda I|=\begin{vmatrix}1-\lambda & 0 & 0\\ 0 & 2-\lambda & 2\\ 0 & 2 & 2-\lambda\end{vmatrix}=(2-\lambda)^2(1-\lambda)-4(1-\lambda)=(1-\lambda)\,\lambda\,(\lambda-4)=0$$

Los autovalores son:

$$\lambda_1=1,\quad \lambda_2=0,\quad \lambda_3=4\Rightarrow \mathrm{sig}(Q)=2$$

además, $\mathrm{rg}(Q)=2$.

c) Matriz asociada a la forma de expresión $Q(x_1,x_2,x_3)=3x_1^2+2x_1x_2+2x_1x_3+4x_2x_3$: $A=\begin{pmatrix}3 & 1 & 1\\ 1 & 0 & 2\\ 1 & 2 & 0\end{pmatrix}$

$$|A-\lambda I| = \begin{pmatrix} 3-\lambda & 1 & 1 \\ 1 & -\lambda & 2 \\ 1 & 2 & -\lambda \end{pmatrix} = \lambda^2(3-\lambda)+4+2\lambda-4(3-\lambda) =$$
$$= (3-\lambda)(\lambda^2-4)+2(\lambda+2) = (3-\lambda)(\lambda+2)(\lambda-2)+2(\lambda+2) =$$
$$= (\lambda+2)(-\lambda^2+5\lambda-4) = 0$$

Luego los autovalores son: $\lambda_1 = -2$; $\lambda_2 = 1$; $\lambda_3 = 4 \Rightarrow \mathrm{sig}(Q) = 2$, además $\mathrm{rg}(Q) = 3$.

Ejercicio 7.8

Clasifíquense las formas cuadráticas del ejercicio anterior.

Solución:

A partir de los resultados obtenidos en el ejercicio anterior, llegamos a que:

a) La forma cuadrática es definida positiva, pues:

$$\mathrm{rg}(Q) = \mathrm{sig}(Q) = 2 = n$$

b) La forma cuadrática es semidefinida positiva, pues:

$$\mathrm{rg}(Q) = \mathrm{sig}(Q) = 2 < n = 3$$

c) La forma cuadrática es indefinida, pues existen autovalores positivos y autovalores negativos.

Ejercicio 7.9

Clasifíquese la forma cuadrática asociada a la matriz: $A = \begin{pmatrix} 1 & x & -x \\ x & 1-x & 0 \\ -x & 0 & 1+x \end{pmatrix}$, para los distintos valores de $x \in \mathbb{R}$.

Solución:

$$|A-\lambda I| = \begin{vmatrix} 1-\lambda & x & -x \\ x & (1-x)-\lambda & 0 \\ -x & 0 & (1+x)-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & x & -x \\ x & (1-\lambda)-x & 0 \\ -x & 0 & (1-\lambda)+x \end{vmatrix} =$$

$$= (1-\lambda)\,[(1-\lambda)-x][(1-\lambda)+x] - x^2[(1-\lambda)-x] - x^2\,[(1-\lambda)+x] =$$
$$= (1-\lambda)\,[(1-\lambda)^2-x^2] - 2x^2(1-\lambda) = (1-\lambda)\,[\lambda^2-2\lambda+(-3x^2+1)] = 0$$

Por tanto es:

$$\lambda_1 = 1; \quad \lambda_2 = 1+\sqrt{3}x; \quad \lambda_3 = 1-\sqrt{3}x$$

Para que la forma cuadrática sea definida positiva debe ser: $\mathrm{rg}(Q) = \mathrm{sig}(Q) = n$.

Luego:

$$\lambda_2 = 1+\sqrt{3}x > 0 \Rightarrow x > \frac{-1}{\sqrt{3}}$$

$$\lambda_3 = 1-\sqrt{3}x > 0 \Rightarrow x < \frac{1}{\sqrt{3}}$$

La forma cuadrática resulta definida positva $\forall\, x \in \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$, ya que, $\mathrm{rg}(Q) = \mathrm{sig}(Q) = n = 3$.

Si $x = \frac{1}{\sqrt{3}}$ ó $x = -\frac{1}{\sqrt{3}}$ la forma cuadrática resulta semidefinida positiva, pues $\mathrm{rg}(Q) = \mathrm{sig}(Q) = 2$ y $n = 3$.

Para los restantes valores de x la forma cuadrática resulta ser indefinida.

Ejercicio 7.10

Estúdiese mediante el criterio de Sylvester el carácter de la forma cuadrática expresada por: $Q(\overline{x}) = x^2 - 2y^2 - z^2 + 2xy + xz$.

Solución:

La matriz asociada a Q es:

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

$$A_1 = 1 > 0;\ A_2 = \begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = -3 < 0;\ |A| = 2 + \frac{1}{2} + 1 = \frac{7}{2}$$

Por tanto, se trata de una forma cuadrática semidefinida.

CAPÍTULO 8

EL PROBLEMA DE LA PROGRAMACIÓN LINEAL

INTRODUCCIÓN

Aunque con la Revolución Industrial ya se buscaban soluciones óptimas a diversos problemas, no es hasta el siglo XX cuando la programación matemática experimenta un mayor desarrollo, debido inicialmente a motivos estratégicos. Ya durante la Primera Guerra Mundial existían diversos grupos, integrados a menudo por científicos, dedicados al análisis de las operaciones tácticas y estratégicas. Por ejemplo, se encargó a Thomas Edison el desarrollo de estrategias para disminuir las pérdidas de navíos mercantes ocasionadas por submarinos, a partir del estudio de sus maniobras. Durante la Segunda Guerra Mundial un equipo de científicos fue el encargado de estudiar y desarrollar tácticas asociadas con la defensa de Inglaterra; su objetivo era determinar la utilización más efectiva de sus limitados recursos militares. La creación de este equipo propició los primeros trabajos formales sobre lo que luego se denominaría la "programación lineal".

Fue Gran Bretaña la iniciadora de los trabajos, pero el liderazgo fue pronto asumido por los EE. UU., donde la Rand Corporation comisionó a un grupo de matemáticos para resolver problemas de logística del ejército. Fue en este período cuando se acuñó el término Programación Lineal. La palabra "programación" proviene de las aplicaciones a problemas que tratan de programar suministros o distribución de cantidades y el término "lineal" indica que las ecuaciones e inecuaciones son lineales.

Por otra parte, la aparición de los ordenadores propició el desarrollo de técnicas y algoritmos para la resolución de este tipo de problemas, contribuyendo al desarrollo de la programación matemática por motivos económicos o industriales.

La más importante técnica matemática específica en este campo fue la denominada "Método Simplex de programación lineal" que fue dada a conocer en 1947 por el matemático norteamericano George B. Dantzig. Se trata de un procedimiento iterativo algebraico que resuelve con exactitud cualquier problema de programación lineal en un número finito de pasos, también indica los casos en que hay solución no acotada o simplemente no existe solución.

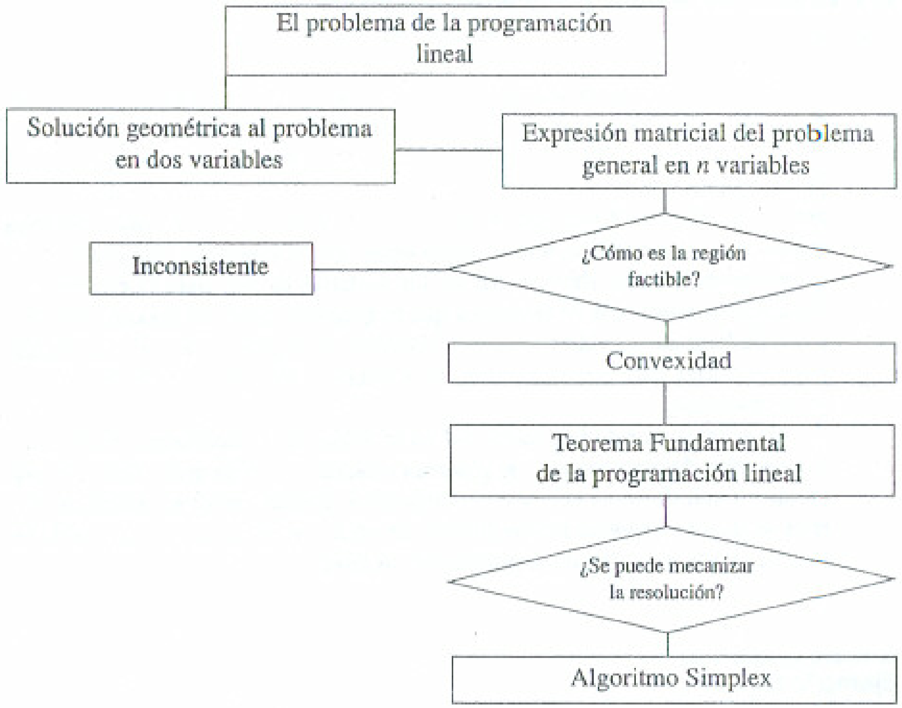
Desde entonces se han desarrollado nuevas técnicas y aplicaciones gracias al esfuerzo y la cooperación entre las instituciones académicas y el mundo empresarial que rápidamente vio la importancia y la eficacia de estas técnicas en la resolución de problemas de índole económico, sobre todo en el campo de la organización y la planificación.

GEORGE DANTZIG

George Dantzig nació el 8 de noviembre de 1914 en Portland, Estados Unidos. Estudió Matemáticas en las Universidades de Maryland y Michigan. Terminó sus estudios en 1937 y desde este año hasta 1939 trabajó en la oficina norteamericana de Estadísticas Laborales. Durante la Segunda Guerra Mundial trabajó en el Cuartel General de Control Estadístico estadounidense. En 1946 se doctoró en Matemáticas en la Universidad de Berkeley en California y fue nombrado asesor matemático del Ejército estadounidense. Desde 1952 trabajó como investigador en RAND Corporation y en diversas Universidades (Berkeley o Stanford). Ha sido distinguido con diversos galardones.

Su mayor aportación a las Matemáticas la hizo en 1947, al desarrollar el método simplex mientras trabajaba en el Ejército. Se convirtió en experto en métodos de planificación resueltos con calculadoras de bolsillo, técnica conocida con el nombre militar de programación. Entonces se refería a problemas de organización de maniobras, suministros logísticos o movimiento de tropas. Dantzig mecanizó el proceso introduciendo la programación lineal. La importancia de estos métodos fue destada por diversos autores, señalando no sólo su impacto en la industria y la economía, sino también que gran parte del tiempo empleado por los ordenadores del mundo está dedicado a programación lineal. Sin embargo, él modestamente escribió "El tremendo poder del método simplex es para mí una sorpresa constante". Tenemos que señalar que, a pesar del tiempo que ha transcurrido desde que se desarrolló el método simplex, el problema original todavía es tema de investigación.

CAPÍTULO 8


El problema de la programación lineal
Solución geométrica al problema en dos variables
Expresión matricial del problema general en n variables
Inconsistente
¿Cómo es la región factible?
Convexidad
Teorema Fundamental de la programación lineal
¿Se puede mecanizar la resolución?
Algoritmo Simplex

8.1 El problema de la programación lineal

Cuando nos enfrentamos a problemas reales, observamos que normalmente los problemas que se pueden plantear como un sistema de ecuaciones lineales (cuya resolución, como ya se ha visto a lo largo de este libro, nos la proporciona el Álgebra) son muchos. Pero es más frecuente que nos encontremos problemas donde interese maximizar o minimizar alguna o algunas cantidades. Por ejemplo, una empresa debe seleccionar a sus proveedores para minimizar costes, debe distribuir las materias primas para la elaboración de diversos productos de tal forma que la ganancia total sea máxima o si dispone de láminas de plástico para la elaboración de placas de varios tamaños, debe cortarlas para que se aproveche al máximo el material y los beneficios sean máximos.

