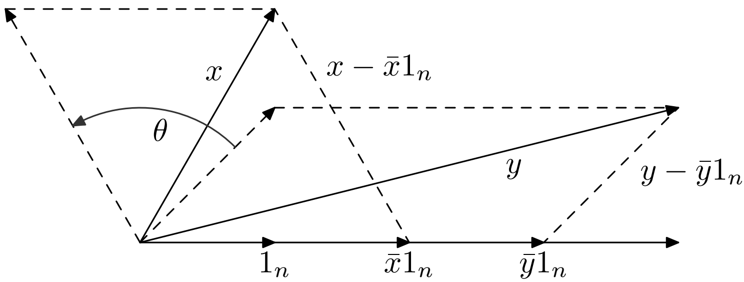
Universidad de Costa Rica

Escuela de Matemática

ALGEBRA LINEAL



Carlos Arce S.

William Castillo E.

Jorge González V.

2003

Algebra Lineal

Carlos Arce S., William Castillo E., Jorge González

Tercera Edición 2002

Edición y Diagramación: Carlos Arce.

Diseño de portada:

Tipografía y compilador: LaTeX

Presentación

Los materiales de este libro han sido desarrollados para el curso introductorio de álgebra lineal, dirigido a estudiantes de ingeniería y otras carreras de la Universidad de Costa Rica. En ellos se resume la experiencia de impartir el curso durante seis años.

Nuestro propósito ha sido dotar a los estudiantes de un folleto con los temas básicos de la teoría del álgebra lineal que, resalte los aspectos geométricos del tema, no oculte algunas demostraciones fundamentales que permiten reconocer las vinculaciones entre distintos conceptos y muestre algunas de sus aplicaciones. El libro incluye además listas abundantes de ejercicios que privilegian la aplicación de conceptos sobre los aspectos púramente algorítmicos.

Esta nueva edición presenta algunos cambios respecto a la anterior. Todos los capítulos fueron revisados, se mejoraron los gráficos y en algunos se agregaron nuevos ejercicios. El capítulo de regresión lineal se reformuló completamente, haciendo una mejor exposición del tema. Por otra parte, el libro incluye ahora, una nueva presentación en formato PDF con hypertexto, que permite ser distribuido por medios electrónicos y que constituye un nuevo recurso para que los profesores hagan sus exposiciones con ayuda de ... o equipos similares.

Agradecemos a los colegas de la cátedra de álgebra lineal sus observaciones sobre errores en la edición anterior y las sugerencias para mejorar el material.

Índice General

Capítulo 1

Sistemas de ecuaciones lineales

Comenzamos el estudio de los temas del algebra lineal con el método de reducción gaussiana para resolver sistemas de ecuaciones lineales. Su introducción se hará de manera operativa y señalando los principales resultados que caracterizan la solución de un sistema. No nos ocuparemos de las demostraciones, pero sí enfatizamos en la lectura de la información que provee la forma escalonada de la matriz aumentada del sistema. Se espera que la experiencia adquirida al conocer dicho método se convierta en una base concreta que permita abordar de mejor manera el estudio de las matrices, sus operaciones y otros conceptos más complejos como la invertibilidad de matrices y la independencia lineal. En la última sección se incluyen algunos ejemplos de redes elementales con el fin de ilustrar una forma de modelación mediante sistemas de ecuaciones lineales.

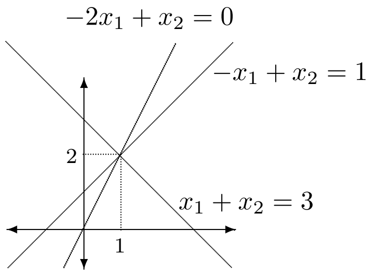
1.1 Sistemas con dos incógnitas

Los sistemas de ecuaciones lineales con dos incógnitas tienen la particularidad de que los puntos (x_1, x_2) que satisfacen cada ecuación pueden ser visualizados como los puntos de una recta, es decir, cada ecuación del sistema es la ecuación de una recta. Por eso, el sistema y su solución tienen una interpretación gráfica como se ilustra en el siguiente ejemplo.

Ejemplo 1.1 Sea el siguiente sistema de 3 ecuaciones lineales en dos incógnitas:

$$\left\{\begin{array}{rcl} x_1 + x_2 & = & 3 \\ -2x_1 + x_2 & = & 0 \\ -x_1 + x_2 & = & 1 \end{array}\right.$$

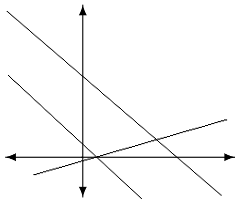
Para resolver este sistema por el método gráfico se dibujan en un mismo sistema de coordenadas, las tres rectas asociadas con las tres ecuaciones, como se ve en la siguiente figura.



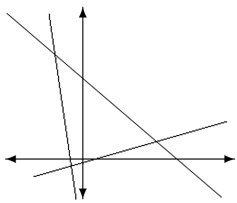
Un punto (a, b) es un punto solución del sistema solo si satisface cada ecuación, o lo que es lo mismo, solo si es un punto de cada una de las rectas (cada recta pasa por (a, b)). En nuestro ejemplo este punto es $(1, 2)$.

Cuando se tiene un sistema de m ecuaciones lineales distintas, con dos incógnitas, se tendrían las siguientes posibilidades:

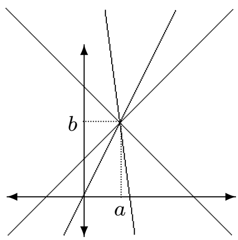
1. Hay dos ecuaciones correspondientes a rectas paralelas como se ilustra en la siguiente figura. En este caso no puede haber un mismo punto que pertenezca a todas las rectas y decimos que el conjunto solución del sistema es: $S = \emptyset$. El gráfico siguiente ilustra esta situación, para el caso de tres ecuaciones.



2. No hay rectas paralelas pero tampoco hay un punto en que concurran todas las rectas, como lo muestra el gráfico. En este caso la solución del sistema también es el conjunto vacío: $S = \emptyset$.

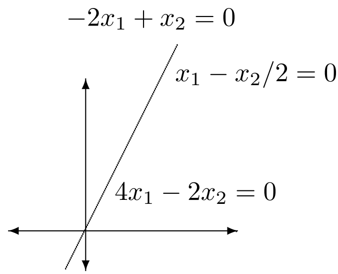


3. Todas las rectas concurren en exactamente un punto. En este caso la solución es el punto de concurrencia, $S = \{(a, b)\}$, como en el gráfico siguiente, para el caso de 4 ecuaciones.



4. Todas las ecuaciones corresponden a la misma recta, en cuyo caso todos los puntos de la recta son solución del sistema. Esto se ilustra con el siguiente sistema:

$$\left\{\begin{array}{rcl} 4x_1 - 2x_2 & = & 0 \\ -2x_1 + x_2 & = & 0 \\ x_1 - x_2/2 & = & 0 \end{array}\right.$$



El conjunto solución es: $S = \{(x_1, x_2) | x_2 = 2x_1\}$.

Para sistemas de ecuaciones lineales con tres incógnitas, cada ecuación en el sistema es la ecuación de un plano. De modo que también en este caso podemos dar un significado geométrico a la solución. Sin embargo, a falta de un mejor conocimiento sobre las ecuaciones de planos en $\mathbb{R}^3$, posponemos su discusión hasta el capítulo 6.

Un sistema de ecuaciones lineales se interpreta también como una formulación matemática de un conjunto de restricciones, una por cada ecuación. Esto se comprende mejor si las variables tienen un significado físico como se ilustra en el ejemplo siguiente.

Ejemplo 1.2 En un acuarium hay dos especies de peces A y B que se alimentan con dos clases de alimentos. En la tabla siguiente se indican el consumo promedio diario de alimento de ambas especies y la cantidad de alimento disponible por clase (columna cuatro), en gramos.

	Especies		
Alimentos	A	B	Total
clase 1	2	1	25
clase 2	4	3	55

El problema a resolver consiste en calcular la cantidad máxima de peces de cada especie que pueden vivir en el acuarium.

Sean x_A y x_B las cantidades máximas de peces de las especies A y B respectivamente, que pueden coexistir en el acuarium. El problema es entonces calcular x_A y x_B. Las restricciones del problema son :

$2x_A + x_B =$ cantidad de alimento de la clase 1, consumida diariamente por la totalidad de los peces. Es decir : $2x_A + x_B = 25$.

$4x_A + 3x_B =$ cantidad de alimento de la clase 2, consumida diariamente por la totalidad de los peces. Es decir : $4x_A + 3x_B = 55$.

La solución al problema planteado se encuentra resolviendo el siguiente sistema de ecuaciones lineales 2×2 :

$$\left\{ \begin{array}{cccc} 2x_A & + & x_B & = & 25 \\ 4x_A & + & 3x_B & = & 55 \end{array} \right.$$

De la primera ecuación se obtiene $x_B = 25 - 2x_A$. Esta expresión se sustituye en la segunda ecuación: $4x_A + 3(25 - 2x_A) = 55$ y despejando x_A, se tiene: $x_A = 10$. Luego

$$x_B = 25 - 2x_A = 25 - 2(10) = 5.$$

Por lo tanto pueden coexistir diariamente 10 peces de la especie A y 5 de la especie B.

1.2 Sistemas $n \times m$

Se dice que un conjunto de ecuaciones, cada una de las cuales restringe los valores que pueden asumir m variables o incógnitas $x_1, \ldots, x_m$, representan un sistema de ecuaciones, cuando nos interesamos por los n-étuplos $(x_1, x_2, \ldots, x_m)$ que satisfacen simultáneamente todas las ecuaciones. Una ecuación en las variables $x_1, \ldots, x_m$, es lineal si es de la forma:

$$a_1x_1 + a_2x_2 + \cdots + a_mx_m = b$$

donde $a_1,\ a_2,\ \ldots, a_m$ y b son números reales dados. En general, escribimos un sistema de ecuaciones lineales como:

Definición 1.1 (Sistema de ecuaciones lineales $n \times m$)

Un sistema de n ecuaciones lineales y m incógnitas $x_1, \ldots, x_m$, se escribe en la forma:

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2m}x_m & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nm}x_m & = & b_n \end{array}\right.$$

donde los números a_{ij} y b_i son conocidos. El símbolo a_{ij} denota el coeficiente en la ecuación i asociado a la variable j.

En adelante usaremos la expresión "sistema $n \times m$" para referirnos a un sistema de n ecuaciones lineales con m incógnitas.

Ejemplo 1.3 El siguiente es un sistema 3×4, en las incógnitas x_1, x_2, x_3 y x_4:

$$\left\{\begin{array}{rcrcrcrcr} & & 2x_2 & + & -4x_3 & + & x_4 & = & 10 \\ -2x_1 & + & 5x_2 & + & -12x_3 & + & -6x_4 & = & 10 \\ x_1 & + & -1x_2 & + & 3x_3 & + & 3x_4 & = & 1 \end{array}\right.$$

Definición 1.2 (Solución de un sistema $n \times m$)

Dado un sistema de n ecuaciones lineales con m incógnitas, se dice que el m-étuplo $(x_1, x_2, \ldots, x_m)$ es una solución del sistema, si satisface cada una de las n ecuaciones. Y al conjunto S de todas las soluciones del sistema, se le llama conjunto solución del sistema.

Ejemplo 1.4 Verifique que $(-1, 4, 0, 2)$ es una solución del sistema en el ejemplo 1.3.

Claramente se tiene que:

$$\left\{\begin{array}{rcrcrcrcr} & & 2(4) & + & -4(0) & + & (2) & = & 10 \\ -2(-1) & + & 5(4) & + & -12(0) & + & -6(2) & = & 10 \\ (-1) & + & -1(4) & + & 3(0) & + & 3(2) & = & 1 \end{array}\right.$$

¿$S = \{(-1, 4, 0, 2)\}$ será el conjunto solución del sistema? Si es así, la solución anterior sería la única solución. Sin embargo, todavía no hemos planteado los resultados necesarios para poder responder.

1.2.1 Sustitución hacia atrás

Algunos sistemas de ecuaciones lineales tienen una forma especial que permite resolverlos fácilmente. El método para hacerlo es conocido como "sustitución hacia atrás" y se introduce con el siguiente ejemplo:

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & -1x_2 & + & x_3 & + & 3x_4 & = & -1 \\ & & x_2 & + & -2x_3 & + & 2x_4 & = & 4 \\ & & & & x_3 & + & -1x_4 & = & 2 \\ & & & & & & x_4 & = & 2 \end{array}\right.$$

Dada la forma de este sistema, de la última ecuación se obtiene el valor para la variable x_4:

$$x_4 = 2$$

Este valor se sustituye en las restantes ecuaciones, lo que permite, a partir de la ecuación 3, determinar el valor de la variable x_3:

$$x_3 - 1(2) = 2 \implies x_3 = 4$$

De nuevo el valor de la variable x_3 se sustituye en las restantes ecuaciones y se obtiene de la ecuación 2 que:

$$x_2 - 2(4) + 2(2) = 4 \implies x_2 = 8$$

Finalmente, conociendo los valores de x_4, x_3 y x_2, de la ecuación 1 se tiene que:

$$x_1 - 1(8) + 4 + 3(2) = -1 \implies x_1 = -3$$

Así $(x_1, x_2, x_3, x_4) = (-3, 8, 4, 2)$ es una solución del sistema.

Al reconocer la forma de aquellos sistemas que se resuelven fácilmente, para resolver un sistema cualquiera, el método de reducción gaussiana se ocupa de transformarlo en uno de este tipo. Sin embargo, claramente, esta transformación debe hacerse de forma que no se cambie el conjunto solución.

1.2.2 Operaciones elementales

Tres tipos de operaciones sobre las ecuaciones de un sistema, llamadas **operaciones elementales**, permiten transformar un sistema a otros, sin modificar el conjunto solución:

a) Multiplicar una ecuación por un número real distinto de cero, esto es, multiplicar a ambos lados de la igualdad por un mismo número.

b) Multiplicar una ecuación por un número real y sumarla a otra ecuación.

c) Intercambiar de posición dos ecuaciones cualesquiera.

Parece bastante claro que aplicando operaciones del tipo a) y c), a un sistema dado, se obtengan otros sistemas con el mismo conjunto solución. Lo mismo ocurre si se aplican operaciones del tipo b), aunque en este caso la justificación tal vez no sea tan evidente. En todo caso, demostrar que las operaciones elementales no cambian el conjunto solución de un sistema es un ejercicio importante, que se pospone hasta disponer del producto de matrices.

1.2.3 Sistemas equivalentes y reducción gaussiana

Definición 1.3 (Sistemas equivalentes)

Dos sistemas de ecuaciones lineales con m incógnitas son equivalentes si tienen el mismo conjunto solución.

De esta definición se tiene que si un sistema de ecuaciones lineales es el resultado de aplicarle operaciones elementales a otro, ambos sistemas son equivalentes.

El método de reducción gaussiana permite resolver sistemas de ecuaciones lineales, explotando la idea de equivalencia. Es decir, transformando el sistema mediante operaciones elementales, hasta obtener uno cuya forma permite resolverlo mediante “sustitución hacia atrás”. El siguiente ejemplo, ilustra la idea.

Ejemplo 1.5 Aplicando la siguiente secuencia de operaciones elementales al sistema del ejemplo 1.3 se obtienen los siguientes sistemas equivalentes:

Intercambiando de posición las ecuaciones 1 y 3 se obtiene:

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & -1x_2 & + & 3x_3 & + & 3x_4 & = & 1 \\ -2x_1 & + & 5x_2 & + & -12x_3 & + & -6x_4 & = & 10 \\ & & 2x_2 & + & -4x_3 & + & x_4 & = & 10 \end{array}\right.$$

Multiplicando la primera ecuación por 2 y sumándola a la segunda:

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & -1x_2 & + & 3x_3 & + & 3x_4 & = & 1 \\ & & 3x_2 & + & -6x_3 & & & = & 12 \\ & & 2x_2 & + & -4x_3 & + & x_4 & = & 10 \end{array}\right.$$

Multiplicando por $\frac{1}{3}$ la segunda ecuación tenemos:

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & -1x_2 & + & 3x_3 & + & 3x_4 & = & 1 \\ & & x_2 & + & -2x_3 & & & = & 4 \\ & & 2x_2 & + & -4x_3 & + & x_4 & = & 10 \end{array}\right.$$

Multiplicando la segunda ecuación por (-2) y sumándola a la tercera, nos da:

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & -1x_2 & + & 3x_3 & + & 3x_4 & = & 1 \\ & & x_2 & + & -2x_3 & & & = & 4 \\ & & & & & & x_4 & = & 2 \end{array}\right.$$

Este último sistema es equivalente al inicial —tiene el mismo conjunto solución— y su solución se puede obtener mediante sustitución hacia atrás, en la siguiente forma:

- de la última ecuación se tiene que $x_4 = 2$
- de la segunda ecuación: $x_2 - 2x_3 = 4 \implies x_2 = 4 + 2x_3$. Observe que en este caso no se conoce el valor de x_3, por lo tanto el valor de x_2 se hace depender del de x_3.
- Y de la primera ecuación:

$$\begin{aligned} x_1 - (4 + 2x_3) + 3x_3 + 3(2) = 1 \quad &\implies \quad x_1 + x_3 + 2 = 1 \\ &\implies \quad x_1 = -1 - x_3. \end{aligned}$$

En resumen

$$\begin{cases} x_1 & = & -1 - x_3 \\ x_2 & = & 4 + 2x_3 \\ x_4 & = & 2 \end{cases}$$

Se observa que tanto el valor de x_1 como el de x_2 dependen de x_3, esto es, se puede dar un valor arbitrario a x_3, lo cual se indica asignádole cualquier número t, $x_3 = t$. Así, $x_1 = -1-t$, $x_2 = 4+2t$ y en todos los casos $x_4 = 2$. Luego se dice que la solución del sistema depende de un parámetro t y es dada por:

$$S = \{(x_1, x_2, x_3, x_4) = (-1 - t, 4 + 2t, t, 2) | t \in I\!R\}$$

1.3 Solución de sistemas representados como matrices

La aplicación del método de reducción gaussiana para resolver un sistema, resulta más sencillo con el empleo de matrices.

1.3.1 Matriz del sistema y matriz aumentada

En un sistema $n \times m$:

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2m}x_m & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nm}x_m & = & b_n \end{cases}$$

se puede distinguir el siguiente arreglo rectangular de números a_{ij} que se conocerá como **matriz del sistema** y se denotará como $A = (a_{ij})$:

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

Una matriz A como esta, se dice que es una matriz $n \times m$, es decir, con n filas y m columnas.

Si a la matriz del sistema se agregan los valores b_i como una última columna, el nuevo arreglo se denomina **matriz aumentada** del sistema, el cual se suele denotar como $(A|b)$:

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{array}\right)$$

Ejemplo 1.6 En el ejemplo 1.3, el sistema 3×4 dado:

$$\left\{\begin{array}{rcrcrcrcr} & & 2x_2 & + & -4x_3 & + & x_4 & = & 10 \\ -2x_1 & + & 5x_2 & + & -12x_3 & + & -6x_4 & = & 10 \\ x_1 & + & -1x_2 & + & 3x_3 & + & 3x_4 & = & 1 \end{array}\right.$$

tiene como matrices del sistema y aumentada:

Matriz del sistema *Matriz aumentada*

$$\begin{pmatrix} 0 & 2 & -4 & 1 \\ -2 & 5 & -12 & -6 \\ 1 & -1 & 3 & 3 \end{pmatrix} \qquad \left(\begin{array}{rrrr|r} 0 & 2 & -4 & 1 & 10 \\ -2 & 5 & -12 & -6 & 10 \\ 1 & -1 & 3 & 3 & 1 \end{array}\right)$$

1.3.2 Operaciones elementales sobre las filas de una matriz

Cuando los sistemas se representan como matrices aumentadas, las operaciones elementales sobre ecuaciones se denotan en la siguiente forma:

	Notación:
af_i	Multiplicar (ecuación) fila i por la constante no nula a.
$af_i + f_j$	Multiplicar la (ecuación) fila i por un número real a y sumarla a la (ecuación) fila j.
f_i, f_j	Intercambiar las (ecuaciones) filas i y j.

Cuando estas operaciones se escriben sobre una flecha, por ejemplo $\xrightarrow{af_1 + f_2}$, se indica que la matriz (sistema) a la derecha es equivalente al anterior, por la aplicación de la operación elemental indicada.

Ejemplo 1.7 El procedimiento de reducción gaussiana aplicado en el ejemplo (1.5) se resume en la siguiente forma:

$$\left(\begin{array}{rrrr|r} 0 & 2 & -4 & 1 & 10 \\ -2 & 5 & -12 & -6 & 10 \\ 1 & -1 & 3 & 3 & 1 \end{array}\right) \xrightarrow{f_1, f3} \left(\begin{array}{rrrr|r} 1 & -1 & 3 & 3 & 1 \\ -2 & 5 & -12 & -6 & 10 \\ 0 & 2 & -4 & 1 & 10 \end{array}\right)$$

$$\xrightarrow{2f_1 + f_2} \left(\begin{array}{rrrr|r} 1 & -1 & 3 & 3 & 1 \\ 0 & 3 & -6 & 0 & 12 \\ 0 & 2 & -4 & 1 & 10 \end{array}\right) \xrightarrow{\frac{1}{3}f_2}$$

$$\left(\begin{array}{rrrr|r} 1 & -1 & 3 & 3 & 1 \\ 0 & 1 & -2 & 0 & 4 \\ 0 & 2 & -4 & 1 & 10 \end{array}\right) \xrightarrow{-2f_2 + f_3} \left(\begin{array}{rrrr|r} 1 & -1 & 3 & 3 & 1 \\ 0 & 1 & -2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

La última matriz tiene la forma especial que se busca al aplicar el método de reducción gaussiana, es decir, representa un sistema con la forma requerida para resolverlo por sustitución hacia atrás, como fue hecho en el ejemplo 1.5.

En general, para aplicar el método de reducción gaussiana es necesario tener claro, cuál es esa forma general de la matriz aumentada que se busca y que permite terminar de resolverlo mediante sustitución hacia atrás.

1.3.3 Matriz escalonada

Para precisar las ideas sobre la forma especial de los sistemas que pueden resolverse por el método de sustitución hacia atrás, se utiliza el concepto de **matriz escalonada**. Así se dirá que un sistema con esta forma especial tiene como matriz aumentada una matriz en la forma escalonada.

Definición 1.4 (Matriz escalonada)

Sea A una matriz $n \times m$. A es escalonada si es nula o si satisface las tres condiciones siguientes:

i. *El* **primer** *elemento no nulo de cada fila, si existe, es un 1.*

ii. *El* **primer 1** *de la segunda fila y sucesivas está a la derecha del primer 1 de la fila anterior.*

iii. *Si tiene filas nulas —compuestas sólo de ceros— estas aparecen en la parte inferior de la matriz, abajo de las filas no nulas.*

Ejemplo 1.8 Las matrices $(A|b)$ y $(B|c)$ siguientes, tienen la forma escalonada:

Matriz $(A|b)$

$$\left(\begin{array}{rrrr|r} 1 & -1 & 3 & 5 & 5 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

Matriz $(B|c)$

$$\left(\begin{array}{rrrr|r} 1 & -1 & 0 & 3 & 1 \\ 0 & 1 & -2 & 0 & 4 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

La matrices escalonadas, además de caracterizar los sistemas que pueden ser resueltos por sustitución hacia atrás, permiten reconocer fácilmente los sistemas que tienen solución de los que no la tienen. Por ejemplo, si consideramos los sistemas de ecuaciones que representan las matrices aumentadas $(A|b)$ y $(B|c)$ anteriores, observamos:

Sistema con matriz aumentada $(A|b)$. El sistema correspondiente a esta matriz no tiene solución, porque hay una ecuación **inconsistente**:

Observe que la última ecuación es:

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 1$$

y claramente, no existen valores x_1, x_2, x_3 y x_4 que multiplicados por cero y sumados puedan dar 1.

Una ecuación de este tipo la llamaremos inconsistente y basta con que un sistema tenga una ecuación inconsistente, para que no tenga solución. Claramente no puede existir una solución que satisfaga **todas** las ecuaciones.

Sistema con matriz aumentada $(B|c)$**.** Este sistema tiene solución puesto que no tiene ecuaciones inconsistentes y se determina por "sustitución hacia atrás":

- de la última fila (ecuación) no nula se tiene que $x_4 = 4$.
- de la tercera ecuación: $x_3 + 4 = -2 \implies x_3 = -6$.
- de la segunda ecuación: $x_2 - 2(-6) = 4 \implies x_2 = -8$.
- Finalmente, de la primera ecuación se obtiene que:

$$x_1 - 1(-8) + 3(4) = 1 \implies x_1 = -19.$$

En conclusión, el sistema tiene una única solución:

$$(x_1, x_2, x_3, x_4) = (-19, -8, -6, 4).$$

La última ecuación de este sistema, la que corresponde a la fila de ceros, es una **ecuación superflua**:

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 0$$

Observe que cualesquiera valores $x_1, x_2, x_3,$ y x_4 la satisfacen, de manera que no constituye ninguna restricción. Si se omite se obtiene un sistema con el mismo conjunto solución.

Un sistema de ecuaciones puede contener **ecuaciones superfluas o redundantes**, en el sentido de que las restricciones que establecen ya están contempladas en las otras ecuaciones. De esta manera no aportan información adicional para resolver el sistema. Caracterizar dichas situaciones es uno de los objetivos más importantes al estudiar los sistemas de ecuaciones lineales, para esto, en el capítulo siguiente se utilizará la idea de dependencia e independencia lineal de las filas de la matriz aumentada. En esta etapa nos conformamos con observar, en cada ejemplo y ejercicio que se resuelva, que las ecuaciones redundantes se convierten en filas nulas, cuando se obtiene la forma escalonada de la matriz aumentada del sistema.

1.3.4 Reducción gaussiana y matriz escalonada

Al hacer operaciones elementales para resolver un cierto sistema, por reducción gaussiana, se busca darle la forma escalonada a la matriz aumentada. En el siguiente ejemplo se verifica que para lograr este propósito no hay una secuencia única de operaciones elementales a realizar. Y aún más, el resultado final de la secuencia de operaciones elementales (la forma escalonada), no siempre será la misma. Es decir, una matriz aumentada puede tener varias formas escalonadas (equivalentes), que representan el mismo sistema de ecuaciones.

Ejemplo 1.9 Resolver el siguiente sistema 4×4, mediante reducción gaussiana:

$$\left\{\begin{array}{rcrcrcrcr} 2x_1 & + & -6x_2 & + & 12x_3 & + & 16x_4 & = & 70 \\ x_1 & + & -2x_2 & + & 6x_3 & + & 6x_4 & = & 26 \\ -1x_1 & + & 3x_2 & + & -3x_3 & + & -7x_4 & = & -30 \\ & & 4x_2 & + & 3x_3 & + & -6x_4 & = & -26 \end{array}\right.$$

Solución: la siguiente secuencia de operaciones elementales transforman la matriz aumentada en una forma escalonada:

$$\left(\begin{array}{rrrr|r} 2 & -6 & 12 & 16 & 70 \\ 1 & -2 & 6 & 6 & 26 \\ -1 & 3 & -3 & -7 & -30 \\ 0 & 4 & 3 & -6 & -26 \end{array}\right) \xrightarrow{\frac{1}{2}f_1} \left(\begin{array}{rrrr|r} 1 & -3 & 6 & 8 & 35 \\ 1 & -2 & 6 & 6 & 26 \\ -1 & 3 & -3 & -7 & -30 \\ 0 & 4 & 3 & -6 & -26 \end{array}\right)$$

$$\xrightarrow{\begin{array}{c} -f_1+f_2 \\ f_1+f_3 \end{array}} \left(\begin{array}{rrrr|r} 1 & -3 & 6 & 8 & 35 \\ 0 & 1 & 0 & -2 & -9 \\ 0 & 0 & 3 & 1 & 5 \\ 0 & 4 & 3 & -6 & -26 \end{array}\right) \xrightarrow{-4f_2+f_4}$$

$$\left(\begin{array}{rrrr|r} 1 & -3 & 6 & 8 & 35 \\ 0 & 1 & 0 & -2 & -9 \\ 0 & 0 & 3 & 1 & 5 \\ 0 & 0 & 3 & 2 & 10 \end{array}\right) \xrightarrow{\begin{array}{c} -f_3+f_4 \\ (1/3)f_3 \end{array}} \left(\begin{array}{rrrr|r} 1 & -3 & 6 & 8 & 35 \\ 0 & 1 & 0 & -2 & -9 \\ 0 & 0 & 1 & 1/3 & 5/3 \\ 0 & 0 & 0 & 1 & 5 \end{array}\right)$$

Pero también, si se inicia con la operación elemental $\xrightarrow{f_1, f_2}$ para obtener el primer 1, en la posición (1,1) y se continúa con las operaciones $\xrightarrow{-2f_1+f_2}$, $\xrightarrow{f_1+f_3}$, $\xrightarrow{f_2, f_3}$, $\xrightarrow{2f_2+f_3}$, $\xrightarrow{-4f_2, f_4}$, $\xrightarrow{1/6 f_3}$ $\xrightarrow{9f_3, f_4}$, se obtiene finalmente otra forma escalonada:

$$\left(\begin{array}{cccc|c} 1 & -2 & 6 & 6 & 26 \\ 0 & 1 & 3 & -1 & -4 \\ 0 & 0 & 1 & 1/3 & 5/3 \\ 0 & 0 & 0 & 1 & 5 \end{array}\right)$$

Sin embargo, observe que las operaciones elementales $\xrightarrow{-3f_3+f_2}$ y $\xrightarrow{-f_2+f_1}$ sobre esta última matriz la reducen a la anterior forma escalonada, es decir, ambas formas escalonadas representan sistemas 4×4 con la misma solución.

Por otra parte, la solución del sistema se obtiene por sustitución hacia atrás:

- de la última ecuación: $x_4 = 5$
- de la tercera ecuación: $x_3 + 5/3 = 5/3 \implies x_3 = 0$.
- de la segunda ecuación: $x_2 + 3(0) - 1(5) = -4 \implies x_2 = 1$
- finalmente: $x_1 - 2(-4) + 6(0) + 6(5) = 26 \implies x_1 = -2$.

Resumiendo, la única solución es $(-2, 1, 0, 5)$.

1.3.5 Reducción de Gauss-Jordan

Cuando se resuelve un sistema de ecuaciones por reducción gaussiana, el proceso que corresponde a la sustitución hacia atrás, también puede ser realizado haciendo más operaciones elementales, hasta obtener un sistema equivalente cuya solución resulta evidente.

En este caso, la forma que debemos buscar en la matriz aumentada del sistema es la denominada **forma escalonada reducida**, definida a continuación y el método de solución de sistemas resultante se conoce como método de Gauss-Jordan.

Definición 1.5 (Matriz escalonada reducida)

Una matriz A, $n \times m$, es escalonada reducida si es escalonada y además todo elemento en una columna, arriba del primer uno de cualquier fila, es cero.

Es decir, la forma escalonada reducida se obtiene de una forma escalonada, haciendo cero los elementos de la matriz arriba de los primeros unos de cada fila.

Ejemplo 1.10 La matriz A siguiente tiene la forma escalonada reducida, sin embargo la matriz B no es escalonada reducida:

$$\begin{array}{cc} \text{Matriz } A & \text{Matriz } B \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 5 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 & 3 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Ejemplo 1.11 Escriba todas las formas escalonadas y escalonadas reducidas, de una matriz 2×2.

Solución: Todas las formas escalonadas son:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix} \text{ y } \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ con } a, b \in I\!R$$

Todas las formas escalonadas reducidas son:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix} \text{ con } a, b \in I\!R$$

■

Ejemplo 1.12 Resolver el sistema del ejemplo 1.9, por el método de Gauss-Jordan.

Solución: Hay que determinar la forma escalonada reducida de la matriz aumentada del sistema. Una secuencia de operaciones elementales para lograr este objetivo, es dada por las que se hicieron en el ejemplo 1.9 para obtener la forma escalonada, a las que se deben agregar las siguientes:

$$\left(\begin{array}{cccc|c} 1 & -3 & 6 & 8 & 35 \\ 0 & 1 & 0 & -2 & -9 \\ 0 & 0 & 1 & 1/3 & 5/3 \\ 0 & 0 & 0 & 1 & 5 \end{array}\right) \xrightarrow{3f_2 + f_1} \left(\begin{array}{cccc|c} 1 & 0 & 6 & 2 & 8 \\ 0 & 1 & 0 & -2 & -9 \\ 0 & 0 & 1 & 1/3 & 5/3 \\ 0 & 0 & 0 & 1 & 5 \end{array}\right)$$

$$\xrightarrow{-6f_3 + f_1} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -2 \\ 0 & 1 & 0 & -2 & -9 \\ 0 & 0 & 1 & 1/3 & 5/3 \\ 0 & 0 & 0 & 1 & 5 \end{array}\right) \xrightarrow{\substack{(-1/3)f_4 + f_3 \\ 2f_4 + f_2}}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 \end{array}\right) \text{ luego } \left\{\begin{array}{cccccc} x_1 & & & & = & -2 \\ & x_2 & & & = & 1 \\ & & x_3 & & = & 0 \\ & & & x_4 & = & 5 \end{array}\right.$$

Así, la solución única del sistema es:

$$(x_1, x_2, x_3, x_4) = (-2, 1, 0, 5).$$

1.4 Matrices equivalentes y rango

Para caracterizar los sistemas de ecuaciones lineales que tienen solución y aquellos que no la tienen se introducirá la noción de rango de una matriz, lo cual requerirá ampliar un poco la notación de sistemas.

1.4.1 Escritura matricial de sistemas

Llamaremos "vector fila" a cualquier n-tuplo

$$(a_1, a_2, \ldots, a_n),$$

compuesto por n números reales. Y cuando este n-tuplo se escribe como una columna: $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ se dirá que es un "vector columna".

Un vector fila y un vector columna, siempre que tengan el mismo número de componentes, pueden ser multiplicados en la siguiente forma:

$$(a_1, a_2, \ldots, a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

Por ejemplo:

$$(2, -1, 3, 5) \begin{pmatrix} 4 \\ 2 \\ 0 \\ -1 \end{pmatrix} = 2(4) - 1(2) + 3(0) + 5(-1) = 1$$

Con esta operación, una ecuación como $2x_1 + 3x_2 - 5x_4 = 10$ puede ser escrita en la forma:

$$(2, 3, 0, -5) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 10.$$

Así, un sistema de ecuaciones $n \times m$

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2m}x_m & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nm}x_m & = & b_n \end{array} \right.$$

se puede expresar como:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

lo que se simplifica escribiendo simplemente

$$Ax = b,$$

donde $A = (a_{ij})$ es la matriz $n \times m$ de coeficientes del sistema,

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \quad \text{y} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ v_n \end{pmatrix},$$

son el vector de constantes en las ecuaciones a la derecha del igual y el vector de incógnitas del sistema, respectivamente.

Esta forma de escritura $Ax = b$ se conocerá como **escritura matricial** del sistema. Y en ella cada ecuación i del sistema se representa como: el producto del vector fila i de la matriz A multiplicado por el vector columna x igual a la componente i del vector b.

Ejemplo 1.13 El siguiente sistema 4×3, en las incógnitas x_1, x_2 y x_3:

$$\left\{ \begin{array}{rcrcrcr} x_1 & & & - & 2x_3 & = & 2 \\ -2x_1 & + & x_2 & + & 5x_3 & = & -1 \\ -x_1 & + & x_2 & + & 3x_3 & = & 1 \\ & & x_2 & + & x_3 & = & 3 \end{array} \right.$$

se escribe en su forma matricial como:

$$\begin{pmatrix} 1 & 0 & -2 \\ -2 & 1 & 5 \\ -1 & 1 & 3 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 3 \end{pmatrix}.$$

Si $Ax = b$ es un sistema de ecuaciones $n \times m$, la matriz A es la matriz del sistema y $(A|b)$ su matriz aumentada.

1.4.2 Equivalencia de matrices

Ahora nos ocuparemos de las matrices de los sistemas de ecuaciones lineales como objetos que existen independientemente de estos. Así una matriz $n \times m$ es un arreglo rectangular de números reales con n filas y m columnas. Y a estos objetos (matrices) les aplicaremos la noción de equivalencia, heredada de los sistemas.

Definición 1.6 (Matrices equivalentes)

Sean A, B matrices $n \times m$, se dice que la matriz A es equivalente por filas a B (o simplemente equivalente), si B se obtiene de A por aplicación de operaciones elementales de renglón. Se escribe $A \to B$.

Aunque la idea de equivalencia en matrices es la misma que en sistemas, presenta una pequeña diferencia. Dos sistemas de ecuaciones lineales en m variables pueden ser equivalentes, aún cuando tengan distinta cantidad de ecuaciones. Sin embargo, dos matrices aumentadas no son equivalentes si tienen un número de filas (ecuaciones) distinto, aún cuando representen sistemas equivalentes. Aunque es claro, en este caso, que agregando filas nulas a la que tenga menos filas, se pueden transformar en matrices equivalentes.

Ejemplo 1.14 Consideremos la matriz aumentada del sistema en el ejemplo 1.13, que denominamos B y apliquemos operaciones elementales hasta obtener una matriz escalonada C:

$$B = \begin{pmatrix} 1 & 0 & -2 & 2 \\ -2 & 1 & 5 & -1 \\ -1 & 1 & 3 & 1 \\ 0 & 1 & 1 & 3 \end{pmatrix} \xrightarrow[\;]{\substack{2f_1 + f_2 \\ f_1 + f_3}} \begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{pmatrix}$$

$$\xrightarrow[\;]{\substack{-f_2 + f_2 \\ -f_2 + f_4}} \begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = C$$

Así se tiene que B es equivalente a C. Y también es cierto que B es equivalente a:

$$\begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{pmatrix}, \text{ y también a } \begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 3 \end{pmatrix}.$$

Ahora se precisará el concepto de rango de una matriz, para lo cual se requiere del siguiente resultado fundamental.

Teorema 1.7 *Si A es una matriz $m\times n$ entonces existe una única matriz B con la forma escalonada reducida que es equivalente a A.*

No proponemos una demostración de este resultado, más bien, observemos con el siguiente ejemplo, que para el caso de matrices escalonadas, el resultado es falso.

Ejemplo 1.15 Sea $A = \begin{pmatrix} 2 & -2 & 3 & 4 \\ 2 & 1 & 0 & 3 \end{pmatrix}$

Calcule dos matrices escalonadas, equivalentes a A.

Solución:

$$\begin{pmatrix} 2 & -2 & 3 & 4 \\ 2 & 1 & 0 & 3 \end{pmatrix} \xrightarrow{-f_1+f_2} \begin{pmatrix} 2 & -2 & 3 & 4 \\ 0 & 3 & -3 & -1 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{2}f_1,\, \frac{1}{3}f_2} \begin{pmatrix} 1 & -1 & \frac{3}{2} & 2 \\ 0 & 1 & -1 & -\frac{1}{3} \end{pmatrix} = B_1$$

Por otra parte tenemos,

$$\begin{pmatrix} 2 & -2 & 3 & 4 \\ 2 & 1 & 0 & 3 \end{pmatrix} \xrightarrow{f_1,f_2} \begin{pmatrix} 2 & 1 & 0 & 3 \\ 2 & -2 & 3 & 4 \end{pmatrix} \xrightarrow{-f_1+f_2}$$

$$\begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & -3 & 3 & 1 \end{pmatrix} \xrightarrow[-\frac{1}{3}f_2]{-\frac{1}{2}f_1} \begin{pmatrix} 1 & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & 1 & -1 & -\frac{1}{3} \end{pmatrix} = B_2$$

Las matrices B_1 y B_2 son escalonadas, $A \to B_1$, $A \to B_2$ y $B_1 \neq B_2$.

∎

1.4.3 Rango de una matriz

Como cada matriz es equivalente a una única matriz en la forma escalonada reducida, se puede definir el rango como:

Definición 1.8 (Rango de A)

Sea A una matriz $n \times m$, se llama rango de A y se denota $Rng(A)$ al número de filas no nulas de la matriz en la forma escalonada reducida equivalente a A.

Así para determinar el rango de una matriz A es necesario calcular su forma escalonada reducida, sin embargo, observe que cualquier matriz escalonada equivalente a A tiene el mismo número de filas no nulas que la escalonada reducida. Esto porque la escalonada reducida se obtiene de la escalonada aplicando más operaciones, las cuales no modifican el número de filas no nulas.

Ejemplo 1.16 En el ejemplo 1.14, se observó que las matrices B y C son equivalentes,

$$B = \begin{pmatrix} 1 & 0 & -2 & 2 \\ -2 & 1 & 5 & -1 \\ -1 & 1 & 3 & 1 \\ 0 & 1 & 1 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = C.$$

Y como C es escalonada reducida (es suficiente con que sea escalonada) y tiene dos filas no nulas, entonces

$$\text{Rng}\,(B) = 2.$$

Recordemos que la matriz B es la matriz aumentada del sistema 4×3 en el ejemplo 1.13. Entonces el rango de B informa que el sistema se puede reducir a uno de dos ecuaciones sin perder información, o sea que las restantes $4 - \text{Rng}\,(B) = 2$ son ecuaciones superfluas.

1.5 Caracterización de los sistemas, por su solución

Seguidamente se dará una caracterización de los sistemas de ecuaciones lineales con solución única, infinito número de soluciones o sin solución, basados en las características de la matriz del sistema y la matriz aumentada, ambas en una forma escalonada, que se resumen en la noción de rango.

1.5.1 Sistemas que no tienen solución

El problema de decidir si un sistema de ecuaciones lineales tiene solución o no, es el problema de reconocer si tiene ecuaciones inconsistentes o no. Y esto se reconoce fácilmente cuando el sistema tiene la forma escalonada y se observa al menos una ecuación de la forma:

$$0x_1 + 0x_2 + \cdots + 0x_n = 1$$

como en el ejemplo 1.8, matriz $(A|b)$.

También resulta fácil de reconocer que hay ecuaciones inconsistentes, en un sistema en su forma inicial, cuando dos de ellas tienen iguales coeficientes asociados a las mismas variables y la constante a la derecha es distinta, por ejemplo, las ecuaciones 2 y 4 del sistema siguiente:

$$\left\{\begin{array}{rcrcrcrcr}
x_1 & & & + & 5x_3 & + & -1x_4 & = & 1\\
2x_1 & + & -1x_2 & + & 5x_3 & + & x_4 & = & 3\\
3x_1 & + & 1x_2 & + & -2x_3 & + & 3x_4 & = & -1\\
2x_1 & + & -1x_2 & + & 5x_3 & + & x_4 & = & 6
\end{array}\right.$$

Obtenga la forma escalonada de este sistema y verifique que en esta forma escalonada aparece una ecuación inconsistente, o sea, el sistema no tiene solución.

Sin embargo, en términos del sistema inicial hay otros tipos de dependencia entre las ecuaciones que las pueden hacer inconsistentes. Este es un problema difícil que se relaciona con los conceptos de rango e independencia lineal, que se estudiará más adelante. Pero en todos los casos, cuando esto ocurre, la forma escalonada del sistema hará evidente la existencia de al menos una ecuación inconsistente con la forma vista.

Finalmente observemos que en un sistema de ecuaciones lineales $Ax = b$, la presencia de al menos una ecuación inconsistente, en la forma escalonada de la matriz aumentada del sistema hará que el rango de A y el rango de $(A|b)$ sean distintos. Por ejemplo, en un sistema 4×4, la forma escalonada de la matriz $(A|b)$ de un sistema inconsistente puede ser:

$$(A|b) \xrightarrow{\quad\cdots\quad} \left(\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

donde el símbolo $*$ representa cualquier número real. En este caso la ecuación que representa la tercer fila de la matriz anterior es inconsistente y hace que $\operatorname{Rng}(A) \neq \operatorname{Rng}(A|b)$. Específicamente,

$$2 = \operatorname{Rng}(A) < \operatorname{Rng}(A|b) = 3.$$

Y en general podemos reconocer que:

$$\operatorname{Rng}(A) < \operatorname{Rng}(A|b) \iff Ax = b \text{ tiene ecuaciones inconsistentes.}$$

1.5.2 Sistemas con solución

Un sistema $Ax = b$ con solución se dice que es consistente y, naturalmente, es un sistema que no tiene ecuaciones inconsistentes. La ausencia de ecuaciones inconsistentes se refleja en que:

> la forma escalonada de la matriz del sistema A y la forma escalonada de la matriz aumentada $(A|b)$ tienen el mismo número de filas no nulas.

Y esto es equivalente a establecer que:

$$\operatorname{Rng}(A) = \operatorname{Rng}(A|b)$$

Por ejemplo, los siguientes esquemas de matrices, que utilizan un $*$ para indicar que en esa posición puede aparecer cualquier número real, corresponden a matrices aumentadas de sistemas consistentes.

$$\left(\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{cccc|c} 0 & 1 & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

En el sistema asociado a matriz de la izquierda

$$\mathrm{Rng}\,(A) = 3 = \mathrm{Rng}\,(A|b).$$

Y para el de la derecha

$$\mathrm{Rng}\,(A) = 2 = \mathrm{Rng}\,(A|b).$$

Sistemas con solución única

Un sistema $n \times m$, $Ax = b$, tiene solución única, si además de la condición anterior, el sistema en la forma escalonada tiene tantas ecuaciones no superfluas, como variables. O lo que es lo mismo:

- en cada columna de la forma escalonada de la matriz del sistema, hay un **primer uno** de alguna fila.
- o, el número de filas no nulas en la matriz del sistema en su forma escalonada es igual a m, el número de variables del sistema.
- o, $\mathrm{Rng}\,(A) = \mathrm{Rng}\,(A|b) = m$.

Los siguientes esquemas de matrices corresponden a sistemas con solución única.

$$\left(\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \end{array}\right) \qquad \left(\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Sistemas con infinito número de soluciones

Finalmente, un sistema $n \times m$, $Ax = b$, tiene un número infinito de soluciones si además de tener solución, el número de filas no nulas de la forma escalonada de la matriz del sistema es menor que m, el número de variables (o columnas de la matriz del sistema). Lo que es equivalente a establecer que:

$$\operatorname{Rng}(A) = \operatorname{Rng}(A|b) < m.$$

Esto significa que hay columnas, en la forma escalonada de la matriz del sistema, que **no** contienen algún **primer uno**. El número de estas columnas corresponde al número de parámetros con que se describe el conjunto solución del sistema y es igual a $m - \operatorname{Rng}(A)$.

Por ejemplo, para los sistemas de ecuaciones $Ax = b$ cuyas matrices aumentadas, en la forma escalonada reducida, son:

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

se tiene, en el primer caso, que

$$\operatorname{Rng}(A) = 3 = \operatorname{Rng}(A|b) < 4.$$

luego el sistema tiene infinitas soluciones que dependen de $4 - \operatorname{Rng}(A) = 1$ parámetro. A partir del sistema que representa la forma escalonada reducida, obtenemos su conjunto solución:

$$\left\{\begin{array}{rcr} x_1 - x_2 \qquad\qquad & = & -1 \\ x_3 \qquad & = & 0 \\ x_4 & = & 2 \end{array}\right. \Longrightarrow \begin{array}{rcl} x_1 & = & -1 + x_2 \\ x_2 & = & \qquad\quad x_2 \\ x_3 & = & 0 \\ x_4 & = & 2 \end{array}$$

A la variable x_2 se asigna un valor arbitrario $t \in I\!R$, decimos que es un parámetro, de manera que el conjunto solución del sistema es:

$$S = \{(-1 + t, t, 0, 2) | t \in I\!R\}.$$

Para el sistema representado por la matriz de la derecha,

$$\operatorname{Rng}(A) = 2 = \operatorname{Rng}(A|b) < 4.$$

luego tiene infinitas soluciones que dependen de $4 - \text{Rng}\,(A) = 2$ parámetros. La solución es dada por:

$$\left\{\begin{array}{rcrcl} x_2 & & + & x_4 = & 3 \\ & x_3 & - & 2x_4 = & 2 \end{array}\right. \Longrightarrow \begin{array}{rcrcr} x_1 & = & x_1 & & \\ x_2 & = & 3 & - & x_4 \\ x_3 & = & 2 & + & 2x_4 \\ x_4 & = & & & x_4 \end{array}$$

De manera que asignando a x_1 un valor arbitrario t, y a x_4 otro valor arbitrario s, se obtiene el conjunto solución del sistema:

$$S = \{(t, 3 - s, 2 + 2s, s) | t, s \in \mathbb{R}\}.$$

El siguiente teorema resume, los anteriores resultados.

Teorema 1.9 *Si $Ax = b$ es un sistema de ecuaciones lineales $n \times m$ y $(A|b)$ su matriz aumentada, entonces*

a) Si $Rng\,(A) < Rng\,(A|b)$ el sistema no tiene solución (el sistema es inconsistente).

b) Si $Rng\,(A) = Rng\,(A|b)$ el sistema es consistente (tiene solución), en este caso:

i) Si $Rng\,(A) = Rng\,(A|b) = m$ el sistema tiene solución única.

ii) Si $Rng\,(A) = Rng\,(A|b) < m$ el sistema tiene infinitas soluciones que dependen de $(m - Rng\,(A))$ parámetros.

Algunos casos particulares, de las caracterizaciones de los sistemas de ecuaciones lineales dadas, merecen especial atención.

1.5.3 Sistemas homogéneos

Si las constantes a la derecha del igual en todas las ecuaciones son cero, se dice que el sistema es homogéneo, es decir, $Ax = b$ se llama homogéneo si $b = 0$ es el vector columna de ceros. En este caso, se tiene que

$$\text{Rng}\,(A) = \text{Rng}\,(A|0)$$

porque ninguna operación elemental sobre las filas de $(A|0)$, puede cambiar los ceros que forman la última columna de $(A|0)$. Así, en un sistema homogéneo no pueden existir ecuaciones inconsistentes.

Además se reconoce fácilmente que $x_1 = x_2 = \ldots x_m = 0$, es una solución, por lo tanto los sistema homogéneos siempre tienen solución.

Teorema 1.10 *Todo sistema homogéneo $Ax = 0$, $n \times m$, es consistente:*

a) *$x = 0_m$, el vector columna de m ceros, es una solución del sistema.*

b) *Si $Rng(A) = m$ entonces el sistema tiene como única solución al vector nulo: $x = 0_m$.*

c) *Si $Rng(A) < m$ entonces el sistema tiene infinitas soluciones que dependen de $m - Rng(A)$ parámetros.*

Demostración: Se deduce directamente del teorema anterior, observando que $\mathrm{Rng}\,(A|0) = \mathrm{Rng}\,(A)$, puesto que las operaciones que transforman $(A|0)$ a su forma escalonada reducida, no alteran la última columna de ceros.

■

1.5.4 Sistemas con menos ecuaciones que variables

Si hay menos ecuaciones que incógnitas el número de filas no nulas, de la matriz aumentada en su forma escalonada, será necesariamente inferior al número de variables, luego un sistema de este tipo no puede tener solución única.

Si además un sistema con menos ecuaciones que variables es homogéneo, entonces tiene un número infinito de soluciones, necesariamente.

1.5.5 Sistemas $n \times n$

Un sistema $Ax = b$, $n \times n$ (con tantas ecuaciones como variables) tiene solución única solamente si al calcular la forma escalonada reducida de su matriz aumentada se obtiene:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right) \xrightarrow{\quad\cdots\quad} \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & * \\ 0 & 1 & & & * \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & * \end{array}\right)$$

O sea, la forma escalonada reducida de A debe ser una matriz $n \times n$, con unos en la diagonal y ceros en las restantes entradas, que llamamos matriz identidad y se denota como I_n.

Teorema 1.11 *Un sistema $n \times n$, $Ax = b$, tiene solución única si y solo si A es equivalente a la matriz identidad $n \times n$.*

1.6 Interpretación del rango de una matriz

Consideremos una matriz A cualquiera $n \times m$, y el sistema homogéneo $Ax = 0$. Observemos que:

1. Cada ecuación que produzca un primer uno, en la forma escalonada de la matriz del sistema, es una ecuación que aporta información para resolver el sistema $Ax = 0$, que no está contemplada en las otras ecuaciones.

2. En tanto que una ecuación que termine como una fila de ceros en la escalonada reducida equivalente a A, es una ecuación que puede omitirse del sistema $Ax = 0$ sin que se pierda nada, es una ecuación redundante o superflua.

3. Por otra parte $\operatorname{Rng}(A)$ es el número de filas no nulas de la escalonada reducida equivalente a A, o sea, el número de primeros 1 de cualquier matriz escalonada equivalente A.

Así, el $\operatorname{Rng}(A)$ informa del número de ecuaciones que realmente aportan información para resolver el sistema homogéneo $Ax = 0$, o en otros términos, El $\operatorname{Rng}(A)$ es:

- el máximo número de ecuaciones del sistema $Ax = 0$ que no incluyen ecuaciones redundantes.
- y el mínimo número de ecuaciones que se deben preservar en el sistema $Ax = 0$ para obtener un sistema equivalente.

Más adelante, con la introducción de los conceptos de combinación lineal e independencia lineal, sección 2.7, se relaciona la idea de redundancia con el concepto de dependencia lineal y el de ecuaciones sin redundancia con el de independencia lineal, momento en que el concepto de rango adquirirá mayor riqueza de interpretaciones. Por ahora, el siguiente teorema resume algunas proposiciones equivalentes y aplicables a matrices cuadradas con rango completo.

Teorema 1.12 *Si A es una matriz $n \times n$, las siguientes proposiciones son equivalentes.*

i) $\mathit{Rng}(A) = n$.

ii) A *es equivalente a la identidad.*

iii) $Ax = 0$ *tiene solución única.*

iv) $Ax = b$ *tiene solución única* $\forall b \in I\!R^n$.

Demostración: Se demostrará que *ii*) $\Longrightarrow$ *i*) $\Longrightarrow$ *iii*) $\Longrightarrow$ *ii*) y *i*) $\Longrightarrow$ *iv*) $\Longrightarrow$ *ii*), lo que es suficiente para establecer las equivalencias.

ii) $\Longrightarrow$i) Si A es equivalente a I_n, la identidad es su forma escalonada reducida, por lo que $\operatorname{Rng}(A) = n$.

i) $\Longrightarrow$iii) Si $\operatorname{Rng}(A) = n$, entonces $\operatorname{Rng}(A|0) = \operatorname{Rng}(A) = n$, lo que significa que $Ax = 0$ tiene solución única, teorema 1.10.

i) $\Longrightarrow$iv) Si $\operatorname{Rng}(A) = n$, entonces la escalonada reducida equivalente a $(A|b)$ es de la forma $(I_n|z)$, que tiene n filas no nulas, luego $\text{Rng}\,(A|b) = \text{Rng}\,(A) = n$, lo que significa que $Ax = b$ tiene solución única $\forall b \in \mathbb{R}^n$.

iv) $\Longrightarrow$ii) Se pospone para la siguiente sección, a fin de disponer de las matrices elementales, que facilitan la escritura de esta demostración.

iii) $\Longrightarrow$ii) Si $Ax = 0$, tiene solución única, la forma escalonada reducida de $(A|0)$ es $(I_n|0)$, (en caso contrario esta forma escalonada reducida tendría al menos una fila de ceros, lo que conduciría a infinitas soluciones). Luego A es equivalente a la identidad.

■■■

1.7 Redes y sistemas de ecuaciones lineales

Se usarán las redes para representar fenómenos tales como el flujo de vehículos, el flujo de una corriente eléctrica, modelos de transporte, etc. De tal modo que el objeto matemático *red* interesa en tanto que una herramienta de ayuda en la resolución de problemas sencillos que involucran sistemas de ecuaciones lineales y programación lineal. Naturalmente queda fuera del objetivo de estas notas el estudio de las propiedades matemáticas de las redes y nos limitamos a dar una descripción, más gráfica, de lo que llamaremos red. Una red es un conjunto de puntos llamados **nodos**, conectados de alguna manera por medio de líneas llamadas **arcos**.

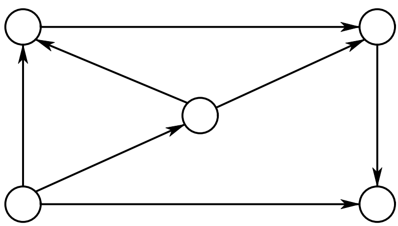


Figura 1.1: Red con cinco nodos y siete aristas.

Como se ve en la Figura 1.1, los nodos se representan con círculos o puntos. Las flechas son usadas para indicar, en ciertos casos, el sentido del "flujo" que se representa.

1.7.1 Redes de flujos

El ejemplo más intuitivo y sencillo es el de redes de carreteras, donde el flujo se refiere a la cantidad de vehículos que circulan por la vía. En el caso de redes que representan flujos supondremos en esta sección que *la cantidad que llega a un nodo es igual a la que sale*:

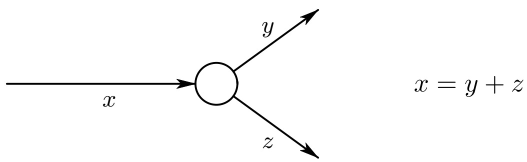


Figura 1.2: Flujo que entra al nodo es igual al que sale.

Ejemplo 1.17 La siguiente red representa el flujo vehicular por un sector de la ciudad donde, los nodos son las intersecciones y los arcos, las carreteras. Las flechas indican el sentido del movimiento de los vehículos.

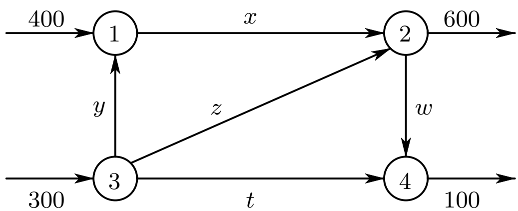


Figura 1.3: Flujo vehicular por un sector de la ciudad.

Las cantidades de vehículos que circulan por la red entre dos nodos consecutivos se indican con las letras x, y, z, w y t las cuales son las variables del problema. En tales condiciones:

i) Formule un sistema de ecuaciones lineales cuya solución aporte todas las opciones posibles de flujo vehicular.

ii) Si el flujo vehicular entre el nodo 1 y 2 es 550 y entre el nodo 2 y 4 es 50, entonces calcule los otros flujos.

Solución

i) De acuerdo con la red el sistema es :

$$\left\{\begin{array}{rrrrrcll} x & -\,y & & & & = & 400 & \text{nodo } 1 \\ x & & +\,z & -\,w & & = & 600 & \text{nodo } 2 \\ & +\,y & +\,z & & +\,t & = & 300 & \text{nodo } 3 \\ & & & +\,w & +\,t & = & 100 & \text{nodo } 4 \end{array}\right.$$

La matriz aumentada del sistema es

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & : & 400 \\ 1 & 0 & 1 & -1 & 0 & : & 600 \\ 0 & 1 & 1 & 0 & 1 & : & 300 \\ 0 & 0 & 0 & 1 & 1 & : & 100 \end{pmatrix}$$

Haciendo operaciones elementales sobre esta matriz se obtiene la solución: $S = \{(700 - z - t, 300 - z - t, z, 100 - t, t)\,|z,\, w \in I\!R\}$.

ii) Como $x = 550$ entonces $y = 150$. Además $w = 100 - t = 50$ entonces $t = 50$. Por otra parte $y = 300 - z - t = 300 - z - 50 = 150$ por lo que $z = 100$.

1.7.2 Redes eléctricas

Considérese un modelo simple de circuito eléctrico que consta solo de resistencias (bombillas eléctricas, electrodomésticos, ...) y fuerza electromotriz o baterías. A través de la red (o circuito) fluye la corriente eléctrica en el sentido que indican las flechas. Los nodos representan puntos de la red donde se redirecciona y distribuye la corriente. Los generadores se simbolizan con dos rayas verticales una más corta que la otra. La corriente entra por la raya corta y sale por la raya larga. El otro símbolo que aparece en la Figura 1.4 es el de las resistencias.

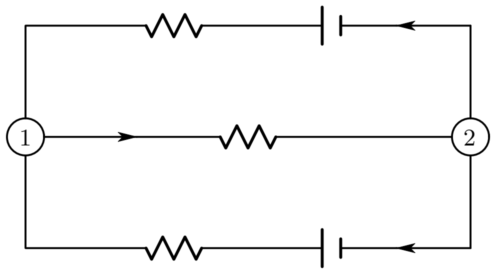


Figura 1.4: Red eléctrica con dos nodos, tres resistencias y dos generadores.

La fuerza electromotriz se mide en *voltios*, la corriente en *amperios* y la resistencia en *ohmios*. El movimiento de la corriente en el circuito se rige por las conocidas leyes de Kirchhoff, a saber:

a) La corriente que fluye hacia un nodo es igual a la que sale.

b) En una trayectoria cerrada la fuerza electromotriz es igual a la suma de las caídas de voltaje.

Una trayectoria cerrada es una parte de la red donde la corriente sale de un nodo y regresa a él. En la figura anterior se tienen dos trayectorias cerradas. una sale del nodo 1 y la otra del nodo 2.

La caída de voltaje E a través de una resistencia es el producto RI donde I es la corriente y R es la resistencia. Es decir $E = RI$.

Ejemplo 1.18 Considere la red siguiente y determine las corrientes I_1, I_2 e I_3

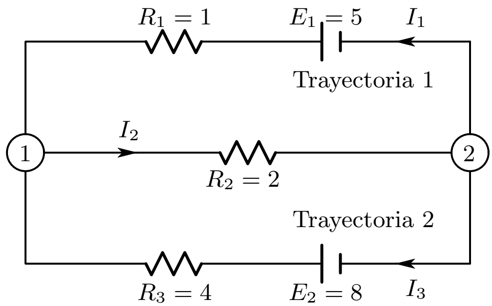


Figura 1.5: Red eléctrica con dos trayectorias.

Solución: A partir de la Figura 1.5, vemos que el sistema de ecuaciones es:

$$\left\{\begin{array}{lllcrr} I_1 & -I_2 & +I_3 & = & 0 & \text{Nodo 1 o 2} \\ I_1 & +2I_2 & & = & 5 & \text{Trayectoria 1} \\ & +2I_2 & +4I_3 & = & 8 & \text{Trayectoria 2} \end{array}\right.$$

Resolviendo el sistema se obtiene la solución: $I_1 = 1$ amperio, $I_2 = 2$ amperios e $I_3 = 1$ amperio.

1.8 Ejercicios

1. Considere el sistema 3×2:

$$\begin{array}{rcr} 2x_1 + 3x_2 & = & 0 \\ 4x_1 + 6x_2 & = & 12 \\ -x_1 + x_2 & = & 2 \end{array}$$

a) Represente gráficamente las tres rectas asociadas al sistema.

b) Del análisis gráfico, ¿qué se puede concluir acerca de la solución del sistema?

c) Resuelva el sistema por el método de Gauss.

2. a) Proponga un ejemplo de un sistema de tres ecuaciones lineales en dos variables que corresponda a la ilustración gráfica de la figura 1.6.

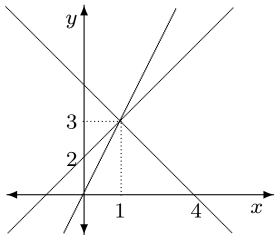


Figura 1.6:

b) La figura 1.7 cambia una de las rectas del gráfico anterior trasladándola paralelamente hasta pasar por el origen. Modifique el sistema propuesto en a) para que corresponda a la ilustración gráfica 1.7.

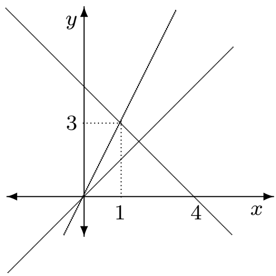


Figura 1.7:

c) En cada caso, sólo observando los gráficos, dé el conjunto solución del sistema.

3. Resuelva los siguientes sistemas

$$\text{a)}\ \left\{\begin{array}{rcl} 2x - y & = & 0 \\ 6x - 5y & = & 0 \\ 2x + 4y & = & 0 \end{array}\right.$$

b) $\begin{cases} 2x + 4y &= 6 \\ 3x + 6y &= 5 \end{cases}$

c) $\begin{cases} x + 5y + 11z &= -5 \\ 2x + 3y + 8z &= 4 \\ -x + 2y + 3z &= -9 \end{cases}$

4. En cada caso, determine el conjunto solución:

1. $\begin{cases} -2x + 6y - 4z = -28 \\ -x + 3y - z = -8 \\ 5x - 15y + 10z = 70 \\ x - 3y = 2 \end{cases}$

2. $\begin{cases} 3x + 9y - 3z = 0 \\ y + z = 1 \\ -2x - 5y + 4z = 4 \\ -2x - 6y + 3z = 4 \end{cases}$

3. $\begin{cases} -x + y + 5z = -36 \\ -2y + 4z = -30 \\ x - 3y = -1 \\ -3x + 9y = 3 \end{cases}$

4. $\begin{cases} 5x_1 - 5x_3 = 0 \\ 3x_2 + 3x_3 - 9x_4 = 3 \\ 2x_1 - 3x_2 - 5x_3 + 10x_4 = 0 \\ -2x_1 + 2x_3 - x_4 = -1 \end{cases}$

5. $\begin{cases} x_1 - x_3 + 2x_4 = 4 \\ -6x_1 + x_2 + 8x_3 - 15x_4 = -32 \\ -4x_1 + x_2 + 6x_3 - 10x_4 = -21 \\ -x_2 - 2x_3 = -1 \\ -2x_1 + 2x_3 - 5x_4 = -11 \end{cases}$

6. $\begin{cases} -3x_1 + 3x_2 + 6x_3 - 15x_4 = 3 \\ x_2 - 3x_3 - 3x_4 = -9 \\ 3x_1 - 3x_2 - 7x_3 + 15x_4 = -6 \\ -2x_2 + 6x_3 + 7x_4 = 18 \\ -3x_1 + 5x_2 + x_3 - 22x_4 = -12 \end{cases}$

5. En cada caso, dé el conjunto de tripletes (x, y, z) que son solución del sistema homogéneo:

1. $\begin{cases} -2x + 6y - 4z = 0 \\ -x + 3y - z = 0 \\ 5x - 15y + 10z = 0 \\ x - 3y = 0 \end{cases}$

2. $\left\{\begin{array}{rrcl} -x & +2z & = & 0 \\ 3x & -6z & = & 0 \\ 2x & -4z & = & 0 \end{array}\right.$

3. $\left\{\begin{array}{rrcl} y & +5z & = & 0 \\ -2y & +4z & = & 0 \\ -3y & & = & 0 \\ -y & +9z & = & 0 \end{array}\right.$

6. Considere el siguiente sistema 2×3:

$$\left\{\begin{array}{rcl} -x +y-5z & = & -2 \\ 3x-3y +z & = & -8 \end{array}\right.$$

a) Determine el conjunto solución.

b) Verifique que si se agrega una ecuación como

$$x - y - 9z = -12$$

que resulta de sumar dos veces la primera ecuación más la segunda, la información que se agrega es redundante, es decir, la nueva restricción ya está contemplada en las dos primeras.

c) Por otra parte, explique porqué si se agrega una ecuación como

$$x - y - 9z = -10$$

se agrega una inconsistencia. Verifique esto último resolviendo el sistema 3×3 resultante.

7. Dado el siguiente sistema 2×3:

$$\left\{\begin{array}{rcl} 2x+3y -z & = & 4 \\ x-2y+2z & = & 1 \end{array}\right.$$

En cada caso, agregue una ecuación de manera que el sistema 3×3 resultante:

a) tenga una única solución.

b) tenga infinitas soluciones.

c) no tenga solución.

8. Sea el sistema cuya matriz aumentada es:

$$\left(\begin{array}{ccc|c} a & 0 & b & 2 \\ a & a & 4 & 4 \\ 0 & a & 2 & b \end{array}\right)$$

Para cuáles valores de a, b el sistema:

a) No tiene solución b) Tiene solución única.

c) Tiene infinitas soluciones dependiendo de:
i) un parámetro, ii) dos parámetros.

9. Considere los sistemas homogéneos cuyas matrices son:

$$\begin{pmatrix} 1 & b & 0 \\ 1 & a & a \\ -1 & -a & b \end{pmatrix} \qquad \begin{pmatrix} b & b & a^2 \\ -b & a & -a^2 \\ -b & a & -1 \end{pmatrix}$$

En cada caso, para cuáles valores de a, b el sistema tiene :

a) solución única.

b) infinitas soluciones dependiendo de un parámetro.

c) infinitas soluciones dependiendo de dos parámetros.

10. Considere el sistema de ecuaciones lineales:

$$\left\{\begin{array}{rcr} -\alpha x - \alpha^2 y + (5\alpha + 4)z & = & -16 \\ x \ + \ \alpha y + \qquad \alpha z & = & \alpha \\ -\alpha x + 4\alpha y + \qquad 4\alpha z & = & 4\alpha \end{array}\right.$$

Determine todos los valores del parámetro α, si existen, tales que el sistema:

a) No tiene solución b) Tiene solución única.

c) Tiene infinitas soluciones dependiendo de:
i) un parámetro, ii) dos parámetros.

11. Considere el sistema de ecuaciones lineales en las variables x_1, x_2 y x_3, que depende del parámetro $k \in \mathbb{R}$:

$$\left\{\begin{array}{rcl} kx_1 + x_2 + x_3 & = & 1 \\ x_1 + kx_2 + x_3 & = & 1 \\ x_1 + x_2 + kx_3 & = & 1 \end{array}\right.$$

a) Determine los valores de k para los cuales el sistema:

i) no tiene solución.

ii) tiene un número infinito de soluciones.

iii) tiene solución única.

b) Para los casos en que hay solución única, determine la solución.

12. a) Decida si la ecuación

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = b,$$

tiene solución en los siguientes casos:

1) $b = \binom{1}{3}$; 2) $b = \binom{1}{2}$; 3) $b = \binom{3}{6}$.

b) En los casos en que hay solución comente sobre el número de soluciones.

c) Para $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y los mismos b responda las preguntas anteriores.

13. Determine un vector x tal que $Ax = b$ si

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \\ -1 & 4 & 0 \end{pmatrix} \quad \text{y} \quad b = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}.$$

Similarmente, determine x tal que $Ax = 0$.

14. Considere el sistema

$$\begin{cases} y + 2w & = & 0 \\ x - z & = & 1 \\ z + w & = & 2 \end{cases}$$

a) Escriba el sistema en la forma matricial.

b) Obtenga la solución del sistema y escríbala en la forma $h + u$ donde u es cualquier vector solución del sistema homogéneo y h es un vector solución fijo, del sistema no homogéneo.

15. Un sistema tiene matriz aumentada

$$\left(\begin{array}{ccr|c} 1 & 0 & -1 & \alpha \\ 0 & 1 & 2 & \alpha \end{array}\right)$$

Calcule α y a sabiendo que un vector solución es $(2,1,a)$. Luego resuelva el sistema.

16. Considere el sistema cuya matriz aumentada es:

$$\left(\begin{array}{rrc|c} 1 & 2 & 0 & 1 \\ -1 & 3 & \alpha & 0 \\ 2 & 1 & 0 & 2 \\ -1 & -\beta & \alpha & 0 \end{array}\right)$$

Con $\alpha, \beta \in I\!R$.

a) ¿Para qué valores de α y β el sistema es equivalente a un sistema no homogéneo 3×3? ¿Para qué valores de α y β el sistema es consistente?

b) Para todos los valores posibles de α y β obtenga la solución del sistema.

17. Sean $a, b \in I\!R$. Considere los planos[1] cuyas ecuaciones son:

$$4x - ay + z + 4 = 0;$$

$$2x - y + z - 1 = 0;$$

$$2x - y + 2bz + 4 = 0.$$

Determine los valores de a y b que corresponden a cada uno de los siguientes casos:

a) La intersección de los tres planos es vacía.

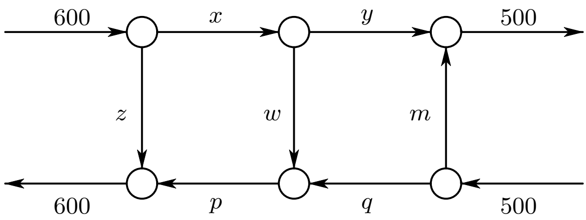
b) La intersección de los tres planos es exactamente un punto.

c) La intersección de los tres planos es exactamente una recta. Es decir, el conjunto solución depende de un parámetro. En este caso, determine el conjunto solución.

[1]LLamamos plano al conjunto de puntos (x, y, z) de $I\!R^3$, que satisfacen una ecuación de la forma $ax + by + cz = d$, donde a, b, c y d son números reales dados.

18. Sean A, B matrices equivalentes. Justifique por qué A y B tienen el mismo rango.

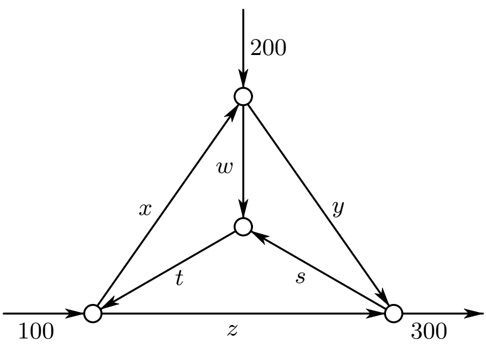
19. En un acueducto fluye agua en miles de metros cúbicos por hora, tal como se muestra en la siguiente red :



a) Resuelva el sistema para el caudal de agua representado por x, y, z, w, m, p, q.

b)Encuentre el patrón de flujo cuando $p = q = 0$.

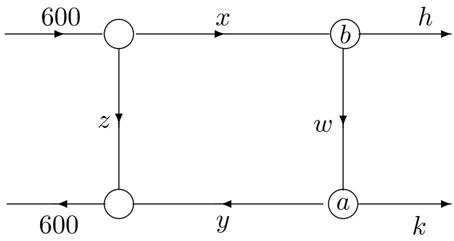
20. Sea la red de flujo



a) Resuelva el sistema correspondiente.

b) Encuentre el flujo cuando $z = 100$, $t = 50 = s$.

21. Considere la siguiente red de flujo vehicular de un sector de la ciudad dependiente de los parámetros h y k.

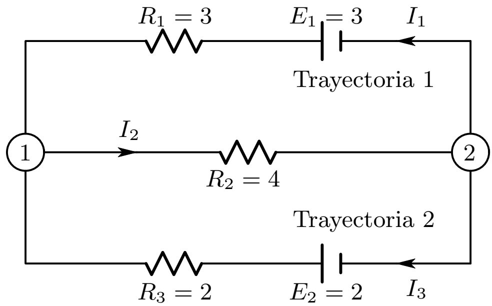


a) Obtenga una relación entre h y k para que el sistema correspondiente tenga solución.

b) Calcule los flujos vehículares en función de h y una variable libre. Calcule los flujos para valores particulares de h, k y la variable libre.

c) Si la cantidad de autos que ingresan al nodo a es 400 y los que salen del nodo b son 550, calcule los flujos restantes.

22. Encuentre las corrientes I_1, I_2 e I_3 de la red eléctrica siguiente:



23. Sea el sistema:

$$\begin{array}{rcrcrcrcr} x & & & - & 2z & + & w & = & 0 \\ 2x & & & + & 2z & - & 2w & = & 4 \\ -x & - & 3y & + & z & - & 2w & = & 1 \\ 2x & + & 3y & & & + & w & = & 1 \end{array}$$

a) Exprese el sistema en la forma matricial $Al = b$, donde $l = (x, y, z, w)^t$.

b) Calcule una columna l_0 tal que $Al_0 = b$.

c) Exprese la solución del sistema en la forma:

$$S = \{l_0 + d | Ad = 0\}.$$

24. Sean $R \neq 0$ y $H \neq 0$ matrices 2×3.

a) Escriba todas las posibles formas escalonadas reducidas de R. Sugerencia: considere los casos Rng $(R) = 1$ y Rng $(R) = 2$, por separado.

b) Pruebe que si R y H son escalonadas reducidas y si los sistemas $Rx = 0$ y $Hx = 0$ tienen igual solución, entonces $R = H$. Ayuda: es equivalente probar que si $R \neq H$ entonces los sistemas $Rx = 0$ y $Hx = 0$ tienen diferente solución.

c) Si R y H son escalonadas, pruebe que el resultado anterior es falso.

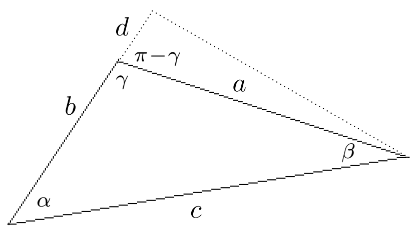
25. Si en un triángulo cualquiera, α, β y γ son las medidas de los ángulos opuestos a los lados de magnitud a, b y c, respectivamente.

a) Muestre que a, b, c, α, β y γ satisfacen el siguiente sistema de ecuaciones lineales

$$\begin{cases} cx + \quad\quad + az &= b \\ \quad\quad cy + bz &= a \\ bx + ay \quad\quad &= c \end{cases}$$

donde $x = \cos\alpha$, $y = \cos\beta$ y $z = \cos\gamma$.

Ayuda: por ejemplo, observe del siguiente gráfico que $b + d = c\cos\alpha$ y $d = a\cos(\pi - \gamma) = -a\cos\gamma$, de donde $b = c\cos\alpha + a\cos\gamma$.



b) Muestre que la matriz escalonada reducida equivalente a la matriz aumentada del anterior sistema, es

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & \dfrac{-a^2+b^2+c^2}{2bc} \\ 0 & 1 & 0 & \dfrac{a^2-b^2+c^2}{2ac} \\ 0 & 0 & 1 & \dfrac{a^2+b^2-c^2}{2ab} \end{array}\right)$$

Observe que a, b y c son magnitudes positivas y por lo tanto no nulas.

c) Muestre que a, b, c, α, β y γ satisfacen

$$\begin{aligned} a^2 &= b^2+c^2-2bc\cos\alpha \\ b^2 &= a^2+c^2-2ac\cos\beta \\ c^2 &= a^2+b^2-2ab\cos\gamma \end{aligned}$$

Capítulo 2

Matrices

Sin duda alguna al lector le resulta familiar los arreglos de datos en filas y columnas para almacenar información de manera ordenada. Por ejemplo, los resultados del fútbol se presentan de modo que las filas dan información por cada equipo y las columnas lo hacen sobre los partidos jugados, ganados, empatados, perdidos, goles a favor, goles en contra y total de puntos. Tal situación la observamos en la siguiente matriz compuesta de 12 filas y 7 columnas:

Posiciones de la Primera División en Costa Rica[1].

Equipos	PJ	PG	PE	PP	GF	GC	Pts
1)Alajuelense	25	17	5	3	59	20	60
2)Herediano	25	14	7	4	45	21	54
3)Cartaginés	25	12	10	3	27	17	50
4)Saprissa	25	11	10	4	42	26	48
5)Pérez Zeledón	25	10	5	10	37	34	39
6)Carmelita	25	8	8	9	26	27	34
7)Belén	24	6	11	7	23	21	34
8)Puntarenas	25	6	10	9	23	33	32
9)Ramonense	25	6	6	13	26	44	27
10)San Carlos	25	4	9	12	24	45	25
11)Turrialba	24	4	7	13	23	42	23
12)Goicoechea	25	4	6	15	23	48	21

[1] Tomado del diario La Nación, 3 de marzo de 1997.

Otro ejemplo lo obtenemos en la matriz con información de la contaminación del agua en algunos puntos de muestreo de los ríos que llegan al embalse La Garita. La primera columna es un índice (entre 0 y 100) que indica la calidad del agua en cada uno de los 9 puntos de muestreo, las restantes columnas son mediciones, en unidades de concentración, que sirvieron para calcular el índice de calidad.

Calidad del agua[2]

	Indice calidad	DBO	PH	Sólidos Totales	Fosfatos
Emb. centro	55.0	15.0	7.15	172.0	0.54
Emb. orilla	58.0	13.7	7.20	164.0	0.52
Emb. represa	58.0	30.6	7.30	180.0	0.68
Emb. salida	59.0	10.0	7.15	157.0	0.55
Emb. Desfogue	58.0	10.5	7.20	165.0	0.66
Río Alajuela	67.0	16.2	8.30	165.0	0.40
Río Ciruelas	70.0	6.1	8.35	187.0	0.62
Río Tizate	77.0	1.5	8.35	192.0	0.64
Río Virilla	62.0	15.0	8.10	295.0	0.92

Un tercer ejemplo de matriz, de nueve filas y cinco columnas, está dado por la tabla siguiente que contiene las notas de un grupo de estudiantes de Algebra Lineal.

Nombre	P1	P2	P3	Q	E
Leonardo	3.6	3.0	4.0	5.0	3.6
Fresia	7.6	9.4	5.6	6.0	7.5
Alejandro	5.5	2.2	0	3.0	2.6
Hernán	7.2	5.4	4.5	7.0	5.8
Michael	8.2	6.3	4.3	7.0	6.3
Karla	8.8	4.5	6.7	8.5	6.7
Norberto	6.3	5.9	5.5	7.7	6.0
Manuel	7.1	7.3	7.1	7.0	7.2
Erick	5.3	3.6	1.3	2.0	3.3

En las filas se lee la situación personal de cada estudiante y en las columnas las notas obtenidas en cada uno de los tres parciales

[2]Datos del Instituto Costarricense de Electricidad (ICE) 1984.

y el promedio de los quices. La última columna: escolaridad, es el promedio ponderado con un 32% cada parcial y un 4% el promedio de quices (redondeado a un decimal).

Una definición formal de matriz es la siguiente :

Definición 2.1 (Matriz)

Sean m y n enteros positivos. Un arreglo rectangular en m filas y n columnas de números reales:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

se llama matriz de orden (o tamaño) $m \times n$ sobre $I\!R$.

En forma abreviada se escribe $A = (a_{ij})_{m \times n}$ donde a_{ij} denota la entrada de A en la fila i y la columna j.

El conjunto de las matrices de orden $m \times n$ con entradas reales se denota con $M(m, n, I\!R)$. En el caso de matrices cuadradas, es decir con igual número de filas y columnas, se usa la notación $M(n, I\!R)$ y se dice que son de orden n.

Ejemplo 2.1 Si $A \in M(2, 3, I\!R)$ es la matriz de dos filas y tres columnas con entradas reales:

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

entonces $a_{12} = 2$ *y* $a_{23} = 6$

2.1 Algunos tipos de matrices

Ciertas matrices con formas especiales aparecerán muy frecuentemente, en los desarrollos siguientes, por lo cual conviene identificarlas con nombres específicos.

Definición 2.2 (Matriz Diagonal)

Una matriz $D = (d_{ij})_{n\times n}$ se llama matriz diagonal si todos los elementos fuera de la diagonal son nulos, esto es $d_{ij} = 0$ si $i \neq j$.

$$D = \begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_{nn} \end{pmatrix}$$

Ejemplo 2.2 La siguiente es una matriz diagonal, cuyas entradas son los pesos usados para calcular la Escolaridad en el ejemplo de la página 48:

$$D = \begin{pmatrix} 0.32 & 0 & 0 & 0 \\ 0 & 0.32 & 0 & 0 \\ 0 & 0 & 0.32 & 0 \\ 0 & 0 & 0 & 0.04 \end{pmatrix}$$

Las matrices diagonales con únicamente unos en la diagonal jugarán un rol importante, una vez que se haya introducido el producto matricial, y se llamarán matrices identidad, lo cual se justifica más adelante.

Definición 2.3 (Matriz Identidad)

La matriz $I_n = (d_{ij})_{n\times n}$, con $d_{ij} = \begin{cases} 1 & si\ i = j \\ 0 & si\ i \neq j \end{cases}$, es decir,

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

se denominará matriz identidad de orden n.

Algunas veces, una matriz identidad se denotará con I, sin hacer referencia a su tamaño.

Definición 2.4 (Matriz Triangular)

Una matriz $T = (t_{ij})_{n\times n}$ *se llama matriz triangular inferior (superior) si todos los elementos arriba (abajo) de la diagonal son cero, esto es* $t_{ij} = 0$ *si* $i < j$ *(si* $i > j$*). Y se llama triangular si es triangular inferior o superior.*

$$T = \begin{pmatrix} t_{11} & 0 & \dots & 0 \\ t_{21} & t_{22} & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ t_{n1} & t_{n2} & \dots & t_{nn} \end{pmatrix} \quad \text{o} \quad T = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ 0 & t_{22} & & t_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & t_{nn} \end{pmatrix}$$

Un tipo muy importante de matrices tanto por las propiedades que poseen como por su utilidad, principalmente en Estadística y el Análisis de Datos, son las matrices simétricas, caracterizadas por la simetría de sus entradas respecto a su diagonal.

Definición 2.5 (Matrices simétricas)

$S = (s_{ij})_{n\times n}$ *es simétrica si su entrada en la fila* i *y columna* j *es igual a la entrada en la fila* j *columna* i*, para* $i, j = 1, 2, \dots, n$*. Es decir, si* $s_{ij} = s_{ji}$ *para todo* i *y* j*.*

Un ejemplo de matriz simétrica es la siguiente:

Ejemplo 2.3 (Matriz de distancia.) Consideramos tres ciudades C_1, C_2, C_3, cuyas distancias entre ellas, en cientos de kilómetros, son dadas por las entradas de la matriz $D = (d_{ij})_{3\times 3}$, donde d_{ij} denota la distancia entre las ciudades C_i y C_j.

$$D = \begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 5 \\ 3 & 5 & 0 \end{pmatrix}$$

Así: 5 es la distancia entre C_2 y C_3 que naturalmente es la misma que la distancia de C_3 a C_2 y por ello la matriz resulta simétrica.

Observe que un caso particular de matrices simétricas son las matrices diagonales.

Por otra parte, la simetría es una propiedad que se caracteriza muy apropiadamente mediante la transposición de matrices.

Definición 2.6 (Matriz transpuesta)

Sea $A \in M(m, n, I\!R)$, la matriz transpuesta de A se denota por A^t y se define como la matriz $n \times m$ que se obtiene al escribir las filas de A como columnas, (o equivalentemente, poner las columnas de A como filas). Simbólicamente:

$$\textit{Si} \quad A = (a_{ij})_{m\times n} \quad \textit{entonces} \quad A^t = (a_{ji})_{n\times m}.$$

Ejemplo 2.4 Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ entonces $A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

Usando la transposición de matrices, observe que una matriz S, $n \times n$, es simétrica si y solo si $S^t = S$.

Matrices escalonadas y escalonadas reducidas.

Debemos recordar que en el capítulo anterior, aparecieron las **matrices escalonadas** y las **escalonadas reducidas**, dos tipos de matrices de gran utilidad en la teoría de sistemas de ecuaciones lineales.

Vectores columna y vectores fila.

Y para terminar, se hará especial énfasis en las matrices que constan de una sola fila o una sola columna. Muy frecuentemente, tales matrices se llamarán también vectores y con ellos se suele representar muchos tipos de información.

Una matriz constituida por una sola columna se escribe como

$$c = \begin{pmatrix} c_{11} \\ \vdots \\ c_{n1} \end{pmatrix}_{n\times 1}, \text{pero usualmente la denotamos por } c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Definición 2.7 (Conjunto de vectores columnas)

Al conjunto de todas las matrices columnas de tamaño $n \times 1$ lo denotamos como $\mathbb{R}^n$ y nos referimos a él como el espacio vectorial $\mathbb{R}^n$ y a sus elementos como vectores.

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| x_i \in \mathbb{R} \right\}$$

Así, cuando nos referimos a un vector [3] en $\mathbb{R}^n$, se debe pensar en una matriz de una columna y n filas, a menos que se diga explícitamente que se trata de un vector fila, en cuyo caso se trata de la matriz $c = (c_{11} \; c_{12} \; \ldots \; c_{1n})$ que, en forma similar al caso de las columnas, se denota como:

$$c = (c_1, c_2, \ldots, c_n)$$

en este caso, empleando comas para separar sus entradas.

Si I_n es la matriz identidad de orden n, sus columnas son vectores de $\mathbb{R}^n$ y se llamarán vectores canónicos de $\mathbb{R}^n$, denotándolos con $e_1, e_2, \ldots, e_n$, respectivamente. $I_n = (e_1, e_2, \ldots, e_n)$. Así, por ejemplo, para $\mathbb{R}^4$ los vectores canónicos son las columnas de las siguiente matriz identificadas [4] por e_1, e_2, e_3 y e_4.

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (e_1, e_2, e_3, e_4).$$

Ejemplo 2.5 Las filas de la matriz $\begin{pmatrix} 1 & 4 \\ 5 & 2 \\ 3 & 1 \end{pmatrix}$ son vectores del

[3]Más adelante se justifican los nombres: vectores y espacio vectorial.

[4]Con esta notación, el vector e_2 puede ser $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ó $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ ó $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ ó $\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, según sea el espacio a que pertenezca.

espacio $I\!\!R^2$:

$$\begin{pmatrix}1\\4\end{pmatrix}, \begin{pmatrix}5\\2\end{pmatrix}, \begin{pmatrix}3\\1\end{pmatrix} \text{ y sus columnas } \begin{pmatrix}1\\5\\3\end{pmatrix}, \begin{pmatrix}4\\2\\1\end{pmatrix}$$

son vectores del espacio $I\!\!R^3$.

Ejemplo 2.6 Cuando se trata de matrices de datos como en el caso de la tabla sobre calidad del agua, en la página 48, es usual referirse a las filas de matriz como "individuos" y a las columnas como "variables". Así, el individuo VIRILLA = $(62.0, 15.0, 8.10, 295, 0.92)^t$ es un vector de $I\!\!R^5$ cuyas entradas son los valores que las cinco variables: calidad, DBO, PH, Sólidos totales y Fosfato asumieron en el punto de muestreo: Río Virilla. Por otro lado la variable CALIDAD = $(55, 58, 58, 59, 58, 67, 70, 77, 62)^t$ es un vector en $I\!\!R^9$ cuyas entradas son la calidad del agua (número entre 0 y 100) en cada uno de los nueve puntos de muestreo.

2.2 Operaciones con matrices y propiedades

En esta sección se definirán las operaciones: suma matricial, multiplicación de una matriz por un escalar y producto matricial. La idea es operar con tablas numéricas de datos en forma similar a como se opera con los números reales, pero con ciertos cuidados especiales, porque hay propiedades muy importantes de los números reales que trasladadas a las matrices dejan de ser válidas.

Definición 2.8 (Igualdad, suma y producto por escalar)

Sean $A, B \in M(m, n, I\!\!R)$, $\alpha \in I\!\!R$, $A = (a_{ij})$, $B = (b_{ij})$. *Definimos:*

1. $A = B \Longleftrightarrow a_{ij} = b_{ij} \quad \forall\ i, j$
2. $A + B = (a_{ij} + b_{ij})$
3. $\alpha A = (\alpha a_{ij})$

Teorema 2.9 *Sean* $A, B, C \in M(m, n, I\!R)$, $\alpha, \beta \in I\!R$, $A = (a_{ij})$, $B = (b_{ij})$, *entonces:*

1. $(A + B) + C = A + (B + C)$ *(La suma es asociativa)*
2. $A + B = B + A$ *(La suma es conmutativa)*
3. $0 + A = A$ *(0 matriz con todas sus entradas iguales a cero)*
4. $-A + A = 0$ *donde* $-A = (-a_{ij})$
5. $\alpha(A + B) = \alpha A + \alpha B$
6. $(\alpha + \beta)A = \alpha A + \beta A$
7. $\alpha(\beta A) = (\alpha\beta)A$

Demostración: (Asociatividad) se demostrará que para todo i y j, la entrada i, j de $(A + B) + C$ es igual a la entrada i, j de $A + (B + C)$:

$$\begin{array}{rcl} ((a+b)+c)_{ij} & = & (a+b)_{ij} + c_{ij} \\ & = & (a_{ij} + b_{ij}) + c_{ij} \\ & = & a_{ij} + (b_{ij} + c_{ij}) \\ & = & a_{ij} + (b+c)_{ij} \\ & = & (a + (b+c))_{ij} \end{array}$$

Observe que en esta demostración se denotan las entradas i, j de de $A+B$ como $(a+b)_{ij}$ y las de $(A+B)+C$ como $((a+b)+c)_{ij}$.

■

Definición 2.10 (Producto de matrices)

Sea $A \in M(m, n, I\!R)$, $B \in M(n, p, I\!R)$. *El producto* AB *es la matriz* $C \in M(m, p, I\!R)$, $C = (c_{ij})$, *definida por:*

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

El diagrama siguiente ilustra que cada entrada c_{ij} de la matriz

AB se obtiene al operar la fila i de A con la columna j de B:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \boxed{a_{i1} \quad \cdots \quad a_{in}} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \boxed{c_{ij}} & \vdots \\ c_{m1} & \cdots & c_{mp} \end{pmatrix}$$

donde $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$

Ejemplo 2.7 Sean

$$A = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 3 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 0 & 2 & 4 \\ -1 & -2 & 1 & 0 \\ 3 & 5 & 6 & 1 \end{pmatrix}$$

El producto AB está bien definido porque A tiene tantas columnas como filas tiene B: A es 2×3 y B es 3×4 por lo tanto AB es una matriz 2×4:

$$AB = \begin{pmatrix} 3 & 2 & 3 & 8 \\ 1 & -1 & 11 & 5 \end{pmatrix}$$

En el ejemplo anterior se puede reconocer que cada entrada i, j de la matriz AB se obtiene al operar la fila i de A con la columna j de B.

Ahora, si elegimos una columna j de B que se mantiene fija e i varía sobre todas las filas de A, el producto de las filas de A por la columna j de B genera todas las entradas de la columna j de AB. Por ejemplo si B_3 es la columna 3 de B, $B_3 = \begin{pmatrix} 2 \\ 1 \\ 6 \end{pmatrix}$, AB_3 es un producto matricial bien definido y

$$AB_3 = \begin{pmatrix} 3 \\ 11 \end{pmatrix}$$

lo que corresponde a la columna 3 de AB.

En general, si A y B son de orden $m \times n$ y $n \times p$, respectivamente, y si B se escribe como $B = (B_1, B_2, \ldots, B_p)$, donde $B_1, B_2, \ldots, B_p$ denotan las p columnas de B entonces

$$AB = A(B_1, B_2, \ldots, B_p) = (AB_1, AB_2, \ldots, AB_p).$$

De manera similar si A_i denota la fila de i de A entonces

$$AB = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix} B = \begin{pmatrix} A_1B \\ A_2B \\ \vdots \\ A_mB \end{pmatrix}$$

En este último caso, A_i es un vector fila, en tanto que en el anterior B_j es un vector columna.

2.3 Algunas interpretaciones para el producto matricial

Inicialmente, la definición del producto matricial tal vez no resulte natural, sin embargo tiene ricas interpretaciones en distintos contextos.

Por ejemplo, como se vio en el capítulo anterior, un sistema de ecuaciones lineales:

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}\right.$$

en su forma matricial se escribe como:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

lo que se simplifica como $Ax = b$, con $A = (a_{ij})$, la matriz $m \times n$ de coeficientes del sistema, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ y $x = \begin{pmatrix} x_1 \\ \vdots \\ v_n \end{pmatrix}$, los vectores de constantes en las ecuaciones a la derecha del igual y de incógnitas del sistema, respectivamente.

También, las operaciones suma vectorial[5] y producto de un escalar por un vector, con la igualdad entre vectores permiten escribir el sistema en la forma:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

lo que se conocerá como **escritura columnar** del sistema.

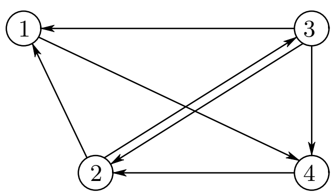
En el siguiente ejemplo se muestra, por otra parte, que el producto matricial también permite describir importantes relaciones en aplicaciones de la matemática, en este caso, en modelos de comunicaciones o teoría de redes.

Ejemplo 2.8 Consideramos un conjunto de n estaciones entre las cuales puede o no haber comunicación. Y convengamos en que la matriz $A = (a_{ij})$, se define como:

$$a_{ij} = \begin{cases} 1 & \text{si hay comunicación de } i \text{ a } j. \\ 0 & \text{si no hay comunicación de } i \text{ a } j. \\ 0 & \text{si } i = j. \end{cases}$$

describiendo así las posibles comunicaciones directas entre estaciones.

Por ejemplo, la comunicación entre cuatro estaciones que se ilustra en el siguiente diagrama, se describe con la matriz A, a la derecha.



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

[5]Recuerde que un vector es una matriz $n \times 1$, luego la suma vectorial y el producto de un escalar por un vector, son las mismas operaciones para matrices ya definidas

Si se multiplica A consigo misma se obtiene:

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

La entrada i, j de A^2, a^2_{ij}, resulta de multiplicar la fila i de A por su columna j:

$$a^2_{i,j} = a_{i1}a_{1j} + a_{i2}a_{2j} + \cdots + a_{in}a_{nj}$$

Y como cada a_{ik} es 0 o 1, los productos $a_{ik}a_{kj}$ son iguales a 0 o a 1, esto es:

$$a_{ik}a_{kj} = \begin{cases} 1 & \text{si hay comunicación de } i \text{ a } k \text{ y de } k \text{ a } j. \\ 0 & \text{si no hay comunicación de } i \text{ a } k \text{ o de } k \text{ a } j. \end{cases}$$

Entonces en la suma de a^2_{ij}, se acumula 1 con las estaciones k que permiten el enlace de i a k y de k a j y cero en otro caso, luego a^2_{ij} es la cantidad de estaciones que permiten un enlace de i a j.

Luego si $a^2_{i,j} \neq 0$, significa que hay comunicación de i a j, usando un intermediario. Y la entrada $a^2_{i,j} = 0$, indica que no hay forma de comunicar i con j, a través de un intermediario. Por ejemplo, $a^2_{24} = 2$, lo cual significa que hay dos formas de lograr la comunicación de la estación 2 a la 4 usando un intermediario.

2.4 Propiedades del producto matricial

La suma de matrices hereda todas las propiedades de la suma en los números reales, lo que no ocurre con el producto matricial. En este caso, hay diferencias sustanciales que se reflejan por su ausencia en el siguiente teorema.

Teorema 2.11 (Propiedades del producto) *Suponga que en cada caso, A, B, C, D son matrices de entradas reales con los tamaños apropiados para que los productos y sumas abajo indicados estén bien definidos, y que $\lambda \in I\!R$. Entonces el producto matricial tiene las siguientes propiedades:*

1. *Asociatividad:* $A(BC) = (AB)C$.
2. *Neutros: si A es $n \times m$,* $I_n A = A I_m = A$.
3. *Distributividad del producto respecto a la suma:*

$$A(B + C) = AB + AC \quad \text{y} \quad (B + C)D = BD + CD.$$

4. $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

Demostración: (Asociatividad) suponga que B y C son de orden $n \times p$ y $p \times m$ respectivamente, de donde BC es de orden $n \times m$. Como $A(BC)$ esta definido A debe tener orden $q \times n$, por lo tanto AB está definido y es de orden $q \times p$, se sigue que $(AB)C$ también lo está y es de orden $q \times m$; además

$$\begin{aligned}
(a(bc))_{ij} &= \sum_{k=1}^{n} a_{ik}(bc)_{kj} & &= \sum_{k=1}^{n} a_{ik} \sum_{r=1}^{p} b_{kr}c_{rj} \\
&= \sum_{k=1}^{n} \sum_{r=1}^{p} a_{ik}b_{kr}c_{rj} & &= \sum_{r=1}^{p} \sum_{k=1}^{n} a_{ik}b_{kr}c_{rj} \\
&= \sum_{r=1}^{p} \left(\sum_{k=1}^{n} a_{ik}b_{kr} \right) c_{rj} & &= \sum_{r=1}^{p} (ab)_{ir}c_{rj} \\
&= ((ab)c)_{ij}
\end{aligned}$$

■■■

Teorema 2.12 *La transposición de matrices tiene las siguientes propiedades.*

i) *Si* $A, B \in M(n, m, I\!R)$, $\qquad (A + B)^t = A^t + B^t$.

ii) *Si* $A \in M(n, m, I\!R)$ *y* $B \in M(m, p, I\!R)$ *entonces* $(AB)^t = B^t A^t$.

Demostración: Ver ejercicio 45 en la página 92.

■■■

Es importante reafirmar que en el teorema 2.11 no aparece la conmutatividad, o sea, el producto de matrices no es conmutativo. Observe que si A es $n \times m$ y B $m \times p$ con $n \neq p$, el producto AB está definido pero BA no se define. Pero aún más, aunque AB y BA están definidos, en general no se tendrá que $AB = BA$. Por ejemplo:

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

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

Por otra parte, si consideramos sólo las matrices cuadradas de orden n, el producto matricial tiene un elemento neutro: la matriz identidad,

$$AI_n = A \ \text{ y } \ I_n A = A \ \ \forall A \in M(n, I\!R)$$

Observe que en este caso particular, I_n conmuta con toda matriz A cuadrada de su mismo orden.

La existencia de elemento neutro para el producto matricial, nos lleva a la pregunta: ¿cada matriz $A \in M(n, I\!R)$, no nula, tendrá elemento inverso multiplicativo?, es decir, ¿existirá una matriz $A^{-1} \in M(n, I\!R)$ tal que $AA^{-1} = I_n$? Y si existiera, ¿se tendrá también que $A^{-1}A = I_n$?

La posible matriz inversa de A se ha denotado como A^{-1}, siguiendo la misma convención con que se denotan los inversos multiplicativos en $I\!R$, pero para buscar respuesta a las preguntas anteriores, momentáneamente cambiemos la notación de esta posible inversa de A a X y ocupémonos del siguiente problema:

> dada la matriz $A \in M(n, I\!R)$ buscar una matriz X de orden n tal que $AX = I$.

Si las n columnas de X se denotan con $X_1, X_2, \ldots, X_n$ y las columnas de la matriz identidad I_n como $e_1, e_2, \ldots, e_n$, entonces $AX = I$ también se puede escribir en la forma:

$$A(X_1, X_2, \ldots, X_n) = I = (e_1, e_2, \ldots, e_n)$$

o sea,

$$(AX_1, AX_2, \ldots, AX_n) = (e_1, e_2, \ldots, e_n)$$

Así, el problema de encontrar X tal que $AX = I$ se puede interpretar como el de encontrar n columnas $X_1, X_2, \ldots, X_n$ tales que

$$AX_i = e_i \quad \text{para } i = 1, 2 \ldots, n.$$

Y dicho problema se resuelve, si todos los sistemas $AX_i = e_i$ tienen solución. Ahora, si aplicando operaciones elementales por fila a A se puede obtener I_n, es decir, si A es equivalente a la identidad, entonces todos estos sistemas tienen solución y, específicamente, tienen solución única.

Por otra parte, como para resolver cada uno de los sistemas, $Ax = e_i$, hay que reducir la matriz $(A|e_i)$ a su forma escalonada reducida y esta tarea demanda exactamente las mismas operaciones elementales para cada i, entonces si se aplican dichas operaciones elementales a la siguiente matriz:

$$(A|e_1, e_2, \ldots, e_n) = (A|I)$$

se resuelven simultáneamente los n sistemas $AX_i = e_i$, y cada solución X_i obtenida queda en la columna que ocupaba e_i. Es decir, si A es equivalente a la identidad, al calcular la forma escalonada reducida de $(A|I)$ se obtiene:

$$(I|X_1, X_2, \ldots, X_n) = (I|X) = (I|A^{-1}).$$

Observe entonces que para que exista X tal que $AX = I$, es suficiente con que A sea equivalente a la identidad, ¿será esta una condición necesaria, también?

Ejemplo 2.9 Para la matriz

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

si existe, determine una matriz X, 3×3, tal que $AX = I$.

$$(A|I_3) = \left(\begin{array}{rrr|rrr} 2 & 6 & 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 7 & 5 & 0 & 0 & 1 \end{array}\right) \xrightarrow{f_1, f_2}$$

$$\left(\begin{array}{rrr|rrr} -1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 6 & 1 & 1 & 0 & 0 \\ 2 & 7 & 5 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\substack{2f_1+f_2 \\ 2f_1+f_3}} \left(\begin{array}{rrr|rrr} -1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 6 & 3 & 1 & 2 & 0 \\ 0 & 7 & 7 & 0 & 2 & 1 \end{array}\right)$$

$$\xrightarrow{\substack{-f_1 \\ (1/6)f_2}} \left(\begin{array}{rrr|rrr} 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 7 & 7 & 0 & 2 & 1 \end{array}\right) \xrightarrow{-7f_2+f_3}$$

$$\left(\begin{array}{rrr|rrr} 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{7}{2} & \frac{-7}{6} & \frac{-1}{3} & 1 \end{array}\right) \xrightarrow{(2/7)f_3} \left(\begin{array}{rrr|rrr} 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 0 & 1 & \frac{-1}{3} & \frac{-2}{21} & \frac{2}{7} \end{array}\right)$$

$$\xrightarrow{\substack{(-1/2)f_3+f_2 \\ f_3+f_1}} \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -1/3 & -23/21 & 2/7 \\ 0 & 1 & 0 & 1/3 & 8/21 & -1/7 \\ 0 & 0 & 1 & -1/3 & -2/21 & 2/7 \end{array}\right)$$

Luego, como resulta que A es equivalente a I_3, cada sistema $AX_i = e_i$ tiene solución única —por ejemplo, la solución de $AX_2 = e_2$ es $(-23/21, 8/21, -2/21)^t$— y la matriz buscada, $X = (X_1, X_2, X_3)$, es:

$$X = \begin{pmatrix} -1/3 & -23/21 & 2/7 \\ 1/3 & 8/21 & -1/7 \\ -1/3 & -2/21 & 2/7 \end{pmatrix}$$

2.5 Matrices inversas

En la sección anterior, observamos que dada una matriz cuadrada A de orden n, si es equivalente a la identidad se podrá encontrar una matriz X tal que $AX = I$, pero no sabemos si esta condición es necesaria para garantizar la existencia de X. Por otra parte, como el producto no es conmutativo, si existiera una matriz X tal que $AX = I$ tampoco es razonable esperar que para tal matriz se tenga también que $XA = I$.

Definición 2.13 (Matrices inversas)

Sea $A \in M(n, I\!R)$. Una matriz $B \in M(n, I\!R)$ se llama inversa izquierda de A si $BA = I$, y es inversa derecha de A si $AB = I$. Ahora, cuando $AB = I = BA$, B se llama una inversa de A y decimos que A es invertible.

Ya se justificó que, dada una matriz A $n \times n$, si para cada $i = 1, 2, \ldots, n$ hay un vector columna X_i que sea solución del sistema $Ax = e_i$, entonces

$$X = (X_1, X_2, \ldots, X_n)$$

es una inversa derecha de A, es decir, $AX = I$.