Problemas de este tipo surgieron en relación con la Economía, Medicina o Ingenierías, donde a menudo se plantean problemas con un gran número de variables y con diversas condiciones impuestas a estas variables. Estos problemas se pueden expresar matemáticamente como un conjunto de desigualdades lineales y son el origen de la programación lineal.

Ejemplo 8.1.1

Una fábrica de CD-ROM produce dos tipos distintos de discos grabables: de 74 y 80 minutos. Para la fabricación de cada CD, éste debe pasar por tres máquinas distintas, en el mismo orden, que llamaremos A, B y C. Las máquinas A y B tienen un bajo coste de mantenimiento, por lo que su tiempo anual (máximo) de funcionamiento depende de la disponibilidad de los operarios de cada una de ellas. Por el contrario, el mantenimiento de la máquina C es muy costoso, por lo que es necesario que esté funcionando un número mínimo de minutos. El máximo anual del tiempo de funcionamiento en esta máquina está condicionado, por los minutos que hay en un año. Estos datos están, obviamente, incluidos en la siguiente tabla, donde además se muestra el número de minutos que necesita cada máquina para fabricar un CD de cada uno de los tipos.

	Minutos empleados en la fabricación		Minutos de funcionamiento
	74 min.	80 min.	
Máquina A	7	1	≤ 230400
Máquina B	3	3	≤ 259200
Máquina C	1	2	≥ 115200 ≤ 525600

Además, sabemos que la venta de cada CD de 74 minutos deja un margen de beneficio de 13 céntimos y el margen que proporciona la venta de cada uno de 80 minutos es de 12 céntimos de euros, ya que ambos se venden al mismo precio y los costes para elaborar un CD de 74 minutos son menores. Teniendo en cuenta estos datos, se quiere determinar cuántos discos de cada tipo hay que fabricar al año para maximizar el beneficio total.

Una primera etapa para la resolución del problema es su formulación en términos matemáticos. Hacemos notar que todos los problemas de este tipo presentan unas características comunes. En primer lugar, tenemos una función (en el ejemplo anterior, beneficio) cuyos valores extremos (máximo, en este caso) queremos determinar. Por otra parte, está sujeta a ciertas restricciones, que en este caso son el tiempo que emplea cada una de las máquinas en la fabricación de cada tipo de CD y los minutos que está funcionando cada una al año.

En resumen, el problema principal de la programación lineal es optimizar una función lineal, lo que significa calcular los valores <u>óptimos</u> (máximos o mínimos) de una función lineal, llamada <u>función objetivo</u>, sujeta a ciertas <u>restricciones</u> (igualdades o desigualdades lineales). Introduciendo una formulación matemática, el problema se traduce en calcular máximos o mínimos de la función lineal:

óptimo $F = c_1x_1 + c_2x_2 + \ldots + c_nx_n$ (1)

si se imponen las restricciones:

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \ldots \quad \ldots \quad \ldots \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n = b_s \\ a_{s+1,1}x_1 + a_{s+1,2}x_2 + \ldots + a_{s+1,n}x_n \le b_{s+1} \\ \ldots \quad \ldots \quad \ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n \le b_r \\ a_{r+1,1}x_1 + a_{r+1,2}x_2 + \ldots + a_{r+1,n}x_n \ge b_{r+1} \\ \ldots \quad \ldots \quad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \ge b_m \end{array}\right. \qquad (2)$$

siendo las b_i no negativas, y $x_1 \ge 0$, $x_2 \ge 0$, ..., $x_n \ge 0$ (3)

Señalamos que tanto la función F como las desigualdades son expresiones lineales de las variables; éste es el origen del término "lineal". En general, las desigualdades del tipo $\le$ suelen representar limitaciones de recursos. Las del tipo $\ge$ normalmente significan condiciones técnicas, legales o relativas a la calidad del servicio. Podemos suponer, sin pérdida de generalidad, que $b_i \ge 0$, ya que en caso de no serlo, basta multiplicar la desigualdad por -1.

En este punto, y como está siendo habitual a lo largo de este libro, como paso previo para poder resolver matemáticamente este problema, es necesario unificar notación y definir formalmente diversos conceptos (ya hemos utilizado algunos implícitamente).

8.1.1. Definición de función objetivo

Función objetivo es la función lineal F que se quiere optimizar.

8.1.2. Definición de restricciones

Restricciones son las desigualdades lineales (2) y (3) que deben verificar las soluciones de un problema de programación lineal.

Estamos trabajando con puntos de la forma $(x_1, x_2, ..., x_n)$, que pueden ser considerados vectores de $\mathbb{R}^n$. Claramente existe una región en $\mathbb{R}^n$ que verifica cada una de las restricciones. Por lo tanto, la intersección de todas ellas, será el conjunto en el que hay que buscar las posibles soluciones que optimizan el problema. Definimos así solución factible, conjunto factible y solución óptima.

8.1.3. Definición de solución factible

Los puntos $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$ que satisfacen las restricciones (2) y (3) son denominados soluciones factibles.

8.1.4. Definición de conjunto factible

Conjunto factible es el conjunto de $\mathbb{R}^n$ cuyos puntos $(x_1, x_2, ..., x_n)$ son soluciones que satisfacen las restricciones (2) y (3). También es denominado región de posibilidad o convexo solución.

El conjunto factible puede ser vacío si las restricciones son incompatibles. En este caso, el problema de programación lineal no tiene solución. Además, puede ser acotado o no acotado. Intuitivamente, podemos decir que en el plano o en el espacio, un conjunto es acotado si está contenido en un círculo o en una esfera de radio R, respectivamente, y es no acotado, en caso contrario.

Por otra parte, se puede suponer que no hay restricciones redundantes o superfluas, es decir, que no hay restricciones cuya eliminación no significa una modificación del conjunto factible. Más adelante explicaremos por qué el conjunto factible se llama también convexo solución.

8.1.5. Definición de solución óptima

Solución óptima es aquella solución factible que optimiza (maximiza o minimiza) la función objetivo.

La solución óptima no tiene por qué ser única. Como veremos, un problema de programación lineal puede alcanzar las soluciones óptimas en un segmento o en otro conjunto.

Vamos a explicar estos conceptos en un ejemplo.

Ejemplo 8.1.2

Continuando con el ejemplo *[8.1.1]*, sean x e y el número de discos de 74 y 80 minutos que dicha fábrica produce al año, respectivamente. El beneficio total, en céntimos de euro es

$$F = 0.13x + 0.12y$$

que es la función objetivo, que queremos maximizar. Además, los minutos empleados para la fabricación de estos discos en las máquinnas A, B y C serán

$$\begin{array}{r} 7x + y \\ 3x + 3y \\ x + 2y \end{array}$$

Teniendo en cuenta el número máximo de minutos que pueden estar en funcionamiento cada una de estas máquinas al año, se deben satisfacer las restricciones

$$\begin{array}{r} 7x + y \leq 230400 \\ 3x + 3y \leq 259200 \end{array}$$

$$\begin{array}{r} x + 2y \leq 525600 \\ x + 2y \geq 115200 \end{array}$$

Obviamente, las cantidades de discos que se fabrican son mayores o iguales que 0, por lo que también se debe verificar

$$\begin{array}{r} x \geq 0 \\ y \geq 0 \end{array}$$

8.2. Resolución geométrica del problema de la programación lineal

Aunque los problemas de programación lineal pueden tener ciertos, e incluso miles, de variables, si consideramos problemas donde en la función objetivo F sólo intervienen dos variables (como el ejemplo *[8.1.2]*), para su resolución se sigue un procedimiento geométrico que resulta muy intuitivo y que pasamos a describir a continuación. En este caso, $F = c_1x + c_2y$.

Sabemos que la representación gráfica de una ecuación en las variables x e y es una recta en el plano, de ecuación $y = mx + b$. Consecuentemente, una desigualdad del tipo $y \leq mx + b$, con $m \geq 0$ describe los puntos (x_0, y_0) situados en uno de los semiplanos delimitados por la recta anterior (sombreado, en la figura 1). De la misma forma, una desigualdad del tipo $y \geq mx + b$ describe los puntos (x_0, y_0) situados en el otro semiplano (en blanco en la figura 1). A partir de estos argumentos, observamos que cada una de las restricciones $a_{i1}x + a_{i2}y \leq$ $\leq b_i$, $a_{i1}x + a_{i2}y \geq b_i$ tienen como conjunto solución uno de los dos semiplanos delimitados por la recta $a_{i1}x + a_{i2}y = b_i$.

Una forma sencilla de determinar, a partir de la recta correspondiente, cuál de los dos semiplanos es el buscado es elegir un punto en uno de los semiplanos y comprobar si satisface la desigualdad. Por ejemplo, podemos elegir el punto (0, 0) y estudiar si $0 \leq b_i$ o, por el contrario, $0 \geq b_i$. A partir de estas desigualdades, seleccionamos el semiplano correspondiente.

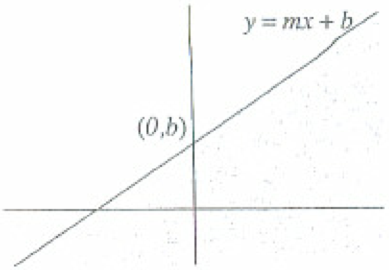


Figura 1

Dado un sistema de desigualdades lineales, considerando dos variables x e y, como los expresados en (2) y (3) de forma general, buscamos el conjunto de puntos que verifique todas las restricciones, es decir, el conjunto factible. Desde el punto de vista intuitivo, claramente la solución óptima debe estar en la región del plano que es la intersección de todos los semiplanos definidos por las desigualdades, ya que debe verificarlas todas simultáneamente.

Ejemplo 8.2.1

Vamos a determinar geométricamente el conjunto del ejemplo *[8.1.2]* Las restricciones son

$$\begin{aligned} 7x + y &\leq 230400 \\ 3x + 3y &\leq 259200 \\ x + 2y &\geq 115200 \\ x + 2y &\leq 525600 \\ x &\geq 0 \\ y &\geq 0 \end{aligned}$$

Los semiplanos correspondientes a las cuatro primeras desigualdades se muestran en la figura 2. Como hemos indicado antes, una forma sencilla de determinar, a partir de la recta correspondiente, cuál de los dos semiplanos es el buscado es elegir un punto en uno de los semiplanos y estudiar si satisface la desigualdad. Por ejemplo, si escogemos (0, 0) observamos que satisface $7x + y$ ≤ 230400, por lo que sombreamos el semiplano en el que está.

Tenemos que señalar que la última de las restricciones es superflua, es decir, su eliminación no modifica la región factible, ya que los puntos del primer cuadrante que verifican la primera restricción también verifican ésta.

Por otra parte, las restricciones $x \geq 0$, $y \geq 0$ significan que las soluciones van a estar en el primer cuadrante, que será el que consideremos. Así, la intersección de los cuatro semiplanos de la figura 2 con este cuadrante nos proporciona la región correspondiente al conjunto factible (figura 3). Observamos que el segmento de recta que une dos puntos cualesquiera del conjunto factible está dentro de dicho conjunto. Esta característica la verifican siempre los conjuntos factibles, como veremos más adelante.