Aplicando este mismo procedimiento a A^t, en algunos casos será posible encontrar una matriz Y tal que $A^tY = I$, o lo que es lo mismo, tal que

$$Y^tA = I.$$

Es decir, Y^t es un inversa izquierda de A. Y aunque el principal problema es conocer las condiciones necesarias para que X y Y^t existan, si suponemos que existen, el siguiente teorema muestra que ambas matrices son la misma.

Teorema 2.14 *Si B es una matriz inversa izquierda de A y C una inversa derecha de A, entonces $B = C$.*

Demostración: Sólo se requiere de la asociatividad del producto matricial. Por hipótesis $BA = I$ y $AC = I$,

$$\text{luego } B = B(AC) = (BA)C = C.$$

▪▪■

Entonces la existencia de inversa derecha e inversa izquierda garantiza la existencia de una matriz inversa, ahora se verá que si existe una matriz inversa esta debe ser única.

Teorema 2.15 *Sea $A \in M(n, I\!R)$. Si A es invertible su inversa es única.*

Demostración: Nuevamente el resultado deriva de la asociatividad del producto matricial. Si tanto B como C son matrices inversas de A entonces: $AB = BA = I$ y $AC = CA = I$,

$$\text{luego } B = BI = B(AC) = (BA)C = IC = C.$$

▪▪■

Corolario 2.16 *Si $A \in M(n, I\!R)$ es invertible entonces los sistemas $Ax = e_i$ tienen solución única para $i = 1, \ldots, n$.*

Demostración: Si A es invertible existe B tal que $BA = I = AB$. Como $AB = I$ entonces la columna B_i de B es solución de $Ax = e_i$ para $i = 1, \ldots, n$, (todos los sistemas son consistentes). Por otro parte si algún sistema $Ax = e_j$ tiene infinitas soluciones, entonces existen B_j y d distintos tales que $AB_j = e_j = Ad$. Ahora, sea C es una matriz igual que B, pero con su columna j cambiada por d entonces C es una inversa derecha de A distinta de B. Lo cual contradice el teorema (2.14) porque B es inversa izquierda y C es inversa derecha luego $B = C$.

▪▪■

Este corolario da una condición necesaria para la existencia de la inversa derecha de A:

$$Ax = e_i \text{ tiene solución única para cada } i = 1, \ldots, n$$

lo que significa que la condición de que A sea equivalente a la identidad es una condición necesaria y suficiente para que exista X tal que $AX = I$. Por otra parte, si A es equivalente a la identidad, utilizando matrices elementales —las que se introducen en la siguiente sección— resulta fácil mostrar que A también tiene una inversa izquierda. Entonces se deduce que una condición necesaria y suficiente para la existencia de la inversa de A es que A sea equivalente a la identidad, lo cual se resume en el siguiente teorema agregando un nuevo enunciado equivalente.

Teorema 2.17 *Si A es una matriz $n \times n$, las siguientes proposiciones son equivalentes:*

i) A es invertible.

ii) A es equivalente a la identidad.

iii) El rango de A es n.

Los anteriores teoremas justifican que para calcular la matriz inversa de A sea suficiente determinar una inversa derecha, con el procedimiento ya establecido.

Teorema 2.18 *Sea $A, B \in M(n, \mathbb{R})$*

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

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

ii) Si A, B son invertibles, entonces AB es invertible y

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

Demostración: i) $AA^{-1} = I = A^{-1}A \Rightarrow A^{-1}A = I = AA^{-1}$, luego $(A^{-1})^{-1} = A$.

ii) $(AB)B^{-1}A^{-1} = A(BB^{-1})A^{-1} = AA^{-1} = I$. Luego $B^{-1}A^{-1}$ es una inversa derecha de AB. Similarmente se puede probar que $B^{-1}A^{-1}$ es una inversa izquierda de AB, luego se tiene que AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.

■

Teorema 2.19 *Si $A \in M(n, \mathbb{R})$ es invertible, entonces A^t es invertible y*

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

Demostración: Ver ejercicio 45.

■

2.6 Matrices elementales

Con la introducción de la matrices elementales de alguna manera volvemos al tema de sistemas de ecuaciones lineales, aunque esta vez, interesados en expresar el proceso de aplicar operaciones elementales a una matriz como un producto de matrices, lo cual será un valioso recurso para deducir nuevos resultados y equivalencias.

Cada operación elemental fila, se puede asociar a una matriz llamada matriz elemental fila o simplemente matriz elemental, en la siguiente forma:

Definición 2.20 (Matriz elemental)

Se llama matriz elemental de orden m a toda matriz que se obtiene aplicando sobre la matriz identidad de orden m, una única operación elemental. Estas matrices se denotan:

$$E(af_i),\ E(f_i, f_j),\ E(af_i + f_j).$$

utilizando la operación elemental que la origina.

Ejemplo 2.10 Las siguientes son matrices elementales 3×3:

$$E(af_1) = \begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad E(af_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{pmatrix}$$

$$E(f_1, f_2) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad E(f_2, f_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$E(af_1 + f_2) = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad E(af_3 + f_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & a \\ 0 & 0 & 1 \end{pmatrix}$$

2.6.1 Propiedades de las matrices elementales

Las siguientes propiedades de las matrices elementales se comprueban directamente haciendo las multiplicaciones matriciales respectivas.

Teorema 2.21 *1. Sea $A \in M(m, n, I\!R)$. Hacer una operación elemental fila sobre A, es lo mismo que multiplicar esta matriz por la izquierda con la respectiva matriz elemental de orden m. Es decir,*

$$\begin{aligned} E(af_i)A &= B \iff A \xrightarrow{af_i} B \\ E(f_i, f_j)A &= B \iff A \xrightarrow{f_i, f_j} B \\ E(af_i + f_j)A &= B \iff A \xrightarrow{af_i + f_j} B \end{aligned}$$

2. La matrices elementales son invertibles:

$(E(af_i))^{-1} = E(\frac{1}{a} f_i)$, siempre que $a \neq 0$.

$(E(f_i, f_j))^{-1} = E(f_i, f_j)$.

$(E(af_i + f_j))^{-1} = E(-af_i + f_j)$.

Ejemplo 2.11 Sea $A = \begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & 3 \\ 0 & -1 & -4 \end{pmatrix}$ y las matrices elementales

$$E(f_1, f_3) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ y } E(2f_1 + f_2) = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Haciendo las multiplicaciones matriciales comprobamos que

$$E(f_1, f_3)A = \begin{pmatrix} 0 & -1 & -4 \\ -2 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{y}$$

$$E(2f_1 + f_2)A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 7 \\ 0 & -1 & -4 \end{pmatrix}$$

que es equivalente a hacer respectivamente las operaciones elementales que siguen:

$$A \xrightarrow{f_1, f_3} \begin{pmatrix} 0 & -1 & -4 \\ -2 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{y}$$

$$A \xrightarrow{2f_1 + f_2} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 7 \\ 0 & -1 & -4 \end{pmatrix}$$

La parte 1) del teorema anterior, permite describir el proceso de aplicar operaciones elementales a una matriz de una manera bastante compacta que lo traslada al álgebra de matrices. Así la frase:

"B es el resultado de aplicar operaciones elementales a A"

se escribe también como:

"existen matrices elementales $E_1, \cdots, E_r$,
tales que $(E_r \cdots E_1)A = B$."

Con esto tenemos, que las tres frases siguientes expresan la misma idea.

i) B es equivalente a A.

ii) B se obtiene de A por medio de una secuencia de operaciones elementales.

iii) Existen matrices elementales $E_1, \cdots, E_r$, tales que

$$(E_r \cdots E_1)A = B.$$

Ejemplo 2.12 Sea $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & 2 & 3 \\ 0 & 1 & -1 & 2 \end{pmatrix}$

a. Calcular una matriz B escalonada y una matriz C escrita como producto de matrices elementales 3×3, tales que $B = CA$.

b. Lo mismo que en a, pero B es escalonada reducida.

Solución:

a. Es suficiente aplicar operaciones elementales sobre A hasta obtener una matriz escalonada. En efecto:

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & 2 & 3 \\ 0 & 1 & -1 & 2 \end{pmatrix} \xrightarrow{\quad f_1, f_2 \quad} \begin{pmatrix} 1 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 2 \end{pmatrix}$$

$$\xrightarrow{\quad f_2, f_3 \quad} \begin{pmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = B$$

Por lo tanto

$$C = E(f_2, f_3)E(f_1, f_2) =$$

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

b. A partir de la matriz B se obtiene la forma escalonada reducida R:

$$B = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{-2f_2 + f_1} \begin{pmatrix} 1 & 0 & 4 & -1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{-2f_3 + f_2} \begin{pmatrix} 1 & 0 & 4 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{f_3 + f_1} \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = R$$

Entonces $R = C_1 B$ donde C_1 viene dada por

$$C_1 = E(f_3 + f_1)E(-2f_3 + f_2)E(-2f_2 + f_1)$$

y como $B = CA$ se tiene que $R = C_1 CA$ con

$$C_1 C =$$

$$E(f_3 + f_1)E(-2f_3 + f_2)E(-2f_2 + f_1)E(f_2, f_3)E(f_1, f_2) =$$

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

∎

Ahora volvemos a las condiciones necesarias y suficientes para la existencia de la matriz inversa, utilizando las matrices elementales.

Teorema 2.22 *Sea $A \in M(n, \mathbb{R})$. Las siguientes afirmaciones son equivalentes*

a. A es invertible.

b. A es equivalente a I_n.

c. A es un producto de matrices elementales.

d. A tiene una inversa derecha.

e. A tiene una inversa izquierda.

Demostración: La estrategia de la prueba consiste en lo siguiente: primero se prueba $a \Rightarrow b \Rightarrow c \Rightarrow a$. Luego $d \Longleftrightarrow c$ y $e \Longleftrightarrow c$.

> $a \Rightarrow b$: existen R escalonada reducida y $R_1, \ldots, R_t$ matrices elementales tales que $R = (R_1 \ldots R_t)A$. Luego R es invertible y por tanto no tiene filas nulas, así $R = I_n$ y A es equivalente a I_n.
>
> $b \Rightarrow c$: A equivalente a $I_n \Rightarrow$ existen $R_1, \ldots, R_s$ matrices elementales tales que
> $I_n = (R_1 \cdots R_s)A$. Se concluye que $A = R_s^{-1} \cdots R_1^{-1}$.
>
> $c \Rightarrow a$: cada matriz elemental es invertible, luego A lo es.
>
> $d \Rightarrow c$: sea B tal que $AB = I$ y $E_1, \cdots, E_s$ matrices elementales tales que $A = (E_1 \ldots E_s)H$ con H escalonada reducida. Por tanto $AB = (E_1 \cdots E_s)(HB) = I$. Si H no es la matriz identidad, entonces su fila n-ésima es nula, luego HB no es invertible, lo cual es una contradicción. Así, $H = I$, con lo que $A = E_1 \cdots E_s$.

Similarmente se prueba que $e \Rightarrow c$. Además es claro que $c \Rightarrow d$ y que $c \Rightarrow e$.

∎

Este teorema muestra que la existencia de una matriz inversa derecha, es condición suficiente para garantizar la existencia de la inversa. También, la demostración de la equivalencia c) justifica de una nueva forma el método de cálculo de la matriz inversa, de la siguiente manera:

Otra deducción del algoritmo de cómputo de A^{-1}

Como se estableció en el teorema 2.22, basta que A tenga una inversa derecha o una izquierda, para que sea invertible. Además tal inversa derecha o izquierda será la inversa de A. Además como A debe ser equivalente a la identidad, entonces

$$(E_t \cdots E_1)A = I \Longrightarrow A^{-1} = E_t \cdots E_1$$

donde $E_1, \ldots, E_t$ son las matrices elementales que reducen a A a la identidad. Luego para determinar A^{-1} se procede de la siguiente forma:

1) Se construye la matriz $(A|I_n)_{n\times 2n}$

2) Si $O_1, \ldots, O_t$ son las operaciones elementales asociadas a las matrices elementales $E_1, \cdots, E_t$, que reducen A a la identidad, al aplicar estas operaciones sobre $(A|I_n)$ se obtiene:

$$(A|I_n) \xrightarrow{\quad O_1 \quad} (E_1A|E_1I_n) = (E_1A|E_1) \xrightarrow{\quad O_2 \quad}$$

$$(E_2E_1A|E_2E_1)\cdots \xrightarrow{\quad O_{n-1} \quad} (E_{t-1}\cdots E_1A|E_{t-1}\cdots E_1)$$

$$\xrightarrow{\quad O_t \quad} (E_tE_{t-1}\cdots E_1A|E_tE_{t-1}\cdots E_1) = (I_n|A^{-1}).$$

3) Naturalmente si el proceso de calcular la escalonada reducida de A, no conduce a I_n, es porque A no es equivalente a la identidad y por lo tanto no es invertible.

Ejemplo 2.13 Sea $A = \begin{pmatrix} 2 & 0 & 0 \\ 4 & 3 & 0 \\ 6 & 2 & 1 \end{pmatrix}$

a) Escriba A^{-1} como un producto de matrices elementales.

b) Escriba A como un producto de matrices elementales.

Solución:

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ 4 & 3 & 0 & 0 & 1 & 0 \\ 6 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\frac{1}{2}f_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 4 & 3 & 0 & 0 & 1 & 0 \\ 6 & 2 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{\substack{-4f_1+f_2 \\ -6f_1+f_3}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 3 & 0 & -2 & 1 & 0 \\ 0 & 2 & 1 & -3 & 0 & 1 \end{array}\right) \xrightarrow{\frac{1}{3}f_2}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 2 & 1 & -3 & 0 & 1 \end{array}\right) \xrightarrow{-2f_2+f_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{5}{3} & -\frac{2}{3} & 1 \end{array}\right)$$

Por lo tanto

$A^{-1} =$

$$E(-2f_2+f_3)E(\frac{1}{3}f_1)E(-4f_1+f_2)E(-6f_1+f_3)E(\frac{1}{2}f_1)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -6 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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

De acuerdo con el resultado anterior se tiene que

$A =$

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -6 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 6 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

$$= E(2f_1)\ E(6f_1 + f_3)\ E(4f_1 + f_2)\ E(3f_1)\ E(2f_2 + f_3)$$

■

2.7 Independencia lineal de vectores

Si A es una matriz $n \times m$, recordemos que el rango de A establece el máximo número de ecuaciones sin redundancia del sistema homogéneo $Ax = 0$ y el mínimo número de ecuaciones a que se puede reducir el sistema $Ax = 0$, sin alterar su conjunto solución.

Esta interpretación para el $\operatorname{Rng}(A)$ extrae información sobre las filas de A, ahora nos ocuparemos principalmente de sus columnas introduciendo conceptos que darán nuevas interpretaciones para el rango de A y más información sobre los sistemas $Ax = 0$ y $Ax = b$.

2.7.1 Combinación lineal de vectores

Consideremos un sistema de ecuaciones lineales, $n \times m$, $Ax = b$ y su escritura columnar:

$$x_1A_1 + x_2A_2 + \cdots + x_mA_m = b$$

donde $A = (A_1A_2 \ldots A_m)$, con $A_i \in I\!R^n$ la i-ésima columna de A y $x = (x_1, x_2, \ldots, x_m)^t$. Si para determinado $b \in I\!R^n$, el vector $x = (x_1, x_2, \ldots, x_m)^t$ es solución de $Ax = b$, entonces b puede expresarse como:

$$b = x_1A_1 + x_2A_2 + \cdots + x_mA_m$$

y se dirá que b es combinación lineal de los vectores columna $A_1, A_2, \ldots, A_m$.

Ejemplo 2.14 Resolviendo el sistema

$$\begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & -4 \\ 0 & 5 & -10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 13 \\ 15 \end{pmatrix}$$

se obtiene la siguiente forma escalonada reducida, de su matriz aumentada:

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

Luego la solución del sistema es $S = \{(5+s, 3+2s, s)^t | s \in I\!R\}$.

Si se elige una solución de $Ax = b$ como $(6,5,1)^t$, $s = 1$, se tiene entonces que:

$$\begin{pmatrix} -1 \\ 13 \\ 15 \end{pmatrix} = 6 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 5 \begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ -4 \\ -10 \end{pmatrix}$$

Es decir, $(-1, 13, 15)^t$ es combinación lineal de los vectores columna de la matriz del sistema.

Definición 2.23 (Combinación lineal de vectores)

Sean $v_1, v_2, \ldots, v_r$ vectores de $I\!R^n$. Se llama combinación lineal de los vectores $v_1, v_2, \ldots, v_r$, a cualquier vector v de la forma:

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_r v_r$$

donde $\alpha_1, \ldots, \alpha_r$ son cualesquiera números reales.

Al conjunto de todas las combinaciones lineales de $v_1, v_2, \ldots, v_r$ lo denotamos como:

$$\mathcal{C}\ell\{v_1, v_2, \ldots, v_r\} = \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_r v_r | \alpha_1, \ldots \alpha_r \in I\!R\}$$

Ejemplo 2.15 El vector $v = (-10, 4)^t$ es combinación lineal de los vectores $u = (1,5)^t$ $w = (4,2)^t$ porque

$$\begin{pmatrix} -10 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 5 \end{pmatrix} + (-3) \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

note que lo anterior se puede escribir como:

$$\begin{pmatrix} 1 & 4 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} -10 \\ 4 \end{pmatrix}$$

Esto muestra que al multiplicar una matriz por un vector , el vector que resulta es una combinación lineal de las columnas de la matriz.

Ejemplo 2.16 En la matriz del ejemplo de la página 47 "Posiciones de la Primera División", se tiene que los vectores columna de esta tabla: $PJ,\ PG,\ PE,\ PP \in I\!R^{12}$, satisfacen la relación

$$PJ = PG + PE + PP,$$

es decir, PJ es combinación lineal de PG, PE y PP.

También en la tabla del ejemplo Notas (página 48), tenemos que los vectores columna: P_1 "Parcial 1", P_2 "Parcial 2", P_3 "Parcial 3", Q "Quices", E "Escolaridad" de $I\!R^9$, satisfacen :

$$E = 0.32P_1 + 0.32P_2 + 0.32P_3 + 0.04Q.$$

Es decir la nota "Escolaridad" es una combinación lineal de los parciales y los quices.

2.7.2 Dependencia e independencia lineal

Definición 2.24 (Vectores linealmente independientes)

Los k vectores $u_1, u_2, \ldots, u_k$ de $I\!R^n$ son:

- *linealmente independientes, y se escribe l.i., si ningún vector u_i, $1 \leq i \leq k$, es combinación lineal de los restantes vectores.*
- *linealmente dependientes, y se escribe l.d., si alguno de ellos es combinación lineal de los otros, o sea, si no son l.i.*

Ejemplo 2.17 En el caso ejemplo (Posiciones de la Primera División página 47), tenemos que los conjuntos$\{PG, PE, PP, PJ\}$ es l.d. y $\{PG, PE, PP\}$ es l.i. En la tabla de notas $\{P_1, P_2, P_3, Q, E\}$ es l.d.

Aunque la anterior definición refleja muy bien la principal característica de los vectores linealmente independientes, no provee de un método apropiado para verificar dicha característica en un conjunto específico de vectores. Los siguientes resultados remedian esta situación.

Sean $u_1, u_2, \ldots, u_k$ vectores de $I\!R^n$ y supongamos que uno de ellos es combinación lineal de los restantes. Para simplificar suponga que u_k es combinación lineal de $u_1, u_2, \ldots, u_{k-1}$. Entonces:

$$\begin{aligned} & u_k = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_{k-1} u_{k-1} \\ \Longrightarrow \quad & \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_{k-1} u_{k-1} - u_k = 0_n \\ \Longrightarrow \quad & \begin{pmatrix} \\ u_1 u_2 \ldots u_k \\ \\ \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{k-1} \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \end{aligned}$$

Luego el sistema homogéneo $Ax = 0$, donde $A = (u_1 u_2 \ldots u_k)$, tiene una solución $x = (\alpha_1, \ldots, \alpha_{k-1}, -1)^t$ que claramente es distinta de 0_n. Esto significa que $Ax = 0$ tiene infinitas soluciones, o sea, tiene soluciones no nulas.

Inversamente, si partimos aceptando que $Ax = 0$ tiene soluciones no nulas, entonces supongamos que $x = (\alpha_1, \ldots, \alpha_k)^t$ es una de estas soluciones y, por ejemplo, que $\alpha_1 \neq 0$. Entonces:

$$\begin{aligned} & Ax = 0_n \\ \Longrightarrow \quad & \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k = 0_n \\ \Longrightarrow \quad & u_1 = \frac{-\alpha_2}{\alpha_1} u_2 + \ldots + \frac{-\alpha_k}{\alpha_1} u_k. \end{aligned}$$

Luego u_1 es combinación lineal de $u_2, \ldots, u_k$.

En resumen, los dos siguientes enunciados son equivalentes:

- Uno de los vectores $u_1, u_2, \ldots, u_k$ es combinación lineal de los restantes.
- El sistema homogéneo $Ax = 0$, donde $A = (u_1 u_2 \ldots u_k)$, tiene soluciones no nulas.

Y también son equivalentes las respectivas proposiciones negativas:

- Los vectores $u_1, u_2, \ldots, u_k$ son l.i.
- El sistema homogéneo $Ax = 0$, donde $A = (u_1 u_2 \ldots u_k)$, tiene solución única.

Por otra parte, la frase "$Ax = 0$ tiene solución única" es equivalente a la proposición: $0 = Ax \Rightarrow x = 0$, la que se suele escribir como:

$$0_n = x_1u_1 + x_2u_2 + \cdots + x_ku_k \Longrightarrow x_1 = x_2 = \ldots = x_k = 0$$

Con lo que se obtiene la clásica proposición equivalente para el enunciado "$\{u_1, u_2, \ldots, u_k\}$ es l.i." dada como definición o teorema en la mayoría de libros de Algebra Lineal.

Teorema 2.25 *Un conjunto $\{u_1, u_2, \ldots, u_k\}$ de vectores en $I\!R^n$ es linealmente independiente si y solo si*

$$0_n = a_1u_1 + a_2u_2 + \cdots + a_ku_k \Longrightarrow a_1 = a_2 = \ldots = a_k = 0$$

En otros términos, $u_1, u_2, \ldots, u_k$ son l.i. si y solo si el vector cero es combinación lineal de dichos vectores solamente de la manera trivial, es decir, con todos los escalares $a_i = 0, \ \ i = 1, 2, \ldots, k$.

Ejemplo 2.18 Demostrar que los vectores $u_1 = (1, 3, 2)^t, u_2 = (-3, 0, 1)^t$ y $u_3 = (-1, 2, 0)^t$ son linealmente independientes.

Demostración: Usando el teorema anterior, se probará que la proposición

$$0_3 = a_1u_1 + a_2u_2 + a_3u_3 \Longrightarrow a_1 = a_2 = a_3 = 0$$

es cierta.

$$0_3 = a_1u_1 + a_2u_2 + a_3u_3$$

$$\Longrightarrow \quad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + a_2 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} + a_3 \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

$$\Longrightarrow \quad \begin{pmatrix} 1 & -3 & -1 \\ 3 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Reduciendo a la forma escalonada reducida, la matriz aumentada de este sistema se obtiene:

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

Por lo que el sistema tiene solución única: $a_1 = a_2 = a_3 = 0$ y por lo tanto

$$0_3 = a_1u_1 + a_2u_2 + a_3u_3 \Longrightarrow a_1 = a_2 = a_3 = 0.$$

Luego el conjunto $\{u_1, u_2, u_3\}$ dado es linealmente independiente.

Como se ha visto:

$Ax = 0$ tiene solución única $\Longleftrightarrow$ las columnas de A son l.i.

Esto nos lleva a agregar una nueva proposición equivalente a las dadas en los teoremas 1.12 en página 31 y 2.17 en la página 66, y a formular un nuevo teorema que resume estas equivalencias:

Teorema 2.26 *Si A es una matriz $n \times n$, las siguientes proposiciones son equivalentes.*

i) *A es invertible.*

ii) *A es equivalente a la identidad.*

iii) *$Rng(A) = n$.*

iv) *$Ax = 0$ tiene solución única.*

v) *$Ax = b$ tiene solución $\forall\ b \in I\!R^n$.*

vi) *Las columnas de A son l.i.*

2.7.3 Más interpretaciones para el rango

En esta sección nos proponemos reconocer que el rango de una matriz A, además de corresponder al máximo número de ecuaciones de Ax = 0, sin redundancia, también es el máximo número de columnas de A l.i. Y para simplificar las notaciones requeridas en una demostración formal, trabajaremos con una situación particular, pero que ilustra las ideas generales.

Supongamos que $A = (A_1 A_2 A_3 A_4 A_5)$ es una matriz 4×5 cuyas columnas se denominan A_1, A_2, A_3, A_4 y A_5. Y que la siguiente matriz R es una forma escalonada equivalente a A,

$$R = \begin{pmatrix} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

donde cada $*$ puede ser sustituido por cualquier número real.

Veamos en primer lugar que las columnas de una forma escalonada, que contienen algún primer 1 (de alguna fila), son l.i. En el caso ilustrado esta columnas son la 1, 3 y 5. Y son l.i. porque el siguiente sistema tiene solución única, lo que se verifica fácilmente resolviendo mediante sustitución hacia atrás.

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

En segundo lugar las restantes columnas de la escalonada son combinación lineal de las columnas a su izquierda que contengan un primer uno. Por ejemplo, la columna 4 de R es combinación lineal de las columnas 1 y 3. Lo que puede reconocerse porque es claro que el siguiente sistema tiene solución, sin importar los números que se sustituyen con *.

$$\begin{pmatrix} 1 & * \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}$$

De esto se deduce que entre las columnas de R hay a un máximo de 3 ($\mathrm{Rng}\,(A)$) columnas l.i.:

- primero, hay tres columnas que son l.i. (las que contienen un primer 1)
- y segundo, si se agrega cualquier otra columna, se agrega una que es combinación lineal de las anteriores, con lo que el conjunto resultante sería l.d.

La idea entonces es que en una matriz escalonada: a) hay tantos primeros unos en sus filas como filas no nulas (el rango de A), b) hay tantas columnas l.i. como primeros 1 y c) si al conjunto de columnas con primeros 1 se agrega cualquier otra, el conjunto resultante es l.d., entonces el $\text{Rng}\,(A)$ es el máximo número de columnas l.i. de una escalonada equivalente a A.

Ahora, lo que se dice para las columnas de R también se puede decir para las respectivas columnas de A. Esto porque si R es una forma escalonada equivalente a A, entonces existen $E_1, E_2, \ldots, E_k$ matrices elementales tales que

$$E_k \cdots E_1 A = R = (R_1, R_2, R_3, R_4, R_5)$$

(aquí se supone nuevamente que R y A tienen la forma elegida anteriormente) luego para cada columna R_i de R se tiene que:

$$E_k \cdots E_1 A_i = R_i.$$

De esto se deduce que para cualquier elección de columnas de A, por ejemplo A_1, A_3, A_5, los sistemas de ecuaciones homogéneos:

$$(A_1 A_3 A_5)x = 0 \quad \text{y} \quad (R_1 R_3 R_5)x = 0$$

son equivalentes, es decir, ambos tienen solución única o ambos tienen infinitas soluciones.

Luego si $R_1,\ R_3,\ R_5$ son l.i. entonces $A_1,\ A_3,\ A_5$ son l.i. y si $R_1,\ R_3,\ R_4$ son l.d. entonces $A_1,\ A_3,\ A_4$ son l.d.

Finalmente podemos establecer que:

El Rng (A) es el máximo número de columnas l.i. de A.

Esta nueva interpretación para el rango de A extrae información sobre las columnas de A, sin embargo, los siguientes teoremas permiten dar la misma interpretación respecto a las filas de A.

Teorema 2.27 *Para toda matriz $A \in M(n, m, I\!R)$,*

$$Rng\,(A) \leq Rng\,(A^t).$$

Demostración: Supongamos que $\text{Rng}\,(A) = r$ y que H es la matriz escalonada reducida equivalente con A, luego tenemos:

- H tiene rango r.
- Existen $E_1, \ldots, E_p$ matrices elementales tal que
$$E_p \cdots E_1 A = H \;\text{ con }\; E_p \cdots E_1 = E.$$
- H tiene en sus columnas r-vectores canónicos diferentes de $I\!R^n$ que denotamos como $e_{s(i)}$ con $s(i) \in \{1, \ldots, r\}$, $i = 1, \ldots, r$.

De la igualdad $EA = H$ se sigue que $EA_{s(i)} = e_{s(i)}$ con $A_{s(i)}$ la $s(i)$-ésima columna de A y como E es invertible tenemos que el conjunto de las columnas de A

$$\{A_{s(1)}, \ldots, A_{s(r)}\} = \{E^{-1}e_{s(1)}, \ldots, E^{-1}e_{s(r)}\}$$

es linealmente independiente (ver ejercicios 42 y 43), esto es que A^t tiene al menos r filas linealmente independientes, por lo tanto:

$$\text{Rng}\,(A) \leq \text{Rng}\,(A^t).$$

■

Corolario 2.28 *Para toda matriz $A \in M(n, m, I\!R)$,*

$$Rng\,(A) = Rng\,(A^t).$$

Demostración: Aplicando el teorema anterior a A y A^t tenemos que $\text{Rng}\,(A) \leq \text{Rng}\,(A^t)$ y $\text{Rng}\,(A^t) \leq \text{Rng}\,((A^t)^t) = \text{Rng}\,(A)$ por lo tanto:

$$\text{Rng}\,(A) = \text{Rng}\,(A^t).$$

■

Así el rango de A también puede interpretarse como el máximo número de filas l.i. en A, con lo cual obtenemos la igualdad entre:

- Máximo número de ecuaciones sin redundancia en $Ax = 0$.
- Máximo número de filas l.i. de A.

Y se ligan las ideas de "independencia lineal" con "información sin redundancia" y la de "dependencia lineal" con "existencia de información redundante".

2.8 Ejercicios

1. Considere las siguientes matrices:

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

Si las siguientes operaciones están bien definidas, calcule la matriz resultante:

a) $A + 3I_3$ b) $(A - B)C$

c) $(A - B)(A + B)$ d) $A^2 - B^2$

e) CC^t f) C^tC

2. Sean $A = \begin{pmatrix} -1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, y $B = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$.

a) Efectúe las siguientes operaciones:

1) $A - 3I_3$ 2) $(A^2 + A - 3I_3)$

3) $(A - 3I_3)(A^2 + A - 3I_3)$ 4) $\frac{1}{9}A(-A^2 + 2A + 6I_3)$

b) Verifique que $-B^3 + B^2 + 5B - I_3 = 0$. Y deduzca que $C = -B^2 + B + 5I_3$ tiene la propiedades: $CB = I$ y $BC = I$.

3. Para qué valores de x se cumple que:

$$(x,4,1)\begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 0 & 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ 4 \\ 1 \end{pmatrix} = 0$$

4. ¿Cuál es el efecto sobre las filas (o columnas) de una matriz A si la multiplicamos por la izquierda (o derecha) por una matriz diagonal?

5. Demuestre que la suma de dos matrices de la forma $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, donde a, b son números reales, es una matriz de la misma forma.

6. Sea E_{pq} la matriz de 2×2 que contiene un 1 en el lugar $pq-$ésimo y el elemento 0 en los demás lugares.

a) Obtenga $E_{11}, E_{12}, E_{21}, E_{22}$.

b) Encuentre los reales a, b, c, d tales que:

$$aE_{11} + bE_{12} + cE_{21} + dE_{22} = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}.$$

7. Sea $S \in M(n, I\!R)$ tal que $Se_i = e_i, i = 1, \cdots, n$, donde $e_1, \cdots, e_n \in I\!R^n$, son los vectores canónicos[6]. Deduzca que $S = I_n$.

8. Sea A y B de orden n y sean

$$C_1 = \alpha_1 A + \beta_1 B; \qquad C_2 = \alpha_2 A + \beta_2 B$$

donde $\alpha_1, \alpha_2, \beta_1, \beta_2$ son escalares tales que $\alpha_1\beta_2 \neq \alpha_2\beta_1$. Demuéstrese que $C_1C_2 = C_2C_1$ sí y solo sí $AB = BA$

9. Proponga un ejemplo, no trivial, de matrices A y B de orden 3, tales que

$$A^3 - B^3 = (A - B)(A^2 + AB + B^2).$$

[6] e_i vector con un 1 en la posición $i-$ésima y el resto de sus entradas iguales a cero

10. Si $AB = BA$, se dice que las matrices A y B son conmutativas o que conmutan. Demuéstrese que para todos los valores de a, b, c, d, las matrices

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \text{ y } B = \begin{pmatrix} c & d \\ -d & c \end{pmatrix} \text{ conmutan.}$$

11. a) Demuéstrese que cualesquiera dos matrices diagonales del mismo orden conmutan.

b) Suminístrese una fórmula para D^p donde D es diagonal y p es un entero positivo.

12. Consideramos un conjunto de n estaciones entre las cuales puede o no haber comunicación. Denotemos esta situación en una matriz $A = (a_{ij})$ donde:

$$a_{ij} = \begin{cases} 1 & \text{si hay comunicación de } i \text{ a } j. \\ 0 & \text{si no hay comunicación.} \\ 0 & \text{si } i = j. \end{cases}$$

A continuación se indica la matriz de comunicación entre cuatro estaciones.

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

a.) Escriba el diagrama correspondiente.

b.) Calcule $A + A^2$ e interprete el resultado.

13. Calcule las matrices inversas derechas de:

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

14. Si es posible, determine una inversa derecha de A, y si existe verifique que también es una inversa izquierda.

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

15. Verifique que para cualquier valor de θ

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ -\operatorname{sen}\theta & -\cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ -\operatorname{sen}\theta & -\cos\theta \end{pmatrix}$$

16. Demuestre que si una matriz cuadrada tiene una fila de ceros, entonces no puede tener inversa. Análogamente si tiene una columna de ceros.

17. Si A es invertible, B escalonada y A equivalente a B, ¿qué forma general tiene B?

18. Sea A una matriz triangular. Probar que A es invertible si y sólo si $a_{ii} \neq 0$ para todo i.

19. Encuentre la inversa de cada una de las siguientes matrices, donde k_1, k_2, k_3, k_4 y k son escalares diferentes de cero.

$$\begin{pmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & k_2 & 0 \\ 0 & k_3 & 0 & 0 \\ k_4 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} k & 0 & 0 & 0 \\ 1 & k & 0 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & k \end{pmatrix}$$

20. Encuentre la matriz B que satisface la ecuación

$$A(B^t + C) = D, \quad \text{con}$$

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 5 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 1 \\ 3 & 0 \\ 1 & 3 \end{pmatrix}.$$

21. a) Sea $A \in M(n, \mathbb{R})$ tal que $A^3 = 0$. Demuestre que $I + A + A^2$ es la matriz inversa de $I - A$.

b) Sea $A = \begin{pmatrix} 0 & 1 & 6 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$. Verifique que $A^3 = 0$ y utilice el resultado en a) para determinar la inversa de $B = \begin{pmatrix} 1 & -1 & -6 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}$.

22. Suponga que A es una matriz 2×1 y B una matriz 1×2. Demuestre que $C = AB$ no es invertible

23. Sea $A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 0 & 5 & 3 \end{pmatrix}$. Calcule una matriz C invertible tal que CA sea triangular con los elementos en la diagonal iguales a uno.

24. Sea $A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 1 & 2 & 2 & 3 \\ 0 & 1 & -1 & 2 \end{pmatrix}$

a. Calcular una matriz B escalonada y una matriz C escrita como producto de matrices elementales 3×3, tales que $B = CA$.

b. Lo mismo que en (a), pero con B es escalonada reducida.

25. a) Pruebe que $A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 6 \\ 6 & -3 & -4 \end{pmatrix}$ no tiene inversa.

b) Calcule una matriz E escalonada y exprésela como $E = (F_k \cdots F_1)A$ donde $F_1, \cdots, F_k$ son matrices elementales.

26. a) Exprese en la forma $Ax = b$ con A matriz, x, b vectores columnas, la relación siguiente:
$\begin{pmatrix} 2 \\ 8 \\ 13 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 2 \\ 5 \end{pmatrix}$

b) Idem con $xA = b$, x, b vectores filas
$2(-1, 3, 5) + (1, 1, 1) - (2, 0, 3) = (-3, 7, 8)$.

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

a) Decida si los siguientes vectores columnas pertenecen al conjunto de las combinaciones lineales de las columnas de A:
$x_1 = (1, 5)^t \;\; x_2 = (a, b)^t$

b) ¿Son los vectores $x_1 = (5, 7, 9)^t \;\; x_2 = (5, 7, 10)^t$ combinaciones lineales de las filas de A?

c) Encuentre condiciones sobre a, b, c de modo que el vector $x = (a, b, c)^t$ pertenezca a las combinaciones lineales de las filas de A.

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

a) Encuentre una combinación lineal de las columnas de A que sea igual al vector $b = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$

b) Usando $b = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ muestre que el problema anterior no siempre tiene solución .

c) Deduzca, sin hacer cálculos, el rango de las matrices:
$(A, b) = \begin{pmatrix} 1 & -1 & 3 \\ 2 & -2 & 6 \end{pmatrix}$, $(A, b) = \begin{pmatrix} 1 & -1 & 3 \\ 2 & -2 & 0 \end{pmatrix}$

29. a) ¿Cuál es la combinación lineal de las columnas de $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$ que es igual al vector $b = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$?

b) ¿Cuál es el rango de $(Ab) = \begin{pmatrix} 1 & -1 & 5 \\ 2 & 3 & 0 \end{pmatrix}$?

30. a) Exprese el vector $(-5, 4, -9)$ como combinación lineal de los vectores $(1, 2, -1)$, $(5, 3, 2)$ y $(11, 8, 3)$, de dos maneras diferentes.

b) Si $A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \\ -1 & 4 & 0 \end{pmatrix}$ y $b = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$. Exprese b como combinación lineal de las columnas de A. ¿Cuántas maneras hay de hacer esto? Considere el caso $Ax = 0$, ¿son los vectores columna de A, l.i.? Justifique.

31. Sea A una matriz 3×3, cuyas columnas A_1, A_2, A_3 satisfacen que:

$$A_1 + A_2 + A_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; A_1 + 2A_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ y } A_1 - A_2 + 2A_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

i) Determine una matriz B tal que $AB = I_3$.

ii) ¿Las columnas de A son l.i.? Justifique su respuesta.

32. ¿Para cuáles valores de a el conjunto de vectores

$$\{(1, 2a, 0)^t, (1, -3, a+1)^t, (0, 1, -4)^t\}$$

es linealmente dependiente?

33. Considere los vectores en $I\!R^3$: $u = (k, 1, 1)$, $v = (1, k, 1)$ y $w = (1, 1, k)$.

¿Para qué valor o valores del parámetro k, si existen,

a) $(1,1,1)$ no es combinación lineal de u, v y w?

b) el vector $(1,1,1) \in \mathcal{C}\ell\{u,v,w\}$?

c) el vector $(1,1,1)$ se expresa como combinación lineal de u, v y w de manera única?

34. Sean u, v, w vectores de $\mathbb{R}^3$ tales que $\{u,w\}$ es l.i. y $u+v-w=0$.

a) ¿Es el conjunto $\{2u-w, w+2u\}$ l.i.? Justifique.

b) ¿Es el conjunto $\{2u-w, w-2u\}$ l.i.? Justifique.

c) Sea W la matriz con columnas iguales a los vectores $-u, 2v, -w$. ¿Cuál es el rango de W?

35. Sean $u=(1,2,-1)$, $v=(0,1,2)$, y $w=(-1,0,5)$. Responda con cierto o falso y justifique su respuesta.

a) $(0,2,4)$ es combinación lineal de $\{u,v\}$.

b) $(0,0,0)$ y w pertenecen a $\mathcal{C}\ell\{u,v\}$.

c) $\{u,v,w\}$ es un conjunto de vectores l.i.

d) Existen escalares α,β y γ, no todos nulos, tales que $(0,0,0)=\alpha u+\beta v+\gamma w$.

36. Sea A una matriz cuyas columnas son los vectores de $\mathbb{R}^4$, v_1, v_2, v_3, y v_4. Suponga que $S=\{(s,-s,0,2s)^t | s \in \mathbb{R}\}$, es el conjunto solución de $Ax=0$. Especifique si las siguientes afirmaciones son ciertas, falsas o no hay suficiente información para decidir y justifique su respuesta.

a) $v_2 = v_1 + 2v_4$.

b) El sistema $Ax=b$ tiene infinitas soluciones $\forall\ b \in \mathbb{R}^4$.

c) Toda matriz escalonada equivalente a A tiene al menos una fila de ceros.

d) Existe un vector columna $x \in \mathbb{R}^4$, $x \neq 0$, tal que $x^tA=0$.

37. Sea $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix} = (u\ v\ w\ 0_3)$, donde u, v, w, y 0_3 son los vectores columna de A. En cada caso, responda con cierto, falso o no hay suficiente información para decidir. Justifique sus respuestas.

a) El conjunto de vectores $\{v, w, 0_3\}$ es l.i.

b) $(2, 2, -1)^t$ y $(0, 0, 1)^t$ pertenecen a $\mathcal{C}\ell\{u, v, w\}$.

c) $\mathrm{Rng}\,(A) < 3$ y por lo tanto $Ax = 0$ tiene soluciones no nulas.

d) A es equivalente, por operaciones fila, a la matriz
$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

e) $(0, 0, 0, 1)^t$ y $u+v-w$ son elementos en $\{x \in I\!R^4 | Ax = 0\}$ (núcleo de A).

38. Dé ejemplos de una matriz 3×4 de rango 2 y otra de 3×2 de rango 1. Verifique que sus transpuestas tienen el mismo rango.

39. Sea $\{v_1, v_2, v_3\}$ un conjunto l.i. de vectores de $I\!R^n$, $A = (v_1 v_2 v_3)$ la matriz cuyas columnas son los vectores dados y $b = 2v_2 - v_1 + v_3$.

(a) ¿Cuál es el conjunto solución de $Ax = 0$? Razone su respuesta sin usar los teoremas 1.10 y 1.9.

(b) ¿Cuál es el conjunto solución de $Ax = b$? Razone su respuesta sin usar el teorema 1.9.

40. Si x, y son vectores no nulos, demuestre que $\{x, y\}$ es l.d. si y solo si, existe $\alpha \in I\!R$ tal que $x = \alpha y$.

41. Sea el conjunto $\{x, y\}$ l.i., con $x, y \in I\!R^n$. Demuestre que:

i) los conjuntos: $\{x, \alpha x + y\}, \{x - y, x + y\}$, son l.i.

ii) el conjunto $\{x - y, y - x\}$ es l.d.

42. Demuestre que cualquier subconjunto de $I\!R^n$ formado por vectores canónicos diferentes es linealmente independiente.

43. Sea $\{v_1, \ldots, v_r\}$ un conjunto de vectores de $I\!R^n$ linealmente independiente. Si $A \in M(n, I\!R)$ es invertible, entonces el conjunto $\{Av_1, \ldots, Av_r\}$ es linealmente independiente.

44. Sea $A \in M(n, I\!R)$. Se llama traza de A a la suma de los elementos de su diagonal principal. Denotamos

$$\text{Traza de } A = \mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$$

Demuestre que:

i) $\operatorname{tr}(A) = \operatorname{tr}(A^t)$

ii) $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$

iii) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ con $B \in M(n, I\!R)$

iv) $\operatorname{tr}(aA) = a \operatorname{tr}(A)$ donde $a \in I\!R$.

45. a) Sea $A = (1, 2, 3)$ Calcule AA^t, A^tA.

b) Demuestre que

i) $(A+B)^t = A^t + B^t$ con $A, B \in M(n, m, I\!R)$.

ii) $(AB)^t = B^tA^t$ con $A \in M(n, m, I\!R),\ B \in M(m, p, I\!R)$.

ii) $(A^t)^{-1} = (A^{-1})^t$ con $A \in M(n, I\!R)$, A invertible.

46. Demuestre que si $x, y \in I\!R^n$ entonces la matriz $xy^t + yx^t$ es simétrica.

47. Sea $P \in M(n, I\!R)$, P es ortogonal si $PP^t = P^tP = I_n$.

a) Demuestre que P es una matriz ortogonal, donde

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

b) Bajo qué condiciones para a y b la matriz siguiente es ortogonal.

$$X = \begin{pmatrix} a+b & b-a \\ a-b & b+a \end{pmatrix}$$

c) Demuestre que si $x \in I\!R^n,\ x \neq 0$

$$I_n - \frac{2}{x^tx}xx^t$$

es una matriz simétrica y ortogonal.

48. a) Sea $A = \begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}$. Resuelva el sistema $Ax = 0$ y defina una matriz B, 2×3, no nula tal que $AB = 0$.

b) Sea A una matriz $n \times n$, no invertible. Pruebe que existe una matriz B, $n \times m$ tal que $B \neq 0$ y $AB = 0$.

49. Supongamos conocida la matriz

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

a) Pruebe que el problema de calcular r, s, x, y, t, w tales que

$$\begin{pmatrix} 1 & 0 & r \\ 0 & 1 & s \end{pmatrix} = \begin{pmatrix} x & y \\ t & w \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

es equivalente a resolver por separado un par de sistemas de la forma $Lz_1 = C_1$ y $Lz_2 = C_2$ donde C_1 y C_2 no dependen de la matriz A.

b) Encuentre una condición sobre las entradas de A de modo que el problema tenga solución única[7].

50. Sea A una matriz $n \times n$. Pruebe que las siguientes proposiciones son equivalentes

a) A es invertible.

b) El conjunto solución de $Ax = 0$ es $S = \{0\}$.

51. Sea $A \in M(n, I\!R)$, demuestre que:

$$A \text{ es invertible} \iff \operatorname{Rng}(A) = n.$$

52. Sea $X \in M(n, m, I\!R)$.

(i) Si $X^t X = 0$ entonces $X = 0$.

(ii) Sean $P, Q \in M(q, m, I\!R)$ tales que $PX^tX = QX^tX$, entonces $PX^t = QX^t$. Sugerencia: si $Y = X(P - Q)^t$ verifique que $Y^tY = 0$.

[7]Para resolver esta parte del ejercicio se recomienda usar determinantes, por lo que se sugiere estudiar primero el siguiente capítulo de este texto.

53. Demuestre que si $A, D \in M(n, I\!R)$, D es diagonal con elementos no negativos, entonces

$$AD^p = D^p A \;\text{ para algún } p \text{ si y solo si }\; AD = DA.$$

Indic: Calcule la entrada ij de AD^p , $D^p A$, AD y DA

Capítulo 3

Determinantes

Aún cuando los determinantes reducen las matrices cuadradas a un simple número, y se podría pensar que en dicha reducción se pierde toda información útil aportada por la matriz, resulta sorprendente la relación de este número con otros conceptos. Para comenzar, si este número es cero el conjunto de las columnas de la matriz son vectores l.d. (la matriz no es invertible), cuando es distinto de cero los vectores columna son l.i. y aún más, puede interpretarse como el volumen del paralelepípedo engendrado por sus columnas. Esta última interpretación se verá, para $n = 3$, cuando se estudie el producto cruz, en el capítulo siguiente.

Por otra parte, los determinantes dan fórmulas explícitas para resolver los problemas $Ax = b$ y A^{-1}, que aunque no son importantes desde el punto de vista computacional, sí resultan de gran utilidad para resolver otros problemas en el cálculo y las ecuaciones diferenciales. Finalmente, tal vez, su ligamen más importantes sea con los valores propios: para una matriz A, los valores propios de A son los escalares λ tales que el determinante de $A - \lambda I$ es cero. Estos escalares juegan un papel muy importante en el algebra lineal y a ellos se dedicará el último capítulo de estos materiales.

3.1 Propiedades definitorias de la función determinante

El cálculo de la función determinante no es un problema difícil para matrices de orden pequeño, sin embargo, las distintas formas generales para definirlos resultan complejas. Nosotros hemos elegido definirlos, mediante su "desarrollo por menores", pero antes de hacerlo conviene ilustrar un caso particular y familiarizarnos con las propiedades esenciales que caracterizan la función determinante.

Ejemplo 3.1 Para matrices $A \in M(2, \mathbb{R})$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, la función determinante, D, se define por:

$$\begin{array}{rcl} D \ : M(2, \mathbb{R}) & \longrightarrow & \mathbb{R} \\ A & \longrightarrow & D(A) = ad - cb. \end{array}$$

Otras notación usuales para D son:

$$D(A) = \det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Fácilmente se verifica, mediante el cálculo directo, que la función D posee tres propiedades básicas:

1. D es lineal en cada fila. Esto significa que para toda matriz A y $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2, \alpha \in \mathbb{R}$; se debe cumplir :

$$\begin{vmatrix} x_1 + \alpha y_1 & x_2 + \alpha y_2 \\ c & d \end{vmatrix} = \begin{vmatrix} x_1 & x_2 \\ c & d \end{vmatrix} + \alpha \begin{vmatrix} y_1 & y_2 \\ c & d \end{vmatrix}$$

$$\begin{vmatrix} a & b \\ x_1 + \alpha y_1 & x_2 + \alpha y_2 \end{vmatrix} = \begin{vmatrix} a & b \\ x_1 & x_2 \end{vmatrix} + \alpha \begin{vmatrix} a & b \\ y_1 & y_2 \end{vmatrix}.$$

Para el caso de la primera fila se tiene:

$$\begin{vmatrix} x_1 + \alpha y_1 & x_2 + \alpha y_2 \\ c & d \end{vmatrix} = (x_1 + \alpha y_1)d - c(x_2 + \alpha y_2) =$$

$$(x_1 d - c x_2) + \alpha(y_1 d - c y_2) = \begin{vmatrix} x_1 & x_2 \\ c & d \end{vmatrix} + \alpha \begin{vmatrix} y_1 & y_2 \\ c & d \end{vmatrix}.$$

Similarmente se deduce la linealidad de D en la segunda fila.

2. D es alternada. Es decir, si dos filas adyacentes son iguales, entonces D vale cero:

$$\begin{vmatrix} x & y \\ x & y \end{vmatrix} = xy - xy = 0.$$

3. $D(I_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \cdot 1 - 0 \cdot 0 = 1.$

Estas tres propiedades caracterizan la función determinate. Es decir, las restantes propiedades enunciadas en los teoremas 3.2, 3.3 y 3.4, se pueden deducir de éstas. Y cualquier función de $M(n, \mathbb{R})$ a $\mathbb{R}$ que tenga estas tres propiedades es la función determinante.

3.2 Determinante de una matriz de orden n

La generalización de la definición de determinantes para matrices de orden n, se hará de manera inductiva:

Definición 3.1 (Función Determinante)

$D\colon M(n, \mathbb{R}) \longrightarrow \mathbb{R}$, se define por:

- *para $n = 1$, $D(a) = a$.*
- *para $n > 1$, y para cualquier fila i de A, $D(A)$ es la suma ponderada de n determinantes de submatrices de orden $n-1$ (denotadas A_{ij}):*

  $$D(A) = (-1)^{i+1} a_{i1} D(A_{i1}) + \cdots + (-1)^{i+n} a_{in} D(A_{in})$$

  donde A_{ij} es la submatriz que se obtiene de A al eliminar la fila i y la columna j.

En particular, para $n = 2$ observe que:

$$D(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (-1)^{1+1} a_{11} D(A_{11}) + (-1)^{1+2} a_{12} D(A_{12})$$

$$= a_{11} D(a_{22}) - a_{12} D(a_{21}) = a_{11}a_{22} - a_{12}a_{21}.$$

Lo cual coincide con la definición dada en el ejemplo anterior. Por otra parte, se puede verificar que, D, satisface las tres propiedades fundamentales:

1. D es lineal en cada fila. Si las filas de A se denotan como vectores fila $A_i \in I\!R^n$ y la fila r se expresa, para ciertos vectores filas x y y de $I\!R^n$, como: $A_r = \alpha x + y$, entonces:

$$D\begin{pmatrix} A_1 \\ \vdots \\ A_{r-1} \\ \alpha x + y \\ A_{r+1} \\ \vdots \\ A_n \end{pmatrix} = \alpha D \begin{pmatrix} A_1 \\ \vdots \\ A_{r-1} \\ x \\ A_{r+1} \\ \vdots \\ A_n \end{pmatrix} + D \begin{pmatrix} A_1 \\ \vdots \\ A_{r-1} \\ y \\ A_{r+1} \\ \vdots \\ A_n \end{pmatrix}.$$

2. D es alternada. Es decir, si dos filas adyacentes son iguales, entonces D vale cero:

3. $D(I_n) = 1$.

Dado que las tres propiedades anteriores caracterizan la función determinante, con solo mostrar que para cualquier elección de fila i, la función D definida en 3.1 tiene estas propiedades, se muestra que tal función es la función determinante. Es decir, en el cálculo del determinante con la definición 3.1, se puede elegir cualquier fila y obtener el mismo resultado.