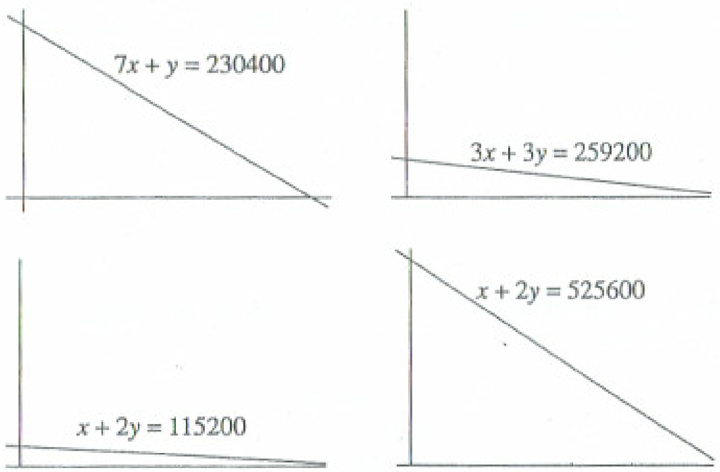


Figura 2

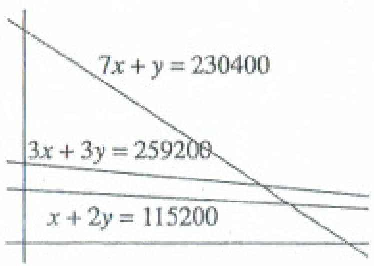


Figura 3

Una vez que sabemos cuál es el conjunto factible de un problema de programación lineal, necesitamos saber cómo "elegir" la solución óptima dentro del conjunto. En este caso, los conjuntos factibles van a estar delimitados por poligonales abiertas o cerradas. Vamos a llamar vértices (concepto que generalizaremos más adelante) a la intersección de las rectas que forman la poligonal. Cuando tenemos los vértices que delimitan el conjunto factible, a partir del siguiente teorema podemos determinar la solución óptima o las soluciones óptimas entre todas las posibles.

8.2.1. Teorema

Si una función objetivo F de un problema de programación lineal definido en el plano alcanza su valor extremo (máximo o mínimo) en una solución factible, entonces este valor extremo se asume en un vértice del conjunto, o en el segmento que une dos vértices

La demostración de este resultado se escapa de los objetivos de esta sección, pero señalamos que se basa en que los valores de la función objetivo son constantes a lo largo de cada recta de $\mathbb{R}^2$ de la forma $c_1x + c_2y = k$, denominadas curvas de nivel. Estas rectas son todas paralelas entre sí y, por lo tanto, los valores extremos se alcanzarán al menos en uno de los vértices del conjunto factible (figura 4). Para una demostración detallada, consultar (Fraleigh y Beauregard, 1989).

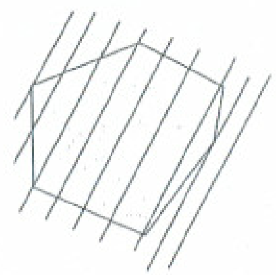


Figura 4

El resultado anterior sugiere el siguiente procedimiento para hallar el máximo o mínimo de una función objetivo, sujeta a unas restricciones:

1. Determinar el conjunto factible.
2. Determinar si se alcanza un valor extremo en este conjunto.
3. En caso de que se alcance, determinar las coordenadas de los vértices del conjunto factible.
4. Calcular el valor de la función objetivo F en los vértices. El valor mayor (o menor) obtenido es el máximo (o, respectivamente, mínimo) de F en el conjunto factible.

El segundo punto es el que, en principio, plantea mayores problemas. En general, podemos afirmar que si el conjunto factible es acotado, entonces la función objetivo alcanza en él su máximo y su mínimo. Si el conjunto no es acotado, entonces puede no alcanzarse ninguno, uno de ellos o los dos, como se verá en los ejercicios.

Ejemplo 8.2.2

Siguiendo con los ejemplos *[8.1.2]* y *[8.2.1]*, los vértices del conjunto son los puntos (0,57600), (0,86400), (24000, 62400) y (26584.6,44307.7), como se aprecia en la figura 5.

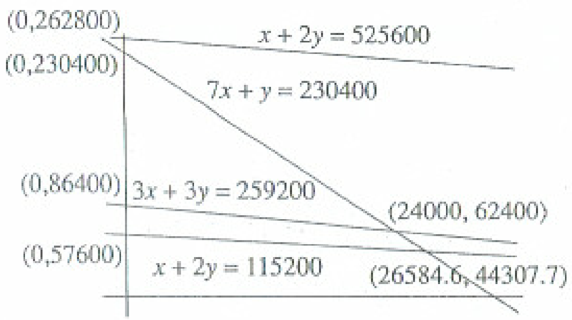


Figura 5

La función factible $F(x, y) = 0.13x + 0.12y$ en los vértices toma los valores

$$F(0,57600) = 6912 \quad F(0,86400) = 10368$$
$$F(24000, 62400) = 10608 \quad F(26584.6, 44307.7) = 8772.9$$

Luego para obtener el máximo beneficio se deben fabricar 24000 discos de 74 minutos y 62400 discos de 80 minutos al año.

8.3. Expresión matricial del problema general en *n* variables

Para problemas donde intervienen dos variables, la resolución geométrica es un método muy intuitivo y relativamente sencillo. Sin embargo, con gran frecuencia se plantean problemas de mayor complejidad, donde el número de variables que intervienen es mayor y su resolución no es posible empleando este método. Para resolverlos, intentaremos reducir nuestro problema a otro cuya resolución conozcamos, técnica habitual en Matemáticas.

Hemos estudiado cómo resolver sistemas de ecuaciones lineales. Por tanto, nos puede interesar transformar las desigualdades en igualdades, ya que así tenemos un método para determinar las soluciones. Esta transformación es posible introduciendo unas nuevas variables, denominadas variables de holgura.

8.3.1. Definición de variables de holgura

Las variables de holgura son unas variables $x_{n+1}, ..., x_{n+m} \geq 0$, que se introducen en el conjunto de restricciones de un problema de programación lineal, de tal forma que las desigualdades se transforman en igualdades.

La introducción de las variables de holgura se realiza, introduciendo la variable x_{n+i} de forma que se obtenga una igualdad, teniendo en cuenta que deben ser siempre mayores o iguales a cero:

$a_{k1}x_1 + a_{k2}x_2 + ... a_{kn}x_n \leq b_k$ es equivalente a $a_{k1}x_1 + a_{k2}x_2 + ... + a_{kn}x_n + x_{n+k} = b_k$
$a_{k1}x_1 + a_{k2}x_2 + ... a_{kn}x_n \geq b_k$ es equivalente a $a_{k1}x_1 + a_{k2}x_2 + ... + a_{kn}x_n - x_{n+k} = b_k$

Obviamente, si las restricciones son igualdades, no es necesario introducir en ellas variables de holgura. Por otra parte, el coeficiente de estas nuevas variables en la función objetivo debe ser siempre $c_{n+i} = 0$ para $i = 1, ..., m$. En una restricción del tipo $\leq$, la variable de holgura significa cantidad de recursos no

utilizada y en una desigualdad del tipo $\geq$ representa el exceso con que las condiciones son satisfechas.

A partir de la introducción de las variables de holgura, las ecuaciones (1), (2) y (3) se pueden escribir como

$$\text{óptimo } F = c_1x_1 + c_2x_2 + ... + c_nx_n \qquad (4)$$

$$\left\{\begin{array}{r} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ ... \qquad ... \qquad ... \\ a_{s1}x_1 + a_{s2}x_2 + ... + a_{sn}x_n = b_s \\ a_{s+1,1}x_1 + a_{s+1,2}x_2 + ... + a_{s+1,n}x_n + x_{n+1} = b_{s+1} \\ ... \qquad ... \qquad ... \\ a_{r1}x_1 + a_{r2}x_2 + ... + a_{rn}x_n + x_{n+r} = b_r \\ a_{r+1,1}x_1 + a_{r+1,2}x_2 + ... + a_{r+1,n} - x_{n+r+1} = b_{r+1} \\ ... \qquad ... \qquad ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n - x_{n+m} = b_m \end{array}\right. \qquad (5)$$

con las b_i no negativas y siendo, de nuevo, $x_1 \geq 0, x_2 \geq 0, ..., x_{n+m} \geq 0$ (6)

8.3.2. Expresión matricial del problema de programación lineal

Incluyendo las variables de holgura, el sistema anterior se puede expresar en forma matricial como

$$\left\{\begin{array}{l} \text{óptimo } F = c_1x_1 + c_2x_2 + ... + c_{n+m}x_{n+m} \\ A\overline{x} = \overline{b} \end{array}\right. \qquad (7)$$

$\overline{x} = (x_1, x_2, ..., x_{n+m})^t$ es el vector de las variables.
$\overline{c} = (c_1, ..., c_{n+m})$ es el vector de costes.
A es una matriz de dimensiones $n \times (n + m)$.
$\overline{b} = (b_1, ..., b_m)^t$ es un vector de n componentes.

Obsérvese que el vector de costes incluyen los coeficientes correspondientes a las variables de holgura y que toman en valor $c_i = 0$ si $i > n$. Por otra parte, las filas de A representan las restricciones y sus columnas las variables totales (es decir, las del problema más las de holgura) y todas las componentes de b son mayores o iguales a cero.

Ejemplo 8.3.1

Dado el siguiente problema de programación lineal, introducir las variables de holgura y expresarlo en forma matricial.

$$\begin{cases} \text{óptimo } F = 3x_1 - x_2 + x_3 \\ x_1 + 2x_2 = 2 \\ x_2 + 4x_3 \leq 8 \\ x_1 - x_2 - x_3 \geq 5 \end{cases}$$

Hay dos desigualdades, por lo que introducimos dos variables de holgura: x_4 y x_5. Las restricciones se pueden expresar como

$$\begin{cases} x_1 + 2x_2 = 2 \\ x_2 + 4x_3 + x_4 = 8 \\ x_1 - x_2 - x_3 - x_5 = 5 \end{cases}$$

La expresión matricial del problema es

$$\begin{cases} \text{óptimo } F = 3x_1 - x_2 + x_3 \\ \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 1 & -1 & -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 5 \end{pmatrix} \end{cases}$$

Observamos que en este problema aparecen no sólo las variables del proceso, sino también las variables de holgura que hemos introducido. Una vez que

hemos transformado el problema real en un sistema ya estudiado y resuelto por el Álgebra Lineal, se va a buscar un algoritmo para su resolución.

Como al introducir las variables de holgura hemos aumentado el número de variables, las definiciones que hemos dado a lo largo del capítulo siguen siendo válidas, con la salvedad de que, en este último caso, las variables consideradas son tanto las iniciales del proceso como las nuevas variables. Resulta conveniente definir ahora otros tipos de soluciones, que se van a utilizar en la resolución del problema general de la programación lineal.

8.3.3. Definición de solución básica

Una solución factible es solución básica si no más de n componentes (m es el número de restricciones) son positivas. Si el número de componentes positivas es m, se dice que es una solución básica no degenerada.

Ejemplo 8.3.2

Si las restricciones son

$$\begin{cases} 2x_1 + x_2 \leq 9 \\ x_1 + x_2 - 2x_3 \geq 3 \end{cases} \quad \text{es equivalente a} \quad \begin{cases} 2x_1 + x_2 + x_4 = 9 \\ x_1 + x_2 - 2x_3 - x_5 = 3 \end{cases}$$

una solución básica no degenerada sería $(0, 0, 0, 9, 3)$.

El nombre de solución básica se debe a que las incógnitas que toman valores distintos de 0 en la solución básica, van a formar una base del problema. Como veremos más adelante, disponer de una solución básica es muy importante para poder resolver un problema de programación lineal por el método simplex.

8.4. Conjuntos convexos

El siguiente paso en la resolución de problemas de optimización es la generalización de diversos conceptos y resultados que se expusieron al resolver geométricamente el problema en el plano. Los problemas, a pesar de tener distinto número de variables, presentan conjuntos factibles con propiedades muy similares, que vamos a estudiar a continuación.

8.4.1. Definición de combinación lineal convexa

Dados los puntos (vectores) $P_1, P_2, \ldots, P_k$ de un espacio vectorial de dimensión n, diremos que un punto P es combinación lineal convexa de ellos si existen números reales $\alpha_1, \alpha_2, \ldots, \alpha_k$ tales que

$$P = \alpha_1 P_1 + \alpha_2 P_2 + \ldots + \alpha_k P_k \text{ con } \sum_{i=1}^{k} \alpha_i = 1$$

En la definición anterior, cada punto está definido por sus coordenadas al ser considerado un elemento de un espacio vectorial (podemos suponer que es $\mathbb{R}^n$). Por tanto, no supone ningún problema realizar $P = \alpha_1 P_1 + \alpha_2 P_2 + \ldots + \alpha_k P_k$, ya que tenemos definidos el producto de vectores por números reales (escalares) y la suma de vectores.

Ejemplo 8.4.1

Un ejemplo de combinación lineal que vimos en capítulos anteriores es considerar los colores que se pueden formar a partir de tres colores primarios. En un ordenador, los colores son tratados como una combinación lineal convexa de rojo (R), verde (G) y azul (B).

8.4.2. Definición de segmento

Un segmento de extremos P_1 y P_2 es el conjunto de todas las combinaciones lineales convexas de los puntos P_1 y P_2.

Esta defínción coincide con la idea geométrica que tenemos de segmento. Si tenemos dos puntos en el plano $P_1 = (x_1, y_1)$ y $P_2 = (x_2, y_2)$, una característica común a cualquier punto P del segmento que los une, con coordenadas (x, y) es que $x_1 \leq x \leq x_2$ e $y_1 \leq y \leq y_2$, como se muestra en la figura 6. Para que se verifique esto y, además, que los puntos estén en la recta que los une, debe ser $x = \alpha x_1 + (1 - \alpha)x_2$ e $y = \alpha y_1 + (1 - \alpha)y_2$ para $0 \leq \alpha \leq 1$. Esto es lo mismo que decir que P es una combinación lineal convexa de P_1 y P_2.

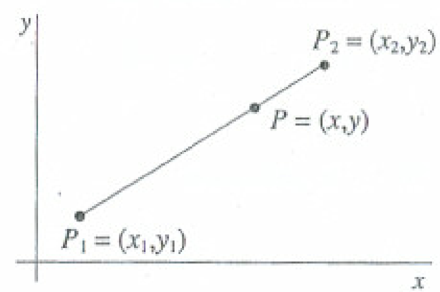


Figura 6

8.4.3. Definición de conjunto convexo

C es un conjunto convexo si dados dos puntos de él cualesquiera, P_1 y P_2, cualquier combinación lineal convexa de ellos está incluida en el conjunto.

A partir de la definición de segmento, podemos afirmar que un conjunto convexo es aquel que contiene el segmento que une dos puntos cualquiera de él. Para aclarar geométricamente el significado de los conjuntos convexos, en la figura 7 hemos representado un conjunto convexo (a) y dos conjuntos no convexos (b) y (c). Señalamos que el conjunto (c) no es convexo porque está formado por dos componentes, a pesar de que cada componente sí es convexa.

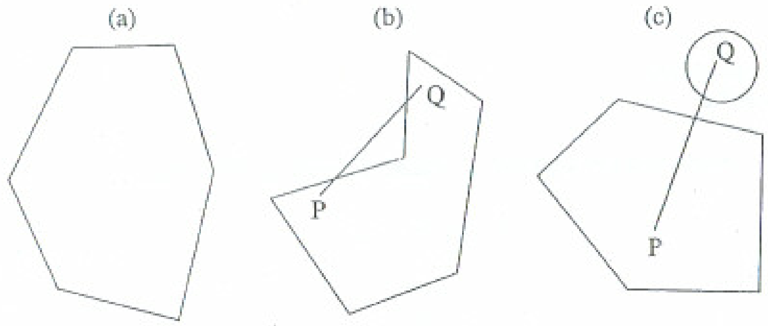


Figura 7

Ejemplo 8.4.2

Ejemplos de conjuntos convexos son: una recta, un círculo, los puntos interiores a un rectángulo, a una esfera o a un paralelogramo. Otro ejemplo de conjunto convexo es un segmento.

Demostremos analíticamente que un círculo es un conjunto convexo. Los puntos (x, y) del círculo centrado en el origen verifican $x^2 + y^2 \le R^2$, donde R es su radio. Sean (x_1, y_1) y (x_2, y_2) dos puntos del círculo. Entonces un punto (x, y) que pertenece al semento que los une verifica

$$x = \alpha x_1 + (1-\alpha)x_2$$
$$y = \alpha y_1 + (1-\alpha)y_2$$

para algún α, $0 \le \alpha \le 1$. Hay que demostrar que $x^2 + y^2 \le R^2$:

$$x^2+y^2=(\alpha x_1+(1-\alpha)x_2)^2+(\alpha y_1+(1-\alpha)y_2)^2=$$
$$=\alpha^2x_1^2+(1-\alpha)^2x_2^2+2\alpha(1-\alpha)x_1x_2+\alpha^2y_1^2+(1-\alpha)^2y_2^2+2\alpha(1-\alpha)y_1y_2=$$
$$=\alpha^2(x_1^2+y_1^2)+(1-\alpha)^2(x_2^2+y_2^2)+2\alpha(1-\alpha)(x_1x_2+y_1y_2)\le$$
$$\le\alpha^2R^2+(1+\alpha^2-2\alpha)R^2+2\alpha(1-\alpha)R^2=R^2$$

Hemos utilizado las propiedades del producto escalar de dos vectores (x_1, y_1) y (x_2, y_2) y que el coseno del ángulo que forman, está comprendido entre -1 y 1

$$x_1x_2+y_1y_2\le(x_1,y_1)\cdot(x_2,y_2)=\|(x_1,y_1)\|\,\|(x_2,y_2)\|\cos\alpha\le\sqrt{x_1^2+y_1^2}\,\sqrt{x_2^2+y_2^2}\le R^2.$$

Luego un círculo es un conjunto convexo. Suponiendo tres variables se demuestra, igualmente, que una esfera sólida es un conjunto convexo.

8.4.4. Definición de vértice

Un punto P de un conjunto convexo C es un vértice si no se puede expresar como combinación lineal convexa de otros dos puntos cualquiera de C, distintos de P. También se denomina punto extremo de C.

De nuevo, para un polígono en el plano, la idea geométrica de vértice coincide con esta definición de vértice, ya que los puntos de los lados o el interior del polígono son combinaciones lineales convexas de los puntos de los vértices geométricos. Los únicos puntos que no son combinación lineal convexa son ellos mismos.

8.4.5. Propiedades de los conjuntos convexos

1. Los puntos que verifican una desigualdad del tipo $a_1x_1 + a_2x_2 + \ldots + a_nx_n \le b$ son conjuntos convexos.
2. La intersección finita de conjuntos convexos es convexo.

Demostración:

1. Sean $(x_1, x_2, ..., x_n)$ e $(y_1, y_2, ..., y_n)$ dos puntos que verifican la desigualdad. Consideramos un punto $(z_1, z_2, ..., z_n)$ que sea combinación lineal convexa de ellos. Se verifica que existe $0 \le \lambda \le 1$ tal que, para $i = 1, ..., n$ es

$$z_i = \lambda x_i + (1-\lambda)y_i$$

Entonces

$$a_1z_1 + a_2z_2 + ... + a_nz_n = a_1(\lambda x_1 + (1-\lambda)y_1) + ... + a_n(\lambda x_n + (1-\lambda)y_n) =$$
$$= \lambda(a_1x_1 + ... + a_nx_n) + (1-\lambda)(a_1y_1 + ... + a_ny_n) \le \lambda b + (1-\lambda)b = b$$

por lo que el punto $(z_1, z_2, ..., z_n)$ también verifica la desigualdad.

2. Sean C_1 y C_2 dos conjuntos convexos. Dados dos puntos P y Q cualesquiera de la intersección, obviamente los puntos están en los dos conjuntos. Por tanto, también el segmento que los une está en los dos conjuntos y, consecuentemente, está en la intersección. Luego el conjunto intersección es convexo. Por inducción se demuestra para cualquier intersección finita. Recordamos que la unión de conjuntos convexos no tiene por qué ser convexa, como mostramos en la figura 7(c).

OBSERVACIÓN

Como no hemos impuesto ninguna condición a los valores de b, entonces los puntos que verifican una desigualdad del tipo $a_1x_1 + a_2x_2 + ... + a_nx_n \ge b$ también son un conjunto convexo. Y, consecuentemente, también lo serán los puntos que verifican $a_1x_1 + a_2x_2 + ... + a_nx_n = b$.

Teniendo en cuenta que los puntos que verifican cada una de las restricciones de un problema de programación lineal forman un conjunto convexo, y que la intersección de conjuntos convexos es convexa, podemos enunciar el siguiente teorema. Este resultado será clave en la resolución de problemas de programación lineal.

8.4.6. Teorema

El conjunto factible de un problema de programación lineal es un conjunto convexo.

Ejemplo 8.4.3

Consideremos el siguiente problema de programación lineal.

$$\begin{cases} \text{óptimo } F = x + y \\ x + y \ge 8 \\ x - y \le 2 \\ x \ge 0, y \ge 0 \end{cases}$$

En primer lugar, representamos gráficamente los puntos (x, y) que verifican las restricciones, es decir, representamos el conjunto convexo definido por

$$x + y \ge 8, x - y \le 2, x \ge 0, y \ge 0$$

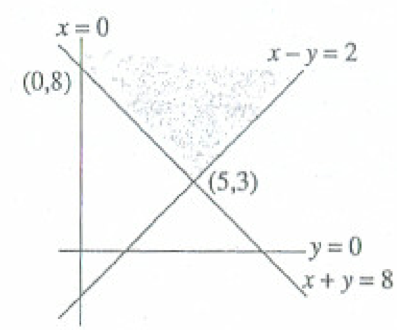


Figura 8

Como se aprecia en la figura, este conjunto es convexo y sus vértices son los puntos definidos por la intersección de las rectas

$$(x + y = 8) \cap (x - y = 2),\ (x + y = 8) \cap (x = 0)$$

es decir

$$\begin{cases} x + y = 8 \\ x - y = 2 \end{cases} \Rightarrow x = 5,\ y = 3$$

$$\begin{cases} x + y = 8 \\ x = 0 \end{cases} \Rightarrow x = 0,\ y = 8$$

Además, observamos que este conjunto no es acotado.

8.5. Teorema fundamental de la programación lineal

Las definiciones y propiedades que hemos enunciado en el apartado anterior nos van a permitir generalizar el método de resolución geométrica que hemos estudiado en el plano. Para poder resolver problemas de optimización de cualquier dimensión, nos falta especificar qué puntos van a ser los candidatos a extremos. El Teorema fundamental de la programación lineal nos los proporciona.

8.5.1. Teorema fundamental de la programación lineal

1. Toda función lineal F definida sobre un conjunto poligonal convexo C acotado toma sus valores máximo y mínimo sobre los vértices de dicho conjunto.
2. Si el extremo se alcanza en más de un vértice, entonces la función también alcanza este valor en cualquier combinación lineal convexa de dichos vértices.

Demostración:

1. Vamos a esbozar la demostración de este resultado, ya que ilustra la importancia de los conjuntos convexos para problemas de programación lineal. Para la demostración de la primera afirmación, vamos a partir de que cada punto del conjunto considerado C se puede expresar como combinación lineal convexa de sus vértices $V_1, V_2, ..., V_k$ (cuya demostración vamos a omitir en este texto). Si P es un punto del convexo, esto significa que se verifica

$$P = \alpha_1 V_1 + \alpha_2 V_2 + ... + \alpha_k V_k \text{ con } \sum_{i=1}^{k} \alpha_i = 1$$

Además, si F es la función lineal que queremos optimizar, podemos suponer sin que esto suponga una pérdida de validez del razonamiento, que están ordenados de forma creciente, es decir, que se verifica

$$F(V_1) \le F(V_2) \le \dots \le F(V_k)$$

Entonces, por ser F lineal es

$$F(P) = F(\alpha_1 V_1 + \alpha_2 V_2 + \dots + \alpha_k V_k) = \alpha_1 F(V_1) + \alpha_2 F(V_2) + \dots + \alpha_k F(V_k)$$

Como $\sum_{i=1}^{k} \alpha_i = 1$, se verifica que

$$\begin{aligned} F(V_1) = \alpha_1 F(V_1) + \alpha_2 F(V_1) + \dots + \alpha_k F(V_1) \le F(P) \le \\ \le \alpha_1 F(V_k) + \alpha_2 F(V_k) + \dots + \alpha_k F(V_k) = F(V_k) \end{aligned}$$

para cualquier punto P de C, hemos demostrado que los máximos y mínimos de F en C se alcanzan en los vértices del conjunto C.