Ejemplo 3.2 Calcule $D(A)$, para la matriz

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

Solución: Desarrollando por la segunda fila tenemos:

$$\begin{aligned} D(A) &= \textstyle\sum_{j=1}^{3}(-1)^{2+j}a_{2j}D(A_{2j}) \\ &= (-1)^{1+2}(1)\begin{vmatrix} -2 & 2 \\ 2 & 0 \end{vmatrix} + (-1)^{2+2}(0)\begin{vmatrix} 3 & 2 \\ -1 & 0 \end{vmatrix} \\ &\qquad + (-1)^{3+2}(1)\begin{vmatrix} 3 & -2 \\ -1 & 2 \end{vmatrix} \\ &= -1[(-2)0-(2)2]+0+(-1)[(3)2-(-2)2] \\ &= 0. \end{aligned}$$

∎

Ejemplo 3.3 Calcule el determinante de la matriz elemental

$$E = E(af_i)$$

Usando las propiedades (1) y (3) se tiene que

$$D(E) = aD(I_n) = a.$$

3.3 Propiedades del determinante

Los siguientes teoremas recogen las propiedades más importantes de la función determinante.

Teorema 3.2 *Si $A \in M(n, I\!R)$ y los vectores filas $A_1, A_2, \ldots, A_n$ son las filas de A, entonces:*

1. *$D(A) = 0$ si alguna fila de la matriz A es nula.*
2. *$D(\alpha A) = \alpha^n D(A) \quad \forall \alpha \in I\!R$.*
3. *D cambia de signo si dos filas k, r se intercambian:*

$$D\begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_r \\ \vdots \\ A_n \end{pmatrix} = -D\begin{pmatrix} A_1 \\ \vdots \\ A_r \\ \vdots \\ A_k \\ \vdots \\ A_n \end{pmatrix}.$$

   Como caso particular, se tiene que si E es la matriz elemental de permutación de las filas i y j, entonces

$$D(E) = -D(I_n) = -1.$$

4. *$D(A) = 0$ si dos filas de A son iguales.*

Demostración: Demostración de las propiedades 1 y 2:

1. Sea $A \in M(n, I\!R)$ con la fila $A_r = 0_n$. Entonces por la linealidad en la fila r, se concluye que

$$D\begin{pmatrix} A_1 \\ \vdots \\ 0_n \\ \vdots \\ A_n \end{pmatrix} = D\begin{pmatrix} A_1 \\ \vdots \\ 0_n + 0_n \\ \vdots \\ A_n \end{pmatrix} = D\begin{pmatrix} A_1 \\ \vdots \\ 0_n \\ \vdots \\ A_n \end{pmatrix} + D\begin{pmatrix} A_1 \\ \vdots \\ 0_n \\ \vdots \\ A_n \end{pmatrix}.$$

   Luego $D(A) = 2D(A)$ de donde se deduce que $D(A) = 0$.

2. $D(\alpha A) = D\begin{pmatrix}\alpha A_1\\ \alpha A_2\\ \alpha A_3\\ \vdots\\ \alpha A_n\end{pmatrix} = \alpha D\begin{pmatrix}A_1\\ \alpha A_2\\ \alpha A_3\\ \vdots\\ \alpha A_n\end{pmatrix} =$

$$\alpha^2 D\begin{pmatrix}A_1\\ A_2\\ \alpha A_3\\ \vdots\\ \alpha A_n\end{pmatrix} = \cdots = \alpha^n D\begin{pmatrix}A_1\\ A_2\\ A_3\\ \vdots\\ A_n\end{pmatrix}.$$

■■■

El siguiente teorema resume una propiedad de los determinantes especialmente importante, porque permite definir una manera de simplificar su cálculo, reduciendo la matriz a una forma triangular, mediante operaciones elementales del tipo $\alpha f_k + f_r$. Esto se verá en los ejemplos 3.5 y 3.6.

Teorema 3.3 *Si $A \in M(n, \mathbb{R})$ y los vectores filas $A_1, A_2, \ldots, A_n$ son las filas de A, entonces la operación elemental $\alpha f_k + f_r$ no altera el valor del determinante. Es decir, para todo $\alpha \in \mathbb{R}$ se tiene:*

$$D\begin{pmatrix}A_1\\ \vdots\\ A_k\\ \vdots\\ A_r\\ \vdots\\ A_n\end{pmatrix} = D\begin{pmatrix}A_1\\ \vdots\\ A_k\\ \vdots\\ A_r + \alpha A_k\\ \vdots\\ A_n\end{pmatrix}.$$

En particular si E es la matriz elemental $E = E(af_i + f_j)$, entonces $D(E) = D(I_n) = 1$.

Demostración: De la linealidad del determinante y la parte (4) del teorema anterior, sigue que:

$$D\begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_r + \alpha A_k \\ \vdots \\ A_n \end{pmatrix} = D\begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_r \\ \vdots \\ A_n \end{pmatrix} + \alpha D\begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_k \\ \vdots \\ A_n \end{pmatrix} = D\begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_r \\ \vdots \\ A_n \end{pmatrix} + 0.$$

∎

Teorema 3.4 *Si* $A, B \in M(n, I\!R)$ *entonces:*

1. $D(A) \neq 0 \iff A$ *es invertible; o lo que es equivalente:* A *no es invertible* $\iff D(A) = 0$.
2. $D(A) = D(A^t)\ \ \forall A \in M(n, I\!R)$.
3. $D(AB) = D(A)D(B)\ \ \forall A, B \in M(n, I\!R)$.

De la propiedad 2, en el teorema 3.4 se deduce que el determinante se puede calcular desarrollándolo por una columna. Y que todas las propiedades enunciadas sobre las filas de A también valen para sus columnas.

Ejemplo 3.4 Usando sólo las propiedades del determinante, deduzca:

$$\begin{vmatrix} a & b \\ b+c & a+c \end{vmatrix} = (a+b+c)\begin{vmatrix} a & b \\ 1 & 1 \end{vmatrix}$$

Solución: Realizando la operación elemental $f_1 + f_2$ se tiene :

$$\begin{vmatrix} a & b \\ b+c & a+c \end{vmatrix} = \begin{vmatrix} a & b \\ a+b+c & a+b+c \end{vmatrix} = (a+b+c)\begin{vmatrix} a & b \\ 1 & 1 \end{vmatrix}$$

(por la linealidad).

∎

Ejemplo 3.5 Calcular el determinante de $A = \begin{pmatrix} 0 & -1 & 3 \\ 1 & 2 & -2 \\ 2 & 0 & 3 \end{pmatrix}$.

Solución: Aplicando operaciones elementales por filas, se tiene:

$$A \xrightarrow{f_1, f_2} \begin{pmatrix} 1 & 2 & -2 \\ 0 & -1 & 3 \\ 2 & 0 & 3 \end{pmatrix} \xrightarrow{-2f_1 + f_3}$$

$$\begin{pmatrix} 1 & 2 & -2 \\ 0 & -1 & 3 \\ 0 & -4 & 7 \end{pmatrix} \xrightarrow{-4f_2 + f_3} \begin{pmatrix} 1 & 2 & -2 \\ 0 & -1 & 3 \\ 0 & 0 & -5 \end{pmatrix}$$

Desarrollando por la primera columna y tomando en cuenta que la primera operación produjo un cambio de signo se sigue que

$$D(A) = -\begin{vmatrix} 1 & 2 & -2 \\ 0 & -1 & 3 \\ 0 & 0 & -5 \end{vmatrix} = -\begin{vmatrix} -1 & 3 \\ 0 & -5 \end{vmatrix} = -5.$$

■

Ejemplo 3.6 Demuestre que

$$D(A) = \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = (y-x)(z-x)(z-y).$$

Solución: Haciendo operaciones elementales se tiene:

$$A \xrightarrow{-f_2 + f_3} \begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 0 & z-y & z^2-y^2 \end{pmatrix} \xrightarrow{-f_1 + f_2}$$

$$\begin{pmatrix} 1 & x & x^2 \\ 0 & y-x & y^2-x^2 \\ 0 & z-y & z^2-y^2 \end{pmatrix} = B.$$

Es decir, $B = E(-f_1 + f_2)E(-f_2 + f_3)A$. Por lo tanto:

$$D(B) = D(A) = \begin{vmatrix} y-x & y^2-x^2 \\ z-y & z^2-y^2 \end{vmatrix}$$

$$= (y-x)(z-y)\begin{vmatrix} 1 & y+x \\ 1 & z+y \end{vmatrix} = (y-x)(z-y)(z-x).$$

■

3.4 Regla de Cramer

Una de las aplicaciones más conocida para los determinante es la "regla de Cramer", que da una fórmula explícita para la solución de sistemas cuadrados de ecuaciones lineales, con solución única.

Teorema 3.5 *Si $A \in M(n, I\!R)$ y $D(A) \neq 0$ entonces el sistema $Ax = b$, donde $b \in I\!R^n$, tiene una única solución $x = (x_1, \ldots, x_n)^t$ dada por:*

$$x_i = \frac{D(A_1, \cdots, A_{i-1}, b, A_{i+1}, \cdots, A_n)}{D(A)}; \quad i = 1, \cdots, n$$

expresión en la cual $A_1, \cdots, A_n$ son las columnas de A.

Observe que la matriz $(A_1 \cdots A_{i-1} b\, A_{i+1} \cdots A_n)$, en el numerador de la fracción que determina a x_i, se obtiene sustituyendo en A la columna i-ésima por el vector b.

Demostración: Como $D(A) \neq 0$ se tiene que A es invertible y por lo tanto la solución del sistema $Ax = b$ es única. Utilizando las propiedades del determinante, se tiene:

$$\begin{aligned}
& D(A_1, \cdots, A_{i-1}, b, A_{i+1}, \cdots, A_n) \\
= \quad & D(A_1, \cdots, A_{i-1}, Ax, A_{i+1}, \cdots, A_n) \\
= \quad & D(A_1, \cdots, A_{i-1}, \textstyle\sum_{j=1}^n x_j A_j, A_{i+1}, \cdots, A_n) \\
= \quad & \textstyle\sum_{j=1}^n x_j D(A_1, \cdots, A_{i-1}, A_j, A_{i+1}, \cdots, A_n) \\
= \quad & x_i D(A_1, \cdots, A_{i-1}, A_i, A_{i+1}, \cdots, A_n) \\
= \quad & x_i D(A).
\end{aligned}$$

Los otros términos de la suma anterior son cero pues la matriz tiene dos columnas iguales. Despejando x_i se tiene:

$$x_i = \frac{D(A_1, \cdots, A_{i-1}, b, A_{i+1}, \cdots, A_n)}{D(A)}; \quad i = 1, \cdots, n.$$

■■■

Ejemplo 3.7 Resolver el sistema $\begin{cases} 2x + y + z &= 0 \\ x - y + 5z &= 0 \\ y - z &= 4 \end{cases}$

Solución:

$$D(A) = \begin{vmatrix} 2 & 1 & 1 \\ 1 & -1 & 5 \\ 0 & 1 & -1 \end{vmatrix} = -6; \quad x = -(\frac{1}{6}) \begin{vmatrix} 0 & 1 & 1 \\ 0 & -1 & 5 \\ 4 & 1 & -1 \end{vmatrix} = -4;$$

$$y = -(\frac{1}{6}) \begin{vmatrix} 2 & 0 & 1 \\ 1 & 0 & 5 \\ 0 & 4 & -1 \end{vmatrix} = 6; \quad z = -(\frac{1}{6}) \begin{vmatrix} 2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 4 \end{vmatrix} = 2.$$

∎

3.5 Ejercicios

1. En cada caso obtenga una matriz C triangular tal que A es equivalente a C por filas y $|A| = |C|$ y calcule $|A|$.

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 0 & 2 & 1 & 1 \\ 1 & -2 & 1 & 0 \\ 2 & 0 & -3 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

2. Sean A y B matrices cuadradas, conociendo que:

$$\left| \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array} \right) \right| = |A||B|,$$

calcule el determinante de las matrices:

$$\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 2 & 1 & 3 \\ 1 & -2 & 1 & 0 \\ 0 & 0 & -3 & 2 \\ 0 & 0 & 3 & 1 \end{pmatrix}.$$

3. Use sólo las propiedades del determinante para verificar que:

(a) $\begin{vmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & b+a \end{vmatrix} = 0.$

(b) $\begin{vmatrix} a+b & b+c & c+a \\ b+c & c+a & a+b \\ c+a & a+b & b+c \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}.$

4. Calcule $\begin{vmatrix} 2a+2b & 2b+2c & 2c+2a \\ 2b+2c & 2c+2a & 2a+2b \\ 2c+2a & 2a+2b & 2b+2c \end{vmatrix}$ si $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} = 3$

5. Sea $A \in M(n, I\!R)$. Demuestre que :

(i) A es invertible $\Rightarrow |A^{-1}| = \dfrac{1}{|A|}$.

(ii) $A^2 = A \Rightarrow |A| = 1$ o $|A| = 0$.

(iii) $A^t A = I_n \Rightarrow |A| = 1$ o $|A| = -1$.

(iv) $\text{Rng}\,(A) = n \Longleftrightarrow |A| \neq 0$.

6. Si $B = P^{-1}AP$, con A, B, P matrices en $M(n, I\!R)$ y P invertible, entonces

a) Muestre que $\det(B) = \det(A)$.

b) Muestre que $\det(\lambda I - B) = \det(\lambda I - A)$, con $\lambda \in I\!R$.

c) Si A es invertible y si $A^{-1} = A^t$, demuestre que $\det(2A) = \pm 2^n$.

7. Compruebe que los siguientes determinantes no dependen de a.

(i) $\begin{vmatrix} 1 & a+3 & (a+2)(a+3) \\ 1 & a+4 & (a+3)(a+4) \\ 1 & a+5 & (a+4)(a+5) \end{vmatrix}.$

(ii) $\begin{vmatrix} 1 & a+2 & (a+2)^2 \\ 1 & a+3 & (a+3)^2 \\ 1 & a+4 & (a+4)^2 \end{vmatrix}.$

8. Si $A \in M(n, I\!R)$ muestre que:

1. $|A| = 0$ si y solo si el conjunto de las filas de A es linealmente dependiente.

2. $|A| = 0$ si y solo si $\text{Rng}\,(\text{A}) < n$

9. Sea $A = \begin{pmatrix} x+1 & x & x \\ x & x+1 & x \\ x & x & x+1 \end{pmatrix}$. Determine todos los valores de x para los cuales A es invertible

10. Sea $A = \begin{pmatrix} 4-x & 2\sqrt{5} & 0 \\ 2\sqrt{5} & 4-x & \sqrt{5} \\ 0 & \sqrt{5} & 4-x \end{pmatrix}$. Determine todos los valores de x para los cuales $\text{Rng}\,(A) < 3$.

11. Decida si el siguiente conjunto es linealmente independiente $\{(1,2,1),(2,1,-1),(-1,1,1)\}$.

12. Demuestre que la recta que pasa por los puntos $P = (a,b)$ y $Q = (c,d)$ es determinada por la ecuación :

$$\begin{vmatrix} 1 & 1 & 1 \\ x & a & c \\ y & b & d \end{vmatrix} = 0$$

13. Sea

$$A = \begin{pmatrix} -1 & 0 & \cdots & 0 & x_1 \\ -1 & -1 & \cdots & 0 & x_2 \\ \vdots & & & & \vdots \\ -1 & -1 & \cdots & -1 & x_{n-1} \\ 0 & 0 & \cdots & 0 & -1 \end{pmatrix}$$

Compruebe que $|A| = 1$ si n es par y que $|A| = -1$ si n es impar.

14. Definición: Sea A matriz de $n \times n$

(1) $\lambda \in I\!R$ es un valor propio (valor característico) de A si y sólo si, λ es solución de la ecuación.

$$\det(\lambda I - A) = 0$$

(2) Se llama vector propio de A asociado al valor propio λ a todo vector $v \in I\!R^n$, $v \neq 0$ que verifica la propiedad.

$$Av = \lambda v$$

Encuentre los valores y vectores característicos (si existen), de las siguientes matrices.

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

$$C = \begin{pmatrix} 2 & -5 \\ -2 & -1 \end{pmatrix} \qquad D = \begin{pmatrix} 6 & 10 & -28 \\ -2 & -3 & 4 \\ 1 & 2 & -7 \end{pmatrix}$$

$$E = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

15. Use la regla de Cramer (si es posible), para resolver los sistemas siguientes:

$$(a) \quad \left\{ \begin{array}{rcr} 2x + y - 4z & = & 9 \\ x + 2y + 3z & = & -1 \\ x - y + 2z & = & -2 \end{array} \right.$$

$$(b) \quad \left\{ \begin{array}{rcr} x + z - w & = & -4 \\ 2x + y - z + w & = & 8 \\ -x + 2y - 2w & = & -5 \\ x + 2z + 2w & = & 3 \end{array} \right.$$

$$(c) \quad \left\{ \begin{array}{rcr} x - 3y - 2w & = & 0 \\ x - 3y + z + w & = & 0 \\ -3y + w & = & 0 \end{array} \right.$$

16. Considere el sistema lineal $Au = v$, con $A \in M(3, I\!R)$ y $u, v \in I\!R^3$, dado por $\left\{ \begin{array}{l} x + ay + a^2 z = a^2 \\ x + by + b^2 z = b^2 \\ x + cy + c^2 z = c^2 \end{array} \right.$, donde a, b y c son constantes cualesquiera de $I\!R$.

a) Muestre que

$$\det(A) = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (b-a)(c-a)(c-b).$$

b) Pruebe que el sistema dado tiene solución única si y solo si a, b y c son todas distintas.

c) Para el caso en que el sistema tiene solución única, use la regla de Cramer para calcular dicha solución.

17. Si el sistema

$$\left\{\begin{array}{rcl} a_1x + a_2y + a_3z + a_4w & = & a_1 \\ b_1x + b_2y + b_3z + b_4w & = & b_1 \\ c_1x + c_2y + c_3z + c_4w & = & c_1 \\ d_1x + d_2y + d_3z + d_4w & = & d_1 \end{array}\right.$$

tiene solución única, demuestre que ésta es:
$x = 1, y = z = w = 0$.

18. Demuestre que el sistema

$$\left\{\begin{array}{rcl} x + y + z & = & ax \\ 2x + y - z & = & ay \\ x - 2z & = & az \end{array}\right.$$

tiene soluciones no nulas si y sólo si $a(a - \sqrt{6})(a + \sqrt{6}) = 0$.

19. Estudie la solución de los sistemas siguientes con parámetro $a \in I\!R$.

(a) $\left\{\begin{array}{rcl} x + ay - z & = & a \\ ax + y + z & = & 0 \\ x + y + az & = & a^2 \end{array}\right.$ (b) $\left\{\begin{array}{rcl} ax + y + z & = & 1 \\ x + ay + z & = & 1 \\ x + y + az & = & 1 \end{array}\right.$

20. Demuestre que para todo $r \in I\!R$ y para todo $a, b, c, d \in I\!R$ el sistema siguiente tiene solución

$$\begin{array}{rcl} (1 + r)x + 2y + 2z + w & = & a \\ -x + (r - 1)y - z + 2w & = & b \\ (1 + r)z + 2w & = & c \\ -z + (r - 1)w & = & d \end{array}$$

21. Encuentre condiciones en términos de los parámetros, para que los tres planos siguientes se intersequen en un único punto.

$$\begin{array}{rcl} x - y + az & = & d_1 \\ x \quad + bz & = & d_2 \\ x + y + cz & = & d_3 \end{array}$$

22. Dados los puntos A, B y C en $I\!R^2$.

a) Establezca condiciones necesarias y suficientes sobre los puntos A, B y C, de modo que exista un único polinomio de grado menor o igual que 2, que "pasa" por esos tres puntos.

b) Establezca condiciones necesarias y suficientes en términos de determinantes de modo que exista un único polinomio de grado 2 que pasa por los tres puntos.

c) Extienda los resultados anteriores al caso de $n+1$ puntos en $I\!R^2$.

Capítulo 4

Programación Lineal

Este capítulo se inicia con la exposición de los problemas clásicos de producción y de transporte, en una versión simplificada, para ilustrar el tipo de problema que resuelve la programación lineal. El estudio de la solución geométrica a estos problemas de programación lineal deja ver la estrategia de solución que emplea el simplex y se reconoce una interesante aplicación del álgebra lineal. Con la presentación de las ideas que sustentan el método simplex se hace un intento por justificar cada paso, acudiendo sólo a la teoría conocida de los sistemas de ecuaciones lineales. Naturalmente, algunos resultados teóricos deberán ser aceptados sin demostración. El capítulo cierra con un esbozo de la estrategia de variables artificiales para iniciar la aplicación del simplex, cuando la formulación del problema de programación lineal no tiene la forma canónica.

4.1 Dos modelos clásicos de programación lineal

Los conocidos problemas del transporte y de asignación de recursos (producción), servirán para exponer el tipo de problemas que la programación lineal permite resolver.

4.1.1 Modelo de transporte

Supongamos que en m lugares (orígenes) hay $O_1, \ldots, O_m$ cantidades de un producto que deben ser transportadas hasta n lugares (destinos), con el fin de satisfacer una demanda conocida $D_1, \ldots, D_n$ de estos recursos. Nos interesa determinar las cantidades de cada producto que se deben llevar desde cada origen hasta cada destino, de manera que el costo de transporte sea mínimo. Para fijar ideas consideremos el siguiente caso particular:

Ejemplo 4.1 Dos fábricas de papel A y B, deben satisfacer la demanda semanal de tres imprentas etiquetadas con los números 1, 2 y 3. En el cuadro siguiente se indican los costos de transporte en dólares por tonelada, la producción semanal de cada fábrica en la quinta columna y la demanda semanal de cada imprenta en la cuarta fila.

	Imprenta			
Fábrica	1	2	3	Prod/sem.
A	17	22	15	350
B	18	16	12	550
Dem/sem	300	400	200	

Las condiciones de producción y transporte se rigen por las siguientes restricciones:

1. Cada fábrica envía toda su producción semanal a las imprentas.
2. Cada imprenta recibe semanalmente de las fábricas, una cantidad de papel que es igual a la demanda semanal.

El problema de programación lineal aquí es determinar la cantidad de toneladas de papel que se deben enviar semanalmente de cada fábrica a cada imprenta, de manera que el costo de transporte sea mínimo, respetando las dos restricciones anteriores.

Solución:

Definición de variables: Sea a_j (respect. b_j) el número de toneladas de papel que envía la fábrica A (respect. B) a la imprenta j, $a_j \geq 0$, $b_j \geq 0$; $j = 1,\ 2,\ 3$.

Restricciones: Este problema se puede representar por medio del diagrama de la Figura 4.1, en la cual las flechas van del origen al destino.

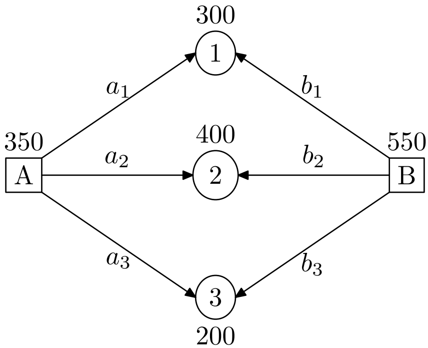


Figura 4.1: Representación del modelo de transporte con dos orígenes y tres destinos.

Tenemos las siguiente relaciones:

- La cantidad de papel enviado a la imprenta 1 es: $a_1 + b_1$, a la imprenta 2 es: $a_2 + b_2$ y a la imprenta 3: $a_3 + b_3$. Estas cantidades deben ser iguales a la demanda de las imprentas.

- La cantidad de papel enviada por la fábrica A a las imprentas es $a_1 + a_2 + a_3$ que debe ser igual a su producción y en el caso de la fábrica B es $b_1 + b_2 + b_3$.

Por lo tanto las restricciones del modelo se expresan mediante el sistema de ecuaciones (1):

$$(1) \quad \begin{cases} a_1 + b_1 & = & 300 \\ a_2 + b_2 & = & 400 \\ a_3 + b_3 & = & 200 \\ a_1 + a_2 + a_3 & = & 350 \\ b_1 + b_2 + b_3 & = & 550 \\ a_j \geq 0 \ ; \ \ b_j \geq 0 & & j = 1, 2, 3. \end{cases}$$

Función objetivo: se busca minimizar el costo total de transporte, C, el cual es:

$$C = 17a_1 + 22a_2 + 15a_3 + 18b_1 + 16b_2 + 12b_3$$

La solución de este problema de programación lineal consiste en encontrar un vector $(a_1^*, a_2^*, a_3^*, b_1*, b_2^*, b_3*)$ que minimice la función objetivo C, y que satisfaga las restricciones (1).

∎

4.1.2 Modelo de producción

Consideramos que una fábrica es un sistema cuya "entrada" la constituyen recursos e insumos tales como materias primas, fuerza de trabajo, tiempo de máquina y otros. La "salida" del sistema es el conjunto de bienes producidos con estos recursos. El problema básico es operar el sistema en condiciones óptimas.

Ejemplo 4.2 Una fábrica produce dos artículos, A y B. Los recursos que utiliza para producir cada artículo son: materia prima (MP), tiempo de máquina (TM) y fuerza de trabajo (FT). El sistema se puede representar mediante el esquema de la Figura 4.2.

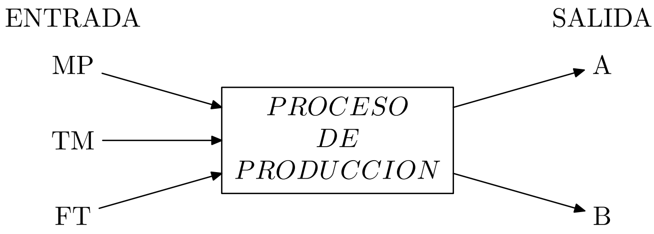


Figura 4.2: Representación del modelo de producción con tres insumos y dos productos.

En el cuadro siguiente se indican el gasto de insumos por unidad de artículo producido, así como las reservas para la producción durante tres meses y la ganancia por unidad de artículo vendido.

	Gasto de insumos por und/prod.			
Artículo	MP(lbs)	TM(min)	FT(hrs)	Ganancia
A	50	6	3	50
B	30	5	5	60
Reservas	2000	300	200	

Formule el problema de programación cuya solución es la cantidad de producto del tipo A y del tipo B que se debe producir en tres meses, para obtener la máxima ganancia.

Solución:

Definición de variables:

$x_1 \geq 0$: número de unidades producidas del artículo A.

$x_2 \geq 0$: número de unidades producidas del artículo B.

Restricciones: El gasto total de cada insumo no puede exceder el total disponible de este insumo. Por lo tanto si se producen x_1 unidades de A y x_2 unidades de B tenemos:

Gasto total de materia prima: $50x_1 + 30x_2$, luego

$$50x_1 + 30x_2 \leq 2000$$

Gasto total de tiempo de máquina: $6x_1 + 5x_2$, luego

$$6x_1 + 5x_2 \leq 300$$

Gasto total en fuerza de trabajo: $3x_1 + 5x_2$, luego

$$3x_1 + 5x_2 \leq 200$$

Función objetivo: Si se venden x_1 unidades de A y x_2 unidades de B, la ganancia total es $z = 50x_1 + 60x_2$.

La solución al problema de programación lineal será un vector $(x_1^*,\ x_2^*)$ que minimice la función objetivo

$$-z = -50x_1 - 60x_2$$

y que satisfaga las restricciones:

$$\begin{array}{rcl} 50x_1 + 30x_2 & \leq & 2000 \\ 6x_1 + 5x_2 & \leq & 300 \\ 3x_1 + 5x_2 & \leq & 200 \\ x_1 \geq 0 & , & x_2 \geq 0 \end{array}$$

Como el máximo de z se alcanza en el mismo (o mismos) punto(s) donde se alcanza el mínimo de $-z$, el problema de programación lineal puede formularse como un problema de minimización.

∎

4.2 Solución del problema de programación lineal

El método por excelencia para resolver un problema de programación lineal es el *método simplex* y como se verá, la geometría y el álgebra lineal juegan un papel esencial en la fundamentación de este método. Primero se ilustrarán los aspectos geométricos del problema mediante un ejemplo que permita exhibir la relación entre las propiedades geométricas y algebraicas de las soluciones al problema de programación lineal.

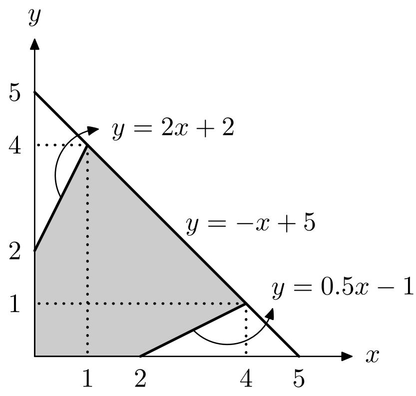


Figura 4.3: Región de soluciones factibles

4.2.1 Método geométrico

Se considera el problema de programación lineal:

$$\text{Min } z = y - x + 1$$

$$\begin{array}{lrcl} \text{Sujeto a:} & -2x + y & \leq & 2 \\ & x - 2y & \leq & 2 \\ & x + y & \leq & 5 \\ & x \geq 0 & , & y \geq 0 \end{array}$$

Los puntos del plano de coordenadas no negativas que satisfacen las restricciones, son los que conforman la región poligonal limitada por las rectas $y = 2x + 2$, $y = -x + 5$, $y = 0.5x - 1$ y los ejes coordenados, como se ilustra en la Figura 4.3. Dicha región se llama región de **soluciones factibles**.

Por otra parte, la función objetivo $z = y - x + 1$ se escribe como $y = x + z - 1$ que es una familia infinita de rectas paralelas dependientes del parámetro z. En la Figura 4.4 se muestran estas rectas para valores de z iguales a 2, 1, 0 y -2 y, en la tabla siguiente, las ecuaciones respectivas.

z	2	1	0	-2
ecuación	$y = x + 1$	$y = x$	$y = x - 1$	$y = x - 3$

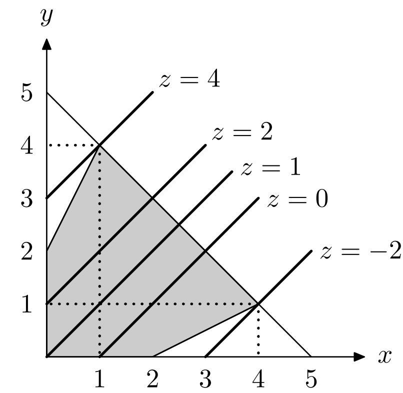


Figura 4.4: Corrimientos paralelos de la función objetivo.

Cuando se fija un valor de z, por ejemplo $z = 2$, y se traza la respectiva recta de la función objetivo, $y = x + 1$, se determinan los puntos (x, y) que evaluados en la función objetivo, $z = y-x+1$, producen el valor $z = 2$. Esto permite, por inspección del gráfico, determinar el mínimo (o máximo) valor de z cuya respectiva recta de la función objetivo interseca la región de soluciones factibles en al menos un punto. Observe que una de estas rectas que no corte la región de soluciones factibles, no determina puntos (x, y) que sean soluciones factibles al problema de programación lineal. Así, los desplazamientos de la rectas objetivos serán limitados a que dichas rectas intersequen la región de soluciones factibles.

En el ejemplo, a medida que z disminuye, se desplazan las rectas $y = x + z - 1$ hacia la parte inferior del polígono, alcanzándose el valor mínimo de z en el vértice (4,1), para el cual $z = -2$.

Siempre en referencia a la Figura 4.4, obsérvese que el máximo de z que es 4, se alcanza en el punto (1,4) que corresponde al máximo desplazamiento posible de las rectas paralelas de ecuación $y = x + z - 1$, $(z = -x + y + 1)$, hacia la parte superior de la Figura 4. El punto (1,4) es el mismo vértice donde la función $-z = x-y-1$ alcanza su mínimo, que es -4. Se ve que un problema de maximización se puede convertir en uno de minimización y recíprocamente, con solo *cambiar el signo* de la función objetivo.

Ambos problemas tienen la misma solución. Es decir, tanto el mínimo como el máximo se alcanzan en el mismo punto.

El enfoque geométrico, además de lo anterior, pone en evidencia que el valor óptimo de la función objetivo se alcanza en un vértice del polígono (región de soluciones factibles) o eventualmente en una arista, en cuyo caso el problema tiene infinitas soluciones óptimas. Desde luego que tal resultado vale y su generalización a problemas con más variables requiere caracterizar, entre otros, el concepto de vértice del conjunto de soluciones factibles, más comúnmente conocido como punto extremo.

Estas propiedades geométricas dan origen al desarrollo y fundamentación matemática del *método simplex*, cuya estrategia de solución podemos describir como:

- Caracterizar las soluciones de punto extremo (vértices).
- A partir de una solución de punto extremo, generar otra solución de punto extremo que disminuya el valor de la función objetivo.
- Este último paso se repite hasta que:
  - no es posible mejorar la solución y se evidencia que se está en una solución óptima,
  - o se reconoce que se pueden encontrar soluciones factibles que hagan tan "pequeño" como se quiera el valor de la función objetivo, en cuyo caso, el problema no tiene una solución óptima (la función objetivo no está acotada inferiormente).

Estas ideas serán las conductoras en el desarrollo siguiente de la solución algebraica al problema de programación lineal, aunque en algunos casos se omitirá la respectiva fundamentación teórica.

4.2.2 Solución algebraica: método simplex

El enfoque geométrico es irrealizable cuando intervienen más de tres variables, puesto que el ojo humano no ve más allá de tres dimensiones. Por eso es necesario desarrollar y fundamentar un procedimiento algebraico teórico, que resuelva nuestro problema.

Definición 4.1 (Programa lineal)

Un programa lineal consta de una función $z = c_1x_1+\cdots+c_nx_n+z_0$ llamada función objetivo, la cual debe ser maximizada o minimizada, y un sistema de ecuaciones lineales $m \times n$ (m restricciones):

$$\left\{\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{m1}x_n & = & b_m \end{array}\right.$$

donde $x_i \geq 0$ y la matriz de coeficientes $A = (a_{ij})_{m\times n}$ tiene rango m.

En forma más resumida, el problema de programación lineal lineal se plantea como:

$$\text{Minimizar } \; z = cx, \text{ sujeto a } Ax = b \text{ y } x \geq 0$$

donde $c = (c_1, \ldots, c_n)$, $x = (x_1, \ldots, x_n)^t$ y $b = (b_1, \ldots, b_m)^t$.

Definición 4.2 (Solución factible)

Una solución factible de un programa lineal

$$\textit{Minimizar } \; z = cx, \textit{ sujeto a } Ax = b \textit{ y } x \geq 0$$

es cualquier vector x, tal que $Ax = b$ y cuyas coordenadas son no negativas, es decir, $x \geq 0$.

Definición 4.3 (Solución óptima de un programa lineal)

Una solución óptima de un programa lineal es una solución factible x^* para la cual la función objetivo alcanza su valor mínimo. Es decir $z(x^*) \leq z(h)$ para todo h tal que $Ah = b$, $h \geq 0$.

Observaciones:

- Un programa lineal puede ser de la forma

$$\max\ z = cx \ \text{ sujeto a } \ Ax = b,\ x \geq 0.$$

  En tal caso es equivalente resolver el programa lineal

$$\min\ -z = -cx \ \text{ sujeto a } \ Ax = b,\ x \geq 0.$$

  puesto que el máximo de z (si existe) se alcanza en los mismos puntos en los cuales $w = -z$ alcanza su mínimo. Y min w = - max z.

- Las restricciones pueden ser, no solo ecuaciones, si no también inecuaciones. En cualquier caso, como se verá más adelante, las inecuaciones de un programa lineal siempre se pueden escribir como ecuaciones.

Definición 4.4 (Programa lineal canónico)

La formulación de programa lineal tiene la forma canónica y decimos que es un **programa lineal canónico***, si satisface las tres condiciones siguientes:*

1. *Los m vectores canónicos $e_1, e_2, \ldots, e_m$ de $\mathbb{R}^m$, son columnas de la matriz A, en algún orden. Las variables asociadas a estas columnas se denominan* **variables básicas** *y las restantes* **variables no básicas**.

2. *La función objetivo no depende de las variables básicas, es decir, los coeficientes asociados a las variables básicas son cero.*

3. *Los lados derechos de las restricciones son no negativos, esto es, $b_j \geq 0$, para $j = 1, \ldots, m$.*

Ejemplo 4.3 Considere el programa lineal

$$\min\ z = 5x_1 + 3x_4 - 2x_5 + 1$$

sujeto a las restricciones:

$$\left\{ \begin{array}{rrrrrrrl} -6\,x_1 & & +\,x_3 & -\,2x_4 & +\,2x_5 & = & 6 & \\ -3\,x_1 & +\,x_2 & & +\,6x_4 & +\,3x_5 & = & 15 & \\ & & & & x_i & \geq & 0 & \forall\, i \end{array} \right.$$

Las columnas 3 y 2 de la matriz del sistema de estas dos restricciones, son los vectores canónicos e_1 y e_2 de $\mathbb{R}^2$. Luego las variables x_2 y x_3 son variables básicas y la función objetivo no depende de estas variables. Además los valores b_1 y b_2 son no negativos: 6 y 15. Por lo tanto este programa lineal tiene la forma canónica.

Definición 4.5 (Solución básica factible)

Una solución básica factible de un programa lineal en la forma canónica es una solución factible, para la cual las variables **no básicas** *valen cero.*

Dos aspectos de las definiciones dadas hasta ahora, merecen resaltarse:

- Las soluciones básicas factibles corresponden a soluciones de puntos extremos (o vértices) de la región de soluciones factibles, lo cual no entramos a justificar.
- La formulación canónica de un problema de programación lineal permite conocer, de manera trivial, una solución básica factible.

En el ejemplo anterior, haciendo cero las variables no básicas, $x_1 = x_4 = x_5 = 0$, se reduce el sistema de restricciones a:

$$\left\{\begin{array}{ccc} & x_3 & = & 6 \\ x_2 & & = & 15 \end{array}\right.$$

Por lo tanto $x_2 = 15$, $x_3 = 6$ y $x_1 = x_4 = x_5 = 0$, constituye una solución factible. Así, $(0, 15, 6, 0, 0)$ es una solución **básica factible** porque es factible y las variables no básicas valen 0. Además, el valor de la función objetivo en esta solución básica factible es

$$z = (5, 0, 0, 3, -2)(0, 15, 6, 0, 0)^t + 1 = 1$$

Ejemplo 4.4 El programa lineal asociado con el modelo de producción (página 114) es min $z = 50x_1 + 60x_2$ sujeto a

$$\left\{\begin{array}{cccc} 50\ x_1 & +30x_2 & \leq & 2000 \\ 6\ x_1 & +\ 5x_2 & \leq & 300 \\ 3x_1 & +\ 5x_2 & \leq & 200 \end{array}\right.$$

con $x_i \geq 0 \; \forall \; i$. Para formular este programa lineal como un programa lineal canónico se convierten las inecuaciones en ecuaciones agregando variables de holgura: x_4 materia prima sobrante, x_5 tiempo de máquinas no utilizado y x_6 fuerza de trabajo no empleada. Así el problema se plantea como:

$$\min \; z = 50x_1 + 60x_2 \text{ sujeto a:}$$

$$\left\{ \begin{array}{rrrrrcr} 50x_1 & + 30x_2 & + x_3 & & & = & 2000 \\ 6x_1 & + \;\; 5x_2 & & + x_4 & & = & 300 \\ 3x_1 & + \;\; 5x_2 & & & + x_5 & = & 200 \end{array} \right.$$

con $x_i \geq 0 \; \forall \; i = 1, \ldots 6$. Observe que en este caso, se obtiene una formulación canónica con variables básicas x_3, x_4 y x_5. De la cual se deduce una primera solución básica factible: $(0, 0, 2000, 300, 200)$, que corresponde a no producir nada, $x_1 = 0$ y $x_2 = 0$, con lo cual sobran la totalidad de los recursos disponibles.

En general un programa lineal escrito en la forma:

$$\min \; z = \sum_{i=1}^{n} c_i x_i + z_0 \text{ sujeto a}$$

$$\left\{ \begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & \leq & b_1 \\ a_{21}x_1 & + & \cdots & + & a_{2n}x_n & \leq & b_2 \\ \vdots & & \vdots & & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{m1}x_n & \leq & b_m \end{array} \right.$$

donde $x_i \geq 0$ y $b_j \geq 0$, es equivalente al programa lineal canónico:

$$\min \; z = \sum_{i=1}^{n} c_i x_i + z_0 \text{ sujeto a}$$

$$\left\{ \begin{array}{lllllcl} a_{11}x_1 + \cdots + a_{1n}x_n & + x_{n+1} & & & & = & b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n & + & x_{n+2} & & & = & b_2 \\ \vdots & \vdots & & \ddots & & & \vdots \\ a_{m1}x_1 + \cdots + a_{m1}x_n & + & & & x_{n+m} & = & b_m \end{array} \right.$$

con $x_i \geq 0 \; \forall \; i = 1, \ldots, n+m$ y $b_j \geq 0 \; \forall \; j$.

Cuando el valor de alguna variable básica, en una solución básica factible se anula, se dice que la solución es **degenerada**. En estos casos, la teoría de la programación lineal debe contemplar situaciones especiales. En el marco de estas notas se supondrá que los problemas planteados no conducen a soluciones básicas factibles degeneradas, salvo mención contraria.

Con la caracterización de las soluciones extremas como soluciones básicas factibles y la obtención de estas mediante la formulación del problema de programación lineal en la forma canónica, podemos centrarnos en el problema de buscar una solución óptima.

Resultados que conducen la búsqueda del óptimo

Consideremos un problema de programación lineal con la forma canónica:

$$\text{Min } z = cx, \quad \text{sujeto a} \quad Ax = b \text{ y } x \geq 0.$$

Y supongamos que la matriz A, de orden $m \times n$ con $n > m$, tiene los vectores canónicos, $e_1, \ldots, e_m$, ocupando ordenadamente las primeras m columnas. Así, el sistema de restricciones $Ax = b$ se escribe en su forma columnar como:

$$x_1 e_1 + \cdots + x_m e_m + x_{m+1} A_{m+1} + \cdots + x_n A_n = b$$

donde $A_{m+1}, \ldots, A_n$ son las últimas $n - m$ columnas de A. Con esto se tiene también que:

- $x_1, x_2, \ldots, x_m$ son variables básicas.
- $x_{m+1}, \ldots, x_n$ son variables no básicas.
- $c_1 = c_2 = \cdots = c_m = 0$. (Porque se supone que el problema tiene la forma canónica).
- $x^* = (b_1, \ldots, b_m, 0, \ldots, 0)^t \in I\!R^n$ es una solución básica factible.

Como se verá, la argumentación siguiente no depende de que las variables básicas sean las primeras m, de manera que la única hipótesis importante es que el problema tiene la forma canónica. La elección anterior —el orden en las variables básicas— sólo facilita la escritura de ideas.

El cuestión fundamental es: **si el problema de programación lineal tiene la forma canónica, sólo uno de los siguientes casos se puede presentar:**

Caso 1: Todos los coeficientes de la función objetivo, asociados a variables no básicas, son no negativos.
Es decir, $c_i \geq 0 \ \ \forall i = m+1, \ldots, n$.

En esta situación, para toda solución factible $x \geq 0$,

$$z = cx = c_1x_1 + \cdots + c_mx_m + \cdots + c_nx_n \geq 0.$$

Además, para la solución básica factible

$$x^* = (b_1, \ldots, b_m, 0, \ldots, 0) \in I\!R^n$$

se tiene que

$$z(x*) = 0b_1 + \cdots + 0b_m + c_{m+1}0 + \cdots + c_n0 = 0$$

Luego en esta solución básica factible se alcanza el mínimo de la función z.

Caso 2: Existe por lo menos un coeficiente c_s de la función objetivo que es negativo, es decir, existe $c_s < 0$ con x_s una variable no básica.

Si este es el caso, se puede determinar una nueva solución factible que mejora la función objetivo, es decir, existe $y \in I\!R^n$ que es solución factible y $z(y) < z(x^*)$. Para determinar una nueva solución y se consideran dos subcasos.

Caso 2.a: Todas las entradas de la columna A_s son no positivos. Esto es : $a_{ks} \leq 0, \ \ \forall \ k = 1, \ldots, m$.

En este caso las restricciones $Ax = b$ se escriben

$$x_1e_1 + \cdots + x_me_m + x_{m+1}A_{m+1} + \cdots + x_sA_s + \cdots + x_nA_n = b$$

o equivalentemente:

$$\begin{aligned} x_1e_1 + \cdots + x_me_m + x_{m+1}A_{m+1} + \cdots & & \\ +x_{s-1}A_{s-1} + x_{s+1}A_{s+1} + \cdots + x_nA_n &= b - x_sA_s. \end{aligned} \tag{4.1}$$

Sea $d = b - x_sA_s$. Observe que $\forall \ x_s > 0, \ \ d \geq 0$, es decir, todas sus entradas son positivas o cero, porque las entradas de A_s son negativas o cero.

Por otra parte, si hacemos cero las variables no básicas, excepto x_s que se deja como parámetro, se tiene que

$$y = (d_1, \ldots, d_m, 0, \ldots, 0, x_s, 0, \ldots, 0) \in I\!R^n$$

es una solución del sistema 4.1, es decir, es una solución factible para cualquier elección de $x_s > 0$. Además, si la función objetivo se evalúa en y se tiene:

$$z(y) = c_s x_s < 0 = z(x^*)$$

porque $c_s < 0$. Entonces, claramente, se pueden elegir soluciones factibles para hacer $z(y)$ tan pequeña como se quiera. Es decir, la función objetivo no es acotada inferiormente y por lo tanto no existe una solución óptima.

Caso 2.b: Existe una entrada en la columna A_s que es positiva. Es decir, existe $a_{ks} > 0$.

En este subcaso expresando el conjunto de restricciones como en (4.1) para obtener una nueva solución factible, que dependa de x_s, se debe elegir x_s de manera que todas las entradas de $d = b - x_s A_s$ sean positivas o cero, esto es:

$$d = b - x_s A_s = \begin{pmatrix} b_1 \\ \vdots \\ b_k \\ \vdots \\ b_m \end{pmatrix} - x_s \begin{pmatrix} a_{1s} \\ \vdots \\ a_{ks} \\ \vdots \\ a_{ms} \end{pmatrix} = \begin{pmatrix} b_1 - x_s a_{1s} \\ \vdots \\ b_k - x_s a_{ks} \\ \vdots \\ b_m - x_s a_{ms} \end{pmatrix} \geq 0$$

Cuando $a_{ks} \leq 0$, claramente se tiene que $b_k - x_s a_{ks} \geq 0$ para cualquier elección $x_s > 0$. Y para las entradas k de A_s en las que $a_{ks} > 0$, x_s debe cumplir:

$$\begin{array}{ll} & b_k - x_s a_{ks} \geq 0 \\ \Longrightarrow & b_k \geq x_s a_{ks} \\ \Longrightarrow & x_s a_{ks} \leq b_k \\ \Longrightarrow & x_s \leq \frac{b_k}{a_{ks}} \end{array}$$

Entonces eligiendo x_s como:

$$x_s = \frac{b_r}{a_{rs}} = \min\{\frac{b_k}{a_{ks}} \mid a_{ks} > 0, k = 1, \ldots, m\}$$

se obtiene que x_s es positivo y es el mayor valor tal que $d_k = b_k - x_s a_{ks} \geq 0$ para cada $k = 1, \ldots, m$. Luego, como en el caso anterior:

$$y = (d_1, \ldots, d_m, 0, \ldots, 0, x_s, 0, \ldots, 0) \in I\!R^n$$

es una nueva solución factible que al evaluarla en la función objetivo mejora la solución x^* que le dio origen:

$$z(y) = c_s x_s < 0 = z(x^*)$$

Por otra parte, se observa que $d_r = b_r - \frac{b_r}{a_{rs}} a_{rs} = 0$ con lo que puede probarse que y es una solución básica factible, con variables básicas $x_1, \ldots, x_{r-1}, x_s, x_{r+1}, \ldots, x_m$. Así, en la nueva solución x_r deja de ser variable básica y se sustituye por la nueva variable básica x_s.

Tabla del simplex

Para efectos de calcular la nueva solución y, en el caso 2.b, es suficiente con darle al problema de programación lineal una nueva forma canónica que convierta, mediante operaciones elementales, la columna A_s en el vector canónico e_r. O lo que es lo mismo, se deben efectuar las siguientes operaciones elementales:

$$\tfrac{-a_{ks}}{a_{rs}} f_r + f_k \ \forall\ k = 1, \ldots, m, k \neq r \quad \text{y} \quad \tfrac{1}{a_{rs}} f_r$$

sobre la matriz aumentada del sistema de restricciones.

Para hacer esto, convenimos en representar el problema

$$\text{Min } z = cx, \quad \text{sujeto a } \ Ax = b \text{ y } x \geq 0.$$

mediante la matriz aumentada de las restricciones, agregando la función objetivo como una ecuación más que depende del parámetro z:

$$\begin{array}{|ccccc|c|}
\hline
a_{11} & \ldots & a_{1s} & \ldots & a_{1n} & b_1 \\
a_{21} & \ldots & a_{2s} & \ldots & a_{2n} & b_2 \\
\vdots & & \vdots & & \vdots & \vdots \\
a_{r1} & \ldots & [a_{rs}] & \ldots & a_{rn} & b_r \\
\vdots & & \vdots & & \vdots & \vdots \\
a_{m1} & \ldots & a_{ms} & \ldots & a_{mn} & b_m \\
\hline
c_1 & \ldots & c_s & \ldots & c_n & z \\
\hline
\end{array}$$

Aunque la tabla anterior no lo refleje, todavía estamos considerando que sus primeras m columnas corresponden a los vectores canónicos $e_1, e_2, \ldots, e_m$.

Cuando se hacen las operaciones elementales indicadas —para producir un 1 en la fila r y columna s y ceros en las restantes entradas de dicha columna— al elemento a_{rs} se le llama **pivote** y suele identificarse en la tabla anterior encerrándolo por un círculo, en nuestro caso lo marcamos con [ ].

Observe que las mencionadas operaciones elementales no modifican las columnas $e_1, \ldots, e_{r-1}, e_{r+1}, \ldots, e_m$, porque estas tienen un 0 en la fila r, luego lo que hacen es cambiar de posición el vector e_r, de la columna r (en nuestro caso) a la columna s. Además si $x_s = \frac{b_r}{a_{rs}}$, los valores d_k de la nueva solución y son:

$$d_k = b_k - x_s a_{ks} = b_k - \frac{b_r}{a_{rs}} a_{ks} = b_k - \frac{a_{ks}}{a_{rs}} b_r$$

que es precisamente el resultado de la operación elemental $\frac{-a_{ks}}{a_{rs}} f_r + f_k$ sobre la última columna de la tabla anterior. Así, las operaciones elementales indicadas dan la forma canónica al problema de programación lineal, cuya respectiva solución básica factible corresponde a la solución y obtenida en el caso 2.b. Aunque falta hacer que el coeficiente, c_s, de la función objetivo asociado a la nueva variable básica sea cero, para lo que se requiere aplicar la operación elemental $\frac{-c_s}{a_{rs}} f_r + f_{m+1}$ en la tabla anterior.

Por otra parte, una vez elegida la columna s, donde c_s es negativo y es el menor de los coeficientes de la función objetivo, para seleccionar la fila r que corresponderá al pivote, se suele agregar una última columna a la tabla anterior, con los coeficientes $\frac{b_k}{a_{ks}}$ tales que $a_{ks} > 0$:

$$\begin{array}{|ccccc|c|c|}
\hline
a_{11} & \cdots & a_{1s} & \cdots & a_{1n} & b_1 & \frac{b_1}{a_{1s}} \\
a_{21} & \cdots & a_{2s} & \cdots & a_{2n} & b_2 & \frac{b_2}{a_{2s}} \\
\vdots & & \vdots & & \vdots & \vdots & \vdots \\
a_{r1} & \cdots & [a_{rs}] & \cdots & a_{rn} & b_r & \frac{b_r}{a_{rs}} \\
\vdots & & \vdots & & \vdots & \vdots & \vdots \\
a_{m1} & \cdots & a_{ms} & \cdots & a_{mn} & b_m & \frac{b_m}{a_{ms}} \\
\hline
c_1 & \cdots & c_s & \cdots & c_n & z & \\
\hline
\end{array}$$

Así, la fila r del elemento pivote es la fila del menor cociente, $\frac{b_k}{a_{ks}}$, considerando sólo aquellos en los que $a_{ks} > 0$.

Algoritmo Simplex

El procedimiento conocido como Simplex, para resolver un problema de programación, parte de un problema formulado en la forma canónica, y representado en una tabla como se indicó. En estas tablas se aplican recursivamente los resultados anteriores en la siguiente forma:

P.1: ¿Todos los coeficientes c_i de la función objetivo son positivos o cero? Si esto es cierto, la solución correspondiente a esta forma canónica es la óptima y el algoritmo termina.

Si es falsa, continúa en el paso **P.2**.

P.2: Sea s el índice de la columna tal que

$$c_s < 0 \ \text{ y } \ c_s = \min_{i=1,\ldots,n} c_i.$$

Si todas las entradas a_{ks} de la columna s de A son negativas o cero, entonces el problema de programación lineal no alcanza un valor mínimo. La función objetivo no es acotada inferiormente. El algoritmo termina.

En caso contrario, continúa en el paso **P.3**.

P.3: Si al menos una entrada a_{ks} de la columna s de A es positiva, considere el índice r de la fila de A tal que

$$\frac{b_r}{a_{rs}} = \min\{\frac{b_k}{a_{ks}} | a_{ks} > 0, k = 1, \ldots, m\}$$

Transforme el problema de programación lineal, a otra forma canónica haciendo 1 la entrada a_{rs} de la tabla del simplex y cero las restantes entradas de esta columna, mediante operaciones elementales.

Regrese al paso **P.1**.