2. Demostremos la segunda afirmación. Supongamos que el óptimo de la función se alcanza en más de un vértice; llamenos $A_1, \dots, A_s$ a estos vértices y M al valor óptimo de F en estos vértices. Sea Q un punto que es combinación lineal convexa de estos vértices

$$Q = \beta_1 A_1 + \dots + \beta_s A_s \text{ con } \sum_{i=1}^{s} \beta_i = 1$$

Entonces, como $\sum_{i=1}^{s} \beta_i = 1$ es

$$F(Q) = F(\beta_1 A_1 + \dots + \beta_s A_s) = \beta_1 F(A_1) + \dots + \beta_s F(A_s) = \beta_1 M + \dots + \beta_s M = M$$

por lo que el extremo también se alcanza en cualquier punto Q que es combinación lineal convexa de estos vértices.

Finalmente, añadimos que si el conjunto factible no es acotado, no tiene por qué verificarse este Teorema fundamental, como veremos con un ejemplo.

Ejemplo 8.5.1

Vamos a determinar los valores extremos del problema de programación lineal

$$\begin{cases} F = x + 2y + 3z \\ x + y + z \le 3 \\ x \ge 0,\ y \ge 0,\ z \ge 0 \end{cases}$$

Primero tenemos que determinar los vértices del conjunto factible. Representamos gráficamente este conjunto, que es un triedro (figura 9).

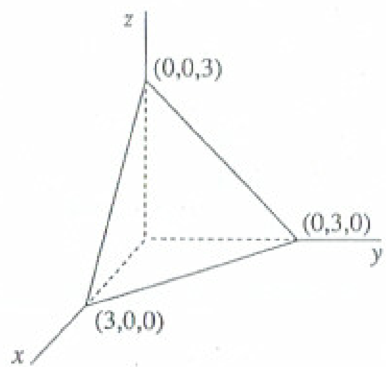


Figura 9

Observamos que sus vértices son los puntos $V_1 = (3, 0, 0)$, $V_2 = (0, 3, 0)$, $V_3 = (0, 0, 3)$ y $V_4 = (0, 0, 0)$, que podemos determinar como la intersección de los planos en $\mathbb{R}^3$

$$\begin{cases} x + y + z = 3 \\ x = 0,\ y = 0,\ z = 0 \end{cases}$$

Como el conjunto delimitado por las restricciones es un convexo acotado, sabemos por el Teorema fundamental que alcanza en sus vértices los valores máximos y mínimos. Por tanto, para determinar en qué puntos alcanza estos valores, basta obtener el valor de F en estos puntos.

$$F(0, 0, 0) = 0,\ F(3, 0, 0) = 3,\ F(0, 3, 0) = 6,\ F(0, 0, 3) = 9$$

Luego el mínimo se alcanza en el punto $V_4 = (0, 0, 0)$ y el máximo en $V_3 =$ $= (0, 0, 3)$.

Ejemplo 8.5.2

En el ejemplo 8.4.3 representamos el conjunto dado por las restricciones y determinamos sus vértices. Vamos a estudiar dónde alcanza la función F sus óptimos. Como F es creciente para x e y crecientes no va a tener un máximo en el conjunto factible. Sin embargo, sí es posible determinar dónde alcanza el mínimo, ya que va a ser en uno de sus vértices. Calculamos

$$F(5,3) = 5 + 3 = 8$$
$$F(0,8) = 0 + 8 = 8$$

Por tanto, como alcanza el mínimo en los dos vértices, también va a ser mínimo el valor de F en el segmento que une estos dos puntos, es decir, en los puntos que son combinación lineal convexa de los dos vértices. Que F no tenga máximo en el conjunto factible no contradice el Teorema fundamental, ya que el conjunto factible no es acotado.

8.6. Algoritmo simplex

El propósito de esta sección es exponer un proceso algorítmico para calcular la solución óptima de un problema de programación lineal dado. El método se denomina Método Simplex, como indicamos en la introducción del capítulo, fue ideado por G. B. Dantzig. Se trata de un proceso iterativo tal que en cada bucle mediante ciertos cambios en las variables que tomamos como básicas, se pasa de un vértice a otro contiguo que es una solución factible, y que mejora la función objetivo o la deja como estaba. La repetición de este proceso conduce a la solución óptima.

Una de sus principales virtudes es que permite la resolución de un problema de más de dos variables sin tener que determinar todos los vértices de la región factible. Además como es un proceso algorítmico, su resolución está implementada en diversos programas de cálculo simbólico, como por ejemplo Maple; también se puede encontrar en paquetes comerciales de software dadas sus innumerables aplicaciones.

El problema de programación lineal que resuelve el método simplex es

$$\begin{cases} \max F = c_1x_1 + c_2x_2 + \ldots + c_nx_n \\ A\overline{x} \leq \overline{b} \\ x_i \geq 0, \text{ para } i = 1, \ldots, n \end{cases}$$

Existen problemas más generales (como los que contempla cualquier tipo de restricción), que ya hemos estudiado y resuelto por el método geométrico en el caso de dos variables. La idea del método simplex para su resolución es la misma que la que se expone a continuación. Sin embargo, el algoritmo es más complicado, ya que requiere la introducción de unas variables (denominadas ficticias o artificiales) que proporcionan de forma inmediata una solución básica no degenerada. Por este motivo, sólo vamos a considerar en este capítulo el problema donde todas las restricciones son del tipo $\leq$. Para una descripción de la resolución de problemas más generales remitimos al lector a la bibliografía (por ejemplo, Fraleigh y Beauregard, 1989; de la Villa, 1994).

Obsérvese que este problema resuelve el problema equivalente del mínimo, ya que

$$\min F \Leftrightarrow -\max(-F)$$

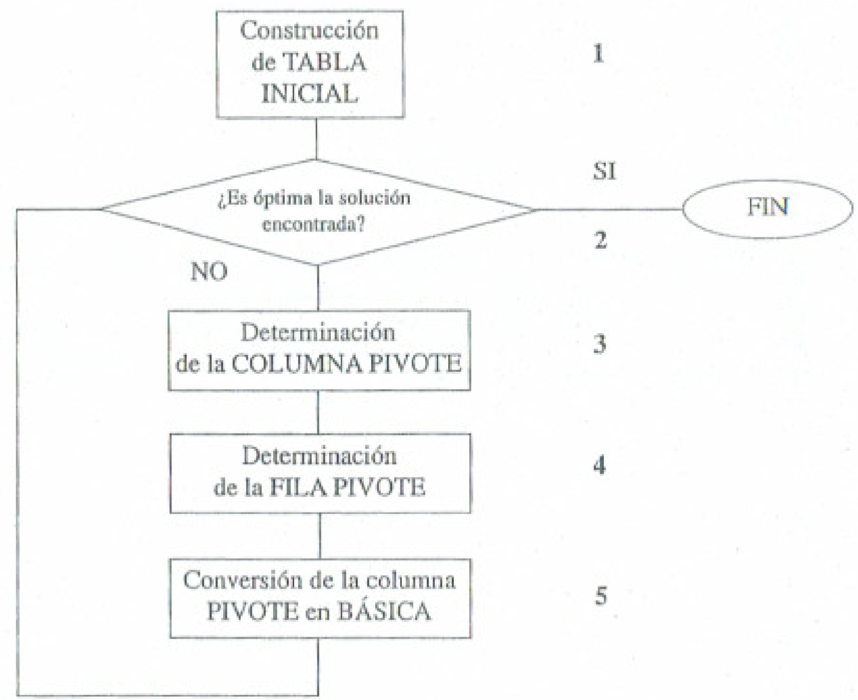
El primer paso en la resolución de estos problemas aplicando el método simplex es normalizar el problema con las variables de holgura $x_{n+1}, \ldots, x_{n+m}$

$$\begin{cases} \max F = c_1x_1 + c_2x_2 + \ldots + c_nx_n \\ A\overline{x} = \overline{b} \\ x_i \geq 0, \text{ para } i = 1, \ldots, n+m \end{cases}$$

La resolución práctica que vamos a describir del Algoritmo Simplex, se basa en la utilización de diversas tablas. Se parte de un vértice que es solución básica, generalmente (0, ... ", ..., 0) para las variables iniciales, se estudia si esa solución es óptima; si no es, se pasa a otra solución básica que mejora la función objetivo F, es decir se pasa a un vértice contiguo y se estudia si se alcanza en él el máximo. Si no es así, se repite de nuevo el proceso, hasta alcanzar el máximo.

Vamos a describir detalladamente el Algoritmo del Método Simplex y analizar a continuación cada una de las fases utilizadas.

ALGORITMO DEL MÉTODO DEL SIMPLEX



1. Construcción de la tabla inicial

	x_1	...	x_n	x_{n+1}	...	x_{n+m}	$\overline{b}$
x_{n+1}	a_{11}	...	a_{1n}	1	...	0	b_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_{n+m}	a_{m1}	...	a_{mn}	0	...	1	b_m
FO	$-c_1$	...	$-c_n$	0		0	0

Como ya indicamos, las variables $x_{n+1}, \ldots, x_{n+m}$ son las variables de holgura que normalizan el problema inicial (convierten en igualdades las desigualdades $\leq$).

La última columna $\overline{b}$ es el vector de componentes los términos independientes de cada una de las ecuaciones dadas.

La fila j- ésima tiene como entradas los coeficientes de las variables de la restricción j -ésima.

La última fila FO es la fila objetivo y tiene como entradas los coeficientes de la función objetivo escrita de la forma:

$$F - c_1x_1 - c_2x_2 - \ldots - c_nx_n = 0$$

La fila objetivo contendrá datos que indiquen qué variables no básicas deberán modificarse para convertirse en básicas.

2. ¿Es óptima la solución encontrada?

Consideramos la solución $(x_1, \ldots x_n, x_{n+1}, \ldots, x_{n+m}) = (0, \ldots, 0, b_1, \ldots, b_m)$, para este punto las variables no básicas son $x_1, \ldots, x_n$ y las variables básicas son $x_{n+1}, \ldots, x_{n+m}$. El valor que toma la función a optimar F, en la solución óptima considerada, viene dado por la última entrada de la fila objetivo FO.

En este paso podemos utilizar el TEST de MAXIMALIDAD:

"Si la FO no tiene valores negativos hemos llegado a una solución óptima".

3. Determinación de la COLUMNA PIVOTE

La columna pivote o variable de entrada es la columna a la que pertenece el elemento c_i más negativo en la fila objetivo FO.

- No se tiene en cuenta el elemento correspondiente de $\overline{b}$.
- Si hay dos que son igualmente los más negativos, se puede elegir uno cualquiera.

- Puede ocurrir que ningún coeficiente de la columna pivote sea positivo, es decir, todas las entradas son negativas o cero. Esto significa que F no está acotada, probablemente el problema esté mal modelado o el problema no tiene solución.

4. Determinación de la FILA PIVOTE

Partiendo de la columna pivote (columna i -ésima) en ella consideramos

$$\left\{ \frac{b_k}{a_{ki}} \text{ tal que } a_{ki} \geq 0,\ 1 \leq k \leq m \right\}$$

y calculamos:

$$\min_{1\leq k\leq m} \left\{ \frac{b_k}{a_{ki}} \right\} = \frac{b_j}{a_{ji}}$$

El índice j nos determina la fila pivote buscada, que es la j -ésima.

- No se tiene en cuenta la fila objetivo, FO.
- Si hay dos cocientes iguales, se elige al azar y si la selección nos lleva a un ciclo sin fin se vuelve al lugar de empate y se selecciona otro.

El elemento (j, i) que pertenece a la columna pivote y a la fila pivote se llama elemento pivote.

5. Conversión de la COLUMNA PIVOTE en BÁSICA

El elemento pivote lo convertimos en 1, dividiendo la fila pivote por a_{ji} y aplicando el método de Gauss-Jordan (es decir, realizando operaciones elementales con las filas de la tabla), obtenemos 0 en todos los coeficientes de la columna pivote, excepto en el elemento pivote que es 1.

Mediante las transformaciones oportunas en las filas la nueva tabla que resulta es de la forma

	x_1	...	x_i	...	x_n	x_{n+1}	...	x_{n+m}	$\overline{b}$
x_{n+1}	a'_{11}	...	0	...	a'_{1n}		...		b'_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$			$\vdots$		$\vdots$	$\vdots$
x_i	a'_{j1}		1	...	a'_{jn}				b'_j
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_{n+m}	a'_{m1}	...	0		a'_{mn}				b'_m
FO	c'_1	...	c'_i	...	c'_m				

- La variable x_i entra como variable básica, y la variable x_{n+j} sale como variable básica.
- La nueva solución que consideraremos será:

$$(x_1, ..., x_i, ..., x_n, x_{n+1}, ..., x_{n+j}, ..., x_{n+m}) = (0, ..., b'_j, ..., 0, b'_1, ..., 0, ..., b'_m)$$

- El proceso se repite hasta que la función objetivo FO no tenga entradas negativas.
- Si existen columnas con todas las entradas negativas, la solución no está acotada y el problema no tiene solución.

Veamos el método simplex aplicado a un ejemplo.

Ejemplo 8.6.1

Calcular mediante el Método simplex, la solución óptima del problema de programación lineal siguiente

$$\begin{cases} \max F = 3x_1 + 2x_2 + x_3 \\ x_1 + x_2 - x_3 \le 10 \\ 2x_1 + x_2 + x_3 \le 6 \\ x_1 \ge 0,\ x_2 \ge 0, x_3 \ge 0 \end{cases}$$

El primer paso es normalizarlo, introduciendo las variables de holgura, x_4, x_5.

$$\begin{cases} \max z = 3x_1 + 2x_2 + x_3 + 0x_4 + 0x_5 \\ x_1 + x_2 - x_3 + x_4 = 10 \\ 2x_1 + x_2 + x_3 + x_5 = 6 \\ x_i \ge 0 \ \text{ para } i = 1 \ldots 5 \end{cases}$$

Estamos en las condiciones idóneas para aplicar algoritmo simplex. Los pasos serían:

1. Construcción de la tabla inicial

	x_1	x_2	x_3	x_4	x_5	$\overline{b}$
x_4	1	1	–1	1	0	10
x_5	2	1	1	0	1	6
FO	–3	–2	–1	0	0	0

2. La solución encontrada es (0, 0, 0, 10, 6) y no es óptima como se puede comprobar (el valor que toma F es 0), aunque no haría falta, ya que en la FO hay entradas negativas.

3. Determinación de la COLUMNA PIVOTE.

   Nos fijamos en la FO y el elemento más negativo es –3, por tanto, la primera columna es la columna pivote.

4. Determinación de la FILA PIVOTE

   Los cocientes a considerar son:

$$\frac{10}{1} \text{ y } \frac{6}{2}$$

   La fila pivote buscada es la segunda porque $\frac{6}{2} < \frac{10}{1}$

   El elemento pivote es la entrada (2, 1). Por lo tanto la variable x_1 entra como variable básica y sale la variable x_5, pasando a ser variable no básica.

5. Conversión de la columna PIVOTE en BÁSICA.

   Dividimos por 2 la fila pivote, $F_2 \leftrightarrow \frac{1}{2} F_2$ y seguimos llamando F_2 a esa fila.
   Aplicando el método de Gauss-Jordan, transformamos las entradas de la columna pivote distintas del elemento pivote en ceros.
   Los otros cambios a efectuar en las filas son:

$$F_1 \leftrightarrow F_1 - F_2$$
$$F_3 \leftrightarrow F_3 + 3F_2$$

   La tabla que resulta es:

	x_1	x_2	x_3	x_4	x_5	$\overline{b}$
x_4	0	$\frac{1}{2}$	$\frac{-3}{2}$	1	$-\frac{1}{2}$	10-3
x_1	1	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	3
FO	0	$-\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{3}{2}$	9

Si nos fijamos en la nueva FO vemos que todavía existen entradas negativas, por tanto la solución obtenida (3, 0, 0, 7, 0) mejora la anterior ($F = 9$) pero no es la solución óptima buscada.

Entonces repetimos los pasos 3, 4, 5.

3. Determinación de la COLUMNA PIVOTE.

   El elemento más negativo de FO es $-\frac{1}{2}$, luego la columna pivote es la segunda.

4. Determinación de la FILA PIVOTE.
   Los cocientes que consideramos son:

$$\frac{7}{\frac{3}{2}} \text{ y } \frac{3}{\frac{1}{2}}$$

   por tanto, la fila pivote es la segunda. El elemento pivote es la entrada (2, 2).
   En este caso, la variable x_2 entra como variable básica y sale la variable x_1.

5. Conversión de la COLUMNA PIVOTE en BÁSICA.

   Las transformaciones son, primero en la fila pivote, $F_2 \leftrightarrow 2F_2$, y luego en las demás filas:

$$F_1 \leftrightarrow F_1 - \frac{1}{2} F_2$$
$$F_3 \leftrightarrow F_3 + \frac{1}{2} F_2$$

	x_1	x_2	x_3	x_4	x_5	$\bar{b}$
x_4	-1	0	-2	1	-1	4
x_2	2	1	1	0	1	6
FO	1	0	1	0	2	12

La fila FO tiene todas sus entradas no negativas, por tanto el método simplex nos asegura que la solución (0, 6, 0, 4, 0) es óptima y el valor máximo que alcanza la función es 12. Por tanto, la solución para el problema inicial es (0, 6, 0).

EJERCICIOS

Ejercicio 1

Una azucarera produce diariamente 600 kg de azúcar moreno y 1100 kg de azúcar blanco. Dicha azucarera vende dos tipos de azúcar mezcla: la primera mezcla (A) que tiene 3 partes de azúcar blanco por cada parte de azúcar moreno y la segunda mezcla B tiene 2 partes de moreno por cada parte de azúcar blanco. Se pretende que los beneficios sean máximos sabiendo que las ganancias de la azucarera son de 80 céntimos de euro por kg de azúcar mezcla (A), y en el azúcar mezcla de tipo (B) se gana 11 céntimos de euros/kg. Además, el azúcar sobrante de ambas mezclas se vende para animales con un beneficio neto de 8 céntimos de euro por cada kg. Se pide la función objetivo, las restricciones y los vértices de la región factible.

Solución:

La siguiente tabla resume la información que tenemos de la azucarera

Tipos de azúcar	**Kg para A**	**Kg para B**	**Sobrante**
Moreno	x	$2y$	$600-(x+2y)$
Blanco	$3x$	y	$1100-(3x+y)$
Total	$4x$	$3x$	$1700-(4x+3y)$

Las ganancias en euros se obtienen de la ecuación:

$$0.8 \cdot 4x + 0.11 \cdot 3y + 0.08(1700-4x-3y) =$$
$$= (0.8-0.08)4x + (0.11-0.08)3y + 1700 \cdot 0.08$$

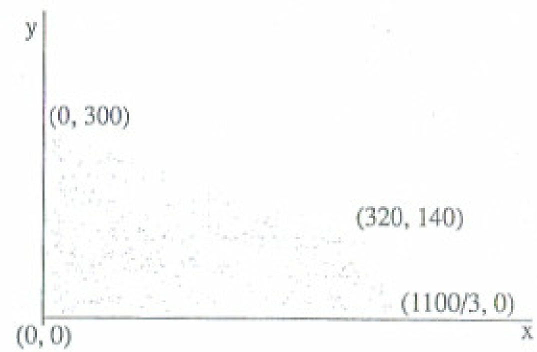
Simplificando se obtiene la función objetivo.

$$F = 2.88x + 0.09y + 136$$

Las restricciones son:

$$\begin{cases} x+2y \leq 600 \\ 3x+y \leq 1100 \\ x \geq 0,\ y \geq 0 \end{cases}$$

y los vértices son (0, 0), (0, 300), (1100/3, 0), (320, 140). Estos vértices obtienen representando la región y calculando los cortes correspondientes.



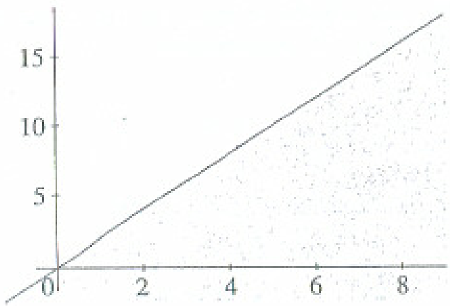
Ejercicio 2

Dado el sistema $\begin{cases} 2x \geq y \\ 3x - y \geq 8 \\ x + y \leq 6 \\ y \leq 3 \end{cases}$, se pide representar gráficamente la región del plano que verifica todas las desigualdades que lo componen. ¿Es superflua alguna desigualdad? ¿Se puede añadir alguna desigualdad para que la región sea acotada:

Solución:

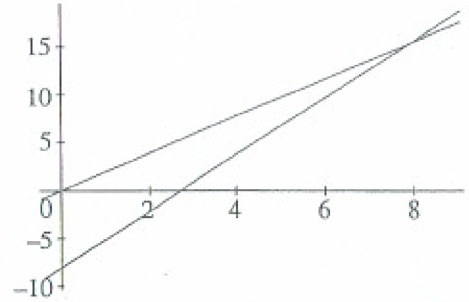
Para resolver el problema vamos a representar los convexos que determinan cada una de las restricciones, considerando el convexo que se obtiene cada vez que añadimos una desigualdad.

La gráfica para $2x \geq y$ es el convexo

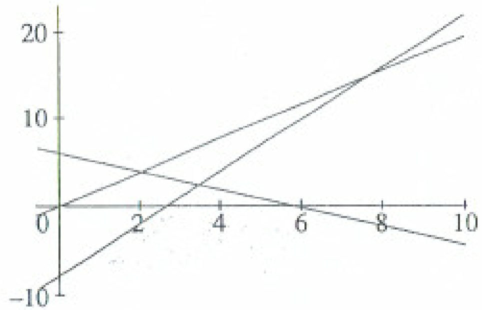


Si consideramos la desigualdad siguiente quedará el convexo definido por

$$\begin{cases} 2x \geq y \\ 3x - y \geq 8 \end{cases}$$

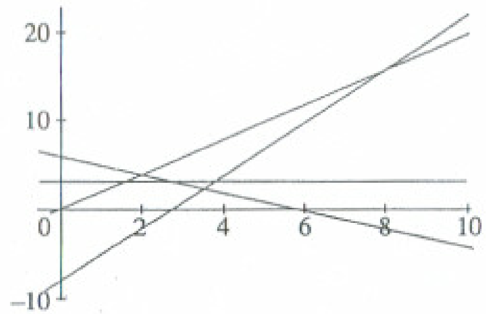


Si añadimos una desigualdad nueva, $\begin{cases} 2x \geq y \\ 3x - y \geq 8 \\ x + y \leq 6 \end{cases}$ resultará:



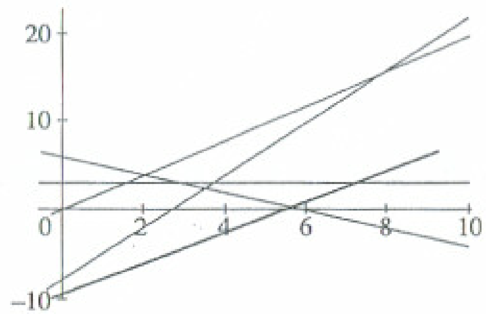
De momento como se puede comprobar gráficamente, no hay ninguna superflua porque el conjunto convexo cambia con cada nueva desigualdad que hemos considerado.