Ejemplo 4.5 Sea el programa lineal

$$\min z = x_2 - x_1 + 1, \text{ sujeto a:}$$

$$\begin{cases} -2x_1 & + \; x_2 & \leq & 2 \\ x_1 & - \; 2x_2 & \leq & 2 \\ x_1 & + \; x_2 & \leq & 5 \\ & x_i & \geq & 0 \quad \forall\, i \end{cases}$$

Si la función objetivo se interpreta como una ecuación con parámetro z, $-x_1 + x_2 = z - 1$, el programa lineal se escribe como un sistema de ecuaciones lineales con un parámetro z:

$$\begin{cases} -2x_1 & + \; x_2 + x_3 & & & = & 2 \\ x_1 - 2x_2 & & + \; x_4 & & = & 2 \\ x_1 & + \; x_2 & & + \; x_5 & = & 5 \\ -x_1 & + \; x_2 & & & = & z - 1 \end{cases}$$

con $x_i \geq 0 \; \forall\, i$. La primera solución básica factible se calcula haciendo cero las variables no básicas, $x_1 = x_2 = 0$, por lo que $x_3 = 2$, $x_4 = 2$ y $x_5 = 5$. El valor de la función objetivo en esta solución es $z = 1$ y se obtiene al despejar el parámetro z en la última ecuación después de evaluarla.

Como nota al margen del método simplex, si se hace una representación geométrica de la región de soluciones factibles de este problema, en su formulación inicial (con dos variables), se observará que $(0, 0)$ es uno de sus vértices, al que corresponde esta primera solución básica factible $(0, 0, 2, 2, 5)$, cuando se omiten las variables de holgura x_3, x_4 y x_5.

Primera iteración:

Partiendo de la tabla inicial del simplex:

$$\begin{array}{|rrrrr|c|} -2 & 1 & 1 & 0 & 0 & 2 \\ 1 & -2 & 0 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 & 1 & 5 \\ \hline -1 & 1 & 0 & 0 & 0 & z-1 \\ \hline \end{array}$$

a la pregunta del paso **P.1** respondemos con falso, porque $c_1 = -1$. Y en el paso **P.2** se observa que $a_{21} = a_{31} = 1 > 0$, luego debe efectuarse el paso **P.3**:

En este caso se tiene que $s = 1$ y calculando la columna de cocientes $\frac{b_k}{a_{ks}}$ obtenemos:

$$\begin{array}{|rrrrr|c|c|}
\hline
-2 & 1 & 1 & 0 & 0 & 2 & \\
[1] & -2 & 0 & 1 & 0 & 2 & 2/1 \\
1 & 1 & 0 & 0 & 1 & 5 & 5/1 \\
\hline
-1 & 1 & 0 & 0 & 0 & z-1 & \\
\hline
\end{array}$$

El cociente mínimo aparece en la fila $r = 2$, luego la posición del pivote es la fila $r = 2$ y columna $s = 1$, el cual se señala en la tabla anterior encerrado por paréntesis cuadrados. Aplicando las operaciones elementales

$$2f_2 + f_1, \ -f_2 + f_3, \ f_2 + f_4$$

a la tabla anterior, se transforma en 1 el elemento pivote (en este caso ya se tiene) y en cero las restantes entradas de la columna del pivote:

$$\begin{array}{|rrrrr|c|}
\hline
0 & -3 & 1 & 2 & 0 & 6 \\
1 & -2 & 0 & 1 & 0 & 2 \\
0 & 3 & 0 & -1 & 1 & 3 \\
\hline
0 & -1 & 0 & 1 & 0 & z+1 \\
\hline
\end{array}$$

La última fila de esta tabla corresponde a los nuevos coeficientes de la función objetivo y las nuevas variables básicas son x_3, x_5 y x_1. La nueva solución básica factible, en este caso, es:

$$(x_1, x_2, x_3, x_4, x_5) = (2, 0, 6, 0, 3)$$

y si observamos la representación gráfica de la región de soluciones factibles, se reconoce que $(x_1, x_2) = (2, 0)$ es otro de sus vértices.

Utilizando la nueva formulación canónica, el algoritmo continúa en el paso **P.1**.

Segunda iteración

P.1: Es falso que todos los nuevos coeficientes de la función objetivo, sean no negativos. Entonces se continúa en el paso **P.2**.

P.2: $s = 2$, porque el coeficiente de la función objetivo en la columna 2, es negativo y es el menor de todos.

Además es falso que todas la entras en la columna $s = 2$ sean negativas o cero (no positivas), luego se sigue en el paso **P.3**.

P.3: $r = 3$ puesto que en la columna $s = 2$ de esta matriz todas las entradas son negativas, excepto la de la fila 3.

$$\left|\begin{array}{ccccc|c} 0 & -3 & 1 & 2 & 0 & 6 \\ 1 & -2 & 0 & 1 & 0 & 2 \\ 0 & [3] & 0 & -1 & 1 & 3 \\ \hline 0 & -1 & 0 & 1 & 0 & z+1 \end{array}\right|$$

El pivote es la entrada en la fila 3 y columna 2, como se señala en la última tabla. Luego para obtener una nueva formulación canónica del problema con una mejor solución básica factible, deben efectuarse las operaciones:

$$f_3 + f_1, \quad \tfrac{2}{3} f_3 + f_2, \quad \tfrac{1}{3} f_3 + f_4 \;\; \text{y} \;\; \tfrac{1}{3} f_3$$

con lo que se obtiene:

$$\left|\begin{array}{ccccc|c} 0 & 0 & 1 & 1 & 1 & 9 \\ 1 & 0 & 0 & 1/3 & 2/3 & 4 \\ 0 & 1 & 0 & -1/3 & 1/3 & 1 \\ \hline 0 & 0 & 0 & 2/3 & 1/3 & z+2 \end{array}\right|$$

La nueva solución básica factible es $(4, 1, 9, 0, 0)$, la que corresponde con el vértice $(4, 1)$ de la región de soluciones factibles.

Tercera iteración

Al continuar en el paso **P.1**, con la nueva formulación canónica del problema, se observa que todos los coeficientes en la función objetivo son números no negativos, entonces el mínimo de la función objetivo se alcanza en la correspondiente solución básica factible $(x_1, x_2, x_3, x_4, x_5) = (4, 1, 9, 0, 0)$.

El valor mínimo de la función objetivo se puede obtener al evaluar la solución en la última forma de la función objetivo:

$$z + 2 = \frac{2}{3} x_4 + \frac{1}{3} x_5 = 0, \quad \text{luego} \quad z = -2.$$

o naturalmente, en la forma original de la función objetivo:

$$z = -1x_1 + x_5 + 1 = -1(4) + 1 + 1 = -2.$$

Observación: En la fundamentación del método simplex no se han considerado los programas lineales que tengan algún lado derecho de una restricción igual a cero. Es decir evitamos las soluciones degeneradas. Sin embargo, el algoritmo también logra su objetivo en estos casos, aunque teóricamente si hay soluciones degeneradas no se puede garantizar la convergencia del método. El ejercicio 2 en la página 140 permite comprobar que el simplex opera bien en presencia de soluciones degeneradas.

Cuando la formulación inicial del problema de programación lineal no tiene la forma canónica, no es claro qué operaciones elementales deben realizarse para transformalo en uno con la forma canónica. Y sin la forma canónica no podemos aplicar el método simplex. La siguiente técnica resuelve este problema.

4.3 Variables artificiales

Las variables artificiales son un artificio que permiten utilizar el mismo simplex, para encontrar una primera forma canónica a un problema de programación lineal, siempre que el problema tenga soluciones factibles. Enseguida se ilustra el procedimiento.

4.3.1 Un ejemplo

Ejemplo 4.6 Considere el programa lineal

$$\max z = -2x_1 + x_2 + 1, \text{ sujeto a:}$$

$$\begin{cases} \begin{array}{rcrcr} 2x_1 & + & x_2 & \geq & 4 \\ -x_1 & + & x_2 & \leq & 4 \\ -3x_1 & + & x_2 & \geq & -15 \\ x_1 & & & \leq & 7 \\ & & x_1, \; x_2 & \geq & 0 \end{array} \end{cases}$$

Como en todo problema de programación lineal, primero se efectúan las operaciones elementales $-f_j$ necesarias para que los coeficientes b_j sean no negativos. Luego se convierten las inecuaciones en ecuaciones utilizando variables de holgura.

$$(1)\ \left\{\begin{array}{llllllll} 2x_1 & + x_2 & - x_3 & & & & = & 4 \\ -x_1 & + x_2 & & + x_4 & & & = & 4 \\ 3x_1 & - x_2 & & & + x_5 & & = & 15 \\ x_1 & & & & & + x_6 & = & 7 \\ x_i & \geq 0 & \forall\, i & & & & & \end{array}\right.$$

Observe que cuando se tienen inecuaciones con $\geq$ y a la derecha el número es positivo, la introducción de la variable de holgura no produce un vector canónico, luego la formulación resultante no siempre tendrá la forma canónica.

En esta situación se agregan artificialmente variables adicionales suficientes para que aparezca la base canónica, entre las columnas de la matriz aumentada del sistema modificado. En el ejemplo sólo es necesario agregar la variable artificial x_7 en la ecuación 1.

$$(2)\ \left\{\begin{array}{lllllllll} 2x_1 & + x_2 & - x_3 & & & & + x_7 & = & 4 \\ -x_1 & + x_2 & & + x_4 & & & & = & 4 \\ 3x_1 & - x_2 & & & + x_5 & & & = & 15 \\ x_1 & & & & & + x_6 & & = & 7 \\ x_i & \geq 0 & \forall\, i & & & & & & \end{array}\right.$$

Con el conjunto de restricciones (2), denotado como $Ax = b$, consideramos un nuevo problema de programación lineal:

$$\text{Minimizar } \ w = x_7, \text{ sujeto a } Ax = b \text{ y } x \geq 0$$

donde, en general, la función objetivo será minimizar la suma de las variables artificiales. Nos referiremos a este problema como el problema ampliado con variables artificiales o simplemente como el problema con variables artificiales.

Aplicamos el método simplex al nuevo problema, sin olvidar la función objetivo del problema original, a fin de actualizarla con las tranformaciones que se produzcan sobre el problema ampliado con variables artificiales. Así la primera tabla del simplex será:

2	1	-1	0	0	0	1	4
-1	1	0	1	0	0	0	4
3	-1	0	0	1	0	0	15
1	0	0	0	0	1	0	7
2	-1	0	0	0	0	0	$-z+1$
0	0	0	0	0	0	1	w

Sin embargo, observe que el coeficiente de la función objetivo w en la columna 7 (correspondiente a una variable básica) no es cero, luego todavía la formulación del problema con variables artificiales no tiene la forma canónica. Para corregir esto basta con efectuar la operación elemental $-f_1 + f_6$:

2	1	-1	0	0	0	1	4
-1	1	0	1	0	0	0	4
3	-1	0	0	1	0	0	15
1	0	0	0	0	1	0	7
2	-1	0	0	0	0	0	$-z+1$
-2	-1	1	0	0	0	0	$w-4$

Ahora sí se puede aplicar el simplex al problema con variables artificiales. Y se observa que la columna con coeficiente menor y negativo en la función objetivo (del problema con variables artificiales) es la columna $s = 1$. Por otra parte, según la columna de cocientes en la tabla siguiente, se tiene que fila del pivote es la 1.

[2]	1	-1	0	0	0	1	4	2
-1	1	0	1	0	0	0	4	
3	-1	0	0	1	0	0	15	5
1	0	0	0	0	1	0	7	7
2	-1	0	0	0	0	0	$-z+1$	
-2	-1	1	0	0	0	0	$w-4$	

Haciendo las operaciones elementales necesarias para convertir en 1 el pivote y en cero las restantes entradas de la primera columna, obtenemos:

1	0.5	-0.5	0	0	0	0.5	2
0	1.5	-0.5	1	0	0	0.5	6
0	-2.5	1.5	0	1	0	-1.5	9
0	-0.5	0.5	0	0	1	-0.5	5
0	-2.0	1.0	0	0	0	-1.0	$-z-3$
0	0.0	0.0	0	0	0	1.0	w

Todos los coeficientes en la última fila son no negativos por lo que el método simplex ha concluido con una solución óptima para el problema ampliado con variables artificiales. La solución óptima es $(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = (2, 0, 0, 6, 9, 5, 0)$ y como la variable artificial x_7 es no básica, su valor es cero, lo cual permite descartarla, para reducirla a la siguiente solución básica factible del problema original: $(x_1, x_2, x_3, x_4, x_5, x_6) = (2, 0, 0, 6, 9, 5)$.

Más específicamente, si se elimina la columna 7 y la fila 6 de la última tabla, se reconoce que lo que se ha hecho es transformar el problema original mediante operaciones elementales en uno que tiene la forma canónica.

1	0.5	−0.5	0	0	0	2
0	1.5	−0.5	1	0	0	6
0	−2.5	1.5	0	1	0	9
0	−0.5	0.5	0	0	1	5
0	−2.0	1.0	0	0	0	$-z-3$

Así, el programa lineal canónico correspondiente a esta tabla es equivalente al programa original:

$$\min\ -z = 2x_1 - x_2 - 1 \ \text{ sujeto a las restricciones (1),}$$

lo cual permite resolverlo mediante el simplex.

En esta nueva aplicación del simplex, la iteración 1 ubica el pivote en la columna $s = 2$ y los cocientes correspondientes son:

1	[0.5]	−0.5	0	0	0	2	$2/(0.5) = 4$
0	1.5	−0.5	1	0	0	6	$6/(1.5) = 4$
0	−2.5	1.5	0	1	0	9	
0	−0.5	0.5	0	0	1	5	
0	−2.0	1.0	0	0	0	$-z-3$	

Luego el pivote puede ser tanto el elemento de (1,2) como el (2,2). Si se elige el segundo, puede verificarse que conduce más rápido a la solución, nosotros escogemos como pivote el elemento de la fila 1 y columna 2 y efectuamos las operaciones elementales necesarias para hacer 1 en la entrada (1,2) y 0 en las restantes posiciones de la columna 2.

$$
\begin{array}{|rrrrrr|c|}
\hline
2 & 1 & -1 & 0 & 0 & 0 & 4 \\
-3 & 0 & [1] & 1 & 0 & 0 & 0 \\
5 & 0 & -1 & 0 & 1 & 0 & 19 \\
1 & 0 & 0 & 0 & 0 & 1 & 7 \\
\hline
4 & 0 & -1 & 0 & 0 & 0 & -z+5 \\
\hline
\end{array}
$$

Con la elección del elemento (2,3) como pivote una nueva iteración produce la tabla:

$$
\begin{array}{|rrrrrr|c|}
\hline
-1 & 1 & 0 & 1 & 0 & 0 & 4 \\
-3 & 0 & 1 & 1 & 0 & 0 & 0 \\
2 & 0 & 0 & 1 & 1 & 0 & 19 \\
1 & 0 & 0 & 0 & 0 & 1 & 7 \\
\hline
1 & 0 & 0 & 1 & 0 & 0 & -z+5 \\
\hline
\end{array}
$$

Y obtenemos una formulación canónica para el problema original con todos los coeficientes de la función objetivo no negativos, por lo que hemos encontrado una solución óptima:

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (0, 4, 0, 0, 19, 7)$$

con valor en la función objetivo $z = 5$.

Observe que el valor mínimo es $-z = -5$ y el valor máximo de la función objetivo es $z = 5$.

Por otra parte, la solución óptima y la solución básica factible en el paso anterior son soluciones degeneradas, porque en ambos casos x_3 es una variable básica y su valor es 0.

4.3.2 Formulación de la técnica de las variables artificiales

Dado el programa lineal min $z = \sum_{i=1}^{n} c_i x_i + z_0$ sujeto a

$$
(1)\left\{
\begin{array}{ccccccc}
a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & \vdots & \vdots & \vdots \\
a_{m1}x_1 & + & \cdots & + & a_{m1}x_n & = & b_m
\end{array}
\right.
$$

con $x_i \geq 0 \ \forall \ i$ y $b_j \geq 0 \ \forall \ j$, la técnica de las variables artificiales, para generar una primera forma canónica de este problema, consiste en aplicar los siguientes pasos:

1. En cada ecuación se suma una variable $y_j \geq 0$, llamada variable artificial. Algunas veces, como ocurre en el ejemplo anterior, hay variables que pueden servir como básicas, lo que hace innecesario sumar una variable artificial en las ecuaciones correspondientes. El sistema (1) así transformado es canónico teniendo como variables básicas las variables artificiales y, eventualmente, algunas de las variables de la lista $x_1, \ldots, x_n$. Denotamos este sistema como (2):

$$(2) \left\{ \begin{array}{llllll} a_{11}x_1 & + \cdots + & a_{1n}x_n & + y_1 & & = & b_1 \\ a_{21}x_1 & + \cdots + & a_{2n}x_n & & + y_2 & = & b_2 \\ \vdots & & \vdots & & & & \vdots \\ a_{m1}x_1 & + \cdots + & a_{m1}x_n & & & + y_m = & b_m \end{array} \right.$$

2. Si $y_1, \ldots, y_p$ son las variables artificiales con $p \leq m$, se considera el nuevo problema de programación lineal:

$$\text{Minimizar} \quad w = y_1 + \cdots + y_p,$$

$$\text{sujeto a las restriciones (2) y} \quad x \geq 0, \quad y \geq 0.$$

   Como la nueva función objetivo depende de las variables básicas, deben aplicarse las operaciones elementales necesarias para hacer cero todos los coeficientes asociados a variables básicas en esta función objetivo. Con esto el programa lineal min w sujeto a las restricciones (2) es canónico y se aplica el método simplex.

3. Una vez obtenida la solución óptima del problema modificado y si en ésta, las variables artificiales son todas variables no básicas. Se eliminan las columnas correspondientes a variables artificiales y la última fila, con lo que se obtiene una formulación canónica equivalente al problema de programación original.

El procedimiento se basa en los siguientes resultados:

- Toda solución factible del problema con variables artificiales en la que las variables artificiales valgan cero, cuando estas se omiten la reducen a una solución factible del problema original.

- E inversamente, toda solución factible del problema original es una solución factible al problema con variables artificiales, tomando como cero las variables artificiales.

- Para toda solución factible x del problema con variables artificiales, $w(x) \geq 0$.

  Además, toda solución factible x del problema original agregando las variables artificiales con valor cero, produce una solución x^* factible para el problema con variables artificiales, que es óptima porque $w(x^*) = 0$.

De esto se deducen los dos principales resultados:

- Una solución básica factible óptima, del problema con variables artificiales, cuyas variables artificiales sean no básicas, se reduce a una solución básica factible del problema original.

  Y la formulación canónica al problema con variables artificiales, del cual se obtuvo esta solución óptima, se reduce a una formulación canónica del problema original, simplemente eliminando las variables artificiales.

- Si la solución óptima x^* del problema con variables artificiales, incluye como variable básica alguna variable artificial con valor no cero, entonces $w(x^*) > 0$ y como es óptima entonces el conjunto de soluciones factibles del problema original debe ser $\emptyset$.

4.4 Ejercicios

1. Identifique la región de soluciones factibles del sistema de restricciones

$$\begin{cases} \frac{1}{2}x_1 + \frac{1}{4}x_2 \leq 30 \\ x_1 + 5x_2 \leq 200 \\ x_1 + x_2 \geq 50 \\ x_1, \ x_2 \geq 0 \end{cases}$$

2. Sea el programa lineal max $3x_1 - x_2$ sujeto a

$$\begin{array}{rcrcl} x_1 & - & x_2 & \leq & 3 \\ 2x_1 & & & \leq & x_2 \\ x_1 & + & x_2 & \geq & 12 \\ & & x_2 & \leq & 10 \\ & x_1, & x_2 & \geq & 0 \end{array}$$

1. Resuelva el programa lineal.
2. ¿Si se quita la restricción $x_1 - x_2 \leq 3$, se modifica la solución? Explique.
3. ¿Si se quita la restricción $x_2 \leq 10$, cuál es la respuesta a la pregunta 1?

3. Resuelva el programa lineal max $3x_1 + 2x_2$ sujeto a

$$\begin{array}{rcrcl} x_1 & & & \leq & 50 \\ & & x_2 & \leq & 100 \\ 2x_1 & + & 4x_2 & \leq & 400 \\ & x_1, & x_2 & \geq & 0 \end{array}$$

4. Considere el siguiente conjunto de restricciones lineales:

$$\begin{array}{rcrcl} -3x_1 & + & 2x_2 & \geq & 12 \\ 5x_1 & - & x_2 & \leq & 10 \\ 3x_1 & + & 2x_2 & \geq & 24 \end{array}$$

$x_1 \geq 0$ y $x_2 \geq 0$.

a) Grafique la región de soluciones factibles, indicando la intersección de las rectas con los ejes.

b) Resuelva el programa lineal: Max $z = -9x_1 + 4x_2$, sujeto a las restricciones dadas, utilizando el método gráfico.

5. Una compañía tiene dos fábricas ensambladoras de automóviles de un mismo modelo, para abastecer dos clientes. Los costos de transporte en dólares por cada automóvil, las demandas de los clientes así como las reservas de automóviles, se dan en el cuadro siguiente:

	Fábrica		
Cliente	A	B	Demanda
1	30	36	400
2	25	30	300
Reservas	a	b	

donde $a > 0$ y $b > 0$. La compañía desea abastecer a sus clientes con el mínimo costo de transporte.

1. Dibuje una red para representar el problema de transporte.
2. Formule el programa lineal correspondiente suponiendo que la compañía puede dejarse un sobrante. Resuelva el programa por el método gráfico y dé las condiciones sobre los parámetros a y b necesarias para que haya solución.

6. Considere el programa lineal siguiente: Min $z = x_1 + x_2 + 4x_3 + 7x_4$ sujeto a las restricciones $x_1 + x_2 + 5x_3 + 2x_4 = 8$, $2x_1 + x_2 + 8x_3 = 14$ y $x_i \geq 0$, $i = 1, 2, 3, 4$.

1. Mediante operaciones elementales, obtenga un programa lineal canónico con variables básicas x_1 y x_2, equivalente al propuesto.
2. Resuelva el programa lineal.

7. Sea el programa lineal min $z = -x_1 - 2x_4 + x_5$ sujeto a las restricciones $x_1 + x_3 + 6x_4 + 3x_5 = 2$, $-3x_1 + x_2 + 3x_4 + x_5 = 2$ y $x_i \geq 0$, $i = 1, 2, 3, 4, 5$. Usando solo operaciones elementales, halle un programa lineal canónico equivalente, en el cual se lea directamente la solución óptima. Calcule esta solución.

8. Resuelva el programa lineal max $z = 10x_1 + 6x_2 - 8x_3$ sujeto a las restricciones $5x_1 - 2x_2 + 6x_3 \leq 20$, $10x_1 + 4x_2 - 6x_3 \leq 30$ y $x_i \geq 0$, $i = 1, 2, 3$.

9. Una compañía cafetalera compra lotes de mezclas de granos de café y después los clasifica en de primera, regulares, más una porción inservible. La compañía necesita al menos 280 ton (toneladas) de primera y 200 ton de clase regular. La compañía puede comprar a los distribuidores A y B, cualquier cantidad de café sin clasificar. Muestras proporcionadas por los distribuidores tienen los siguientes porcentajes de grano de primera, regular e inservible:

Distribuidor	De Primera	Regular	Inservible
A	20%	50%	30%
B	40%	20%	40%

Si A cobra \$125 por tonelada y B cobra \$200 por tonelada, formule el modelo de programación cuya solución provee la cantidad de café que compra la compañía a cada distribuidor para cubrir sus necesidades a un costo mínimo. Resuelva gráficamente el problema lineal.

10. Un granjero tiene 100 acres para plantar dos cultivos A y B. La semilla para el cultivo A cuesta \$4 por acre y la semilla para el cultivo B, \$6 por acre. El costo total de la mano de obra será de \$20 por acre para el cultivo A y \$10 por acre para el cultivo B. El granjero confía obtener un ingreso de \$110 por acre del cultivo A; y del cultivo B, \$150 por acre. Si el granjero no desea gastar más de \$480 para la semilla y \$1400 para la mano de obra , ¿cuántos acres de cada uno de los cultivos debe plantar para obtener: (a) máximo ingreso y (b) máxima ganancia (use el método gráfico).

11. Una compañía fabrica dos productos A y B. Para cada producto es necesario emplear dos máquinas diferentes, X y Y. Para fabricar una unidad del producto A, la máquina X debe usarse 1/2 hora, y la máquina Y, 1 hora. Para fabricar una unidad del producto B, es necesario utilizar tanto la máquina X como la Y, 2 horas. La ganancia en el producto A es \$20 por unidad, y la ganancia en el B es \$50 por unidad. Si la máquina X puede usarse durante 8 horas al día y la máquina Y durante 12 horas al día, determine cuántas unidades de cada producto deberán fabricarse cada día para maximizar la ganancia.

12. En una feria de un día, un hombre tiene un puesto donde venderá bolsas de maní y bolsas de dulces. Tiene \$100 disponibles para adquirir su mercancía, que costará \$0.10 la bolsa de maní y \$0.20 la bolsa de dulces. Pretende vender el maní a \$0.15 y los dulces a \$0.26 la bolsa. Puede acomodar en su puesto 500 bolsas de maní y 400 bolsas de dulces. De experiencias pasadas sabe que no venderá más de un total de 700 bolsas.

1. Formule los modelos de programación cuya solución dan las ventas máximas y la utilidad máxima, donde ésta se define como la diferencia de las ventas menos el costo de la mercancía.
2. Represente gráficamente la región de soluciones factibles.
3. Encuentre el número de bolsas de cada artículo que deberá tener disponibles para lograr: (a) ventas máximas y (b) utilidad máxima. ¿Cuál es el monto de las ventas y utilidades máximas?

13. Resuelva el programa lineal : max $2x_1 + x_2 + x_3$ sujeto a

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & \leq & 60 \\ -4x_1 & + & 2x_2 & + & x_3 & \geq & 52 \\ 2x_1 & & & + & x_3 & \geq & 40 \\ & & & & x_i & \geq & 0 \quad \forall\, i \end{array}$$

14. Resuelva el programa lineal : min $x_1 + x_2 + x_3$ sujeto a

$$\begin{array}{ccccccc} -x_1 & + & 2x_2 & + & x_3 & \leq & 1 \\ -x_1 & + & & & 2x_3 & \geq & 4 \\ x_1 & - & x_2 & + & 2x_3 & = & 4 \\ & & & & x_i & \geq & 0 \quad \forall\, i \end{array}$$

15. Resuelva el programa lineal por dos métodos distintos:
min $2x_1 + 2x_2 - 5x_3$ sujeto a

$$\begin{array}{ccccccc} 3x_1 & + & 2x_2 & - & 4x_3 & = & 1 \\ x_1 & - & x_2 & + & 3x_3 & = & 2 \\ & & & & x_i & \geq & 0 \quad \forall\, i \end{array}$$

16. Resuelva el programa lineal correspondiente al modelo de transporte (ver la página 112).

Capítulo 5

$I\!R^n$: Geometría de vectores

En capítulos anteriores se trabajó con las matrices $M(m, n, I\!R)$, sus operaciones y conceptos relacionados; momento en que se llamó vector fila a las matrices con una sola fila y n columnas, y vector columna a las que tienen una sola columna y n filas. Ahora se estudiarán estos mismos objetos como elementos del espacio $I\!R^n$, sin hacer mayor distinción porque sean vectores fila o columna, y principalmente, para reconocer las interpretaciones geométricas que se les asocia.

Esta nueva aproximación a los vectores de $I\!R^n$, parte del reconocimiento de las operaciones de igualdad entre vectores, suma de vectores, multiplicación de un escalar por un vector, y sus propiedades, todo lo anterior visto ya al observar que

$$I\!R^n = M(1, n, I\!R) \ \text{ ó } \ I\!R^n = M(n, 1, I\!R).$$

Así también se reconocen, los conceptos de combinación lineal e independencia lineal de vectores y sus distintas caracterizaciones, desarrolladas con anterioridad.

5.1 Representación geométrica de vectores

A partir de la representación de $I\!R$ como una recta numérica, los elementos (a_1, a_2) de $I\!R^2$ y (a_1, a_2, a_3) de $I\!R^3$ se asocian con puntos de un plano y puntos del espacio tridimensional, en la forma que es bien conocida y se ilustra seguidamente.

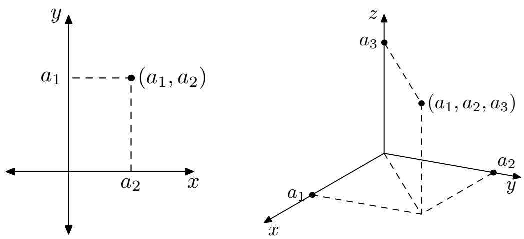


Figura 5.1: Representación de (a_1, a_2) y (a_1, a_2, a_3) como puntos.

En estos gráficos el sistema de coordenadas está determinado por dos rectas numéricas en un plano, dispuestas perpendicularmente, (o tres rectas numéricas en el espacio, mútuamente perpendiculares). El punto de intersección representa a $(0, 0)$ (y $(0, 0, 0)$ respectivamente) y una vez que se elige cuáles de estas rectas identifican el eje X, eje Y (y eje Z), cada elemento $(a_1, a_2) \in I\!R^2$ o $((a_1, a_2, a_3) en I\!R^4)$ se asocia con el punto determinado por la coordenada a_1 sobre el eje X, a_2 sobre el eje Y (y a_3 sobre el eje Z, para el caso de $I\!R^3$), como se muestra en la figura 5.1.

Estas ideas se extienden a $I\!R^n$ y se piensa en cada elemento $A = (a_1, a_2, \ldots, a_n)$ como en un punto de un espacio con n dimensiones, donde cada componente a_i de A corresponde a la i-ésima coordenada medida sobre el i-ésimo eje de un sistema de coordenadas rectangulares, o cartesianas, con n ejes mútuamente perpendiculares. No podemos visualizar n ejes mútuamente perpendiculares si $n > 3$, pero esta idea, imprecisa ahora, se formalizará más adelante echando mano a recursos púramente algebraicos, que permitirán intuir algunas características geométricas de los objetos de $I\!R^n$. Todo esto, teniendo siempre como referencia el mejor conocimiento de los espacios $I\!R^2$ y $I\!R^3$, que el poder de la visualización nos da.

5.1.1 Interpretación geométrica de flecha para vectores

Adicionalmente a la interpretación de "punto" que se ha dado a los elementos de $I\!R^n$, se puede asociar a cada uno de ellos, una nueva idea geométrica, con total independencia a la anterior (lo que no significa que no estén relacionadas). Considere, primero, algunos ejemplos:

Ejemplo 5.1 El vector $(2, -3) \in I\!R^2$ se interpreta como el **desplazamiento resultante** de moverse dos unidades en la dirección positiva del eje X y 3 unidades en la dirección negativa del eje Y.

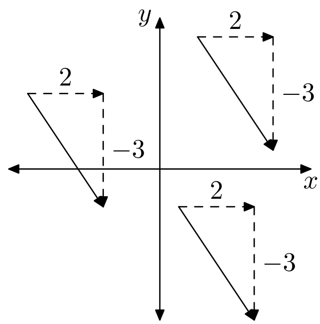


Figura 5.2: Tres posibles representaciones del vector $(2, -3)$ como flecha.

En esta nueva interpretación geométrica, las coordenadas del vector $(2, -3)$ sólo describen un desplazamiento: dos unidades en la dirección positiva del eje X y 3 unidades en la dirección negativa del eje Y, sin indicar el punto donde se origina el movimiento. **Se trata de una nueva interpretación geométrica para** $(2, -3)$**, que también depende del sistema de coordenadas, pero esta vez resumiendo la idea de "flecha": una identidad caracterizada por su magnitud y dirección y que no tiene ubicación**.

Ejemplo 5.2 Similarmente, en $I\!\!R^3$ el vector $(-3, 2, -4,)$ se puede visualizar como una **flecha** o desplazamiento resultante de moverse tres unidades en la dirección negativa del eje X, 2 unidades en la dirección positiva de Y y 4 unidades en la dirección negativa del eje Z.

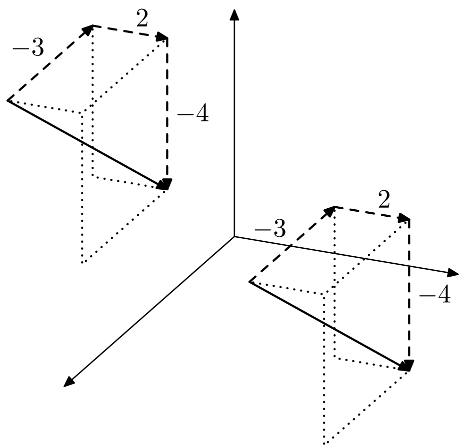


Figura 5.3: dos posibles formas de representar el vector $(-3, 2, -4)$ como flecha.

Y de nuevo, el desplazamiento total descrito por las coordenadas del vector $(-3, 2, -4)$, se visualiza como una flecha de magnitud y dirección que determinan las coordenadas -3, 2 y -4, pero sin especificar el punto inicial del movimiento.

En el caso general de $I\!\!R^n$, esta nueva representación geométrica de sus elementos, encuentra la misma dificultad de visualización que la de punto. Sin embargo, si se puede pensar en n ejes mútuamente perpendiculares que doten a $I\!\!R^n$ de un sistema de coordenadas rectangulares, es posible intuir el desplazamiento resultante de n movimientos en cada una de las direcciones de los ejes, lo cual describe un desplazamiento total con determinada magnitud y dirección e independientemente del punto donde se inicia el movimiento.

En resumen, si $a = (a_1, a_2, \cdots, a_n) \in I\!\!R^n$, se puede dar a este elemento una interpretación geométrica de punto — el que determinan las coordenadas a_i— o de **flecha**: la determinada por los desplazamientos paralelos a los ejes que indican las coordenadas a_i. Ambas interpretaciones geométricas son igualmente válidas y aplicables en cualquier caso, sin embargo, algunas veces conviene tener en mente una de las dos interpretaciones, preferentemente, por la cual se hará el siguiente convenio:

Notación: ¿puntos? ¿flechas?

- Cuando se escribe $A = (a_1, a_2, \cdots, a_n)$ o $A \in I\!R^n$, utilizando letras mayúsculas, se atribuye a A una interpretación geométrica de punto, preferentemente (no exclusivamente).

- Si se utilizan letras minúsculas "techadas" con una flecha como $\vec{a} = (a_1, a_2, \cdots, a_n)$ o $\vec{a} \in I\!R^n$, entonces, geométricamente, se interpreta a $\vec{a}$ como una flecha. También preferentemente y no exclusivamente.

Por otra parte, en este material inicialmente se utilizarán nombres de vectores techados con flechas, para reforzar su nueva interpretación geométrica, como $\vec{x}$ o $\vec{a}$, sin embargo, más adelante se escribirá simplemente x o a con exactamente el mismo significado.

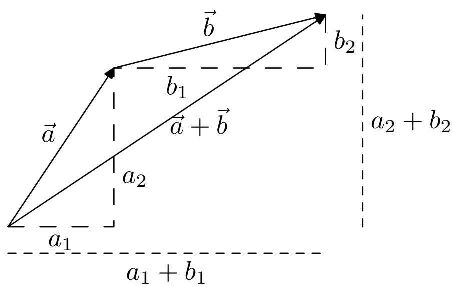


Figura 5.4: $\vec{a} + \vec{b} = (a_1 + b_1, a_2 + b_2)$.

5.1.2 Interpretación geométrica de la suma de vectores

Dados dos vectores $\vec{a}$ y $\vec{b}$ en $I\!R^n$, geométricamente, el vector $\vec{a}+\vec{b}$ corresponde a una nueva flecha que resume los dos desplazamientos totales: el determinado por las coordenadas de $\vec{a}$ seguido del correspondiente a las coordenadas de $\vec{b}$, o inversamente.

Por ejemplo, en $I\!R^2$, si $\vec{a} = (a_1, a_2)$ y $\vec{b} = (b_1, b_2)$, entonces $\vec{a} + \vec{b} = (a_1 + b_1, a_2 + b_2)$ y refleja el desplazamiento total de los dos vectores o flechas $\vec{a}$ y $\vec{b}$ como se ilustra en los gráficos 5.4 y 5.5.

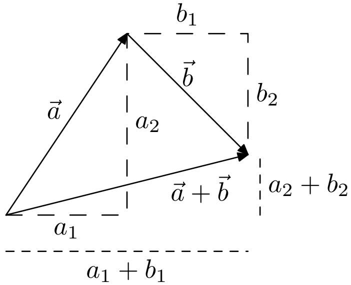


Figura 5.5: $\vec{a} + \vec{b} = (a_1 + b_1, a_2 + b_2)$, con $b_2 < 0$.

En general, si en $I\!R^n$ se representa el vector $\vec{a}$ mediante una flecha y al vector $\vec{b}$ también mediante una flecha pero que comienza en el punto terminal de $\vec{a}$, entonces $\vec{a}+\vec{b}$ se asocia al desplazamiento total resultante de efectuar los desplazamientos determinados por ambas flechas,

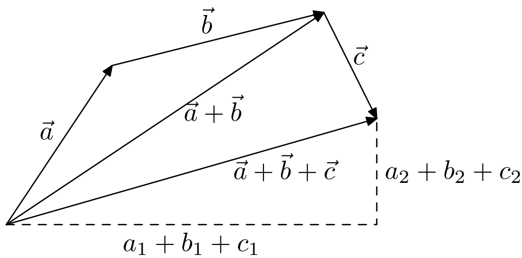


Figura 5.6: Ilustración de la suma de tres vectores

y representado por una nueva flecha que se origina en el punto inicial de $\vec{a}$ y termina en el punto final de $\vec{b}$. La situación se generaliza al caso de la suma de más dos vectores, como se muestra en la figura 5.6.

Si los dibujos de las flechas, no consideran los sistemas de coordenadas rectangulares, la representación geométrica resulta más clara y aplicable a todos los espacios $I\!R^n$, observe la figura 5.7.

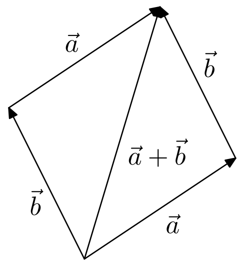


Figura 5.7: Paralelogramo de lados $\vec{a}$ y $\vec{b}$ y diagonal $\vec{a}+\vec{b}=\vec{b}+\vec{a}$.

Dado que $\vec{a}+\vec{b}=\vec{b}+\vec{a}$, la interpretación geométrica para $\vec{a}+\vec{b}$ se asocia con un paralelogramo de lados sucesivos $\vec{a}$ y $\vec{b}$ cuya diagonal une el punto inicial de $\vec{a}$ con el punto terminal de $\vec{b}$, o el punto inicial de $\vec{b}$ con el punto terminal de $\vec{a}$.

5.1.3 Interpretación geométrica del producto de un escalar por un vector

Cuando un vector $\vec{a} \in I\!R^n$, se multiplica por un escalar $t > 0$ el nuevo vector $t\vec{a}$ tiene la misma dirección que $\vec{a}$ y su magnitud aumenta o disminuye en un factor de t. Sin embargo si el escalar es negativo, $r < 0$, entonces la dirección de $r\vec{a}$ es contraria a la de $\vec{a}$.

En la figura 5.8, se representan los vectores $\vec{a}$, $t\vec{a}$ y $r\vec{a}$, suponiendo que $t > 2$ y $-1 < r < 0$. Por ello $t\vec{a}$ tiene la misma dirección que $\vec{a}$ y su magnitud es mayor que la de $2\vec{a}$), y $r\vec{a}$ tiene dirección contraria a $\vec{a}$ y su magnitud es menor que la de $\vec{a}$.

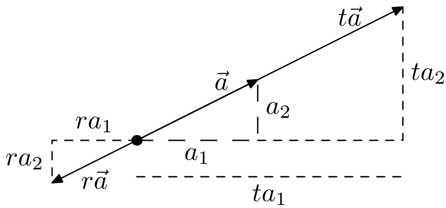


Figura 5.8: $\vec{a}$, $t\vec{a}$ y $r\vec{a}$, con $t > 2$ y $-1 < r < 0$.

Observe que los vectores $\vec{a}$, $t\vec{a}$ y $r\vec{a}$ se alinean sobre una misma recta dado que la multiplicación por los factores t y r no cambia la inclinación (pendiente) de los vectores:

$$\frac{a_2}{a_1} = \frac{ta_2}{ta_1} = \frac{ra_2}{ra_1}.$$

Definición 5.1 (Vectores paralelos)

Dos vectores $\vec{a}$ y $\vec{b}$ no nulos, se dicen paralelos si existe $t \in I\!R$ tal que $\vec{a} = t\vec{b}$. Cuando $t > 0$, $\vec{a}$ y $\vec{b}$ tienen la misma dirección y si $t < 0$, tienen dirección contraria.

5.1.4 Relación entre flechas y puntos

Aunque la interpretación de flecha dada a los vectores se ha establecido de manera independiente a la idea de "punto", estas dos interpretaciones geométricas se relacionan de manera muy conveniente, a través del concepto de **flecha localizada**.

Definición 5.2 (Flechas localizadas $\overrightarrow{AB}$)

Dados A y B en $I\!R^n$, a los que se les atribuye una interpretación geométrica de punto, se llama flecha localizada de A a B al vector $\overrightarrow{AB} = B - A \in I\!R^n$ y que, geométricamente, se asocia con la única flecha que se origina en el punto A y termina en B.

Observe que como $B - A \in I\!R^n$, $B - A$, al igual que cualquier vector en $I\!R^n$, puede ser interpretado como un punto o como una flecha sin ubicación. Sin embargo, a $\overrightarrow{AB}$ sólo le damos una interpretación geométrica: la de la flecha que comienza en A y termina en B. Naturalmente, esta única flecha que se asocia con $\overrightarrow{AB}$ es una de las infinitas que se le pueden asociar a $B-A$ (o que refleja el desplazamiento descrito por las coordenadas de $B-A$). Estas ideas se discuten seguidamente:

Sean $O=(0,0,\cdots,0)\in I\!R^n$ el origen del sistema de coordenadas rectangulares de $I\!R^n$ y A, B puntos de $I\!R^n$. Observe que según la definición de flecha localizada se tiene que:

$$\overrightarrow{OA}=A \quad \text{y} \quad \overrightarrow{OB}=B.$$

Algebraicamente, lo anterior es claro, pero además geométricamente significa que la flecha localizada de O a A, corresponde a la interpretación de A como flecha (cuando A se representa por los desplazamientos determinados por sus coordenadas, cuando estos se originan en O). Lo mismo se tiene para $\overrightarrow{OB}$ y la representación de B como flecha.

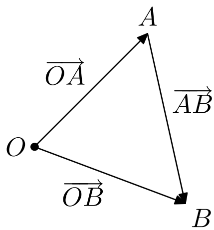


Figura 5.9: Representación de $\overrightarrow{AB}$ y los puntos A y B identificados también como flechas desde O.

Por lo tanto, de la interpretación geométrica de la suma de vectores, se obtiene que:

$$\begin{array}{rrcl} & \overrightarrow{OA}+\overrightarrow{AB} & = & \overrightarrow{OB} \\ \Longrightarrow & A+\overrightarrow{AB} & = & B \\ \Longrightarrow & \overrightarrow{AB} & = & B-A. \end{array}$$

lo que hace consistente la definición de flecha localizada dada.

En el siguiente gráfico, observe la representación geométrica del vector $B - A$, cuando se le representa como una flecha que se inicia en el origen y cuando se identifica como el vector localizado $\overrightarrow{AB}$.

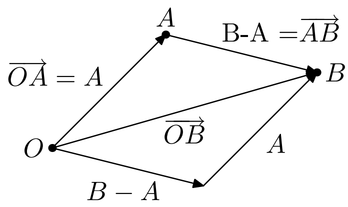


Figura 5.10: $B - A$ representado como la flecha $\overrightarrow{AB}$ y como una flecha que se origina en O.

La flecha localizada $\overrightarrow{AB}$ es un caso particular de flecha que corresponde al vector $B - A$, de entre una infinidad de posibilidades. Por ejemplo, para cualquier $H \in I\!R^n$, si $C = A + H$ y $D = B + H$ entonces la flecha localizada $\overrightarrow{CD}$ también corresponde al vector $B - A$, puesto que

$$\overrightarrow{CD} = D - C = (B + H) - (A + H) = B - A.$$

Ejemplo 5.3 Suponga que A, B, C y D son los vértices de un cuadrilátero de manera que $\overrightarrow{AB}$ y $\overrightarrow{AD}$ son lados adyacentes —y por lo tanto $\overrightarrow{AB}$ no es paralelo a $\overrightarrow{AD}$— como se muestra en la figura 5.11.

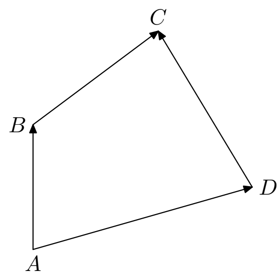


Figura 5.11: Cuadrilátero A, B, C, D.

Demuestre que si $\overrightarrow{BC}$ es paralelo a $\overrightarrow{AD}$ y $\overrightarrow{DC}$ es paralelo a $\overrightarrow{AB}$ entonces:

a) el cuadrilátero $ABCD$ tiene sus lados opuestos iguales.

b) $\overrightarrow{AB} + \frac{1}{2}\overrightarrow{BD} = \frac{1}{2}\overrightarrow{AC}$.

Demostración: Parte a). Como $\overrightarrow{BC}$ es paralelo a $\overrightarrow{AD}$ y $\overrightarrow{DC}$ paralelo a $\overrightarrow{AB}$ entonces existen t y s en $\mathbb{R}$ no nulos tales que:

$$\overrightarrow{BC} = t\overrightarrow{AD} \tag{5.1}$$
$$\overrightarrow{DC} = s\overrightarrow{AB} \tag{5.2}$$

Además, de la interpretación geométrica de la suma de vectores se tiene que:

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AD} + \overrightarrow{DC}$$

Luego, si se sustituye $\overrightarrow{BC}$ y $\overrightarrow{DC}$ según (5.1) y (5.2) se obtiene que $\overrightarrow{AB} + t\overrightarrow{AD} = \overrightarrow{AD} + s\overrightarrow{AB}$ y finalmente:

$$\begin{aligned} \overrightarrow{AB} - s\overrightarrow{AB} &= \overrightarrow{AD} - t\overrightarrow{AD} \\ (1-s)\overrightarrow{AB} &= (1-t)\overrightarrow{AD} \end{aligned}$$

Ahora, si $s \neq 1$ entonces $\overrightarrow{AB} = \frac{1-t}{1-s}\overrightarrow{AD}$ y contradice la hipótesis de que $\overrightarrow{AB}$ no es paralelo a $\overrightarrow{AD}$. De igual manera, si $t \neq 1$ entonces $\overrightarrow{AD} = \frac{1-s}{1-t}\overrightarrow{AB}$ y se contradice también que $\overrightarrow{AD}$ no es paralelo a $\overrightarrow{AB}$. Luego necesariamente debe ocurrir que $s = 1$ y $t = 1$ y de (5.1) y (5.2) se tiene que los lados no consecutivos de un paralelogramo son iguales.

$$\begin{aligned} \overrightarrow{BC} &= \overrightarrow{AD} \\ \overrightarrow{DC} &= \overrightarrow{AB} \end{aligned}$$

∎

Demostración: Parte b). Observe que $\overrightarrow{AD} = \overrightarrow{AB} + \overrightarrow{BD}$ de donde $\overrightarrow{BD} = \overrightarrow{AD} - \overrightarrow{AB}$ y $\frac{1}{2}\overrightarrow{BD} = \frac{1}{2}\overrightarrow{AD} - \frac{1}{2}\overrightarrow{AB}$, entonces

$$\begin{aligned}
\overrightarrow{AB} + \tfrac{1}{2}\overrightarrow{BD} &= \overrightarrow{AB} + \tfrac{1}{2}\overrightarrow{AD} - \tfrac{1}{2}\overrightarrow{AB} \\
&= \overrightarrow{AB} - \tfrac{1}{2}\overrightarrow{AB} + \tfrac{1}{2}\overrightarrow{AD} \\
&= \tfrac{1}{2}\overrightarrow{AB} + \tfrac{1}{2}\overrightarrow{AD} \\
&= \tfrac{1}{2}(\overrightarrow{AB} + \overrightarrow{AD})
\end{aligned}$$

Pero $\overrightarrow{AB} + \overrightarrow{AD} = \overrightarrow{AC}$, dado que $\overrightarrow{AD} = \overrightarrow{BC}$ como se demostró en la parte a), entonces:

$$\overrightarrow{AB} + \frac{1}{2}\overrightarrow{BD} = \frac{1}{2}\overrightarrow{AC}$$

■

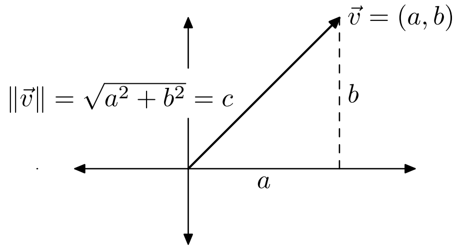


Figura 5.12: $\|\vec{v}\| = \sqrt{a^2 + b^2}$.

5.2 Normas, ángulos y proyecciones

En esta sección nos ocuparemos de precisar los conceptos de magnitud de un vector y ángulo entre dos vectores de $\mathbb{R}^n$. Conceptos muy naturales si se piensan en los espacios $\mathbb{R}^2$ o $\mathbb{R}^3$, pero no tanto si se aplican a vectores de $\mathbb{R}^n$ en general. Este es el momento en que la intuición debe abandonar el sentido de la vista y comenzar a creer más en las descripciones algebraicas, para lograr de esta manera que el ojo humano pueda “penetrar” a $\mathbb{R}^n$.

En $I\!R^2$ o $I\!R^3$, la magnitud de un vector puede determinarse a partir del teorema de Pitágoras. Si la magnitud de un vector $\vec{v}$ se denota por $\|\vec{v}\|$, observando el triángulo rectángulo en el gráfico 5.12, por el teorema de Pitágoras se tiene que:

$$\|\vec{v}\|^2 = a^2 + b^2$$

por lo cual

$$\|\vec{v}\| = \sqrt{a^2 + b^2}.$$

Similarmente, el caso de $I\!R^3$, dada la representación de $\vec{v} = (a, b, c)$, mostrada en la figura 5.13, y si d es la magnitud de la diagonal del triángulo rectángulo con catetos de magnitud a y b, entonces $d^2 = a^2 + b^2$.

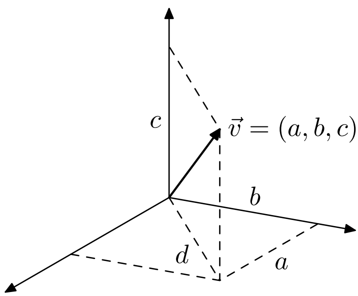


Figura 5.13: $\|\vec{v}\| = \sqrt{a^2 + b^2 + c^2}$.

Y como la magnitud del vector v, $\|\vec{v}\|$, es la magnitud de la diagonal del triángulo rectángulo cuyos catetos miden d y c, en valor absoluto, entonces:

$$\|\vec{v}\|^2 = d^2 + c^2 = a^2 + b^2 + c^2$$

luego $\|\vec{v}\| = \sqrt{a^2 + b^2 + c^2}$.

5.2.1 Producto punto y norma

Para extender el concepto de magnitud de un vector de $I\!R^2$ o $I\!R^3$ a $I\!R^n$, así como el de ángulo entre vectores, se introducirá una nueva operación vectorial que asocia a cada pareja de vectores en $I\!R^n$

con un número, denominada producto punto o producto escalar. Como se verá más adelante, la definición de esta operación puede cambiar, y con ello modificar las ideas de magnitud, distancia y ángulo entre vectores, pero siempre conservando sus propiedades establecidas en los teoremas de esta sección.

Definición 5.3 (Producto punto $\vec{a} \cdot \vec{b}$)

Sean $\vec{a} = (a_1, a_2, \ldots, a_n)^t$ y $\vec{b} = (b_1, b_2, \ldots, b_n)^t$. El producto escalar, o producto punto de $\vec{a}$ y $\vec{b}$ es un número real denotado y expresado en la siguiente forma:

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

En términos de operaciones matriciales, el producto punto $\vec{a} \cdot \vec{b}$ es el producto matricial: vector fila $\vec{a}$ por el vector columna $\vec{b}$, o sea:

$$\begin{aligned} \vec{a} \cdot \vec{b} &= \vec{a}^t \vec{b} \\ &= (a_1, a_2, \ldots, a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\ &= a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \end{aligned}$$

Una de las razones que justifican la introducción del producto punto, se deja ver en el siguiente ejemplo.

Ejemplo 5.4 Si $\vec{u} = (a, b, c)$ entonces

$$\vec{u} \cdot \vec{u} = (a, b, c) \cdot (a, b, c) = a^2 + b^2 + c^2$$

de manera que la norma de $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$.

Este ejemplo sugiere que, en general, si $\vec{u} \in I\!R^n$ entonces se puede definir la norma de $\vec{u}$, como $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$, como efectivamente se hará. Sin embargo, una definición así requiere mostrar que el producto punto tiene las características suficientes para garantizar que el concepto de magnitud de un vector en $I\!R^n$ nace con las propiedades de la magnitud que son bien conocidas en los espacios $I\!R^2$ o $I\!R^3$. Resultado que se logra con el siguiente teorema.

Teorema 5.4 *Si $\vec{u}, \vec{v}$ y $\vec{w}$ son tres vectores cualesquiera de $I\!R^n$ y $a \in I\!R$, entonces el producto punto tiene las siguientes propiedades:*

$$\begin{array}{ll}
(1) & \vec{u}\cdot\vec{u} > 0 \ \textit{si}\ \vec{u} \neq \vec{0} \\
 & \vec{u}\cdot\vec{u} = 0 \ \textit{solo si}\ \vec{u} = \vec{0} \\
(2) & \vec{u}\cdot\vec{v} = \vec{v}\cdot\vec{u} \\
(3) & \vec{u}\cdot(a\vec{v}) = (a\vec{u})\cdot\vec{v} = a(\vec{u}\cdot\vec{v}) \\
(4) & \vec{u}\cdot(\vec{v}+\vec{w}) = (\vec{u}\cdot\vec{v}) + (\vec{u}\cdot\vec{w})
\end{array}$$

Demostración: Ver ejercicio 19.

■■■

Definición 5.5 (Norma de un vector)

Si $\vec{u} = (u_1, u_2, \cdots, u_n)^t$ es un vector de $I\!R^n$, se llama norma o magnitud de $\vec{u}$ y se denota $\|\vec{u}\|$ al siguiente valor:

$$\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

La proposición (1) en el teorema 5.4, garantiza que la anterior definición es consistente con la idea que la magnitud de un vector siempre será positiva y que será cero sólo si se trata del vector cero. Las restantes proposiciones garantizan otras características que se esperan de la magnitud de un vector, y que son establecidas en el siguiente teorema.

Teorema 5.6 *Si $\vec{u}$ y $\vec{v}$ son vectores cualesquiera de $I\!R^n$ y $a \in I\!R$, entonces la norma de vectores tiene las siguientes propiedades:*

$$\begin{array}{lll}
(1) & \|\vec{u}\| \geq 0 \ \ y \ \ \|\vec{u}\| = 0 & \textit{si y solo si}\ \vec{u} = \vec{0} \\
(2) & \|a\vec{v}\| = |a|\,\|\vec{v}\| & \\
(3) & \|\vec{u}-\vec{v}\| = \|\vec{v}-\vec{u}\| & \\
(4) & |\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\,\|\vec{v}\| & \textit{(Desigualdad de Cauchy-Schwarz)} \\
(5) & \|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\| & \textit{(Desigualdad triangular)}
\end{array}$$

Demostración: (1), (2), (3) y (5) ejercicio.

■■■

Demostración: parte (4) –desigualdad de Cauchy-Schwarz–. Observe que para todo $x \in I\!R$, $\vec{u}+x\vec{v}$ es un vector cuya norma al cuadrado es positiva, de manera que

$$\|\vec{u}+x\vec{v}\|^2 = (\vec{u}+x\vec{v})\cdot(\vec{u}+x\vec{v}) \geq 0$$

además fácilmente se muestra que

$$(\vec{u}+x\vec{v})\cdot(\vec{u}+x\vec{v}) = x^2(\vec{v}\cdot\vec{v}) + 2x(\vec{u}\cdot\vec{v}) + \vec{u}\cdot\vec{u}$$

entonces

$$x^2(\vec{v}\cdot\vec{v}) + 2x(\vec{u}\cdot\vec{v}) + \vec{u}\cdot\vec{u} \geq 0 \quad \forall x \in I\!R.$$

Esta desigualdad involucra un polinomio de segundo grado, ax^2+bx+c, en la variable x, cuyos coeficientes son los números reales $a=\vec{v}\cdot\vec{v}, b=2\vec{u}\cdot\vec{v}$ y $c=\vec{u}\cdot\vec{u}$. Además se conoce que $ax^2+bx+c \geq 0$, $\forall x \in I\!R$ sólo si su discriminante es negativo y $a>0$. En este caso como $a=\vec{v}\cdot\vec{v}>0$ debe tenerse que:

$$b^2-4ac = (2\vec{u}\cdot\vec{v})^2 - 4(\vec{v}\cdot\vec{v})(\vec{u}\cdot\vec{u}) \leq 0$$

Y de esta desigualdad se sigue que $(\vec{u}\cdot\vec{v})^2 \leq (\vec{v}\cdot\vec{v})(\vec{u}\cdot\vec{u})$, de manera que:

$$|\vec{u}\cdot\vec{v}| \leq \sqrt{(\vec{v}\cdot\vec{v})(\vec{u}\cdot\vec{u})}.$$

Por lo tanto $|\vec{u}\cdot\vec{v}| \leq \sqrt{\vec{v}\cdot\vec{v}}\sqrt{\vec{u}\cdot\vec{u}} = \|\vec{u}\|\,\|\vec{v}\|$.

∎

Las propiedades para el producto escalar que establece el teorema 5.4 y la desigualdad de Cauchy-Schwarz permiten demostrar los otros enunciados del anterior teorema, cuyos resultados garantizan que la norma de un vector (y los ángulos entre vectores) correspondan con las ideas que se tienen para $I\!R^3$. Y de esa manera, compensan de alguna forma, la imposibilidad de visualizar los objetos de $I\!R^n$.

Desigualdad de Cauchy-Schwarz. Observe que si

$$\vec{a} = (a_1, a_2, \ldots, a_n)^t \text{ y } \vec{b} = (b_1, b_2, \ldots, b_n)^t$$

entonces la desigualdad de Cauchy-Schwarz adquiere la siguiente forma:

$$|a_1b_1+a_2b_2+\cdots+a_nb_n| \leq \sqrt{a_1^2+a_2^2+\cdots+a_n^2}\sqrt{b_1^2+b_2^2+\cdots+b_n^2}$$

Definición 5.7 (Distancia entre los vectores $\vec{u}$ y $\vec{v}$)

Dados dos vectores $\vec{u}=(u_1,u_2,\cdots,u_n)^t$ y $\vec{v}=(v_1,v_2,\cdots,v_n)^t$ en $I\!R^n$ se define la distancia entre estos vectores como la norma del vector diferencia, $d(\vec{u},\vec{v})=\|\vec{u}-\vec{v}\|=\|\vec{v}-\vec{u}\|$, o sea:

$$d(\vec{u},\vec{v})=\sqrt{(v_1-u_1)^2+(v_2-u_2)^2+\cdots+(v_n-u_n)^2}$$

El anterior concepto resulta más natural si se les da a $\vec{u}$ y $\vec{v}$ la interpretación de punto.

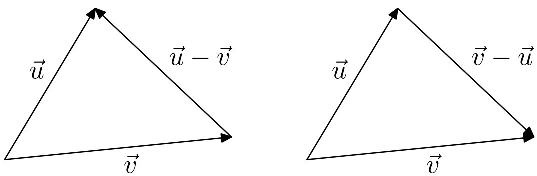


Figura 5.14: $d(u,v)=\|v-u\|=\|u-v\|$

Sin embargo, la distancia entre $\vec{u}$ y $\vec{v}$ —vistos como flechas— se puede entender como la distancia entre sus puntos terminales, si estos han sido representados de manera que el punto inicial de ambas flechas sea el mismo. Lo anterior porque tanto $\vec{u}-\vec{v}$ como $\vec{v}-\vec{u}$ se representan por flechas que unen los puntos terminales de $\vec{u}$ y $\vec{v}$, en uno u otro sentido.

Teorema 5.8 *Sean A,B y C en $I\!R^n$ arbitrarios, entonces:*

$$\begin{array}{ll} (1) & d(A,B)=d(B,A) \\ (2) & d(A,B)\leq d(A,C)+d(C,B) \end{array}$$

Demostración: Ejercicio.

■■■

5.2.2 Angulos en $I\!R^n$

La siguiente ley de cosenos es un resultado conocido para triángulos en $I\!R^2$ y $I\!R^3$ —ver ejercicio 1.8.25 en página 45—, el cual se demuestra utilizando resultados básicos de trigonometría y la noción de magnitud que apoyan nuestros sentidos.

Ley de cosenos: las magnitudes a, b y c, de los lados de cualquier triángulo, satisfacen que:

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

donde θ es la medida del ángulo entre los "lados" a y b.

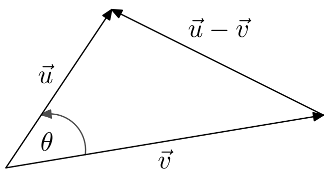


Figura 5.15: $\|\vec{u} - \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\,\|\vec{v}\|\cos\theta$.

Si esta ley se escribe utilizando una notación vectorial, en la que los lados del triángulo son los vectores $\vec{u}, \vec{v}$ y $\vec{u} - \vec{v}$, como se muestra en la figura 5.15 se obtiene que:

$$\|\vec{u} - \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\,\|\vec{v}\|\cos\theta \tag{5.3}$$

con θ la medida del ángulo comprendido entre los vectores $\vec{u}$ y $\vec{v}$. Pero recuerde que, hasta ahora, esto es así siempre que los vectores pertenezcan a $I\!R^2$ o $I\!R^3$, donde las nociones de ángulos y magnitudes están bien precisadas.

Por otra parte, si se continúa desarrollando $\|\vec{u} - \vec{v}\|^2$ en 5.3, se transforma en:

$$\|\vec{u} - \vec{v}\|^2 = (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v}) = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\vec{u} \cdot \vec{v} \tag{5.4}$$

Con lo cual, igualando los lados derechos de (5.3) y (5.4) resulta:

$$\vec{u} \cdot \vec{v} = \|u\| \, \|v\| \cos\theta$$

Entonces de la ley de cosenos se obtiene:

$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|}$$

Reiterando, siempre que $\vec{u}$ y $\vec{v}$ sean vectores no nulos en $I\!R^2$ o $I\!R^3$. Este resultado ofrece una forma de definir el coseno del ángulo entre dos vectores, dependiendo sólo de la noción de producto punto, y que por lo tanto puede extenderse a cualesquiera vectores de $I\!R^n$. Pero antes de hacer esto, se requiere reconocer que tal definición sería una buena definición en $I\!R^n$.

Observe que si $\vec{u}$ y $\vec{v}$ son vectores en $I\!R^n$, la desigualdad de Cauchy-Schwarz

$$|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \, \|\vec{v}\|$$

conduce a:

$$-\|\vec{u}\| \, \|\vec{v}\| \leq \vec{u} \cdot \vec{v} \leq \|\vec{u}\| \, \|\vec{v}\|$$

y que dividiendo por $\|\vec{u}\| \, \|\vec{v}\|$, para vectores no nulos, se obtiene:

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

Entonces se puede garantizar que para cada pareja de vectores $\vec{u}$ y $\vec{v}$ en $I\!R^n$ no nulos, existe un único valor $\theta \in [0, \pi]$ tal que:

$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|}$$

lo cual justifica la siguiente definición.

Definición 5.9 (Angulo entre dos vectores de $I\!R^n$)

Si $\vec{u}$ y $\vec{v}$ son dos vectores no nulos de $I\!R^n$, se dice que el único valor θ, $0 \leq \theta \leq \pi$ tal que

$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|}$$

es la medida del ángulo entre los vectores $\vec{u}$ y $\vec{v}$.

Definición 5.10 (Vectores ortogonales)

Dos vectores $\vec{u}$ y $\vec{v}$ en $I\!R^n$ son ortogonales si el ángulo entre ellos mide $\pi/2$ radianes, o al menos uno de ellos es el vector cero.

Teorema 5.11 *Los vectores $\vec{u}$ y $\vec{v}$ en $I\!R^n$ son ortogonales o perpendiculares si y sólo si $\vec{u} \cdot \vec{v} = 0$*

Demostración: Ejercicio.

5.2.3 Proyecciones ortogonales

Dados dos vectores $\vec{u}$ y $\vec{v}$, el proceso de descomponer el vector $\vec{u}$ como la suma de dos vectores ortogonales, uno de ellos paralelo a $\vec{v}$, tiene gran importancia en el desarrollo de modelos, por ejemplo al determinar la componente de una fuerza en una dirección dada, y en general, como un mecanismo de aproximación optimal de vectores, idea que se discutirá más adelante.

El problema: Dados los vectores $\vec{u}$ y $\vec{v}$ de $I\!R^n$, $\vec{v} \neq \vec{0}$, se quiere determinar un vector $\vec{a}$ paralelo a $\vec{v}$ y otro ($\vec{b}$) ortogonal a $\vec{v}$ y tales que

$$\vec{u} = \vec{a} + \vec{b}$$

como se muestra en la siguiente figura.

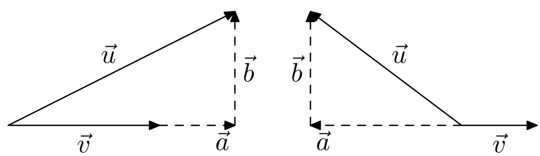


Figura 5.16: $\vec{a}$: proyección de $\vec{u}$ sobre $\vec{v}$.

En términos algebraicos lo anterior significa que se deben determinar $\vec{a}$ y $\vec{b}$ tales que:

$$\begin{array}{rl} 1) & \vec{a} = t\vec{v} \text{ para algún } t \in I\!R \\ 2) & \vec{v}\cdot\vec{b} = 0 \\ \text{y } 3) & \vec{b} = \vec{u} - t\vec{v}. \end{array}$$

lo cual reduce el problema a buscar el valor t apropiado, para definir $\vec{a}$ y $\vec{b}$ como en 1) y 3), de forma que se satisfaga 2):

$$\begin{array}{rrcl} & \vec{v}\cdot\vec{b} & = & 0 \\ \Rightarrow & \vec{v}\cdot(\vec{u} - t\vec{v}) & = & 0 \\ \Rightarrow & \vec{v}\cdot\vec{u} - t\vec{v}\cdot\vec{v} & = & 0 \\ \Rightarrow & \vec{v}\cdot\vec{u} & = & t\vec{v}\cdot\vec{v} \end{array}$$

Por otra parte, como $\vec{v} \neq \vec{0}$ entonces $\vec{v}\cdot\vec{v} \neq 0$ y el valor de t requerido es:

$$t = \frac{\vec{v}\cdot\vec{u}}{\vec{v}\cdot\vec{v}}$$

Finalmente los vectores $\vec{a}$ y $\vec{b}$ buscados son:

$$\begin{array}{lrcl} & \vec{a} & = & \frac{\vec{v}\cdot\vec{u}}{\vec{v}\cdot\vec{v}}\vec{v} \\ \text{y} & \vec{b} & = & \vec{u} - \vec{a} \end{array}$$

Geométricamente, la idea anterior corresponde a proyectar ortogonalmente el vector $\vec{u}$ sobre el $\vec{v}$, para obtener $\vec{a}$, por lo cual al vector $\vec{a}$ se le llama **proyección ortogonal de $\vec{u}$ sobre $\vec{v}$** y a $\vec{b}$ componente de $\vec{u}$ ortogonal a $\vec{v}$.

Definición 5.12 (Proyección ortogonal de $\vec{u}$ sobre $\vec{v}$)

Si $\vec{u}$ y $\vec{v}$ son dos vectores $I\!R^n$ y $\vec{v} \neq \vec{0}$ se llama proyección ortogonal de $\vec{u}$ sobre $\vec{v}$ y se denota $Proy_{\vec{v}}\vec{u}$ al vector

$$Proy_{\vec{v}}\vec{u} = \frac{\vec{v}\cdot\vec{u}}{\vec{v}\cdot\vec{v}}\vec{v} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v}$$

Y el vector $\vec{u} - Proy_{\vec{v}}\vec{u}$ se conoce como componente de $\vec{u}$ ortogonal a $\vec{v}$.

Observe que si $\vec{v}$ es un vector unitario entonces la proyección ortogonal de $\vec{u}$ sobre $\vec{v}$ se reduce a: $\text{Proy}_{\vec{v}}\vec{u} = (\vec{u}\cdot\vec{v})\vec{v}$.

En modelos estadísticos, la idea de ortogonalidad entre vectores, está asociada con la idea de independencia "pura" de la información que resumen ambos vectores. En otras palabras, ortogonalidad supone no correlación entre las variables representadas por los vectores, contexto en el cual la descomposición de un vector $\vec{u}$ en dos componentes ortogonales, $\text{Proy}_{\vec{v}}\vec{u}$ y $\vec{u} - \text{Proy}_{\vec{v}}\vec{u}$ corresponde a la idea de separar la información en $\vec{u}$ que es dependiente de $\vec{v}$, esto es $\text{Proy}_{\vec{v}}\vec{u}$, de aquella que es totalmente independiente de $\vec{v}$ o sea $\vec{u} - \text{Proy}_{\vec{v}}\vec{u}$. Lograda esta separación, se dice que $\text{Proy}_{\vec{v}}\vec{u}$ es la mejor representación de $\vec{u}$ en el espacio generado por $\vec{v}$ y $\|\vec{u} - \text{Proy}_{\vec{v}}\vec{u}\|$ es una medida del error de esa representación.

En términos más generales, se dice que $\text{Proy}_{\vec{v}}\vec{u}$ es el vector paralelo a $\vec{v}$ que mejor aproxima a $\vec{u}$, o la mejor representación de $\vec{u}$ en la recta generada por $\vec{v}$, ideas sobre las que se seguirá trabajando, en los siguientes capítulos.

5.3 Producto cruz

El producto cruz es una operación singular entre vectores de $I\!R^3$, que resulta con algunas propiedades interesantes por sus interpretaciones geométricas. Sin embargo, se debe enfatizar que es un concepto definido sólo para vectores en $I\!R^3$.

Definición 5.13 (Producto cruz)

Sean $\vec{v} = (v_1, v_2, v_3)$ y $\vec{w} = (w_1, w_2, w_3)$ dos vectores en $I\!R^3$, entonces el producto cruz de $\vec{v}$ con $\vec{w}$ es un nuevo vector de $I\!R^3$, denotado por $\vec{v} \times \vec{w}$, que se define como:

$$\vec{v} \times \vec{w} = \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \vec{e}_3$$

donde $\vec{e}_1, \vec{e}_2, \vec{e}_3$ son los vectores canónicos de $I\!R^3$.

Una forma simple de recordar la definición del producto cruz de $\vec{v}$ y $\vec{w}$, es mediante el siguiente "determinante", desarrollándolo por la primera fila:

$$\vec{v} \times \vec{w} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

Aunque esta expresión carece de sentido, porque la primera fila de la matriz no se compone de números sino de vectores, y la expresión resultante es un vector de $I\!R^3$, y no un número como correspondería al determinante. Aún así, ignorando esta aberración, se seguirá empleando por la comodidad que ofrece para recordar la definición de producto cruz.

Ejemplo 5.5 Si $\vec{v} = (1, 3, 2)$ y $\vec{w} = (2, -1, 0)$:

$$\begin{aligned} \vec{v} \times \vec{w} &= \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ 1 & 3 & 2 \\ 2 & -1 & 0 \end{vmatrix} \\ &= \begin{vmatrix} 3 & 2 \\ -1 & 0 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} 1 & 3 \\ 2 & -1 \end{vmatrix} \vec{e}_3 \\ &= 2\vec{e}_1 + 4\vec{e}_2 - 7\vec{e}_3 \\ &= (2, 4, -7). \end{aligned}$$

En particular, para los vectores conónicos, fácilmente se verifica que:

1. $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$.
2. $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$.
3. $\vec{e}_1 \times \vec{e}_3 = -\vec{e}_2$.

Teorema 5.14 (Propiedades del producto cruz) *Sean $\vec{u}$, $\vec{v}$ y $\vec{w}$ vectores de $\mathbb{R}^3$ y a un escalar. El producto cruz verifica las siguientes propiedades:*

a) $\vec{w} \times \vec{0} = \vec{0} \times \vec{w} = \vec{0}$

b) $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$

c) $a\vec{v} \times \vec{w} = a(\vec{v} \times \vec{w})$

d) $\vec{v} \times a\vec{v} = \vec{0}$

e) $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$

f) El vector $\vec{v} \times \vec{w}$ *es ortogonal a* $\vec{v}$ *y* $\vec{w}$

g) $(\vec{u} \times \vec{v}) \cdot \vec{w} = \vec{u} \cdot (\vec{v} \times \vec{w})$

h) $(\vec{u} \times \vec{v}) \times \vec{w} = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{v} \cdot \vec{w})\vec{u}$

Demostración: Cada una de estas propiedades se demuestra en forma directa a partir de la definición de producto cruz. A continuación se presenta la correspondiente a g):

Si $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$ y $\vec{w} = (w_1, w_2, w_3)$, entonces

$$\vec{u} \times \vec{v} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \vec{e}_3.$$

Luego,

$$\begin{aligned} (\vec{u} \times \vec{v}) \cdot \vec{w} &= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} w_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} w_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} w_3 \\ &= \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\ &= \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} u_1 - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} u_2 + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} u_3 \\ &= \vec{u} \cdot (\vec{v} \times \vec{w}). \end{aligned}$$

■■■

Teorema 5.15 *Si θ es el ángulo entre los vectores $\vec{u}$ y $\vec{v}$ de $I\!R^3$ entonces*

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \; sen\,(\theta)$$

Demostración: (Indicación) verifique que

$$\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2 \, \|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2$$

y utilice que $(\vec{u} \cdot \vec{v}) = \|\vec{u}\| \, \|\vec{v}\| \cos(\theta)$.

■■■

El teorema 5.15 provee una interpretación geométrica para la magnitud del producto cruz de dos vectores.

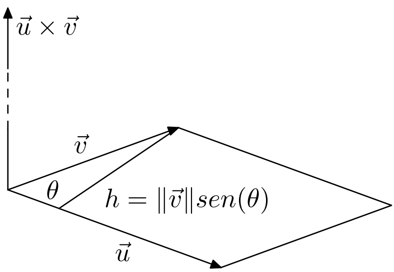


Figura 5.17: $\|\vec{u} \times \vec{v}\|$: área del paralelogramo de lados $\vec{u}$ y $\vec{v}$.

Observe en el anterior gráfico, que si θ es el ángulo entre $\vec{u}$ y $\vec{v}$, la altura h del paralelogramo que determinan estos vectores, se puede expresar como

$$h = \|v\| \; \text{sen}\,(\theta)$$

y utilizando que $\text{sen}\,(\theta) = h / \|v\|$ se tiene:

$$\begin{aligned} \|\vec{u} \times \vec{v}\| &= \|\vec{u}\| \, \|\vec{v}\| \; \text{sen}\,(\theta) \\ &= \|\vec{u}\| \, h, \end{aligned}$$

lo que corresponde al área de paralelogramo de lados $\vec{u}$ y $\vec{v}$.

5.3.1 Relación entre el producto cruz, el volumen de paralelepípedos y los determinantes

Tres vectores $\vec{u}, \vec{v}, \vec{w}$ de $I\!R^3$ que sean l.i. determinan un paralelepípedo: una especie de "caja" o figura cerrada del espacio tridimensional con seis caras planas (con forma de paralelogramos), que se unen en aristas formadas por los vectores $\vec{u}, \vec{v}, \vec{w}$, como se muestra en la siguiente figura.

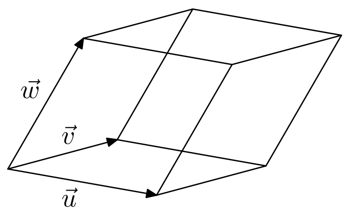


Figura 5.18: Paralelepípedo determinado por los vectores $\vec{u}, \vec{v}, \vec{w}$.

La propiedad g) del teorema 5.14 permite también una interpretación geométrica que la relacionada con los paralelepípedos, específicamente:

> El volumen V del paralelepípedo que determinan $\vec{u}, \vec{v}, \vec{w}$, vectores l.i. de $I\!R^3$, es dado por
>
> $$V = |(\vec{u} \times \vec{v}) \cdot \vec{w}|.$$

Construcción del resultado

Si se considera como base del paralelepípedo, el paralelogramo de lados $\vec{u}$ y $\vec{v}$, como $\vec{u} \times \vec{v}$ es un vector ortogonal a este paralelogramo, la altura h del paralelepípedo corresponde a la magnitud de la proyección ortogonal de $\vec{w}$ sobre $\vec{u} \times \vec{v}$:

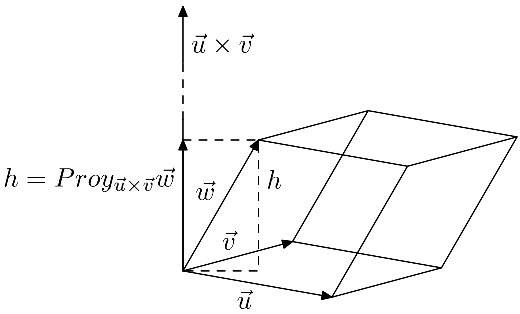


Figura 5.19: Altura: $h = \left\|\text{Proy}_{\vec{u}\times\vec{v}}\vec{w}\right\|$.

Entonces h se calcula en la siguiente forma:

$$\begin{aligned} h &= \left\|\text{Proy}_{\vec{u}\times\vec{v}}\vec{w}\right\| \\ &= \left\|\frac{(\vec{u}\times\vec{v})\cdot\vec{w}}{\|\vec{u}\times\vec{v}\|^2}(\vec{u}\times\vec{v})\right\| \\ &= \frac{|(\vec{u}\times\vec{v})\cdot\vec{w}|}{\|\vec{u}\times\vec{v}\|^2}\,\|(\vec{u}\times\vec{v})\| \\ &= \frac{|(\vec{u}\times\vec{v})\cdot\vec{w}|}{\|\vec{u}\times\vec{v}\|} \end{aligned}$$

De esta manera:

$$\begin{aligned} V &= (\text{área de la base) altura} \\ &= \|\vec{u}\times\vec{v}\|\,h \\ &= \|\vec{u}\times\vec{v}\|\,\frac{|(\vec{u}\times\vec{v})\cdot\vec{w}|}{\|\vec{u}\times\vec{v}\|} \\ &= |(\vec{u}\times\vec{v})\cdot\vec{w}| \end{aligned}$$

con lo que se obtiene el resultado.

Por otra parte, como en la demostración de la propiedad g) del teorema 5.14 se obtuvo que

$$(\vec{u}\times\vec{v})\cdot\vec{w} = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

entonces

$$V = \pm \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix},$$

donde el signo $\pm$ significa que se debe tomar el valor absoluto del determinante. De esta manera se obtiene también una interpretación geométrica para los determinantes de matrices de orden tres por tres.

5.4 Conceptos de distancia y ángulo en el análisis de datos

Las nociones de norma, distancia, ángulo entre vectores y proyección ortogonal vistas, dependen del producto escalar definido en (5.3). Es posible cambiar esta definición de manera que se preserven las características establecidas en el teorema (5.4) y con ello dar origen a nuevos conceptos de normas, distancias, ángulos entre vectores y proyecciones ortogonales sobre vectores.

Definición 5.16 (Producto escalar)

Si u, v y w son vectores en $\mathbb{R}^n$ y $a \in \mathbb{R}$, un producto escalar es una operación, $\langle , \rangle$, que asocia a cada pareja u, v de vectores con un número real denotado $\langle u, v \rangle$ que satisface las siguientes propiedades:

$$\begin{array}{ll} (1) & \langle u, u \rangle > 0 \textit{ siempre que } u \neq 0 \\ & \langle u, u \rangle = 0 \textit{ si y solo si } u = 0 \\ (2) & \langle u, v \rangle = \langle v, u \rangle \\ (3) & \langle u, av \rangle = \langle au, v \rangle = a \langle u, v \rangle \\ (4) & \langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle \end{array}$$

Cualquier producto escalar $\langle , \rangle$ permite definir, de la misma manera que se hizo con el producto escalar clásico:

a) Un concepto de norma para vectores: $\|u\| = \sqrt{\langle u, u \rangle}$.

b) Una distancia entre vectores: $d(u,v) = \|u - v\|$.

c) Una idea de ortogonalidad: u es ortogonal a $v \Leftrightarrow \langle u, v\rangle = 0$

d) Una medida para el ángulo θ entre los vectores u y v :

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\,\|v\|}$$

e) La proyección ortogonal de u sobre un vector v no nulo:

$$\mathrm{Proy}_v u = \frac{\langle u, v\rangle}{\langle v, v\rangle} v = \frac{\langle u, v\rangle}{\|v\|^2} v$$

Definición 5.17 (Matriz definida positiva)

Sea $M \in M(n, I\!R)$, se dice que M es definida positiva si

$$\begin{array}{rlcl} & 1) \quad x^t M x \geq 0 & \forall & x \in I\!R^n \\ y & 2) \quad x^t M x = 0 & \Longleftrightarrow & x = 0_n \end{array}$$

Si M cumple 1) pero no 2) se dice que M se semi-definida positiva.

En el ejercicio 32 en la página 184, se propone demostrar que si $M \in M(n, I\!R)$ es una matriz simétrica y definida positiva entonces

$$\langle x, y\rangle_{M} = x^t M y$$

es un producto escalar para el espacio $I\!R^n$. Lo cual como ya se dijo define una M-norma, una distancia d_M, y un concepto de M-ortogonalidad en $I\!R^n$, ideas que se resumen diciendo que se ha definido una métrica o la métrica M.

En particular, cuando M es una matriz diagonal, cuyos elementos en la diagonal son todos positivos, se tiene claramente que M es una matriz simétrica y definida positiva. Lo segundo porque si $M = \mathrm{diagonal}(m_i)$, con $m_i > 0 \;\; \forall i = 1, \ldots, n$, entonces:

1. $x^t M x = \sum_{i=1}^n m_i x_i^2 \geq 0 \;\; \forall x \in I\!R^n$.

2. $x^t M x = 0 \Longleftrightarrow \sum_{i=1}^n m_i x_i^2 = 0 \Longleftrightarrow x_i = 0 \Longleftrightarrow x = 0_n$.

Y si además $\sum_{i=1}^n m_1 = 1$, la métrica que define la matriz diagonal M tiene especial importancia en el análisis estadístico de datos, y es conocida como una métrica de pesos.

5.4.1 Métricas de pesos y medidas estadísticas

Si $p_1, p_2, \ldots, p_n$ son n números positivos tales que $\sum_{i=1}^{n} p_i = 1$, la siguiente matriz

$$D_p = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & p_n \end{pmatrix}$$

define, para vectores de $I\!R^n$, el producto escalar $\langle , \rangle_{_P}$:

$$\langle x, y \rangle_{_P} = x^t D_p y = \sum_{i=1}^{n} p_i x_i y_i$$

y la correspondiente distancia:

$$d(x, y) = \sqrt{\langle x - y, x - y \rangle_{_P}} = \sqrt{(x - y)^t D_p (x - y)}.$$

A continuación se podrá reconocer la importancia de esta métrica en la descripción de las principales medidas estadísticas para variables cuantitativas.

Media aritmética y desviación estándar

Sea $x^t = (x_1, x_2, \ldots, x_n)$ un vector de $I\!R^n$, que resume n observaciones de una variable estadística cuantitativa x. La principal medida de la tendencia central de x es su media aritmética:

$$\overline{x} = \sum_{i=1}^{n} p_i x_i = \langle 1_n, x \rangle_{_P} = 1_n^t D_P x.$$

La varianza y desviación estándar de x, $\mathrm{Var}\,(x)$ y σ_x, miden la dispersión de las observaciones alrededor de su media:

$$\begin{array}{rclcl} \mathrm{Var}\,(x) & = & \sum_{i=1}^{n} p_i (x_i - \overline{x})^2 & = & \|x - \overline{x} 1_n\|_{_P}^2 \\ \sigma_x & = & \sqrt{\mathrm{Var}\,(x)} & = & \|x - \overline{x} 1_n\|_{_P} \end{array}$$

En particular, cuando todos los pesos son iguales, es decir $p_i = 1/n$ para cada $i = 1, 2, \ldots, n$, se tiene que $\overline{x} = \sum_{i=1}^{n} x_i / n$ y $\mathrm{Var}\,(x) = \sum_{i=1}^{n} (x_i - \overline{x})^2 / n$.

A continuación se obtienen los siguientes descripciones dadas para la media, varianza y desviación estándar de x, en términos de producto escalar $\langle , \rangle_P$:

1. Si $1_n^t = (1, 1, \ldots, 1) \in I\!R^n$ entonces

$$\langle 1_n, x \rangle_P = 1_n^t D_P x = \sum_{i=1}^{n} p_i x_i = \overline{x}$$

2. Por otra parte:

$$\begin{array}{rcl} d_p^2(x, \overline{x}1_n) & = & \|x - \overline{x}1_n\|_P^2 \\ & = & \langle x - \overline{x}1_n, x - \overline{x}1_n \rangle_P \\ & = & (x - \overline{x}1_n)^t D_P (x - \overline{x}1_n) \\ & = & \sum_{i=1}^{n} p_i (x_i - \overline{x})^2 \\ & = & \mathrm{Var}\,(x). \end{array}$$

3. De lo anterior resulta que:

$$d_p(x, \overline{x}1_n) = \sigma_x = \|x - \overline{x}1_n\|_P .$$

4. Y finalmente, $\overline{x}$ se caracteriza también porque:

$$\mathrm{Proy}_{1_n} x = \frac{\langle x, 1_n \rangle_P}{\langle 1_n, 1_n \rangle_P} 1_n = \frac{\overline{x}}{1} 1_n = \overline{x}1_n.$$

Estos resultados se visualizan en la figura 5.20:

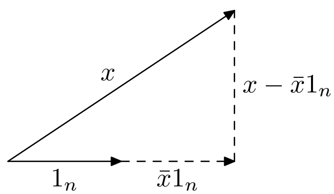


Figura 5.20: Interpretación geométrica para $\overline{x}$ y σ_x .

Por otra parte, como $\mathrm{Proy}_{1_n} x = \overline{x}1_n$ es el vector en la dirección de 1_n más cercano a x, entonces $\overline{x}1_n$ es el vector “más parecido” a x con todas sus componentes iguales. O sea, la mejor manera de resumir las n observaciones de x, $x_1, x_2, \ldots, x_n$, como un sólo número es eligiendo a $\overline{x}$ para que las represente. Y se observa también que ésta simplificación tiene como medida de error, la desviación estándar $\sigma_x = \|x - \overline{x}1_n\|_P$.

Covarianza y coeficiente de correlación

Ahora, sean $y^t = (y_1, y_2, \ldots, y_n)$ y $x^t = (x_1, x_2, \ldots, x_n)$ dos vectores en $\mathbb{R}^n$, que contabilizan n observaciones de las variables estadísticas x y y. Dos medidas estadísticas para la asociación entre estas variables son, la covarianza entre x y y:

$$\mathrm{Cov}\,(x,y) = \sum_{i=1}^{n} p_i(x_i - \overline{x})(y_i - \overline{y})$$

y el coeficiente de correlación lineal

$$r(x,y) = \frac{\mathrm{Cov}\,(x,y)}{\sigma_x \sigma_y}.$$

Utilizando el producto $\langle , \rangle_P$ se puede observar que:

$$\begin{array}{rcl} \langle x - \overline{x}1_n, y - \overline{y}1_n \rangle_P & = & (x - \overline{x}1_n)^t D_P (y - \overline{y}1_n) \\ & = & \sum_{i=1}^{n} p_i(x_i - \overline{x})(y_i - \overline{y}) \\ & = & \mathrm{Cov}\,(x,y). \end{array}$$

O sea, la covarianza entre dos variables es simplemente el producto escalar entre estas dos variables centradas.

Además, si θ es el ángulo entre los vectores $x - \overline{x}1_n$ y $y - \overline{y}1_n$ entonces

$$\begin{array}{rcl} \cos\theta & = & \dfrac{\langle x - \overline{x}1_n, y - \overline{y}1_n \rangle_P}{\|x - \overline{x}1_n\|_P \, \|y - \overline{y}1_n\|_P} \\ \\ & = & \dfrac{\mathrm{Cov}\,(x,y)}{\sigma_x \sigma_y} \\ \\ & = & r(x,y). \end{array}$$

Esto es, la correlación entre dos variables corresponde al coseno del ángulo entre estas variables centradas: $x - \overline{x}1_n$ y $y - \overline{y}1_n$, (llamados vectores de observaciones centradas o simplemente "x centrado" y "y centrado").

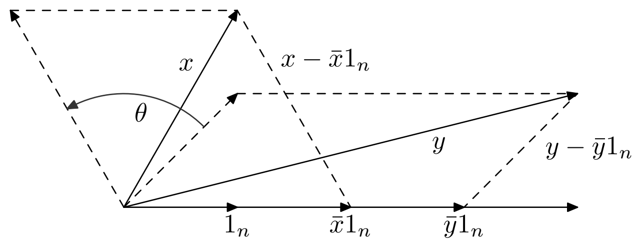
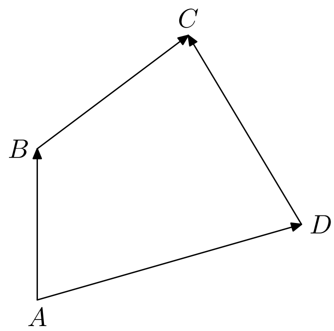


Figura 5.21: $\cos\theta$: correlación lineal entre las variables x y y.

En el gráfico anterior, el ángulo θ entre los vectores $x - \overline{x}1_n$, $y - \overline{y}1_n$ es el mismo que el que se forma entre los planos engendrados por x y 1_n y por y y 1_n. Y como se ha visto, su coseno corresponde al coeficiente de correlación lineal entre x y y. Este coeficiente es una medida de la dependencia lineal entre las variables x y y, cuando $r_{xy} = 0$ se dice que las variables son no correlacionadas, lo que significa en términos del álgebra vectorial que x y y son D_p-ortogonales (ortogonales con el producto escalar inducido por D_p).

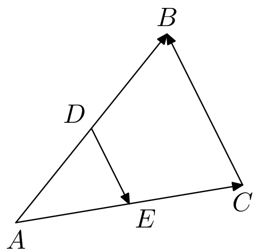
5.5 Ejercicios

1. Demuestre que un cuadrilátero con vértices A, B, C y D



es un paralelogramo si y solo si $A - B + C - D = 0$

2. Considere un triángulo de vértices A, B y C y sean D el punto medio del segmento $\overline{AB}$ y E el punto medio del segmento $\overline{AC}$, como se muestra en la figura:



Demuestre que $\overrightarrow{DE} = \frac{1}{2}\overrightarrow{BC}$, o sea, que el segmento que une los puntos medios de dos lados en un triángulo es paralelo al tercer lado y su longitud es un medio de la de este.

3. Considere los vectores $\vec{u} = (2, -3)$ y $\vec{v} = (-4, 2)$ y determine:

a) Cuatro puntos A, B, C, D distintos tales que $\overrightarrow{AB} = \vec{u} = \overrightarrow{CD}$.

b) Dos puntos A y B tales que $\overrightarrow{AB} = \frac{1}{2}(\vec{u} + \vec{v})$.

c) Los vértices de un paralelogramo con lados consecutivos $\vec{u}$ y $\vec{v}$.

4. Haga el mismo ejercicio anterior, pero con $\vec{u} = (-1, 3, -2)$ y $\vec{v} = (3, 2, 4)$.

5. Sean $A = (-3, 5)$, $B = (2, 1)$ y $C = (3/2, 11/2)$ los vértices de un triángulo.

a) Haga un dibujo que muestre los puntos A, B y C.

b) Determine M, el punto medio del segmento de recta con puntos extremos A y B.

c) ¿El triángulo ABC es isósceles?

d) ¿Es recto el ángulo entre $\overrightarrow{MC}$ y $\overrightarrow{MB}$?

6. En cada caso, haga un dibujo y marque claramente la región que corresponde a todos los puntos que satisfacen la condición dada:

a) los puntos (a, b) tales que $(a, b) = t(-2, 3)$ donde $1 \leq t < 3$.

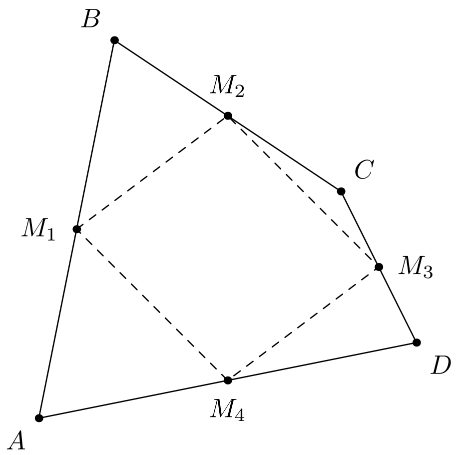
b) los puntos (x, y) tales que $(x, y) = t(-2, 3) + s(1, 4)$ donde $0 \leq t \leq 1$ y $0 \leq s < 1$.

c) los puntos (x, y) tales que $(x, y) = t(-2, 3) + s(1, 4)$ donde $1 \leq t \leq 2$ y $0 \leq s \leq 1$.

d) los puntos (a, b, c) tales que $(a, b, c) = (1, 1, 1) + t(-2, 3, 1)$ donde $0 \leq t \leq 1$.

e) los puntos (x, y) tales que $(x, y, z) = t(0, 3, 1) + s(2, 4, 0)$ donde $0 \leq t \leq 1$ y $0 \leq s \leq 3$.

7. Sean A, B, C y D los vértices de un cuadrilátero cualquiera, y M_1, M_2, M_3 y M_4 los puntos medios de los lados de este cuadrilátero como se muestra en el dibujo. Demuestre que M_1, M_2, M_3 y M_4 son los vértices de un paralelogramo.



8. Calcule un vector unitario, o sea de norma 1, en la dirección de cada vector dado.

1. $(-3, 4)$	2. $(-1, 2, -3)$
3. $(-1, 2, -3, 4)$	4. $(a, b-a, b)$
5. $(x_1, x_2, \ldots, x_n)$	6. $\vec{v}$

9. Demuestre que los vectores

$$(\cos\theta,\ \operatorname{sen}\theta, 0),\ (-\operatorname{sen}\theta, \cos\theta, 0),\ (0, 0, 1)$$

son mútuamente ortogonales y de norma 1, o sea, forman un conjunto ortonormal.

10. Determine el coseno del ángulo entre cada par de vectores.

1. $(-3, 3), (2, 3)$	2. $(1, 0, 3), (-1, 2, 1)$
3. $(1, -3, 2, 2), (-1, 2, 3, -4)$	4. $(a, b, c), (-b, a, 0)$
5. $(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)$	

11. Demuestre que si $\vec{u}$ y $\vec{v}$ en $\mathbb{R}^n$. son paralelos entonces el ángulo entre ellos mide 0 o π radianes.

12. Demuestre que dos vectores $\vec{u}$, $\vec{v}$ de $\mathbb{R}^n$, no nulos, son l.d. si y solo si son paralelos.

13. Sean $\vec{e_i}$, $i = 1, 2, \ldots, n$ los vectores conónicos de $I\!R^n$, es decir, e_i tiene ceros en todas sus componentes excepto en la i-ésima donde tiene un 1. Demuestre que estos vectores son unitarios y ortogonales dos a dos. Demuéstrelo primero para el caso en que $n = 4$.

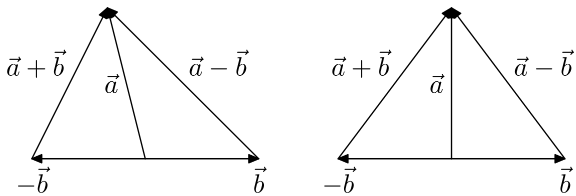
14. En cada caso, considere los vectores dados como vértices de un triángulo:

1. $(-3, 3), (2, 3), (0, -1)$
2. $(1, 0, 3), (-1, 2, 1), (-4, -4, -2)$
3. $(3, -3, 3), (2, -3, 4), (3, 0, 7)$

a) ¿Cuáles de estos triángulos son rectángulos?

b) Determine el área de los triángulos.

15. Sean $\vec{a}$, $\vec{b}$ vectores no nulos en $I\!R^n$. Demuestre que $\left\|\vec{a} + \vec{b}\right\| = \left\|\vec{a} - \vec{b}\right\|$ si y solo si $\vec{a}$ y $\vec{b}$ son ortogonales.



16. Sean $\vec{a}$ y $\vec{b}$ en $I\!R^n$, no nulos y perpendiculares. Demuestre que

$$\left\|\vec{a} + \alpha\vec{b}\right\| \geq \|\vec{a}\| \quad \forall \alpha \in I\!R$$

17. Pruebe que el teorema de Pitágoras es válido en $I\!R^n$, o sea, que si los vectores no nulos $\vec{a}$ y $\vec{b}$ representan catetos en un triángulo de $I\!R^n$ entonces $\left\|\vec{a} - \vec{b}\right\|^2 = \|\vec{a}\|^2 + \left\|\vec{b}\right\|^2$ si y solo si $\vec{a}$ y $\vec{b}$ son ortogonales.

La generalización del teorema de Pitágoras, la ley de los cosenos, ¿será también válida para triángulos en $I\!R^n$?, esto es:

$$\left\|\vec{a} - \vec{b}\right\|^2 = \|\vec{a}\|^2 + \left\|\vec{b}\right\|^2 - 2\,\|\vec{a}\|\,\left\|\vec{b}\right\|\cos\theta$$

donde θ es el ángulo entre $\vec{a}$ y $\vec{b}$. Demuéstrelo.

18. Demuestre que las diagonales de un paralelogramo son ortogonales si y solo si el paralelogramo es un rombo. Y similarmente, que las diagonales tienen la misma magnitud si y solo si el paralelogramo es un rectángulo.

19. Demuestre las propiedades del producto escalar en el teorema 5.4.

20. Demuestre la proposición (5) del teorema 5.6, la desigualdad triangular. Sugerencia: muestre primero que

$$\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2$$

y use la desigualdad de Cauchy-Schwarz para obtener que

$$\|\vec{u}\|^2 + 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2 \leq (\|\vec{u}\| + \|\vec{v}\|)^2$$

21. Dados los vectores $\vec{u}$ y $\vec{v}$ abajo especificados, determine dos vectores $\vec{a}$ y $\vec{b}$ perpendiculares, tales que $\vec{a}$ sea paralelo a $\vec{v}$ y $\vec{u} = \vec{a} + \vec{b}$. Haga una representación geométrica de estos cuatro vectores.

1. $\vec{u} = (-3, 3), \vec{v} = (2, 3)$
2. $\vec{u} = (1, 0, 3), \vec{v} = (-1, 2, 1)$
3. $\vec{u} = (1, -3, 3, 2), \vec{v} = (-1, 2, 3, 4)$
4. $\vec{u} = (a, b, c), \vec{v} = (1, 1, 1)/\sqrt{3}$
5. $\vec{u} = (x_1, x_2, \ldots, x_n), \vec{v} = e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$

22. Demuestre que si $A_1, A_2, \ldots, A_k$ vectores no nulos en $I\!R^n$ son mútuamente perpendiculares, es decir, $A_i \cdot A_j = 0$ si $i \neq j$, para $i, j = 1, 2, \ldots, k$, entonces $A_1, A_2, \ldots, A_k$ son linealmente independientes, o sea:

$$\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_k A_k = \vec{0} \Rightarrow \alpha_i = 0 \text{ para } i = 1, 2, \ldots, k$$

23. Demuestre que si $\vec{a}$ es ortogonal a cada uno de los vectores $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ en $I\!R^n$ entonces $\vec{a}$ es ortogonal a cualquier vector de la forma: $\vec{v} = x_1\vec{v_1} + x_2\vec{v_2} + \cdots + x_k\vec{v_k}$, cualesquiera sean los escalares $x_1, x_2, \ldots, x_k$, o sea, $\vec{a}$ es perpendicular a cualquier vector en $\mathcal{C}\ell\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}\}$.

24. Sean $A_1, A_2, \ldots, A_n$ los vectores columna de una matriz $A \in M(n, I\!R)$, o sea $A = (A_1, A_2, \ldots, A_n)$.

a) Demuestre que $A^t A$ es una matriz diagonal si y solo si los vectores en $A_1, A_2, \ldots, A_n$ son mútuamente ortogonales.

b) Demuestre que A es una matriz ortogonal[1] si y solo si los vectores $A_1, A_2, \ldots, A_n$ son unitarios y ortogonales dos a dos.

Un conjunto de vectores $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$ se llaman ortonormales si son unitarios y ortogonales dos a dos.

25. Considere los vectores $\vec{v} = (a, b)$ y $\vec{w} = (c, d)$ de $I\!R^2$:

1. Demuestre que

$$\|(a, b, 0) \times (c, d, 0)\| = |\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}|$$

2. Elabore una interpretación geométrica para los determinantes de las matrices 2×2 en términos de los vectores que componen sus filas (o sus columnas).

26. Calcule el área de los paralelogramos determinados por los vectores o vértices dados:

a) Los vectores $(2, 3, 1)$ y $(-2, -1, 3)$.

b) Tres de sus vértices son $(1, 2, 3)$, $(0, -3, 1)$ y $(5, -1, -2)$.

c) Los vectores $(1, 3)$ y $(1, 1)$.

d) Los vectores (a, a, a) y $(-a, 0, a)$.

27. Demuestre las propiedades a), b), c), d), e) y f) del teorema 5.14.

28. En cada caso, determine todos los vectores unitarios que son ortogonales a:

a) $(1, 1, 1)$ y $(-1, 0, 1)$.

b) $(1, 1, 1, 1)$ y $(-1, 0, 1, 0)$.

29. En cada caso, calcule el volumen del paralelepípedo que determinan los vectores:

a) $(1, 0, 2)$, $(-2, 0, 3)$ y $(2, 2, 2)$.

b) $(a, 0, 0)$, $(0, b, 0)$ y (c, c, c).

[1] Una matriz cuadrada A es ortogonal si $A^t A = I$.

30. Calcule el área del triángulo de vértices $(1, 1, 0)$, $(-1, 0, 2)$ y $(0, 1, 2)$.

31. Sea A una matriz invertible de orden 3, y $\vec{x}$, $\vec{y}$ y $\vec{z}$ vectores l.i. de $I\!R^3$. Si $\vec{u} = Ax$, $\vec{v} = Ay$ y $\vec{w} = Az$ muestre que

$$V_{uvw} = |\det(A)| V_{xyz}$$

donde V_{uvw} es el volumen del paralelepípedo que determinan $\vec{u}$, $\vec{v}$, $\vec{w}$ y V_{xyz} el correspondiente a $\vec{x}$, $\vec{y}$, $\vec{z}$.

32. Demuestre que si $M \in M(n, I\!R)$ es simétrica y definida positiva entonces

$$\langle x, y \rangle_M = x^t M y$$

es un producto escalar.

33. Considere la siguiente matriz diagonal P:

$$P = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{pmatrix}$$

y el producto escalar $\langle , \rangle_P$

$$\langle x, y \rangle_P = x^t P y.$$

a) Sean $x^t = (x_1, x_2, x_3)$ y $y^t = (y_1, y_2, y_3)$. Determine las fórmulas para la norma de un vector, $\|x\|_P$, la distancia entre vectores, $d_p(x, y)$, y el coseno del ángulo entre vectores, inducido por el producto escalar $\langle , \rangle_P$.

b) Sean $a = (1, 2, 3)^t$ y $b = (1, -2, 1)^t$, calcule $\|a\|_P$, $\|b\|_P$ y $\|b - a\|_P$.

c) Determine el coseno del ángulo entre a y b, según el producto escalar $\langle , \rangle_P$.

d) Observe que con el producto escalar clásico a y b son ortogonales ¿siguen siéndolo con el concepto de ortogonalidad inducido por el producto escalar $\langle , \rangle_P$?, o sea, ¿son a y b P-ortogonales?

e) Con la norma y el concepto de ángulo inducido por el producto escalar $\langle , \rangle_P$, ¿se verifica la ley de los cosenos para el triángulo de lados a, b y $b - a$?:

$$\|b - a\|_P^2 = \|a\|_P^2 + \|b\|_P^2 - 2\,\|a\|_P\,\|b\|_P \cos\theta$$

f) Calcule la proyección de a sobre el vector $1_3 = (1, 1, 1)$, con el concepto de ortogonalidad inducido por $\langle , \rangle_P$.

34. Sea $M = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$

a) M es simétrica claramente, muestre que también es definida positiva.

b) Considere el producto escalar $\langle , \rangle_M$ definido para $I\!R^2$ como $\langle x, y \rangle_M = x^t M y$, la M-norma que induce y calcule $\|e_1\|_M$ y $\|e_2\|_M$, donde $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

c) Determine el M-ángulo entre e_1 y e_2.

d) Determine dos vectores a y b en $I\!R^2$ que sean M-ortogonales y haga una representación gráfica de estos.

35. En la tabla de datos "Calidad del agua" en la página 48 se muestran las mediciones de algunas variables que tienen que ver con la calidad del agua, medidas en varios puntos del embalse La Garita y algunos ríos de Alajuela que lo surten, registradas en el año 1984 por el Laboratorio Químico del ICE.

Considere los siguientes nombres para las variables o columnas de la tabla anterior: x_1, índice de calidad del agua, x_2, DBO, x_3, PH, x_4, sólidos totales y x_5, fosfatos. Considere también la matriz diagonal D_p como se define en el ejercicio 5 con $p_1 = p_2 = \cdots = p_n = \frac{1}{n}$, $n = 9$ y los conceptos de norma, distancia, ángulos y proyecciones que induce.

a) Calcule:

1) $\overline{x}_1, \ \mathrm{Proy}_{1_n} x_1, \ \mathrm{Var}(x_1)$ y σ_{x_1}

2) $\overline{x}_2, \ \mathrm{Proy}_{1_n} x_2, \ \mathrm{Var}(x_2)$ y σ_{x_2}

3) $\overline{x}_5, \ \mathrm{Proy}_{1_n} x_5, \ \mathrm{Var}(x_5)$ y σ_{x_5}

b) Haga una representación geométrica de los vectores x_1, y 1_n como flechas con un origen común y represente los vectores $\overline{x}_1 1_n$, $x_1 - \overline{x}_1 1_n$.

c) Calcule $\text{Cov}(x_1, x_2)$ y $r(x_1, x_2)$ (el coeficiente de correlación). Haga un dibujo de los vectores x_1, $\overline{x}_1 1_n$, $x_1 - \overline{x}_1 1_n$ y x_2, $\overline{x}_2 1_n$, $x_2 - \overline{x}_2 1_n$ representando el ángulo θ tal que $\cos\theta = r(x_1, x_2)$. Justifique que $-1 \leq r(x, y) \leq 1$, cualesquiera sean las variables x y y.