Para terminar, representamos $\begin{cases} 2x \geq y \\ 3x - y \geq 8 \\ x + y \leq 6 \\ y \leq 3 \end{cases}$



Por lo tanto como se puede observar la última desigualdad $y \leq 3$ es superflua, porque añadirla no ha supuesto ninguna variación en el conjunto convexo.

Observamos que el conjunto no es acotado. Sin embargo, podemos añadir cualquier restricción del tipo $kx - y \geq b$ con $k \geq 0$, $b > 8$ y el conjunto delimitado por todas las restricciones sí es acotado, como se puede observar en la figura.



Ejercicio 3

Un agricultor dedicado al cultivo de fresas vende su mercancía exclusivamente a dos distribuidores, que llamaremos A y B. El coste, por tonelada, del transporte para el distribuidor B es tres veces mayor que los costes del transporte para el distribuidor A. Este año la cosecha recogida durante el mes de abril han sido 9 toneladas de fresa. Por la experiencia de otros años, se sabe que la mitad de la cantidad de fresa vendida a A más la cantidad vendida a B siempre oscila entre 5 y 8 toneladas. Se pide plantear el problema y determinar cuántas toneladas deben venderse a cada distribuidor para que los costes sean mínimos.

Solución:

En primer lugar, tenemos que plantear el problema. La función a minimizar es el coste del transporte; como sabemos que el coste, por unidad, del transporte para el distribuidor B es tres veces mayor que los costes del transporte para el distribuidor A, la función objetivo será

$$x + 3y$$

llamando x e y a las toneladas de mercancía vendida a los distribuidores A y B respectivamente. La condición de que la cantidad total de fresa recogida en abril es 9 toneladas se puede expresar como

$$x + y = 9$$

Por otra parte, las otras dos condiciones son

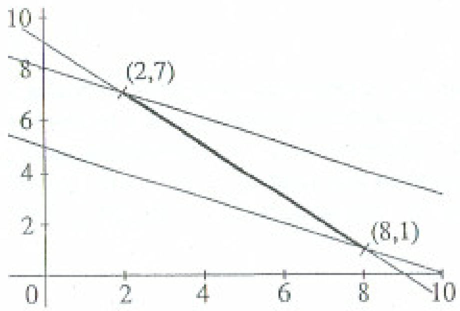
$$\frac{1}{2}x + y \leq 8$$

$$\frac{1}{2}x + y \geq 5$$

Además, cada una de las cantidades que se venden a los distribuidores son mayores o iguales que 0. Luego el problema de programación lineal a resolver es:

$$\begin{cases} \min F = x + 3y \\ x + y = 9 \\ \dfrac{1}{2}x + y \leq 8 \\ \dfrac{1}{2}x + y \geq 5 \\ x \geq 0,\ y \geq 0 \end{cases}$$

Para resolver este problema, representamos primero el conjunto factible.



Observamos que es el segmento comprendido entre los puntos (2, 7) y (8, 1). Según el Teorema fundamental, al ser el conjunto factible acotado, el máximo y el mínimo se alcanzan en los vértices de este conjunto, es decir, en los extremos. Para determinar qué valores extremos se alcanzan en cada uno de ellos, calculamos el valor de z en estos puntos

$$F(2, 7) = 2 + 3 \cdot 7 = 23 \qquad F(8, 1) = 8 + 3 \cdot 1 = 11$$

Luego el máximo se alcanza en el punto (2, 7) y el mínimo en (8, 1), por lo que para minimizar los costes de transporte, deben venderse 8 toneladas al distribuidor A y 1 tonelada al B.

Ejercicio 4

Determínense los extremos (máximo y mínimo) de la función

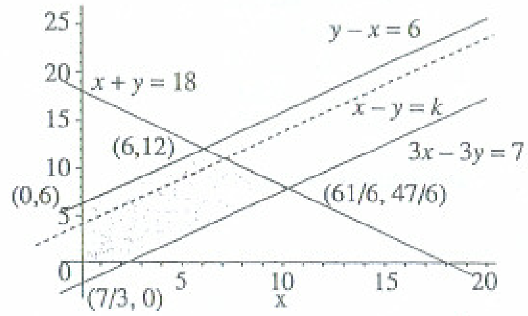
$$F = x - y$$

sujetos a las restricciones

$$\begin{cases} x + y \leq 18 \\ y - x \leq 6 \\ 3x - 3y \geq 7 \\ x \geq 0,\ y \geq 0 \end{cases}$$

Solución:

Representamos gráficamente la región factible y la dirección de la recta $x - y = k$, constante.



Sabemos que los extremos se van a alcanzar en los vértices de la región factible, es decir, en los puntos (6, 12), (0, 6), (0, 0), (7/3, 0) o (61/6, 47/6). Para terminar en qué puntos se alcanzan los extremos, calculamos el valor de F en estos puntos.

$$F(0, 0) = 0, F(0, 6) = -6, F(6, 12) = -6, F(61/6, 47/6) = 7/3, F(7/3, 0) = 7/3$$

Luego el mínimo se alcanza en (0, 6) y (6, 12) y el máximo en otros dos vértices: (7/3, 0) y (61/6, 47/6). Esta misma conclusión la podíamos haber alcanzado tras observar la figura anterior, ya que la dirección de la recta $x - y = k$ es paralela al segmento que une (0, 6) y (6, 12) o al que une (7/3, 0) y (61/6, 47/6). Además, k es menor en el primer caso (por lo que corresponde al mínimo) y mayor en el segundo caso, que es donde se alcanza el máximo.

Por tanto, el mínimo de F es –6 y el máximo es 7/3.

Ejercicio 5

Dado el siguiente problema de programación lineal, se pide introducir las variables de holgura y expresarlo en forma matricial.

$$\begin{cases} \text{óptimo } F = x_1 + 5x_2 - 3x_3 + x_4 + 2x_5 \\ 2x_1 - x_2 + 2x_4 \leq 2 \\ x_1 + x_2 - 4x_4 + x_5 \leq 7 \\ x_2 - x_4 \leq 3 \\ x_3 \leq 6 \\ x_1 + x_2 - x_4 \leq 2 \end{cases}$$

Determínese una solución básica.

Solución:

Hay 5 restricciones, por lo que introducimos 5 variables de holgura: x_6, x_7, x_8, x_9 y x_{10}. Las restricciones se pueden expresar como

$$\begin{cases} 2x_1 - x_2 + 2x_4 + x_6 = 2 \\ x_1 + x_2 - 4x_4 + x_5 + x_7 = 7 \\ x_2 - x_4 + x_8 = 3 \\ x_3 + x_9 = 6 \\ x_1 + x_2 - x_4 + x_{10} = 2 \end{cases}$$

La expresión matricial del problema es

$$\begin{cases} \text{óptimo } F = x_1 + 5x_2 - 3x_3 + x_4 + 2x_5 \\ \begin{pmatrix} 2 & -1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -4 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \end{pmatrix} = \begin{pmatrix} 2 \\ 7 \\ 3 \\ 6 \\ 2 \end{pmatrix} \end{cases}$$

Ahora tenemos que determinar una solución básica, es decir, una solución con no más de 5 componentes positivas. Para ello, nos basta observar que en las igualdades, los coeficientes de las variables de hogura es 1, por lo que podemos hacer que las variables del problema valgan 0 y dar los valores correspondientes a las de holgura. Así, una solución básica sería *(0, 0, 0, 0, 0, 7, 3, 2, 6, 2)*.

Ejercicio 6

Dado el siguiente problema de programación lineal

$$\max F = 4x_1 + 2x_2$$

con las restricciones:

$$\begin{cases} x_1 + 3x_2 \leq 15 \\ 2x_1 - x_2 \leq 8 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

Convertimos las desigualdades en igualdades mediante las variables de holgura x_3 y x_4.

a) ¿Es (8, 0, 7, 0) una solución factible básica? En caso de que lo sea, ¿es no degenerada?
b) ¿Es (7, 1, 5, 0) una solución factible básica?
c) Cualquier punto que pertenezca al segmento [(2, 4), (4, 0)] ¿es solución óptima?

Solución:

La introducción de las variables de holgura transforma las restricciones en el siguiente sistema de ecuaciones lineales

$$\begin{cases} x_1 + 3x_2 + x_3 = 15 \\ 2x_1 - x_2 + x_4 = 8 \\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ x_4 \geq 0 \end{cases}$$

a) (8, 0, 7, 0) es solución factible porque satisface todas las restricciones y es básica porque no tiene más de dos componentes positivas distintas de 0 (dos es el número de restricciones). Es no degenerada porque las componentes distintas de 0 son dos.
b) No es factible básica ya que el número de componentes positivas es tres.
c) Cualquier punto del segmento [(2, 4), (4, 0)] es solución óptima porque la función objetivo es paralela a la frontera de la región de posibilidad a que pertenece el segmento dado. En realidad, los máximos de la función se alcanzan en el segmento [(9/5, 22/5), (4, 0)], es decir, en el segmento comprendido entre los dos vértices (9/5, 22/5) y (4, 0) del conjunto factible.

Ejercicio 7

a) Determínense geométricamente los extremos, si existen, del siguiente problema de programación lineal.

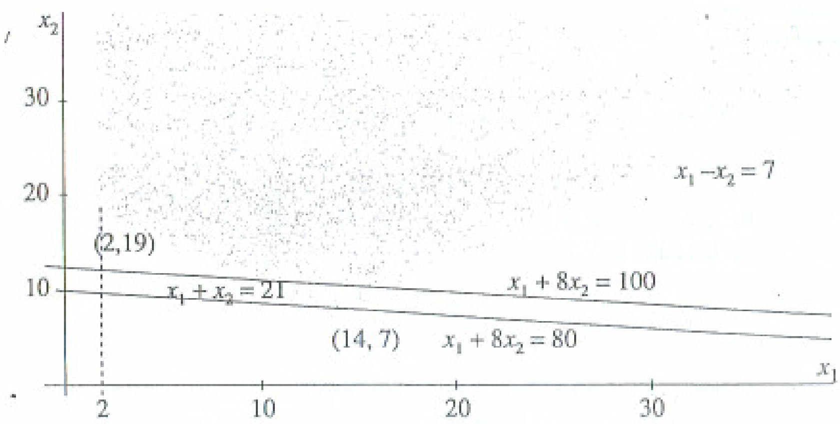
$$\text{Óptimo de } F = x_1 + 8x_2$$

$$\text{Sujeto a } \begin{cases} x_1 - x_2 \leq 7 \\ x_1 + x_2 \geq 21 \\ x_1 \geq 2,\ x_2 \geq 0 \end{cases}$$

b) Mediante un cambio de variable, transfórmese el anterior problema de programación lineal en otro equivalente en el que estén las restricciones $x_1 \geq 0, x_2 \geq 0$.

Solución:

En primer lugar determinamos el conjunto convexo. Para ello representamos el conjunto convexo que representa el conjunto factible. Si representamos la intersección de las desigualdades, resulta:



El conjunto factible no es cerrado. Los vértices del convexo son $A = (14, 7)$ y $B = (2, 19)$, intersección de las rectas $x_1 - x_2 = 7$; $x_1 + x_2 = 21$ y $x_1 = 2$; $x_1 + x_2 = 21$ respectivamente. La función objetivo $F = x_1 + 8x_2$ va creciendo dentro de la región factible.

Además $F(A) = F(14, 7) = 70$ y $F(B) = F(2, 19) = 154$. Por lo tanto no existe vértice que maximice la función objetivo.

Sin embargo, en el vértice (14,7) la función F alcanza el mínimo y es $F(14,7) = 70$.

b) Si consideraremos la primera expresión del problema, con el cambio $x = x_1 - 2$ y (por notación) $x_2 = y$, sustituyendo el problema es equivalente a:

$$\text{Óptimo de } F = x + 8y + 2$$

$$\text{Sujeto a } \begin{cases} x - y \leq 5 \\ x + y \geq 19 \\ x \geq 0,\ y \geq 0 \end{cases}$$

Observamos que en este problema se ha reducido el número de restricciones, respecto a la forma usual de expresar el problema.