36. Considere la misma tabla de datos utilizada en el ejercicio anterior.

La matriz de varianzas y covarianzas, para esta tabla de datos es una matriz $V_{5\times 5} = (v_{i,j})$ tal que $v_{i,j} = \text{Cov}(x_i, x_j)$, $i, j = 1, \cdots, 5$. Observe que si $i = j$ entonces $\text{Cov}(x_i, x_i) = \text{Var}(x_i)$. Muestre que

$$V = X^t D_p X$$

donde X es la tabla de datos centrada, es decir, las columnas de X son $x_i - \overline{x}_i 1_n$, para $i = 1, \cdots, 5$ respectivamente.

Además, si $D_{1/\sigma}$ es la matriz diagonal, con diagonal:

$$\{1/\sigma_1, 1/\sigma_2, 1/\sigma_3, 1/\sigma_4, 1/\sigma_5\}$$

donde σ_i es la desviación estándar de x_i, demuestre que

$$R = D_{1/\sigma} V D_{1/\sigma}$$

es la matriz cuyos elementos r_{ij} son la correlación lineal entre las variables x_i y x_j, $i, j = 1, \cdots, 5$, llamada matriz de correlaciones.

Calcule V y R. Establezca, además, cuáles parejas de variables están más correlacionadas, y cuáles menos correlacionadas.

37. Todas las distancias que se han inducido con la definición de los productos escalares, a partir de matrices definidas positivas, resultan en generalizaciones al concepto de distancia euclídea clásica:

$$d_1(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

por lo cual son llamadas distancias euclídeas. Otras dos distancias, **no euclídeas**, que se utilizan con alguna frecuencia son:

$$\begin{array}{rcl} d_2(x, y) & = & \max_{i=1,\cdots,n}\{|x_i - y_i|\} \\ d_3(x, y) & = & \sum_{i=1}^{n} |x_i - y_i| \end{array}$$

a) Si x, y y z son vectores en $I\!R^n$, demuestre que d_2 y d_3 satisfacen:

$$\begin{array}{ll} (1) & d(x,y) = 0 \text{ si y solo si } x = y \\ (2) & d(x,y) = d(y,x) \\ (3) & d(x,y) \leq d(x,z) + d(z,y) \end{array}$$

b) Considerando las distancias d_1, d_2 y d_3 definidas para vectores en $I\!R^2$, haga una representación gráfica del conjunto $\{(a,b) \in I\!R^2 / d_i((a,b),(0,0)) = 1\}$, para cada distancia d_i, $i = 1, 2, 3$.

Capítulo 6

Rectas y planos

En este capítulo nos ocuparemos de la descripción geométrica de los subconjuntos de $I\!R^n$:

$$\{x = p + t\vec{v} | t \in I\!R\}, \qquad \{x = p + t\vec{u} + s\vec{v} | t, s \in I\!R\} \quad \text{y}$$

$$\{(x_1, x_2, \ldots, x_n) \in I\!R^n | a_1x_1 + a_2x_2 + \ldots + a_nx_n = d\}$$

donde $p, \vec{u}, \vec{v}$ y $(a_1, \ldots, a_n)$ son elementos dados en $I\!R^n$ y $d \in I\!R$. Así se obtendrán interpretaciones geométricas para el conjunto solución de una ecuación lineal en n variables. Finalmente se estudiará el problema del cálculo de distancias entre puntos y los conjuntos citados y entre ellos.

6.1 Descripción vectorial de una recta

Sean P y $\vec{v}$ en $I\!R^3$, P interpretado geométricamente como un punto y $\vec{v}$ como una flecha que describe cierta dirección. Y considere la recta ℓ que contiene a P con la dirección del vector $\vec{v}$, como se muestra en la figura 6.1:

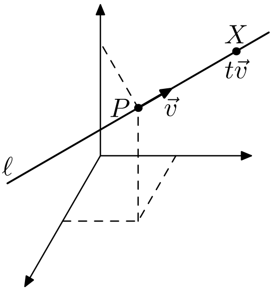


Figura 6.1: Recta $\ell(P, \vec{v})$, que contiene a P en la dirección de $\vec{v}$.

Observe en la figura 6.1 que:

- Para todo punto X en la recta ℓ, el vector $\overrightarrow{PX}$ es paralelo a $\vec{v}$ (tiene la misma dirección o dirección contraria) y como consecuencia:

$$\overrightarrow{PX} = t\vec{v} \quad \text{para algún } t \in I\!R$$

- Además $\overrightarrow{PX} = X - P$, luego:

$$\begin{array}{rrcl} & X - P & = & t\vec{v} \\ \Longrightarrow & X & = & P + t\vec{v} \end{array}$$

Por lo tanto, $\forall X$ en la recta ℓ existe $t \in I\!R$ tal que

$$X = P + t\vec{v}.$$

E inversamente, si un punto Q se puede describir en la forma $Q = P + \alpha\vec{v}$, para cualquier $\alpha \in I\!R$, entonces Q es un punto de la recta ℓ: dado $\alpha \in I\!R$ arbitrario, cualquier punto Q de la forma:

$$Q = P + \alpha\vec{v}$$

es un punto de la recta ℓ, porque $Q - P = \alpha\vec{v}$ de manera que el vector $\overrightarrow{PQ}$ es paralelo a $\vec{v}$ y como P pertenece a ℓ entonces necesariamente $Q \in \ell$. Vea figura 6.2.

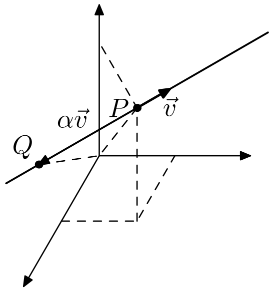


Figura 6.2: Punto Q, en la recta $\ell(P, \vec{v})$.

6.1.1 Ecuación vectorial de una recta

La discusión anterior justifica la siguiente definición.

Definición 6.1 (Rectas en $I\!R^n$)

Se llama recta ℓ que contiene a P en la dirección de $\vec{v}$, y se denota $\ell(P, \vec{v})$, al conjunto de puntos

$$\{X \in I\!R^n / X = P + t\vec{v} \text{ para algún } t \in I\!R\}$$

También se dice que $X = P + t\vec{v}$ es una ecuación vectorial de la recta $\ell(P, \vec{v})$.

Ejemplo 6.1 Determine una ecuación vectorial para la recta que contiene los puntos $A = (-1, 2, 2)$ y $B = (3, -1, 6)$, ilustrada en figura 6.3.

Solución: Como A y B son puntos de la recta, el vector $\overrightarrow{AB}$ tiene la dirección de esta, de manera que:

$$\begin{aligned} \vec{v} = \overrightarrow{AB} &= B - A \\ &= (3, -1, 6) - (-1, 2, 2) \\ &= (4, -3, 4) \end{aligned}$$

Así, una ecuación vectorial para la recta es:

$$(x, y, z) = (-1, 2, 2) + t(4, -3, 4)$$

■

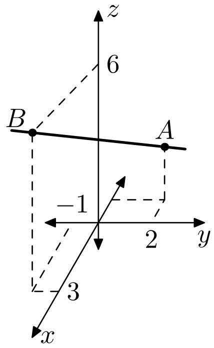


Figura 6.3: Recta por los puntos A y B.

Observación:

- Naturalmente, la descripción de la recta $\ell(P, \vec{v})$ mediante una ecuación vectorial **no es única**, puesto que el punto P y el vector $\vec{v}$ se pueden elegir de infinidad de maneras.

Definición 6.2 (Rectas paralelas y perpendiculares)

Dos rectas $\ell_1(P, \vec{v})$ y $\ell_2(Q, \vec{u})$ son paralelas si $\vec{v}$ y $\vec{u}$ son vectores paralelos y se dicen perpendiculares si $\vec{u}$ y $\vec{v}$ son perpendiculares.

Ejemplo 6.2 Si $\vec{u} = (1, 0, -1)$, la recta $\ell((1, 2, 3), \vec{u})$ es perpendicular a la recta del ejemplo 6.1,

$$(x, y, z) = (-1, 2, 1) + t(4, -3, 4),$$

porque $\vec{v} \cdot \vec{u} = (4, -3, 4) \cdot (1, 0, -1) = 0$.

Nota Observe que dos rectas perpendiculares en $I\!R^n$, $n > 2$, no necesariamente se intersecan.

$\square$

6.1.2 Ecuaciones paramétricas escalares y simétricas de rectas en $I\!R^3$

Dada la ecuación vectorial para una recta $\ell(P, v)$ en $I\!R^3$:

$$(x, y, z) = (p_1, p_2, p_3) + t(v_1, v_2, v_3)$$

si se efectúan las operaciones vectoriales indicadas, se tiene que:

$$(x, y, z) = (p_1 + tv_1, p_2 + tv_2, p_3 + tv_3)$$

Luego, las coordenadas x, y, z de un punto cualquiera de la recta satisfacen:

$$\left\{ \begin{array}{ccc} x & = & p_1 + tv_1 \\ y & = & p_2 + tv_2 \\ z & = & p_3 + tv_3 \end{array} \right.$$

Ecuaciones conocidas con el nombre de **ecuaciones paramétricas escalares** de la recta ℓ. Por otra parte, como el parámetro t es uno y el mismo en las tres ecuaciones anteriores, despejándolo de estas se obtiene:

$$t = \frac{x - p_1}{v_1} = \frac{y - p_2}{v_2} = \frac{z - p_3}{v_3}$$

Todo esto siempre que $v_1 \neq 0$, $v_2 \neq 0$ y $v_3 \neq 0$. A estas últimas ecuaciones se les llama **ecuaciones simétricas** de la recta ℓ.

Ejemplo 6.3 Dadas las siguientes ecuaciones simétricas,

$$\frac{x - 2}{-3} = \frac{y}{2} = z - 1$$

determine dos puntos de esta recta y un vector perpendicular a ella.

a) **Puntos de la recta:** el punto $A = (2, 0, 1)$ pertenece a la recta, puesto que con $x = 2, y = 0$ y $z = 1$ los tres cocientes de las ecuaciones simétricas son iguales a cero. De esta misma manera es fácil determinar otros puntos (x, y, z) de la recta, por ejemplo, los tres cocientes son iguales a 1 si $x = -1, y = 2$ y $z = 2$, luego otro punto de la recta es $B = (-1, 2, 2)$.

b) **Vectores en la dirección de la recta:** cuando se hizo la deducción de las ecuaciones simétricas arriba, se pudo observar que los denominadores de los cocientes son las coordenadas de un vector en la dirección de la recta, de manera que este es $\vec{v} = (-3, 2, 1)$. Por otra parte, también se puede calcular un vector $\vec{v}$ en la dirección de la recta, a partir de los puntos A y B obtenidos anteriormente: $\vec{v} = B - A = (-1, 2, 2) - (2, 0, 1) = (-3, 2, 1)$.

c) **Vectores perpendiculares:** un vector perpendicular a esta recta es cualquier vector $\vec{n} = (a, b, c)$ que satisfaga:

$$(-3, 2, 1) \cdot (a, b, c) = 0$$

o lo que es lo mismo

$$-3a + 2b + c = 0$$

Así, por ejemplo, $(a, b, c) = (1, 1, 1)$ y $(a, b, c) = (-2, -2, -2)$ son vectores perpendiculares a la recta.

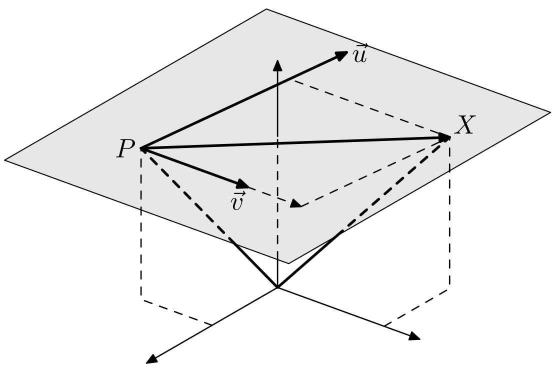


Figura 6.4: Plano $\mathcal{P}(P, \vec{u}, \vec{v})$, que contiene a P en la dirección de $\vec{u}$ y $\vec{v}$.

6.2 Descripción vectorial de los puntos de un plano

El concepto de plano en un espacio $I\!R^n$ nace de manera natural de la idea vivencial de plano que se tiene para el espacio $I\!R^3$. Considere un punto P en un plano y dos vectores $\vec{u}$ y $\vec{v}$ no paralelos y en la dirección del plano, como se muestra en la figura 6.4 en la página 194.

Observe, tanto en la figura 6.4 como en 6.5, que para cualquier punto X en el plano, el vector $\overrightarrow{PX}$ es un vector localizado con punto inicial y final en el plano. Luego el vector $\overrightarrow{PX}$ se puede escribir como combinación lineal de los vectores $\vec{u}$ y $\vec{v}$ dados: $\overrightarrow{PX} = t\vec{u} + s\vec{v}$.

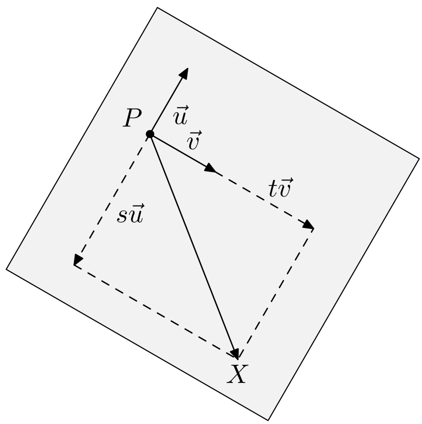


Figura 6.5: $\overrightarrow{PX}$ es combinación lineal de $\vec{u}$ y $\vec{v}$.

Para reconocer mejor este hecho, en el anterior gráfico se ilustra, con otra posición para $\overrightarrow{PX}$, que el vector $\overrightarrow{PX}$ siempre se puede reconocer como la diagonal de cierto paralelogramo con lados $t\vec{u}$ y $s\vec{v}$. de manera que:

$$\overrightarrow{PX} = t\vec{u} + s\vec{v}$$

para algunos valores t y s en $I\!R$.

Como también $\overrightarrow{PX} = X - P$, entonces $X = P + \overrightarrow{PX}$ luego, para todo punto X del plano:

$$X = P + t\vec{u} + s\vec{v}$$

E inversamente, si un punto Q se puede escribir en la forma: $Q = P + t\vec{u} + s\vec{v}$ para algunos valores t y s en $I\!R$, entonces $\overrightarrow{PQ} = Q - P = t\vec{u} + s\vec{v}$ lo que significa que $\overrightarrow{PQ}$ es combinación lineal de dos vectores que describen la dirección del plano y se puede interpretar como la diagonal de un paralelogramo con lados paralelos a $\vec{u}$ y $\vec{v}$, de manera que el vector localizado $\overrightarrow{PQ}$ está en la dirección del plano y como P es un punto del plano necesariamente Q pertenece al plano.

6.2.1 Ecuación vectorial de un plano

Definición 6.3 (Planos en $I\!R^n$, $n > 2$)

Dado un punto P y dos vectores $\vec{u}$ y $\vec{v}$ no paralelos, se llama plano que contiene a P en la dirección de los vectores $\vec{u}$ y $\vec{v}$ al conjunto:

$$\{X \in I\!R^n | X = P + t\vec{u} + s\vec{v} \quad con\ t\ y\ s\ en\ I\!R\}$$

y se denota $\mathcal{P}(P, \vec{u}, \vec{v})$. La ecuación $X = P + t\vec{u} + s\vec{v}$ se denomina ecuación vectorial del plano $\mathcal{P}(P, \vec{u}, \vec{v})$.

Ejemplo 6.4 Dados los puntos $P = (1, 1, -4)$, $Q = (2, -2, 3)$ y $R = (-3, 1, 4)$ determine la ecuación vectorial del plano que los contiene.

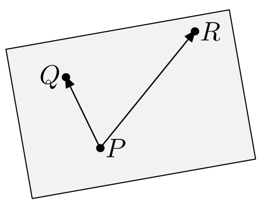


Figura 6.6: Plano por los puntos P, Q y R.

Solución: Observe que los vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$ determinan la dirección del plano:

$$\begin{aligned}\overrightarrow{PQ} &= Q - P \\ &= (2,-2,3)-(1,1,-4) \\ &= (1,-3,7)\end{aligned} \qquad \begin{aligned}\overrightarrow{PR} &= R - P \\ &= (-3,1,4)-(1,1,-4) \\ &= (-4,0,8)\end{aligned}$$

De manera que una ecuación para el plano que contiene los puntos P, Q y R es:

$$(x,y,z)=(1,1,-4)+t(1,-3,7)+s(-4,0,8)$$

La elección del punto P como punto inicial para los vectores que determinan la dirección del plano es arbitraria, igualmente se pudo elegir Q o el punto R, obteniendo en cada caso ecuaciones vectoriales distintas pero que igualmente describen a todos los puntos de este plano.

■

6.2.2 Ecuación normal de un plano en $I\!R^3$

Una nueva forma de describir un plano, válida sólo para planos en $I\!R^3$, se obtiene al reconocer que su dirección también puede ser determinada por un solo vector, pero en este caso, perpendicular al plano.

Sea $P \in I\!R^3$ un punto del plano y $\vec{n}$ un vector perpendicular al plano.

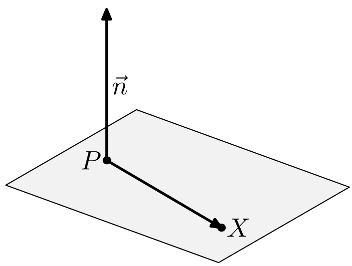


Figura 6.7: Plano que contiene a P, perpendicular a $\vec{n}$.

Todo punto $X \in I\!R^3$ del plano satisface que $\overrightarrow{PX}$ es un vector en la dirección del plano. Y como $\vec{n}$ es perpendicular (al plano), entonces

$$\overrightarrow{PX} \cdot \vec{n} = 0.$$

De esta manera, todo punto X, del plano que contiene a P y es perpendicular a $\vec{n}$, satisface:

$$\begin{array}{rcl} \overrightarrow{PX} \cdot \vec{n} & = & 0 \\ (X - P) \cdot \vec{n} & = & 0 \\ X \cdot \vec{n} & = & P \cdot \vec{n}. \end{array}$$

La ecuación $(X - P) \cdot \vec{n} = 0$ es conocida, a veces, como ecuación punto normal del plano. Además, si se conviene en que $X = (x, y, z)$, $P = (p_1, p_2, p_3)$ y $\vec{n} = (a, b, c)$, la ecuación anterior adquiere la forma:

$$\begin{array}{rcl} X \cdot \vec{n} & = & P \cdot \vec{n} \\ (x, y, z) \cdot (a, b, c) & = & (p_1, p_2, p_3) \cdot (a, b, c) \\ ax + by + cz & = & ap_1 + bp_2 + cp_3 \end{array}$$

donde P y $\vec{n}$ son datos conocidos, entonces el lado derecho de la última ecuación se reduce a una constante $d = ap_1 + bp_2 + cp_3$, y la ecuación adquiere la muy conocida forma:

$$ax + by + cz = d.$$

Definición 6.4 (Ecuación normal de un plano en $I\!R^3$)

Todos los puntos (x, y, z) de un plano que contenga al punto $P = (p_1, p_2, p_3)$ y sea perpendicular al vector $\vec{n} = (a, b, c)$, y sólo estos, satisfacen que:

$$ax + by + cz = d$$

donde $d = ap_1 + bp_2 + cp_3$. Además se dice que el vector $\vec{n}$ es normal al plano, o que el plano contiene a P y es normal a $\vec{n}$.

Ejemplo 6.5 Determine la ecuación normal del plano que contiene los puntos $P(1, 1, -4)$, $Q(2, -2, 3)$ y $R(-3, 1, 4)$.

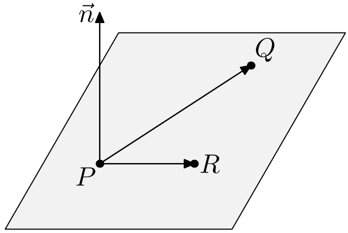


Figura 6.8: Plano por los puntos P, Q y R.

Solución: Con estos datos, para obtener la ecuación normal del plano se requiere determinar un vector $\vec{n} = (a, b, c)$ perpendicular al plano. Como se conoce que los vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$ tienen la dirección del plano, entonces el vector $\vec{n}$ buscado es perpendicular a estos y debe satisfacer:

$$\overrightarrow{PQ} \cdot \vec{n} = 0 \quad \text{y} \quad \overrightarrow{PR} \cdot \vec{n} = 0.$$

En el ejemplo 6.4, se determinó que

$$\overrightarrow{PQ} = (1, -3, 7) \quad \text{y} \quad \overrightarrow{PR} = (-4, 0, 8)$$

entonces $\vec{n} = (a, b, c)$ debe cumplir que:

$$\begin{array}{rcl} (1, -3, 7) \cdot (a, b, c) & = & 0 \\ (-4, 0, 8) \cdot (a, b, c) & = & 0 \end{array}$$

O sea,

$$\begin{array}{rcl} 1a - 3b + 7c & = & 0 \\ -4a + 8c & = & 0 \end{array}$$

Resolviendo este sistema de ecuaciones se llega a que $a = 2c$ y $b = 3c$ donde c se puede elegir libremente. Entonces $(a, b, c) = t(2, 3, 1)$ con $t \in I\!R$ y en particular para $(a, b, c) = (2, 3, 1)$ la ecuación normal del plano buscada es:

$$\begin{array}{rcl} X \cdot \vec{n} & = & P \cdot \vec{n} \\ (x, y, z) \cdot (2, 3, 1) & = & (1, 1, -4) \cdot (2, 3, 1) \\ 2x + 3y + z & = & 1. \end{array}$$

∎

Definición 6.5 (Planos paralelos y perpend. en $I\!R^3$)

Dos planos de $I\!R^3$ se dicen paralelos si sus vectores normales son paralelos y perpendiculares cuando sus vectores normales son perpendiculares.

Observe que la anterior definición es válida sólo para planos en $I\!R^3$. Si $\mathcal{P}_1(P, \vec{v}_1, \vec{u}_1)$ y $\mathcal{P}_2(P, \vec{v}_2, \vec{u}_2$ son dos planos en $I\!R^n$, para que sean paralelos se requiere que

$$\mathcal{C}\ell\{\vec{v}_1, \vec{u}_1\} = \mathcal{C}\ell\{\vec{v}_2, \vec{u}_2\}$$

Esto significa que los subespacios que generan los vectores de direcciones son los mismos. Ahora cuando estos subespacios son ortogonales, ($\vec{v}_1$ es ortogonal a $\vec{v}_2$ y $\vec{u}_2$ así como $\vec{u}_1$). los planos serán perpendiculares. Las ideas de subespacios generados y subespacios ortogonales se introducirán en el siguiente capítulo.

6.3 Hiperplanos

Cuando la anterior forma de describir un plano se aplica a vectores de $I\!R^n$, el resultado son conjuntos de puntos que en $I\!R^2$ constituyen rectas, en $I\!R^3$ planos, y en general se denominan hiperplanos.

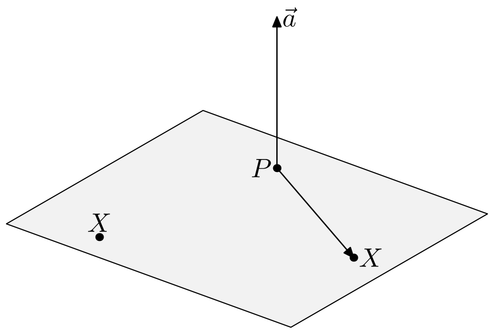


Figura 6.9: Hiperplano que contiene a P ortogonal a $\vec{a}$.

Dado un punto $P \in I\!R^n$ y un vector $\vec{a}$ también en $I\!R^n$, todos los puntos X tales que $\overrightarrow{PX} \perp \vec{a}$, satisfacen:

$$\begin{aligned} \overrightarrow{PX} \cdot \vec{a} &= 0 \\ (X - P) \cdot \vec{a} &= 0 \\ X \cdot \vec{a} &= P \cdot \vec{a} \end{aligned}$$

Si se conviene en que $X = (x_1, x_2, \ldots, x_n)$, $P = (p_1, p_2, \ldots, p_n)$ y $\vec{a} = (a_1, a_2, \ldots, a_n)$, entonces la ecuación última se escribe:

$$\begin{aligned} X \cdot \vec{a} &= P \cdot \vec{a} \\ (x_1, \ldots, x_n) \cdot (a_1, \ldots, a_n) &= (p_1, \ldots, p_n) \cdot (a_1, \ldots, a_n) \\ a_1x_1 + a_2x_2 + \cdots + a_nx_n &= a_1p_1 + a_2p_2 + \cdots + a_np_n \end{aligned}$$

Y como P y $\vec{a}$ son datos dados, el lado derecho de esta última ecuación se reduce a una constante $d = a_1p_1 + a_2p_2 + \cdots + a_np_n$ para obtener la ecuación:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = d$$

que describe al hiperplano de $\mathbb{R}^n$ que contiene a P y es ortogonal a $\vec{a}$.

Definición 6.6 (Hiperplano)

Dado un punto $P \in \mathbb{R}^n$ y un vector $\vec{a} \in \mathbb{R}^n$, se denomina hiperplano que contiene a P ortogonal a $\vec{a}$, al conjunto puntos $X \in \mathbb{R}^n$ tales que $\overrightarrow{PX}$ es un vector perpendicular a $\vec{a}$. O sea, al conjunto

$$\{(x_1, x_2, \ldots, x_n) | a_1x_1 + a_2x_2 + \cdots + a_nx_n = d\}$$

donde $d = a_1p_1 + a_2p_2 + \cdots + a_np_n$.

Al igual que en el caso de los planos, la dirección de un hiperplano está determinada por su vector normal, de manera que si dos hiperplanos tienen vectores normales paralelos, se dice que son paralelos.

Ejemplo 6.6 Determine la ecuación de un hiperplano paralelo a $2x_1 + x_3 - 3x_4 + x_5 = 8$, que contenga el punto $(-1, 1, -1, 0, 6)$.

Solución: Como el hiperplano a determinar es paralelo a

$$2x_1 + x_3 - 3x_4 + x_5 = 8$$

su vector normal puede ser $\vec{a} = (2, 0, 1, -3, 1)$. Además, como $P = (-1, 1, -1, 0, 6)$ es uno de sus puntos, entonces:

$$d = P \cdot \vec{a} = 2 \cdot (-1) + 0 \cdot 1 + 1 \cdot (-1) + -3 \cdot 0 + 1 \cdot 6 = 3$$

y su ecuación es:

$$2x_1 + x_3 - 3x_4 + x_5 = 3$$

■

6.4 Distancias entre puntos, rectas y planos

Los problemas de calcular distancias entre puntos y rectas, puntos y planos o entre dos rectas son problemas de optimización que buscan distancias mínimas entre estos objetos.

Por ejemplo, encontrar la distancia entre el punto Q y un plano $\mathcal{P}(P, \vec{u}, \vec{v})$ requiere determinar el punto R, en este plano, tal que la distancia de Q a R sea mínima. Aunque la solución al problema se puede obtener con la teoría de máximos y mínimos, una vía más directa es reconocer que el punto R, que minimiza la distancia a Q, forma un vector $\overrightarrow{RQ}$ perpendicular al plano, y que la distancia a calcular es la norma de este vector. Dados estos hechos, se utilizan los resultados sobre proyecciones ortogonales para determinar $\overrightarrow{RQ}$, sin necesidad de calcular R.

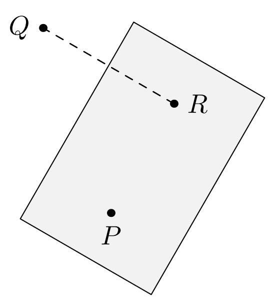


Figura 6.10: Distancia del punto Q a un plano que contiene a P

Esta estrategia, puramente geométrica, resulta ser un valioso procedimiento que puede aplicarse a otros problemas de optimización en $I\!R^n$, como se verá en algunas de las aplicaciones a desarrollar en otros capítulos.

Entonces $d(Q,\mathcal{P})$, la distancia de Q al plano $\mathcal{P}(P,\vec{u},\vec{v})$ es dada por la norma del vector $\overrightarrow{RQ}$, donde R es el punto de intersección entre el plano y la recta que contiene a Q y es perpendicular al mismo plano.

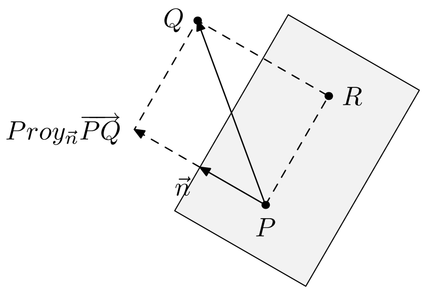


Figura 6.11: Distancia del punto Q al plano $\mathcal{P}(P,\vec{u},\vec{v})$.

En el gráfico 6.11 se puede observar que:

$$\left\|\overrightarrow{RQ}\right\| = \left\|\text{Proy}_{\vec{n}}\overrightarrow{PQ}\right\|.$$

Luego,

$$\begin{aligned}
d(Q,\mathcal{P}) &= \left\|\overrightarrow{RQ}\right\| = \left\|\text{Proy}_{\vec{n}}\overrightarrow{PQ}\right\| \\
&= \left\|\frac{\overrightarrow{PQ}\cdot\vec{n}}{\vec{n}\cdot\vec{n}}\vec{n}\right\| = \left\|\frac{(Q-P)\cdot\vec{n}}{\vec{n}\cdot\vec{n}}\vec{n}\right\| \\
&= \left|\frac{(Q-P)\cdot\vec{n}}{\vec{n}\cdot\vec{n}}\right| \|\vec{n}\| = \frac{|(Q-P)\cdot\vec{n}|}{\|\vec{n}\|^2}\|\vec{n}\| \\
&= \frac{|(Q-P)\cdot\vec{n}|}{\|\vec{n}\|}
\end{aligned}$$

6.5 Ejercicios

1. En cada caso, encuentre ecuaciones a) paramétricas escalares y b) simétricas, para la recta que:

i) Contiene los puntos $(-1, 0, 6)$ y $(2, 3, -5)$.

ii) Contiene el punto $(2, -3, 4)$ con la dirección del vector $(-2, 0, 3)$.

2. Proponga tres ecuaciones vectoriales distintas para la recta que contiene los puntos $(1, -1, 1)$ y $(0, 2, 3)$.

3. En cada caso, determine una ecuación vectorial para la recta con las condiciones dadas:

a) Paralela a $\dfrac{x-2}{2} = y = \dfrac{z+3}{-3}$ y que contenga el punto $(2, 3, -5)$.

b) Que pase por el origen y sea perpendicular a las rectas

$$(x, y, z) = (2 - 3t, -3, 4 + t) \;\; \text{y} \;\; (x, y, z) = (t, 0, 3t)$$

c) Que corte perpendicularmente a la recta $\dfrac{x+1}{3} = \dfrac{3-y}{2}$, $z = 1$, en $(-1, 3, 1)$.

4. En cada caso, dé ejemplos de ecuaciones vectoriales para dos rectas en $I\!R^3$ que cumplan las siguientes condiciones:

i) Que sean perpendiculares y no se corten.

ii) Que sean paralelas y distintas.

5. Obtenga una ecuación normal para el plano con las condiciones dadas, en cada caso:

a) Es perpendicular a la recta $(x, y, z) = (2 - 2t, t, 1 + t)$ y contiene el punto $(2, 1, 2)$.

b) Es paralelo a la recta $(x, y, z) = (2-2t, t, 1+t)$ y contiene a la recta $\dfrac{2-x}{2} = \dfrac{y-1}{-3} = \dfrac{z+3}{4}$.

c) Contiene el punto $(0, -3, 1)$ y es paralelo a plano $(x, y, z) = (2 - 3t + s, t - s, 5t)$.

6. En cada caso, determine una ecuación vectorial para el plano que cumple:

a) Contiene el punto $(2,-6,1)$ y la recta $(x,y,z)=(2-3t,t,5t)$.

b) Es perpendicular a la recta $(x,y,z)=(2-2t,t,1+t)$ y contiene el punto $(1,1,1)$.

c) Contiene las rectas $(x,y,z)=(2-2t,t,1+t)$ y

$$\frac{2-x}{2}=\frac{y}{-3}=\frac{z-1}{4}$$

d) Contiene el punto $(0,-3,1)$ y es paralelo al plano $2x-3y+z=5$.

7. Determine el punto de intersección de la recta

$$\frac{x-2}{2}=y=\frac{z+1}{-5}$$

con el plano $-3x+4y-z=6$.

8. En cada caso, proponga ejemplos de ecuaciones normales para tres planos distintos que satisfagan:

a) Tienen a $(2,-6,1)$ como único punto en común.

b) Su intersección es la recta $(x,y,z)=(2-2t,t,1+t)$.

c) Son mútuamente perpendiculares y contienen el punto $(1,-2,3)$.

d) Son distintos y paralelos.

9. Considere las líneas rectas dadas por las ecuaciones:

$$\left\{\begin{array}{lcl} x & = & 1+t \\ y & = & 1-t \\ z & = & 1+t \end{array}\right. \quad \text{y} \quad \left\{\begin{array}{lcl} x & = & 3-t \\ y & = & -1+t \\ z & = & 1+2t \end{array}\right.$$

a) Pruebe que estas líneas se intersecan ortogonalmente.

b) Muestre que $A=(1,1,1)$, $B=(3,-1,1)$, $C=(7/3,-1/3,7/3)$, forman un triángulo rectángulo y calcule su área.

10. Considere las rectas $l_1=\{(2t,3t-1,t)|t\in I\!R\}$ y $l_2=\{(-t,2t+1,-2t+3)|t\in I\!R\}$.

a) Calcule las ecuaciones cartesianas de dos planos paralelos π_1 y π_2 tales que $l_1 \subset \pi_1$ y $l_2 \subset \pi_2$.

b) Dé una ecuación vectorial para la recta l_3 que es perpendicular a l_1 y a l_2 y contiene el origen.

c) Encuentre el punto M donde l_3 interseca π_1 y el punto N donde l_3 interseca π_2.

d) Utilice c) para calcular la distancia entre las rectas l_1 y l_2.

11. Sea $P = (p_1, p_2)$ y $\vec{a} = (a, b)$.

a) Muestre que todo punto $X = (x, y)$ de la recta que contiene a P y es perpendicular a $\vec{a}$ satisface que:

$$ax + by = ap_1 + bp_2$$

b) Muestre que la ecuación $ax+by = c$ describe una recta de $I\!R^2$ perpendicular a (a, b) y que contiene el punto (x_0, y_0) tal que $c = ax_0 + by_0$. Al vector (a, b) se le llama vector normal de la recta.

c) Demuestre que dos rectas $y = m_1x + b_1$ y $y = m_2x + b_2$ son perpendiculares si y solo si $m_1m_2 = -1$.

12. Demuestre que la distancia de un punto (x_0, y_0, z_0) al plano con ecuación $ax + by + cz + d = 0$ es:

$$\frac{|ax_0 + by_0 + cz_0 + d|}{(a^2 + b^2 + c^2)^{1/2}}$$

Encuentre el punto P del plano $5x - 14y + 2z = -9$ más próximo al punto $Q = (-2, 15, -7)$.

13. Considere la recta $\ell(P, \vec{v})$ de $I\!R^2$ y un vector $\vec{a}$ perpendicular a ℓ.

a) Justifique geométricamente que la distancia de un punto $Q \in I\!R^2$ a la recta ℓ es dada por:

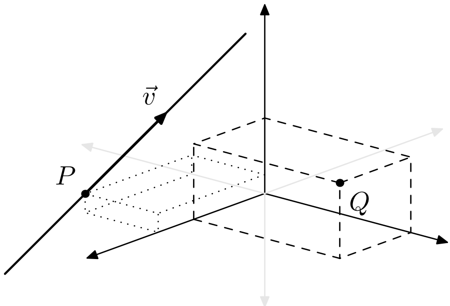
$$d = \frac{|(Q - P) \cdot \vec{a}|}{||\vec{a}||}$$

b) Y si la recta tiene ecuación $ax + by + c = 0$ y el punto $Q = (x_0, y_0)$ demuestre que:

$$d = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$$

c) Muestre también que la distancia de una recta $y = mx+b$ al origen es $|b|/\sqrt{m^2+1}$.

14. Sea $\ell(P, \vec{v})$ una recta en $\mathbb{R}^3$ y Q un punto del mismo espacio.

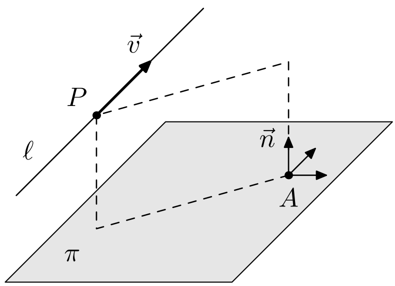


a) Demuestre que la distancia del punto Q a la recta ℓ es dada por:
$d = ||\overrightarrow{PQ} - \text{Proy}_v \overrightarrow{PQ}||$

b) Sea $\vec{a}$ un vector perpendicular a ℓ ¿Porqué $d \neq ||\text{Proy}_a \overrightarrow{PQ}||$?

c) ¿Existe un vector $\vec{a} \perp \ell$ tal que $d = ||\text{Proy}_a \overrightarrow{PQ}||$?

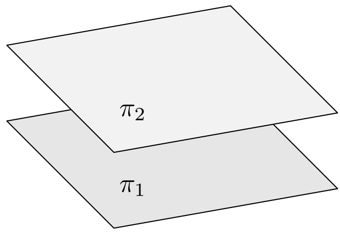
15. Sea $\ell(P, \vec{v})$ una recta paralela al plano $\pi : ax+by+cz+d = 0$.



a) Determine una fórmula para la distancia entre el plano π y la recta ℓ.

b) ¿Cuál es la distancia entre la recta $(x, y, z) = (3-3t, -2+t, 1+2t)$ y el plano $-x+5y-4z = 8$?

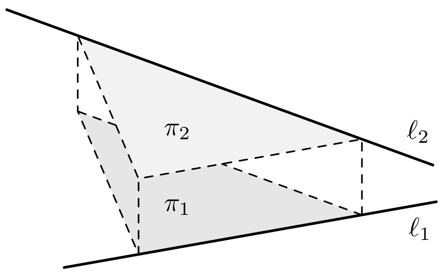
16. Sean $\mathcal{P}_1 : a_1x+b_1y+c_1z+d_1 = 0$ y $\mathcal{P}_2 : a_2x+b_2y+c_2z+d_2 = 0$ dos planos paralelos.



a) ¿Cuál es la distancia del plano $\mathcal{P}_1$ al plano $\mathcal{P}_2$?

b) Considere los planos $2x - 3y + 5z + 6 = 0$ y $-2x + 3y - 5z + 1 = 0$. ¿Qué distancia los separa?

17. Sean ℓ_1 y ℓ_2 dos rectas en $I\!R^n$, $n > 2$. Demuestre que existen dos planos paralelos $\mathcal{P}_1$ y $\mathcal{P}_2$ tales que $\ell_1 \subset \mathcal{P}_1$ y $\ell_2 \subset \mathcal{P}_2$.



Además, considere las rectas:

$$\begin{array}{rcl} \ell_1 & = & \{(2t, 3t-1, t)/t \in I\!R\} \\ \ell_2 & = & \{(-t, 2t+1, -2t+3)/t \in I\!R\} \end{array}$$

a) Calcule la ecuación normal de dos planos con las propiedades indicadas.

b) Determine la distancia entre ℓ_1 y ℓ_2.

c) Dé una ecuación vectorial de la recta perpendicular a ℓ_1 y a ℓ_2, que contiene el origen.

18. Un plano $\mathcal{P}$ tiene las siguientes ecuaciones paramétricas escalares:

$$\begin{aligned} x &= 1 + s + -2t \\ y &= 2 + s + 4t \\ z &= 2s + t \end{aligned}$$

a) ¿Cuáles de los siguientes puntos están en el plano $\mathcal{P}$?: $(0,0,0)$, $(1,2,0)$ y $(2,-3,-3)$.

b) Determine una ecuación vectorial para este plano.

19. En cada caso, determine la intersección de las rectas $\ell_1(P,\vec{v})$ y $\ell_2(Q,\vec{w})$.

a) $P = (1,1,1)$, $\vec{v} = (1,2,3)$, $Q = (2,1,0)$ y $\vec{w} = (2,8,14)$.

b) $P = (0,2,5)$, $\vec{v} = (1,1,-1)$, $Q = (1,2,1)$ y $\vec{w} = (0,1,1)$.

c) $P = (1,-3,2)$, $\vec{v} = (1,2,-1)$, $Q = (4,3,-1)$ y $\vec{w} = (-2,-4,2)$.

20. Sean $P = (1,2,3)$ y $\vec{v} = (1,-2,2)$, observe que todos los puntos X de la recta $\ell(P,\vec{v})$ se pueden expresar en la forma: $X(t) = P + t\vec{v}$. Sea además $Q = (3,3,1)$.

a) Determine $d(t) = ||Q - X(t)||^2$.

b) Demuestre que hay exactamente un punto $X(t_0)$ para el cual $d(t_0)$ es mínima y determínelo.

c) Demuestre que $Q - X(t_0)$ es perpendicular a $\vec{v}$.

21. Sean $X(t) = P + t\vec{v}$ los puntos de una recta $\ell(P,\vec{v})$ en $\mathbb{R}^n$ y Q un punto de $\mathbb{R}^n$ que no pertenece a ℓ. Demuestre que $d(t) = ||Q - X(t)||^2$ es un polinomio en t de grado 2, que alcanza su valor mínimo en un único punto $X(t_0)$ y que $Q - X(t_0)$ es ortogonal a $\vec{v}$.

22. Sean $P = (1,1,1)$, $Q = (1,1,-2)$, $\vec{v} = (2,-1,3)$, $\vec{u} = (2,1,3)$ y $\vec{w} = (0,1,1)$. Demuestre que la recta $\ell(P,\vec{v})$ y el plano $\mathcal{P}(Q,\vec{u},\vec{w})$ se intersecan en un solo punto X_0. Determine X_0.

23. Sea $\mathcal{P}$ el plano determinado por los puntos $(1,1,-2)$, $(3,3,2)$ y $(3,-1,-2)$. Determine:

a) Un vector normal $\mathcal{P}$.

b) Una ecuación cartesiana para $\mathcal{P}$.

c) La distancia del plano al origen.

24. Sean P y Q puntos de un plano $\mathcal{P}$. Demuestre que todo punto de la recta que contiene a P y a Q pertenece al plano $\mathcal{P}$.

25. Encuentre una ecuación para la recta que contiene a $(2, 1, -3)$ y es perpendicular al plano $4x - 3y + z = 5$.

26. Un móvil se mueve en el espacio de modo que en el instante t su posición es $X(t) = (1 - t, 2 - 3t, 2t - 1)$.

a) Demuestre que el móvil se mueve a lo largo de una recta ℓ.

b) ¿En qué instante el móvil toca el plano $2x+3y+2z+1 = 0$?

c) Determine la ecuación cartesiana de un plano perpendicular a ℓ que pasa por $X(2)$.

27. Demuestre que si $\mathcal{P}_1$ y $\mathcal{P}_2$ son planos no paralelos en $\mathbb{R}^3$ entonces su intersección es una recta.

28. Demuestre que las diagonales de un paralelogramo se intersecan en el punto medio.

29. Demuestre que las medianas [1] de un triángulo con vértices $A_1,\ A_2,\ A_3$ se cortan en el punto $\frac{1}{3}(A_1 + A_2 + A_3)$.

[1] Las medianas en un triángulo son segmentos de recta que unen un vértice con el punto medio del lado opuesto.

Capítulo 7

Espacio vectorial $\mathbb{R}^n$

Los espacios vectoriales son un punto de partida, para estudiar muchos conjuntos de objetos de la matemática. Podemos decir que entre estos espacios los más conocidos son $\mathbb{R}$, $\mathbb{R}^2$, y $\mathbb{R}^3$. Sin embargo, tal vez esto podría inducir a una mala interpretación del concepto de espacio vectorial, porque dichos espacios tienen características adicionales que los convierten en espacios vectoriales especiales: en ellos se definen las nociones de distancia y ángulos, conceptos que en general no están presentes en los espacios vectoriales.

Como se verá, lo fundamental que tiene un conjunto para llamarlo espacio vectorial es que sus objetos se puedan sumar y multiplicar por un escalar. Y que estas operaciones tengan las 10 propiedades básicas que se enuncian en la definición siguiente. Así los espacios vectoriales constituyen una plataforma inicial para estudiar ciertos conjuntos relacionados con conceptos como combinación lineal, independencia lineal y bases, entre otros.

Cuando en estos espacios se definen los productos internos y se origina los conceptos de normas, ángulos y proyecciones, los espacios resultantes se denominan espacios con producto interno o espacios vectoriales normados, este es caso de los espacios $\mathbb{R}^n$.

7.1 Definición y propiedades básicas

Definición 7.1 (Espacio vectorial real)

Un conjunto de objetos E se llama espacio vectorial real y sus elementos se llaman vectores, si en este se han definido dos operaciones: a) una suma de vectores (denotada por $+$) y b) un producto de un número real por un vector, que tengan las siguientes propiedades:

$$\begin{array}{rl}
\textit{1.} & \textit{Si } u \in E,\ v \in E \textit{ entonces } u+v \in E \\
\textit{2.} & \textit{Si } u \in E, c \in I\!R \textit{ entonces } cu \in E \\
\textit{Para} & \textit{todo } u, v, w \textit{ en } E \textit{ y todo } a, b \in I\!R \\
\textit{3.} & u+v = v+u \\
\textit{4.} & (u+v)+w = u+(v+w) \\
\textit{5.} & \textit{Existe } 0_e \in E \textit{ tal que: } u + 0_e = u \\
\textit{6.} & \forall u \in E, \textit{ existe } -u \in E \textit{ tal que:} \\
 & u + -u = 0_e \\
\textit{7.} & (a+b)u = au + bu \\
\textit{8.} & a(u+v) = au + av \\
\textit{9.} & (ab)u = a(bu) \\
\textit{10.} & 1u = u
\end{array}$$

7.1.1 Ejemplos: espacio de matrices

Ejemplo 7.1 El conjunto de matrices $M(m, n, I\!R)$, para cada pareja de enteros positivos n y m, es un espacio vectorial real con las operaciones suma de matrices y multiplicación de una matriz por un escalar.

Como se podrá recordar cuando estos conjuntos fueron estudiados, las operaciones mencionadas resultaron con las propiedades enunciadas de (1) a (10) en la definición 7.1.

Ejemplo 7.2 Entre los espacios vectoriales anteriores, los conjuntos $M(n, 1, I\!R)$, constituyen casos particulares especialmente importantes, por lo que recibieron una denominación específica: espacios de vectores columna $I\!R^n$.

Con los 10 axiomas fundamentales para las operaciones de los espacios vectoriales se puede mostrar que el elemento 0_e es único y también es único el elemento inverso aditivo $-u$, para cada $u \in E$. Además se construye una nueva operación, $u - v$, denominada resta o diferencia de vectores, a partir de los inversos aditivos y $+$, en la siguiente forma:

$$\forall \;\; u, v \text{ en } E: \quad u - v = u + (-v).$$

En el siguiente teorema se presentan otras propiedades importantes que se deducen de los 10 axiomas en la definición 7.1 y que son utilizadas con frecuencia al operar con vectores.

Teorema 7.2 *Sea E un espacio vectorial real, $0_e \in E$ el vector cero, y a, b escalares en $I\!R$ entonces para todo $u, v \in E$ se tiene:*

1) $0u = 0_e$
2) $a0_e = 0_e$
3) $(-a)u = -(au) = a(-u)$
4) $au = 0_e \Longrightarrow a = 0$ *o* $u = 0_e$, *o ambos*
5) $au = av$ *y* $a \neq 0 \Longrightarrow u = v$
6) $au = bu$ *y* $u \neq 0_e \Longrightarrow a = b$.

7.1.2 Más ejemplos: espacios de funciones

En los siguientes ejemplos, los vectores son funciones reales de variable real. Para estos casos, la suma de dos vectores, f, g corresponde a la suma de dos funciones, $f + g$, definida corrientemente como:

$$(f + g)(x) = f(x) + g(x).$$

Y el producto de un escalar a por una función f, que define una nueva función af, como:

$$(af)(x) = af(x).$$

Ejemplo 7.3 El conjunto de todas las funciones reales, definidas sobre un intervalo dado.

Ejemplo 7.4 El conjunto de los polinomios en una variable real y con coeficientes reales.

Ejemplo 7.5 El conjunto de los polinomios $P(x)$, de grado menor o igual que n, con coeficientes reales:

$$P_n = \{P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_kx^k | k \leq n\}.$$

La demostración de que cada uno de estos tres conjuntos de funciones, cumplen la definición 7.1, es un ejercicio sencillo aunque largo, donde se verifica que la suma de funciones (o de polinomios), con la multiplicación de un escalar por una función, satisfacen los 10 axiomas de esta definición.

7.2 Subespacios

Algunos subconjuntos de los espacios vectoriales merecen especial atención porque reproducen en sí mismos la estructura del espacio vectorial al que pertenecen.

Definición 7.3 (Subespacio)

Se dice que S es un subespacio de un espacio vectorial E si S es un subconjunto no vacío de E y, con las mismas operaciones de suma y multiplicación por un escalar de E, es en sí mismo un espacio vectorial.

Una caracterización operatoria de los subespacios se logra con el siguiente teorema.

Teorema 7.4 *Sea E un espacio vectorial, y $S \subset E$ un subconjunto no vacío. Si S satisface las dos propiedades:*

$$\begin{array}{rl} (i) & \text{Si } x \in S \text{ y } y \in S \text{ entonces } x + y \in S \\ (ii) & \text{Si } x \in S \text{ y } \alpha \in I\!R \text{ entonces } \alpha x \in S \end{array}$$

entonces S es un subespacio de E.

Demostración: S cumple las 10 propiedades de la definición 7.1 en la página 212: (1) y (2) corresponden a las hipótesis (i) y (ii), las propiedades (3), (4), (7), (8), (9) y (10) se satisfacen porque si x, y y z son elementos de S, como $S \subseteq E$, también lo son de E y en este espacio se sabe que son proposiciones valederas. La propiedad (5) —existe $0_e \in S$ tal que $0_e + x = x, \ \forall\, x \in S$— se deduce de la hipótesis (ii) porque como $S \neq \emptyset$, existe $x \in S$ y por a) del teorema 7.2, $0x = 0_e$ luego $0_e \in S$. Y similarmente se tiene la propiedad(6) — $\forall\, x \in S$ existe $-x \in S$ tal que $-x + x = 0_e$—, esto porque $-x = (-1)x$, según teorema 7.2 tomando $a = -1$ en b), el cual es un elemento de S en virtud de (ii) $\forall\, x \in S$.

∎

Un resultado especialmente importante que se obtiene de la demostración anterior es que si S es un subespacio vectorial entonces $0_e \in S$ (0_e es el cero del espacio vectorial E). Es decir, todo subespacio vectorial contiene al vector cero.

7.2.1 Ejemplos

Ejemplo 7.6 Si E es cualquier espacio vectorial, dos subespacios de E son: $\{0_e\}$ y E, a veces llamados subespacios triviales, Observe que claramente cumplen (i) y (ii) del teorema 7.4.

Ejemplo 7.7 Muestre que si $A \in M(n, m, I\!R)$, entonces

$$S = \{x \in I\!R^m / Ax = 0_n\}$$

es un subespacio vectorial de $I\!R^m$.

Demostración: Observe primero que $S \neq \emptyset$ porque $0_m \in S$. Sean x y y en S entonces $A(x + y) = Ax + Ay = 0_n + 0_n = 0_n$ de manera que $x + y \in S$. Por otra parte, si $x \in S$ y $a \in I\!R$ entonces $A(ax) = aAx = a0_n = 0_n$, por lo tanto $ax \in S$. Luego S es un subespacio de $I\!R^m$.

∎

Nota Las proposiciones

$$\begin{array}{rl} \text{(i)} & \text{Si } x \in S \text{ y } y \in S \text{ entonces } x + y \in S \\ \text{(ii)} & \text{Si } x \in S \text{ y } \alpha \in I\!R \text{ entonces } \alpha x \in S \end{array}$$

en el teorema (7.4) son equivalentes a la proposición

$$\text{Si } x \in S,\ y \in S \text{ y } \alpha \in I\!R \text{ entonces } x + \alpha y \in S$$

lo cual simplifica la escritura de demostraciones sobre subespacios.

Ejemplo 7.8 El subconjunto de:

$$S = \{(t - 2s, -s, t)/t \in I\!R \text{ y } s \in I\!R\}.$$

S es un subespacio vectorial de $I\!R^3$.