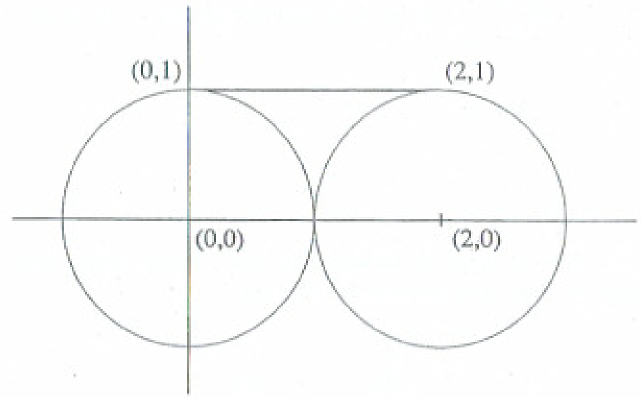
Ejercicio 8

Estúdiese si los siguientes conjuntos son convexos:

a) $C = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \leq 1\} \cup \{(x, y) \in \mathbb{R}^2 / (x-2)^2 + y^2 \leq 1\}$
b) $C = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \leq 1\} \cap \{(x, y) \in \mathbb{R}^2 / (x-2)^2 + y^2 \leq 1\}$
c) Un cubo macizo.

Solución:

a) El conjunto a) $C = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \leq 1\} \cup \{(x, y) \in \mathbb{R}^2 / (x-2)^2 + y^2 \leq 1\}$ es la unión de dos círculos, de radio 1, y centros en (0, 0) y (2, 0) respectivamente, que son conjuntos convexos, como vimos en el ejemplo 8.4.2. Pero su unión no es un conjunto convexo, ya que como se aprecia en la figura (está sombreada la unión de los círculos), el segmento que une los puntos (0, 1) y (2, 1) no está contenido dentro de C. En efecto, este segmento está definido por los puntos $(x, 1)$ con $0 \leq x \leq 2$ y el punto (2, 1), aunque está en el segmento, no está en la unión de los dos círculos.



b) El conjunto C es, en este caso, la intersección de los dos círculos del apartado anterior. Es el punto (1, 0), que es un conjunto convexo (un único punto es siempre un conjunto convexo). También podíamos haber razonado que la intersección de conjuntos convexos es convexo, y haber llegado así al mismo resultado.

c) Podemos suponer, si consideramos el cubo como un subconjunto de $\mathbb{R}^3$ que los vértices del cubo son los puntos (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (1, 0, 1), (0, 1, 1), (1, 1, 1). Es decir, los puntos (x, y, z) del cubo macizo verifican

$$0 \leq x \leq 1,\ 0 \leq y \leq 1,\ 0 \leq z \leq 1$$

Un punto (x, y, z) que es combinación lineal convexa de dos puntos del cubo (x_1, y_1, z_1) y (x_2, y_2, z_2) verifica, para $0 \leq \alpha \leq 1$

$$\begin{array}{ll} x = \alpha x_1 + (1-\alpha)\, x_2 & 0 \leq x = \alpha x_1 + (1-\alpha)\, x_2 \leq \alpha + (1-\alpha) = 1 \\ y = \alpha y_1 + (1-\alpha)\, y_2 \Rightarrow & 0 \leq y = \alpha y_1 + (1-\alpha)\, y_2 \leq \alpha + (1-\alpha) = 1 \\ z = \alpha z_1 + (1-\alpha)\, z_2 & 0 \leq z = \alpha z_1 + (1-\alpha)\, z_2 \leq \alpha + (1-\alpha) = 1 \end{array}$$

Por lo que el cubo macizo es convexo porque el punto (x, y, z) está en el cubo.

Ejercicio 9

Mediante el método simplex, calcúlese la solución óptima para el problema:

$$\begin{cases} \min z = -3x_1 - 5x_2 \\ x_1 + 2x_2 \leq 36 \\ 2x_1 + x_2 \leq 36 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

Solución:

Para seguir los pasos del algoritmo simplex, necesitamos normalizar el problema y transformar el problema de mínimo a un problema de máximo.

Introduciendo x_3, x_4 como variables de holgura, el problema primero es equivalente a

$$\begin{cases} -\max z = 3x_1 + 5x_2 \\ x_1 + 2x_2 + x_3 = 36 \\ 2x_1 + x_2 + x_4 = 36 \\ x_i \geq 0 \text{ para } i = 1, \ldots, 4 \end{cases}$$

Debido al signo negativo que afecta al max z, a el valor máximo obtenido, le debemos cambiar el signo, pero a la solución óptima no le afecta. Aplicando el método, resolveremos el problema paso a paso.

1. Construcción de la tabla inicial

	x_1	x_2	x_3	x_4	$\bar{b}$
x_3	1	2	1	0	36
x_4	2	1	0	1	36
FO	-3	-5	0	0	0

2. La solución encontrada es $(0, 0, 0, 36, 36)$ y no es óptima como se puede comprobar ($z = 0$). Además, en la FO hay entradas negativas.

3. Determinación de la COLUMNA PIVOTE.
   En la fila FO el elemento más negativo es -5, por lo tanto la columna pivote es la segunda.

4. Determinación de la FILA PIVOTE.
   Calculamos los cocientes correspondientes para los elementos de la columna pivote,
   $$\frac{36}{2} \text{ y } \frac{36}{1}$$
   y el mínimo se obtiene para el primer elemento, luego la fila pivote buscada es la primera.
   El elemento pivote es $(1, 2)$. La variable x_2 entra como variable básica y sale x_3.

5. Conversión de la columna PIVOTE en BÁSICA.
   Las transformaciones oportunas que podemos considerar son, en la fila pivote $F_1 \leftrightarrow \frac{1}{2} F_1$ (nombrándola de nuevo F_1) y a continuación:
   $$F_2 \leftrightarrow F_2 - F_1$$
   $$F_3 \leftrightarrow F_3 + 5F_1$$

	x_1	x_2	x_3	x_4	$\bar{b}$
x_2	$\frac{1}{2}$	1	$\frac{1}{2}$	0	18
x_4	$\frac{3}{2}$	0	$-\frac{1}{2}$	1	18
FO	$-\frac{1}{2}$	0	$\frac{5}{2}$	0	90

Si nos fijamos en la nueva fila FO hay elementos negativos, esto quiere decir que no hemos terminado, la solución encontrada que mejora la anterior ($z = 90$) es:

$$(0, 18, 0, 18)$$

Como la solución no es óptima seguimos aplicando el método. De nuevo,

3. Determinación de la columna PIVOTE.
   En la fila FO el único elemento negativo es $-\frac{1}{2}$ por tanto la columna pivote es la primera.
4. Determinación de la FILA PIVOTE

$$\frac{18}{1/2} \text{ y } \frac{18}{3/2}$$

   Luego la fila pivote es la segunda. El elemento pivote es (2, 1). La variable x_1 entra como variable básica y sale x_4.

5. Conversión de la columna PIVOTE en BÁSICA.
   En la fila PIVOTE aplicamos la transformación $F_2 \leftrightarrow \frac{2}{3} F_2$ y con esta nueva fila calculamos las siguientes transformaciones sucesivas.

$$F_1 \leftrightarrow F_1 - \frac{1}{2} F_2$$

$$F_3 \leftrightarrow F_3 + \frac{1}{2} F_2$$

	x_1	x_2	x_3	x_4	$\overline{b}$
x_2	0	1	$\frac{2}{3}$	$-\frac{1}{3}$	12
x_1	1	0	$-\frac{1}{3}$	$\frac{2}{3}$	12
FO	0	0	$\frac{7}{3}$	$\frac{1}{3}$	96

En la fila FO no hay elementos negativos, por lo tanto el método simplex nos asegura que hemos calculado la solución óptima, y ésta es (12, 12).

El valor máximo para $z = 3x_1 + 5x_2$ es $z = 96$, o el mínimo para $z = -3x_1 - 5x_2$ es $z = -96$ y ambos se alcanzan en el punto (12, 12, 0, 0) es decir en $x_1 = 12$, $x_2 = 12$.

Ejercicio 10

Resuélvase por el método simplex el ejercicio 6.

Solución:

Tal como vimos en el ejercicio 6, tras introducir las variables de holgura el problema es

$$\begin{cases} \max F = 4x_1 + 2x_2 \\ x_1 + 3x_2 + x_3 = 15 \\ 2x_1 - x_2 + x_4 = 8 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

Aplicando el método, resolveremos el problema paso a paso.

1. Construcción de la tabla inicial

	x_1	x_2	x_3	x_4	$\overline{b}$
x_3	1	3	1	0	15
x_4	2	−1	0	1	8
FO	−4	−2	0	0	0

2. La solución básica es (0, 0, 0, 15, 8). No es óptima, ya que $F = 0$. Esto concuerda con la predicción del método, ya que en la *FO* hay entradas negativas y, por tanto, debemos continuar.
3. Determinación de la COLUMNA PIVOTE.
En la fila *FO* el elemento más negativo es –4, por lo que la columna pivote es la primera.
4. Determinación de la FILA PIVOTE.
Calculamos los cocientes correspondientes para los elementos de la columna pivote,
$$\frac{15}{1} \text{ y } \frac{8}{2}$$
El mínimo se obtiene para el segundo elemento por lo que la fila pivote buscada es la segunda. El elemento pivote es (2, 1), la variable x_1 entra como variable básica y sale como variable básica x_4.
5. Conversión de la columna PIVOTE en BÁSICA.
Las transformaciones que vamos a realizar son:
   - en la fila pivote se realiza $F_2 \leftrightarrow \frac{1}{2} F_2$ (renombrándola de nuevo F_2)
   - a continuación operamos con las siguientes filas
$$F_1 \leftrightarrow F_1 - F_2$$
$$F_3 \leftrightarrow F_3 + 4F_2$$

	x_1	x_2	x_3	x_4	$\bar{b}$
x_3	0	$\frac{7}{2}$	1	$-\frac{1}{2}$	11
x_1	1	$-\frac{1}{2}$	0	$\frac{1}{2}$	4
FO	0	–4	0	2	16

Si nos fijamos en la nueva fila *FO* todavía hay elementos negativos. Esto significa que no hemos terminado, la solución encontrada que mejora la anterior ($F = 16$) es:
$$(4, 0, 11, 0)$$
Como la solución no es óptima seguimos aplicando el método. De nuevo,

3. Determinación de la columna PIVOTE.
En la fila *FO* el único elemento negativo es –4 por lo tanto la columna pivote es la segunda.

4. Determinación de la FILA PIVOTE. Sólo hay un término positivo, el correspondiente a la primera fila. Luego ésta será la fila pivote. El elemento pivote es (1, 2). La variable x_2 entra como variable básica y sale x_3.

5. Conversión de la columna PIVOTE en BÁSICA.
En la fila PIVOTE aplicamos la transformación $F_1 \leftrightarrow \frac{2}{7} F_1$ y con esta nueva fila calculamos las siguientes transformaciones sucesivas.
$$F_2 \leftrightarrow F_2 + \frac{1}{2} F_1$$
$$F_3 \leftrightarrow F_3 + 4F_1$$

	x_1	x_2	x_3	x_4	$\bar{b}$
x_2	0	1	$\frac{2}{7}$	$-\frac{1}{7}$	$\frac{22}{7}$
x_1	1	0	$\frac{1}{7}$	$\frac{3}{7}$	$\frac{39}{7}$
FO	0	0	$\frac{8}{7}$	$\frac{10}{7}$	$\frac{268}{7}$

En la fila *FO* no hay elementos negativos, por lo tanto el método simplex nos asegura que hemos calculado la solución máxima, y ésta es $\left(\frac{39}{7}, \frac{22}{7}\right)$. Y el valor máximo es $F = \frac{268}{7}$.