Demostración: Si $x \in S$ y $y \in S$ entonces existen $t_1,\ s_1$ y $t_2,\ s_2$ en $I\!R$ tales que

$$x = (t_1 - 2s_1, -s_1, t_1) \text{ y } y = (t_2 - 2s_2, -s_2, t_2)$$

Luego $\forall \alpha \in I\!R$,

$$\begin{array}{rcl} x + \alpha y & = & (t_1 - 2s_1, -s_1, t_1) + \alpha(t_2 - 2s_2, -s_2, t_2) \\ & = & (t_1 - 2s_1 + \alpha t_2 - 2\alpha s_2, -s_1 + -\alpha s_2, t_1 + \alpha t_2) \\ & = & ((t_1 + \alpha t_2) - 2(s_1 + \alpha s_2), -(s_1 + \alpha s_2), t_1 + \alpha t_2) \end{array}$$

De manera que existen $t = t_1 + \alpha t_2$ y $s = s_1 + \alpha s_2$, números reales, tales que

$$x + \alpha y = (t - 2s, -s, t)$$

por lo tanto $x + \alpha y \in S$ y se concluye que S es un subespacio vectorial.

Ejemplo 7.9 Si $\mathcal{F}$ es el espacio vectorial de todas las funciones reales de variable real, el subconjunto F de estas funciones para las que $f(0) = 0$ es un subespacio vectorial de $\mathcal{F}$.

Demostración: Si f y g pertenecen a F, $f(0)=0$ y $g(0)=0$, y si además $a \in I\!R$, entonces

$$(f+ag)(0)=f(0)+ag(0)=0+a0=0$$

luego $f+ag \in F$, por tanto F es un subespacio vectorial de $\mathcal{F}$.

■■■

Ejemplo 7.10 Sea A una la matriz $n\times k$ y V el conjunto de combinaciones lineales de las columnas de A:

$$V=\{b\in I\!R^n/b=Ax,\ \text{para algún } x\in I\!R^k\}.$$

El subconjunto V es un subespacio de $I\!R^n$.

Demostración: Claramente $V\neq\emptyset$, porque $0_n=A0_k$, entonces $0_n\in V$. Ahora, si b y c pertenecen a V y α es cualquier número real existen x y y en $I\!R^k$ tales que $b=Ax$ y $c=Ay$ de manera que

$$b+\alpha c=Ax+\alpha Ay=A(x+\alpha y).$$

Entonces existe $z=x+\alpha y\in I\!R^k$ tal que $b+\alpha c=Az$ por lo tanto $b+\alpha c\in V$.

■■■

El ejemplo 7.8 es un caso particular del ejemplo 7.10, observe que:

$$\begin{pmatrix} t-2s\\ -s\\ t\end{pmatrix}=t\begin{pmatrix}1\\0\\1\end{pmatrix}+s\begin{pmatrix}-2\\-1\\0\end{pmatrix}=\begin{pmatrix}1&-2\\0&1\\1&0\end{pmatrix}\begin{pmatrix}t\\s\end{pmatrix}$$

luego, $S=\{b\in I\!R^3/b=Ax,\ \text{para algún } x\in I\!R^2\}$, donde

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

Describiendo S en la anterior forma, es posible mostrar, tal vez de una manera más sencilla, que el conjunto S, en el ejemplo 7.8, es un subespacio vectorial, usando el mismo esquema de demostración del ejemplo 7.10.

Ejemplo 7.11 Considere una matriz $A \in M(n, m, \mathbb{R})$ y $S = \{x \in \mathbb{R}^m / Ax = b\}$ donde b es un vector no nulo de $\mathbb{R}^n$. Observe que S no es un subespacio vectorial de $\mathbb{R}^m$.

Demostración: Como el vector 0_m no es una solución de $Ax = b$, puesto que $A0_m = 0_n \neq b$, entonces $0_m \notin S$ y S no es subespacio de $\mathbb{R}^m$.

■■■

Ejemplo 7.12 Sea $U = \{A \in M(3, \mathbb{R})/A \text{ es diagonal }\}$. Muestre que U es un subespacio vectorial.

Demostración: La matriz de ceros de orden 3×3 es diagonal, por lo tanto $U \neq \emptyset$. Ahora, sean $A \in U$, $B \in U$ y $t \in \mathbb{R}$, y observe que

$$\begin{aligned} A + tB &= \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} + t \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{pmatrix} \\ &= \begin{pmatrix} a_1 + tb_1 & 0 & 0 \\ 0 & a_2 + tb_2 & 0 \\ 0 & 0 & a_3 + tb_3 \end{pmatrix} \end{aligned}$$

Luego $A + tB$ es también diagonal, lo cual prueba que U es subespacio de $M(3, \mathbb{R})$.

■■■

7.2.2 Tres subespacios típicos de $\mathbb{R}^n$

Entre los subespacios de $\mathbb{R}^n$ tres tipos de ellos merecen especial atención por su interpretación geométrica y la frecuencia con que aparecen en el Algebra Lineal. Estos son:

Las rectas por el origen

Dado un vector $v \in \mathbb{R}^n$ no nulo el subespacio

$$L = \mathcal{C}\ell\{v\} = \{tv/t \in \mathbb{R}\}$$

está constituido por todos los puntos de $\mathbb{R}^n$ en una recta que contiene el origen y la dirección del vector v.

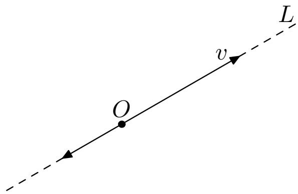


Figura 7.1: Subespacio L: recta por el origen en la dirección de v.

Planos por el origen

Si se consideran dos vectores $u, v \in \mathbb{R}^n$ l.i. el subespacio

$$\mathcal{P} = \mathcal{C}\ell\{u, v\} = \{tu + sv \mid t, s \in \mathbb{R}\}$$

corresponde al conjunto de todos los puntos de un cierto plano de $\mathbb{R}^n$ que contiene el origen en la dirección de los vectores u y v. Es precisamente la descripción vectorial de un plano de puntos $X = P + tv + su$, para el que $P = 0_n$.

Ejemplo 7.13 El subespacio

$$S = \mathcal{C}\ell\{(0,1,1)^t, (1,1,0)^t\}$$

corresponde al plano de $\mathbb{R}^3$ determinado por los puntos $(0,0,0)^t$, $(0,1,1)^t$ y $(1,1,0)^t$, que se extiende en la dirección de los vectores $(0,1,1)^t$, $(1,1,0)^t$. Y por lo tanto, todo punto que pertenezca a este plano se puede escribir en la forma:

$$(x, y, z)^t = t(0,1,1)^t + s(1,1,0)^t$$

Hiperplanos por el origen

Como caso especial entre los subespacios correspondientes al ejemplo 7.7 expuesto en la página 215, se tienen los hiperplanos por el origen, es decir los subespacio de $\mathbb{R}^n$ correspondiente al conjunto solución de una sola ecuación homogénea en n variables:

$$H = \{x \in \mathbb{R}^n | a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}.$$

Este subespacio también se describe en términos del vector $a = (a_1, a_2, \ldots, a_n)^t$ como:

$$H = \{x \in \mathbb{R}^n | a \cdot x = 0\}$$

Lo que nos permite visualizar que el subespacio H está formado por todos los vectores de $\mathbb{R}^n$ que son ortogonales al vector a.

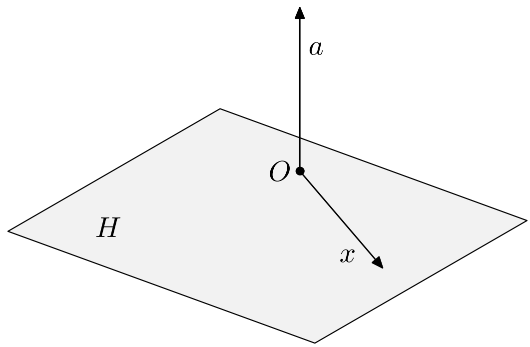


Figura 7.2: Subespacio H: hiperplano por el origen ortogonal a a.

7.3 Combinaciones lineales y conjuntos generadores

Basados, exclusivamente, en la estructura algebraica de los espacios vectoriales —suma de vectores y multiplicación de un escalar por un vector, con sus propiedades— se definen una serie de conceptos como: combinación lineal de vectores, conjuntos generadores, independencia lineal, bases, etc., algunos ya introducidos al estudiar los vectores columna, pero que ahora se retoman en una forma más general.

Definición 7.5 (Combinaciones lineales)

Sea E un espacio vectorial y $\{v_1, v_2, \ldots, v_p\}$, un conjunto de vectores de E. Se llama combinación lineal de los vectores $v_1, v_2, \ldots, v_p$ al vector

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_p v_p$$

cualquiera sea la elección de los escalares $a_1, a_2, \ldots, a_p$. Y al conjunto

$$\mathcal{C}\ell\{v_1, \ldots, v_p\} = \{a_1 v_1 + a_2 v_2 + \cdots + a_p v_p | a_1, a_2, \ldots, a_p \in I\!R\}$$

se le denomina conjunto de combinaciones lineales de $v_1, v_2, \ldots, v_p$.

Esta definición no tiene nada nuevo si se piensa sólo en los vectores columna de $I\!R^n$, sin embargo, se ha repetido para reafirmar que ahora se aplicará a cualquier espacio vectorial, es decir, a vectores v_i que también pueden ser matrices, polinomios o funciones.

Ejemplo 7.14 Sean $v_1, v_2, \ldots, v_k$ vectores de cualquier espacio vectorial E, entonces $S = \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$ es un subespacio de E.

Claramente, cualquiera sea el espacio vectorial E, la suma de dos combinaciones lineales — de un mismo conjunto de vectores — es una nueva combinación lineal de tales vectores, lo mismo que el producto de un escalar por una combinación lineal de dichos vectores. Por lo tanto, la suma de vectores y la multiplicación por escalares son operaciones cerradas en $\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$ y este es un subespacio vectorial.

7.3.1 Conjuntos generadores

Una de las ideas más importantes a asociar con la de espacio vectorial es la de conjunto generador.

Definición 7.6 (Conjunto Generador)

Un conjunto de vectores $\{v_1, v_2, \ldots, v_k\}$, de un espacio vectorial E, se llama un conjunto generador de E si todo $v \in E$ se puede expresar como combinación lineal de los vectores $v_1, v_2, \ldots, v_k$.

Observe que cuando $\{v_1, v_2, \ldots, v_k\}$ genera a E, entonces

$$E = \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}.$$

Porque siempre se tiene que $\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\} \subseteq E$ y cuando ocurre que todo $v \in E$ se puede expresar como combinación lineal de $v_1, v_2, \ldots, v_k$ entonces $E \subseteq \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$.

De esta manera se logra una caracterización importante del espacio vectorial E: todo vector en E es una combinación lineal de un conjunto finito de vectores de E. Pero ... ¿todos los espacios vectoriales tendrán conjuntos finitos con esta característica? La respuesta es no, sin embargo los espacios $\mathbb{R}^n$ y los matriciales si tienen esta particularidad.

Ejemplo 7.15 Considere el espacio vectorial $\mathbb{R}^2$ y sus vectores $e_1 = (1,0)^t, e_2 = (0,1)^t$. Claramente todo vector $(x,y)^t \in \mathbb{R}^2$ se puede escribir como combinación lineal de e_1 y e_2:

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

luego $\mathbb{R}^2 = \mathcal{C}\ell\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ y se dice que $\{e_1, e_2\}$ es un conjunto generador de $\mathbb{R}^2$.

Ejemplo 7.16 Sean $v_1 = (1,2)^t$, $v_2 = (0,3)^t$ y $v_3 = (-1,6)^t$, se puede comprobar que el conjunto $\{v_1, v_2, v_3\}$ es también un conjunto generador de $\mathbb{R}^2$.

Hay que mostrar que para todo vector $(x,y)^t \in \mathbb{R}^2$ existen escalares $a_1,\ a_2,\ a_3$ tales que

$$\begin{pmatrix} x \\ y \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 3 \end{pmatrix} + a_3 \begin{pmatrix} -1 \\ 6 \end{pmatrix}.$$

Esta última proposición de igualdad conduce al sistema de ecuaciones lineales:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & 6 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

en las variables a_1, a_2 ya_3, el cual se transforma a uno equivalente mediante las siguientes operaciones elementales:

$$\begin{pmatrix} 1 & 0 & -1 & x \\ 2 & 3 & 6 & y \end{pmatrix} \xrightarrow{-2f_1 + f_2} \begin{pmatrix} 1 & 0 & -1 & x \\ 0 & 3 & 8 & y-2x \end{pmatrix}$$

$$\xrightarrow{(1/3)f_2} \begin{pmatrix} 1 & 0 & -1 & x \\ 0 & 1 & 8/3 & (y-2x)/3 \end{pmatrix}$$

De lo anterior se sigue que el sistema tiene infinitas soluciones; específicamente, que para cada $(x, y)^t$:

$$\begin{array}{rcl} a_1 & = & x + t \\ a_2 & = & \dfrac{y-2x}{3} - \frac{8}{3}t \qquad \text{con } t \in I\!R \\ a_3 & = & t \end{array}$$

En particular, si se elige $t = 0$ entonces $\forall\ (x, y)^t$

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{y-2x}{3} \begin{pmatrix} 0 \\ 3 \end{pmatrix} + 0 \begin{pmatrix} -1 \\ 6 \end{pmatrix}$$

Conclusión: $I\!R^2 = \mathcal{C}\ell\{(1,2)^t, (0,3)^t, (-1,6)^t\}$.

Ejemplo 7.17 Considere los vectores de $I\!R^3$, $v_1 = (1, 1, -1)^t$, $v_2 = (0, 1, -2)^t$, $v_3 = (-1, 0, -1)^t$. Muestre que el conjunto $\{v_1, v_2, v_3\}$ no genera a $I\!R^3$.

Solución: Nos proponemos demostrar que

$$I\!R^3 \neq \mathcal{C}\ell\{v_1, v_2, v_3\} \tag{7.1}$$

¿Para todo $v = (a, b, c)^t$, existen escalares a_1, a_2, a_3 tales que

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} + a_3 \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} \ ?$$

Reduciendo la matriz aumentada del sistema

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & -2 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

se obtiene:

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 1 & 1 & 0 & b \\ -1 & -2 & -1 & c \end{array}\right) \xrightarrow[]{\substack{-1f_1+f_2 \\ 1f_1+f_3}} \left(\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 0 & 1 & 1 & -a+b \\ 0 & -2 & -2 & a+c \end{array}\right)$$

$$\xrightarrow{2f_2+f_3} \left(\begin{array}{ccc|c} 1 & 0 & -1 & a \\ 0 & 1 & 1 & -a+b \\ 0 & 0 & 0 & -a+2b+c \end{array}\right)$$

Luego, cualquier vector $v = (a, b, c)^t$ tal que $-a + 2b + c \neq 0$ no pertenece a $\mathcal{C}\ell\{v_1, v_2, v_3\}$, por ejemplo $v = (1, 1, 1)^t$. Esto muestra la existencia de vectores de $I\!R^3$ que no pertenecen a $\mathcal{C}\ell\{v_1, v_2, v_3\}$ y por lo tanto demuestra la proposición (7.1).

∎

De este ejemplo se tiene que no es suficiente con tener 3 vectores para generar $I\!R^3$, aunque claramente los tres vectores $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ y $e_3 = (0, 0, 1)$ sí lo generan. Por otra parte, en los ejemplos anteriores se pudo observar que varios conjuntos de vectores pueden generar a un mismo espacio vectorial. De todo esto surgen dos preguntas centrales:

1. ¿Para cada espacio vectorial, existirá un conjunto finito de vectores que lo generan?

2. Cuando hay un conjunto finito que genera el espacio ¿Cuál es el menor número de elementos necesarios para generarlo?

La primera pregunta ya se había formulado y se respondió que no. Específicamente, el espacio de funciones reales de variable real, no tiene un conjunto de funciones que lo generen, y tampoco el espacio de los polinomios en una variable con coeficientes reales tiene un conjunto finito que lo genere.

La segunda pregunta nos lleva al concepto de independencia lineal, ya visto para el caso de vectores columna y se responderá más adelante.

7.3.2 Dependencia e independencia lineal

Igual que en el caso visto de vectores columna, si $\{u_1, \ldots, u_p\}$ son vectores de un espacio vectorial E cualquiera, se dice que son l.d. (linealmente dependientes), si al menos uno de ellos se puede expresar como combinación de los restantes y que son l.i. (linealmente independientes) si no son l.d. lo cual se caracteriza en la siguiente forma:

Definición 7.7 (Dependencia e independencia lineal)

Un conjunto $\{u_1, u_2, \ldots, u_p\}$ de vectores de un espacio vectorial E se llama linealmente dependiente, l.d., si existen escalares $a_1, a_2, \ldots, a_p$ **no todos nulos** *tales que*

$$a_1 u_1 + a_2 u_2 + \cdots + a_p u_p = 0_e$$

Es decir, si uno de ellos es combinación lineal de los restantes. Y se llaman linealmente independientes, l.i., si

$$a_1 u_1 + a_2 u_2 + \cdots + a_p u_p = 0_e \Longrightarrow a_1 = a_2 = \cdots = a_p = 0.$$

Observe que si un conjunto de vectores contiene a 0_e entonces es l.d. y si está compuesto por un único vector no nulo entonces es l.i.

Teorema 7.8 *Sea $V = \mathcal{C}\ell\{v_1, v_2, \ldots, v_p\}$, donde cada v_i pertenece a un espacio vectorial E, entonces cualquier conjunto de $p+1$ vectores en V es l.d.*

De este teorema se deriva que cualquier subconjunto finito de $\mathcal{C}\ell\{v_1, v_2, \ldots, v_p\}$, con más de p elementos es l.d. Su demostración es un ejercicio donde se aplica inducción matemática sobre p y puede ser consultada en [2].

7.4 Bases

Entre los conjuntos generadores de un espacio vectorial E, aquellos para los cuales cada vector de E se expresa como combinación lineal única de ellos, adquieren especial importancia.

Definición 7.9 (Bases de un e.v.)

Un conjunto de vectores $\{v_1, v_2, \ldots, v_k\}$ *de un espacio vectorial* E*, es una base de este espacio si y solo si todo vector* $v \in E$ *se puede expresar como combinación lineal* **única** *de los vectores* $v_1, v_2, \ldots, v_k$*.*

Ejemplo 7.18 Muestre que

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

es una base de $I\!R^4$.

Solución: Para todo $v = (a, b, c, d)^t \in I\!R^4$ deben existir valores $x_1,\ x_2,\ x_3$ y x_4 **únicos** tales que

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -1 \\ -2 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{7.2}$$

o sea, el sistema

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & -1 & 0 \\ -1 & 0 & -2 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

debe tener solución única. Su matriz aumentada es equivalente a la matriz

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2a + c \\ 0 & 1 & 0 & 0 & -3a + b - 2c \\ 0 & 0 & 1 & 0 & -a - c \\ 0 & 0 & 0 & 1 & -3a + b - 2c + d \end{array} \right)$$

de manera que para cada $v = (a, b, c, d)^t \in I\!R^4$, los únicos valores $x_1,\ x_2,\ x_3,\ x_4$ que satisfacen (7.2) son:

$$\begin{array}{rcl} x_1 & = & 2a + c \\ x_2 & = & -3a + b - 2c \\ x_3 & = & -a - c \\ x_4 & = & -3a + b - 2c + d \end{array}$$

■

Ejemplo 7.19 Sean $A_1 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ y $A_4 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$. Muestre que $\{A_1, A_2, A_3, A_4\}$ es una base de $M(2, I\!R)$.

Solución: Se debe demostrar que cualquier matriz

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

de $M(2, I\!R)$ se escribe como combinación lineal de A_1, A_2, A_3 y A_4, de manera única, es decir, existen escalares x_1, x_2, x_3, x_4 únicos tales que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = x_1A_1 + x_2A_2 + x_3A_3 + x_4A_4$$

Sustituyendo por las matrices respectivas y efectuando las operaciones se obtiene:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} x_1 + x_3 & -x_1 + x_2 + x_4 \\ x_2 & x_1 + x_3 + x_4 \end{pmatrix}$$

lo cual, nuevamente, representa un sistema de ecuaciones lineales:

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

En este caso, la forma escalonada reducida de la matriz aumentada es:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -a - b + c + d \\ 0 & 1 & 0 & 0 & c \\ 0 & 0 & 1 & 0 & 2a + b - c - d \\ 0 & 0 & 0 & 1 & d - a \end{array}\right)$$

lo que refleja que la solución es única. Entonces cualquiera sea la matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, esta se escribe de manera única como:

$$A = (-a - b + c + d)A_1 + cA_2 + (2a + b - c - d)A_3 + (d - a)A_4$$

luego $\{A_1, A_2, A_3, A_4\}$ es una base de $M(2, I\!R)$.

∎

Observe que si un conjunto $\{v_1, v_2, \ldots, v_k\}$ es una base de E, entonces el cero de E, 0_e, se expresa como combinación lineal única de $v_1, v_2, \ldots, v_k$,

$$0_e = a_1v_1 + a_2v_2 + \cdots + a_kv_k$$

Y como $0_e = 0v_1 + 0v_2 + \cdots + 0v_k$, por la unicidad se tiene que $a_1 = a_2 = \cdots = a_k = 0$, luego $\{v_1, v_2, \ldots, v_k\}$ es l.i. Esto muestra una de las direcciones del siguiente teorema.

Teorema 7.10 *Un conjunto $\{v_1, v_2, \cdots, v_k\}$ de vectores de un espacio vectorial E es una base de E si y solo si el conjunto es l.i. y genera a E.*

Demostración: "$\Leftarrow$": suponga que $\{v_1, v_2, \cdots, v_k\}$ es l.i. y genera a E. Entonces todo $v \in E$ es combinación lineal de $v_1, v_2, \cdots, v_k$ y suponga que un cierto vector v se expresa como combinación lineal de las dos siguientes maneras:

$$\begin{aligned} & v = a_1v_1 + a_2v_2 + \cdots + a_kv_k \\ \text{y} \quad & v = b_1v_1 + b_2v_2 + \cdots + b_kv_k \end{aligned}$$

luego

$$\begin{aligned} & v - v = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_k - b_k)v_k \\ \Rightarrow \quad & 0_e = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_k - b_k)v_k \\ \Rightarrow \quad & a_1 - b_1 = a_2 - b_2 = \cdots = a_k - b_k = 0 \end{aligned}$$

porque $\{v_1, v_2, \cdots, v_k\}$ es l.i. De esto se deduce que $a_i = b_i$, para $i = 1, \ldots, k$, o sea, v se expresa como combinación lineal única de $v_1, v_2, \cdots, v_k$.

■■■

Ejemplo 7.20 Sean $v_1 = (1, 0, 2, 3)^t$, $v_2 = (0, -3, 0, 2)^t$ y $v_3 = (0, 0, 2, 0)^t$, y la matriz $A = (v_1, v_2, v_3)$. Demuestre que $\{v_1, v_2, v_3\}$ es una base del espacio S:

$$S = \{y \in I\!\!R^4 | y = Ax \text{ y } x \in I\!\!R^3\}.$$

Solución: S es el espacio generado por las columnas de A, o sea, $S = \mathcal{C}\ell\{v_1, v_2, v_3\}$. Por otra parte, fácilmente se verifica que

$$a_1v_1 + a_2v_2 + a_3v_3 = 0_4 \Longrightarrow a_1 = a_2 = a_3 = 0$$

entonces $\{v_1, v_2, v_3\}$ es l.i. y genera S por lo tanto es una base de S.

Teorema 7.11 *Sea E un espacio vectorial que tiene una base finita, entonces toda base de E tiene el mismo número de elementos.*

Demostración: Sean $\mathcal{B}_1$ y $\mathcal{B}_2$ dos bases de E y suponga que tienen k y p elementos, respectivamente. Como $\mathcal{B}_1$ genera a E entonces $E = \mathcal{C}\ell\{\mathcal{B}_1\}$ y cualquier conjunto con más es k elementos en E es l.d. por teorema 7.8, luego como $\mathcal{B}_2$ es l.i. se concluye que $p \leq k$. Intercambiando las bases $\mathcal{B}_1$ y $\mathcal{B}_2$ en la argumentación anterior se concluye que $k \leq p$. Luego $k = p$.

7.4.1 Dimensión

Los teoremas 7.10 y 7.11 responden a la pregunta 2, en la página 224 y motivan la siguiente definición.

Definición 7.12 (Dimensión de un espacio vectorial)

Si E es un espacio vectorial y tiene una base con n elementos, entonces el entero n es la dimensión de E.

Ejemplo 7.21 Sea S el conjunto de los vectores columna $(x_1, x_2, x_3, x_4)^t$ tales que:

$$\begin{array}{lcl} -2x_1 + 3x_2 - x_4 & = & 0 \\ x_1 - 2x_2 - x_3 & = & 0 \end{array}$$

Determine el menor número de vectores que generan a S.

Solución: El menor número de vectores que generan un subespacio, es la dimensión del subespacio y para conocerlo se requiere determinar el número de elementos de una base de S.

Este subespacio es un caso particular del ejemplo (7.7), conjunto solución de un sistema homogéneo de ecuaciones lineales y, en todos estos casos, resolviendo el sistema se obtiene un conjunto de generadores:

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

$$\xrightarrow{f_1 \leftrightarrow f_2} \left(\begin{array}{cccc|c} 1 & -2 & -1 & 0 & 0 \\ -2 & 3 & 0 & -1 & 0 \end{array}\right)$$

$$\xrightarrow{2f_1 + f_2} \left(\begin{array}{cccc|c} 1 & -2 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \end{array}\right)$$

$$\xrightarrow{-1f_2} \left(\begin{array}{cccc|c} 1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow{-2f_2 + f_1} \left(\begin{array}{cccc|c} 1 & 0 & 3 & 2 & 0 \\ 0 & 1 & 2 & 1 & 0 \end{array}\right).$$

De manera que si $x_3 = t$ y $x_4 = s$ las soluciones al sistema son de la forma:

$$\begin{array}{rcrcr} x_1 & = & -3t & - & 2s \\ x_2 & = & -2t & - & s \\ x_3 & = & t & & \\ x_4 & = & & & s \end{array} \qquad \text{con } t \text{ y } s \text{ en } I\!R$$

entonces si $x^t = (x_1, x_2, x_3, x_4)$, S se puede describir como:

$$S = \{x | x = t(-3, -2, 1, 0)^t + s(-2, -1, 0, 1)^t \ t, s \in I\!R\}$$

o equivalentemente,

$$S = \mathcal{C}\ell\{(-3, -2, 1, 0)^t, (-2, -1, 0, 1)^t\}.$$

Como claramente los vectores $(-3, -2, 1, 0)^t, (-2, -1, 0, 1)$ son l.i. entonces una base de S es:

$$\mathcal{B} = \{(-3, -2, 1, 0)^t, (-2, -1, 0, 1)^t\}$$

y su dimensión, o menor número de vectores necesarios para generarlo, es 2.

∎

Teorema 7.13 *Si E es un espacio vectorial de dimensión n y v_i, $i = 1, 2, \ldots, k$, son k vectores l.i. de E:*

a) Si $k < n$, entonces existen $v_{k+1}, \ldots, v_n$ vectores en E tales que $\{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ es una base de E.

b) Si $k = n$ entonces $\{v_1, v_2, \ldots, v_n\}$ generan a E.

Corolario 7.14 *Sea E un espacio vectorial de dimensión n entonces:*

(i) Cualquier conjunto de n vectores que genere a E es una base.

(ii) Cualquier conjunto de n vectores l.i. de E es una base.

Ejemplo 7.22 Sea $T = \{A \in M(2, I\!R)/A$ es triangular inferior$\}$ y $A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Demuestre que $\mathcal{B} = \{A_1, A_2, A_3\}$, es una base de T.

Demostración: Observe que cualquier matriz A, triangular inferior, se puede escribir como:

$$A = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

luego $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ generan T y como son l.i. forman una base, de manera que la dimensión de T es 3.

Por otra parte $A_1 \in T$, $A_2 \in T$ y $A_3 \in T$ y para mostrar que son una base es suficiente con demostrar que son l.i. por el corolario 7.14.

$\mathcal{B}$ es un conjunto l.i.

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = x_1 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + x_3 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Longrightarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} x_1 + x_3 & 0 \\ x_1 + x_2 & x_2 + x_3 \end{pmatrix}$$

$$\Longrightarrow \left\{ \begin{array}{lcl} 0 & = & x_1 + x_3 \\ 0 & = & x_1 + x_2 \\ 0 & = & x_2 + x_3 \end{array} \right.$$

$$\Longrightarrow x_3 = -x_1, \; x_2 = -x_1, \; \text{y} \; -2x_1 = 0$$

$$\Longrightarrow x_1 = x_2 = x_3 = 0$$

Luego $\mathcal{B}$ es l.i. y por lo tanto es una base de T.

■■■

7.4.2 Conjuntos generadores de hiperplanos

Recordemos que los subespacio de $I\!R^n$:

$$H = \{x \in I\!R^n / a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}$$

corresponden a hiperplanos que contienen el origen y todos sus vectores son ortogonales al vector $a = (a_1, a_2, \ldots, a_n)^t$.

Nos interesa determinar una base para estos subespacios: la ecuación

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

permite despejar una de las variables x_i, en términos de las restantes, de manera que cualquier solución $x = (x_1, x_2, \ldots, x_n) \in H$ depende de la elección de $n-1$ parámetros. Para aclarar esta idea, suponga que $a_n \neq 0$, entonces

$$x_n = (-a_1x_1 - a_2x_2 - \cdots - a_{n-1}x_{n-1})/a_n.$$

Luego es posible describir cada solución $(x_1, \ldots, x_n)$ de la ecuación homogénea en la siguiente forma:

$$\begin{array}{lclllll} x_1 & = & x_1 & & & & \\ x_2 & = & & x_2 & & & \\ \vdots & & & & \ddots & & \\ x_{n-1} & = & & & & & x_{n-1} \\ x_n & = & -\frac{a_1}{a_n}x_1 & -\frac{a_2}{a_n}x_2 & -\cdots & & -\frac{a_{n-1}}{a_n}x_{n-1} \end{array}$$

Si se piensa que $x_1, x_2, \ldots, x_{n-1}$ son parámetros que varían libremente en $I\!R$, entonces:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{a_1}{a_n} \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -\frac{a_2}{a_n} \end{pmatrix} + \cdots + x_{n-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{a_{n-1}}{a_n} \end{pmatrix}$$

con lo que se observa que los $n-1$ vectores columna de la derecha, que denotamos $v_1, v_2, \ldots, v_{n-1}$, son vectores de H y generan a H. También resulta fácil verificar que son l.i. y por lo tanto forman una base de H. Así

$$H = \mathcal{C}\ell\{v_1, v_2, \ldots, v_{n-1}\}$$

y por lo tanto los hiperplanos en $I\!R^n$ tienen dimensión $n-1$.

Por otra parte, si $u_1, u_2, \ldots, u_{n-1}$ son $n-1$ vectores l.i. de $I\!R^n$, el subespacio

$$T = \mathcal{C}\ell\{u_1, u_2, \ldots, u_{n-1}\}$$

debe corresponder a un cierto hiperplano que contiene el origen. Y resulta interesante determinar un vector $a = (a_1, a_2, \ldots, a_n)$ que permita describir T como:

$$T = \{x \in I\!R^n / a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}$$

Este problema se propone como ejercicio (16), al final de la siguiente sección de ejercicios 7.6.

Ejemplo 7.23 Considere

$$H = \{(x, y, z, w) \in I\!R^4 / x - 3y + w = 0\}.$$

Observe que $H = \{(x, y, z, w) \in I\!R^4 / x = 3y - w\}$ donde y, z w son variables a las que se les puede asignar cualquier valor real. Entonces si $y = t, z = s$ y $w = r$ un vector cualquiera de H se escribe

$$(x, y, z, w) = t(3, 1, 0, 0) + s(0, 0, 1, 0) + r(-1, 0, 0, 1)$$

Lo que también significa que

$$H = \mathcal{C}\ell\{(3, 1, 0, 0), (0, 0, 1, 0), (-1, 0, 0, 1)\}$$

El vector $a = (1, -3, 0, 1)$ — definido por los coeficientes de de la ecuación $x - 3y + w = 0$— verifica que es ortogonal a cada vector de la base $\{(3, 1, 0, 0), (0, 0, 1, 0), (-1, 0, 0, 1)\}$ de H y por lo tanto, también a cada vector del hiperplano H.

7.4.3 Vector de coordenadas en una base $\mathcal{B}$

Definición 7.15 (Coordenadas de v en la base $\mathcal{B}$)

Si $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ es una base de E entonces $\forall v \in E$, se llaman coordenadas de v en la base $\mathcal{B}$ a los únicos escalares $a_1, a_2, \ldots, a_n$ tales que

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$$

Y se denota con $[v]_{\mathcal{B}}$ al vector columna $[v]_{\mathcal{B}} = (a_1, a_2, \ldots, a_n)^t$ de coordenadas de v en la base $\mathcal{B}$.

Ejemplo 7.24 Sea $\mathcal{B} = \{v_1, v_2, v_3, v_4\}$ una base de $\mathbb{R}^4$ y suponga que

$$v = 2v_1 - 3v_3 + v_4$$

entonces el vector de coordenadas de v en la base $\mathcal{B}$ es:

$$[v]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 0 \\ -3 \\ 1 \end{pmatrix}.$$

Por otra parte, si $\mathcal{B}_1 = \{v_2, v_4, v_3, v_1\}$ es la misma base pero con los vectores considerados en un nuevo orden, entonces

$$[v]_{\mathcal{B}_1} = \begin{pmatrix} 0 \\ 1 \\ -3 \\ 2 \end{pmatrix}.$$

Es decir, el concepto de coordenadas de un vector en una base depende del orden en que se listan los vectores de la base.

$\boxed{\textbf{Nota}}$ Observe que si $B = (v_1, v_2, \ldots, v_n)$ denota la matriz cuyas columnas son los vectores de la base $\mathcal{B}$ entonces

$$v = B[v]_{\mathcal{B}},$$

dado que $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ y $[v]_{\mathcal{B}} = (a_1, a_2, \ldots, a_n)^t$.

$\square$

Ejemplo 7.25 Sea $\mathcal{B} = \{u, v\} = \{(-1, 1), (1, 2)\}$ una base de $I\!R^2$. Puesto que

$$w = (7/2, 1) = -2(-1, 1) + (3/2)(1, 2) = -2u + (3/2)v$$

las coordenadas del vector $w = (7/2, 1)$ en la base $\mathcal{B}$ son:

$$[w]_{\mathcal{B}} = (-2, 3/2).$$

En el siguiente gráfico se muestran los ejes que representan la base canónica y la base $\mathcal{B}$, y las coordenadas de w en la base canónica y en la base $\mathcal{B}$.

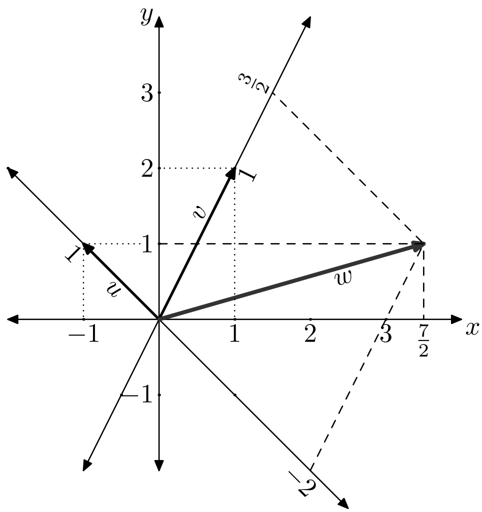


Figura 7.3: Coordenadas canónicas de $w = (7/2, 1)$ y Coordenadas de w en la base $\mathcal{B}$, $[w]_{\mathcal{B}} = (-2, 3/2)$.

7.5 Espacio generado por las filas de una matriz

Si $v_1, v_2, \ldots, v_k$ son vectores fila en $I\!R^n$ y $A^t = (v_1^t, v_2^t, \ldots, v_k^t)$, la matriz $A_{k\times n}$ cuyas filas son los vectores v_i, o sea:

$$A = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix}$$

se denota con $\mathcal{F}_A$ al subespacio vectorial generado por las filas de A, esto es:

$$\mathcal{F}_A = \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}.$$

Observe que en esta situación, si $w \in \mathcal{F}_A$, y se escribe como un vector columna, entonces existen escalares $x_1, x_2, \ldots, x_k$ tales que

$$w = x_1 v_1^t + x_2 v_2^t + \cdots + x_k v_k^t = (v_1^t, v_2^t, \ldots, v_k^t) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = A^t x.$$

O lo que es lo mismo, existe $x = (x_1, x_2, \ldots, x_k)^t \in I\!R^k$ tal que $w = A^t x$ y de esta manera $\mathcal{F}_A$ también se puede describir como

$$\mathcal{F}_A = \{b \in I\!R^n / b = A^t x, \text{ para algún } x \in I\!R^k\}.$$

Por otra parte el espacio generado por las columnas de A es

$$\mathcal{F}_{A^t} = \{c \in I\!R^k / c = Ay, \text{ para algún } y \in I\!R^n\}.$$

7.5.1 Operaciones elementales y espacio generado

Las operaciones elementales sobre las filas de una matriz no modifican el espacio que éstas generan.

Teorema 7.16 *Sean $v_1, v_2, \ldots, v_k$ y $u_1, u_2, \ldots, u_k$ vectores fila en $I\!R^n$ y*

$$A = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} \; y \; B = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{pmatrix},$$

dos matrices $k \times n$, equivalentes, es decir B se obtiene de A mediante operaciones elementales de renglón, entonces

$$\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\} = \mathcal{C}\ell\{u_1, u_2, \ldots, u_k\}.$$

Demostración: Observe que:

$$\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\} = \{b \in I\!R^n / b = A^t x, \text{ para algún } x \in I\!R^k\},$$

$$\mathcal{C}\ell\{u_1, u_2, \ldots, u_k\} = \{c \in I\!R^n / c = B^t y, \text{ para algún } y \in I\!R^k\}.$$

Sea P el producto de las matrices elementales necesarias para reducir A en B, o sea, $B = PA$ y $P^{-1}B = A$, (P es una matriz $k \times k$). Sea $b \in \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$ entonces existe $x \in I\!R^k$ tal que

$$b = A^t x = (P^{-1}B)^t x = B^t[(P^{-1})^t x] = B^t y$$

de manera que b es combinación lineal de las filas de B, por lo cual $b \in \mathcal{C}\ell\{u_1, u_2, \ldots, u_k\}$ y de esto se deduce que

$$\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\} \subset \mathcal{C}\ell\{u_1, u_2, \ldots, u_k\}.$$

En forma similar se muestra que:

$$\mathcal{C}\ell\{u_1, u_2, \ldots, u_k\} \subset \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$$

para concluir finalmente

$$\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\} = \mathcal{C}\ell\{u_1, u_2, \ldots, u_k\}.$$

■■■

Corolario 7.17 *Sean $v_1, v_2, \ldots, v_k$ y $w_1, w_2, \ldots, w_s$ vectores de $I\!R^n$, correspondientes a filas de A y filas no nulas de R, respectivamente:*

$$A = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} \text{ y } R = \begin{pmatrix} w_1 \\ \vdots \\ w_s \\ \vdots \\ 0 \end{pmatrix}.$$

Suponga además que R es una matriz escalonada equivalente a A, con s filas no nulas, $s \leq k$, entonces

$$\mathcal{B} = \{w_1, w_2, \ldots, w_s\}$$

es una base del subespacio, $\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$, que generan las filas de A.

Demostración: Es suficiente con reconocer que:

$$\mathcal{C}\ell\{v_1, v_2, \ldots, v_k\} = \mathcal{C}\ell\{w_1, w_2, \ldots, w_s\}$$

y que las filas no nulas de una matriz escalonada son vectores l.i..

■■■

Ejemplo 7.26 Si $w_1 = (-3, -2, 1, 0)$, $w_2 = (-2, -1, 0, 1)$, $w_3 = (5, 3, -1, -1)$, $w_4 = (0, -1, 0, 1)$ y

$$W = \mathcal{C}\ell\{w_1, w_2, w_3, w_4\}$$

a) Muestre que el subespacio W tiene dimensión 3.

b) Determine una base $\mathcal{B}$ para W y complete $\mathcal{B}$ hasta obtener una base de $I\!R^4$.

Solución: Haciendo operaciones elementales sobre una matriz cuyas filas son los vectores que generan W, se obtienen otras matrices cuyas filas también generan a W, por teorema 7.16:

$$\begin{pmatrix} -3 & -2 & 1 & 0 \\ -2 & -1 & 0 & 1 \\ 5 & 3 & -1 & -1 \\ 0 & -1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 & 1 \\ 0 & -1/3 & -2/3 & -1 \\ 0 & -1 & 0 & 1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2/3 & 1/3 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

En la última matriz, es evidente que sus tres filas no nulas son l.i., y como generan a W, son una base de este subespacio. De lo anterior se tiene que la dimensión de W es 3 y además que:

$$\begin{aligned} W = & \\ & \mathcal{C}\ell\{(-3,-2,1,0),(-2,-1,0,1),(5,3,-1,-1),(0,-1,0,1)\} \\ = & \ \mathcal{C}\ell\{(3,2,1,0),(0,1,2,3),(0,-1,-2,-3),(0,-1,0,1)\} \\ = & \ \mathcal{C}\ell\{(1,2/3,1/3,0),(0,1,2,3),(0,0,2,3)\} \end{aligned}$$

Proponer una base para W, supone elegir tres vectores de W l.i., los cuales, por las igualdades anteriores, pueden escogerse entre las filas de cualquiera de las tres matrices dadas, o entre los generados por estas. Incluso, puede elegirse una fila de cada matriz en el tanto sean l.i.. Dos ejemplos de bases para W son: $\mathcal{B}_1 = \{(-3,-2,1,0),(-2,-1,0,1),(0,-1,0,1)\}$ y $\mathcal{B}_2 = \{(1,2/3,1/3,0),(0,1,2,3),(0,0,2,3)\}$.

Dar una base de $I\!R^4$ supone elegir 4 vectores l.i. de $I\!R^4$, dado que este espacio tiene dimensión 4. Entonces para completar las anteriores bases hasta formar una base de $I\!R^4$, sólo se requiere elegir un nuevo vector de $I\!R^4$ que en conjunto con los de $\mathcal{B}_1$ o $\mathcal{B}_2$ formen un conjunto l.i. Observando las filas de la última matriz es fácil proponer la fila $(0,0,0,1)$, para reemplazar la última de ceros y establecer cuatro vectores fila l.i., de los cuales los tres primeros constituyen la base $\mathcal{B}_2$ de W. Así, este vector completa tanto la base $\mathcal{B}_2$, como la $\mathcal{B}_1$ para formar una base de $I\!R^4$.

∎

7.6 Ejercicios

1. Decida si S es un subespacio de V y justifique su respuesta:

(a) $S = \{(1, x, y) | x, y \in I\!R\}$ con $V = I\!R^3$.
(b) $S = \{(x, y, z) | x - y = y + z\}$ con $V = I\!R^3$.
(c) $S = \{A \in M(n, I\!R) | A$ es simétrica$\}$ con $V = M(n, I\!R)$.
(d) $S = \{(x, y, z, w) | x - y = 0, z - w = 0\}$ con $V = I\!R^4$.
(e) $S = \{x \in I\!R^n | Ax = e_1\}$ donde A es una matriz $m \times n$ de rango m y $V = I\!R^n$.
(f) $S = \{(x, y, z, w) | x + y + z + w + 1 = 0\}$ con $V = I\!R^4$.
(g) $S = \{x \in I\!R^p | x \cdot a = 0\}$ donde $a \in V = I\!R^p$.
(h) $S = \{A \in M(5, I\!R) | A$ es una matriz triangular superior con unos en la diagonal$\}$ y $V = M(5, I\!R)$.
(i) $S = \{A | A$ es una matriz diagonal de orden $n\}$ con $V = M(n, I\!R)$.
(j) $S = \{(x, y, z) | x = 0$ ó $z = 0\}$ con $V = I\!R^3$.

2. Sea $v = (\alpha, \beta, 0, -1)$ y

$$W = \mathcal{C}\ell\{(2, -1, 3, 2), (-1, 1, 1, -3), (1, 1, 9, -4)\}$$

Calcule α y β de modo $v \in W$.

3. Sea $A = \begin{pmatrix} 1 & 1 & -2 & 1 \\ 1 & -1 & 1 & 0 \\ 1 & 5 & -8 & 3 \end{pmatrix}$ y

$$S = \{x \in I\!R^4 / Ax = 0\}$$

Determine una base para S.

4. Sea $\mathcal{B} = \{(1, -1, 2), (1, 0, 1), (0, 1, 1)\}$

(a) Demuestre que $\mathcal{B}$ es una base de $I\!R^3$.
(b) Encuentre $[v]_{\mathcal{B}}$ las coordenadas de $v = (1, 2, -1)$ en la base $\mathcal{B}$.

5. Sean $A = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 2 & -1 & 1 & 7 \\ 1 & 5 & 6 & -2 \end{pmatrix}$,
F el espacio generado por las filas de A y C el espacio de las columnas de A.

(a) Determine una base para F y llámela $\mathcal{B}_1$.

(b) ¿$(1,1,1,1) \in F$? Calcule $[(2,6,8,0)]_{\mathcal{B}_1}$.

(c) Determine una base para C e identifíquela con $\mathcal{B}_2$.

(d) Calcule $[A_2]_{\mathcal{B}_2}$, donde A_2 es la columna 2 de A.

(e) Los subespacios F y C tienen la misma dimensión cualquiera sea la matriz A. ¿Porqué?

6. Sea $S = \mathcal{C}\ell\{v_1, v_2, v_3, v_4\}$ donde $v_1 = (1,-1,2), v_2 = (1,1,0)$, $v_3 = (1,0,1)$ y $v_4 = (0,1,1)$.

(a) Demuestre que $S = I\!R^3$.

(b) Determine una base para S.

(c) ¿El sistema, $x_1v_1+x_2v_2+x_3v_3+x_4v_4 = (a,b,c)$, de ecuaciones lineales, tiene solución única para todo $(a,b,c) \in I\!R^3$?

7. Sean $v_1, v_2, v_3, v_4 \in I\!R^3$ y suponga que

$$I\!R^3 = \mathcal{C}\ell\{v_2, v_3, v_4\}$$

Justifique las respuestas a las siguientes preguntas.

(a) ¿$\{v_2, v_3, v_4\}$ es l.i.?

(b) ¿$I\!R^3 = \mathcal{C}\ell\{v_1, v_2, v_3\}$?

(c) ¿El sistema, $x_1v_1+x_2v_2+x_3v_3+x_4v_4 = (a,b,c)^t$, de ecuaciones lineales, tiene solución única para todo $(a,b,c) \in I\!R^3$?

8. Sea $U = \{(x, x+y, 0)/x, y \in I\!R\}$

(a) Verifique que U es un subespacio de $I\!R^3$.

(b) Encuentre una base para U.

(c) Determine un vector (a_1, a_2, a_3) que sea ortogonal a cada vector de la base de U.

(d) Represente en un gráfico U y el vector (a_1, a_2, a_3).

(e) Justifique geométricamente que

$$U = \{(x,y,z)/a_1x + a_2y + a_3z = 0\}$$

9. Sea $S = \{A \in M(3, I\!R)/A \text{ es simétrica}\}$. Determine una base $\mathcal{B}$ para S y las coordenadas de la matriz

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

en la base $\mathcal{B}$.

10. Calcule la dimensión de

$$U = \{(x, y, z) \in I\!R^3/3x - y + 4z = 0\}$$

11. Determine condiciones sobre a y b para que

$$W = \mathcal{C}\ell\{\begin{pmatrix} 1 \\ a \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 2 \\ a+b \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 6 \\ 3 \end{pmatrix}\} \text{ tenga dimensión 3.}$$

12. Sea $A = \begin{pmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & -1 & -1 & 2 & 0 \\ 2 & -2 & -1 & 4 & 1 \end{pmatrix}$

a) Encuentre una base para el subespacio que es solución de $Ax = 0$.

b) Use el resultado en a), para proponer un ejemplo de una matriz B, 5×2, con columnas no nulas, tal que $AB = 0$.

c) ¿Pertenece el vector $(1, 2, 1, 0)^t$ al espacio generado por las columnas de A? (Justifique)

d) Encuentre una base para el espacio generado por las filas de A e indique cuál es el rango de A.

13. Sean V y W subespacios de un espacio vectorial E.

1. Muestre que $V \cap W$ es un subespacio de E.
2. Sea $E = I\!R^3$, $V = \{(x, y, z)|2x - 3y + z = 0\}$ y $W = \{(x, y, z)|x - y + 2z = 0\}$:
   (a) Determine el subespacio $V \cap W$.
   (b) Muestre con un ejemplo que existen $v, w \in V \cup W$ tales que $v + w \notin V \cup W$.

3. Muestre que $V \cup W$ no siempre es subespacio de E. ¿En qué situación $V \cup W$ es un subespacio?

14. Sea $W = \{(x, x+y+2z, y+z, z)/x, y, z \in I\!R\}$. Encuentre una base para W. ¿Es W un hiperplano de $I\!R^4$? Determine un vector (a_1, a_2, a_3, a_4) tal que

$$W = \{(x, y, z, w)/a_1x + a_2y + a_3z + a_4w = 0\}$$

15. Sea $A = \begin{pmatrix} 1 & 2 & -1 & 0 & 1 \\ 1 & 0 & 1 & 1 & -1 \\ 2 & 2 & 0 & 1 & 0 \\ 0 & -2 & 2 & -2 & 2 \end{pmatrix}$, y

$$W = \{y \in I\!R^4 | y = Ax, \text{ para algún } x \in I\!R^5\}.$$

(a) Determine cuatro vectores v_1, v_2, v_3, v_4 tales que $W = \mathcal{C}\ell\{v_1, v_2, v_3, v_4\}$.

(b) Determine una base $\mathcal{B}$ para W.

(c) Si $R = \{y \in I\!R^5 | y = A^t x, \text{ para algún } x \in I\!R^4\}$. ¿Cuál es la dimensión de R? Conteste sin hacer más cálculos.

16. Si $v_1, v_2, \ldots, v_{n-1}$ son $n-1$ vectores l.i. de $I\!R^n$, y $H = \mathcal{C}\ell\{v_1, v_2, \ldots, v_{n-1}\}$, demuestre que:

(a) Existe $a = (a_1, a_2, \ldots, a_n) \in I\!R^n$ no nulo tal que $a \cdot v_i = 0 \; \forall \; i = 1, \ldots, n-1$. Sugerencia: observe que a es solución de un sistema homogéneo con $n-1$ ecuaciones.

(b) $\{v_1, v_2, \ldots, v_{n-1}, a\}$ es una base de $I\!R^n$.

(c) $H = \{x \in I\!R^n / a \cdot x = 0\}$.

17. Sean V y W subespacios de un espacio vectorial E y considere el conjunto:

$$V + W = \{v + w | v \in V \text{ y } w \in W\}.$$

1. Muestre que $V + W$ es un subespacio de E.
2. Sea $E = I\!R^3$, $V_1 = \mathcal{C}\ell\{(2, -3, 1)^t\}$, $V_2 = \mathcal{C}\ell\{(3, 1, -1)^t\}$ y $W = \{(x, y, z) | x - y + 2z = 0\}$:
   (a) Muestre que $V_1 + V_2 = \mathcal{P}((0, 0, 0), (2, -3, 1), (3, 1, -1))$, el plano que contiene el origen en la dirección de los vectores $(2, -3, 1)$ y $(3, 1, -1)$.
   (b) Muestre que $W + V_1 = I\!R^3$ y que $W + V_2 = W$.

Capítulo 8

Ortogonalidad y proyecciones

La noción de ortogonalidad, así como la de norma, distancia, ángulo y proyección, dependen del producto escalar de vectores. Recuerde que en un espacio vectorial E sólo se disponen de las operaciones, suma de vectores y multiplicación de un escalar por un vector, de manera que la introducción de los anteriores conceptos es posible únicamente en aquellos espacios vectoriales donde se ha definido un producto escalar o interior, llamados espacios vectoriales normados.

En este capítulo trabajaremos en el espacio $I\!R^n$ para presentar, fundamentalmente, los conceptos de conjuntos y subespacios ortogonales y construir bases ortonormales.

8.1 Conjuntos ortogonales

Las ideas que se expondrán son continuación de los conceptos sobre ortogonalidad que se introdujeron en el capítulo 5.

Definición 8.1 (Conjunto ortogonal)

Un conjunto $T = \{u_1, u_2, \ldots, u_k\}$ de vectores, no nulos, en $I\!R^n$ es ortogonal si $u_i \cdot u_j = 0$, para cualquier pareja de índices i, j en $\{1, 2, \ldots, k\}$, siempre que $i \neq j$.

Ejemplo 8.1 El conjunto

$$T = \{u_1, u_2, u_3\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix} \right\}$$

es un conjunto ortogonal porque:

$$\begin{array}{rcl c c} u_1 \cdot u_2 & = & (1)2 + (0)1 + (-1)2 & = & 0 \\ u_1 \cdot u_3 & = & (1)1 + 0(-4) + (-1)1 & = & 0 \\ u_2 \cdot u_3 & = & 2(1) + 1(-4) + (2)1 & = & 0 \end{array}$$

Teorema 8.2 *Si un conjunto $B = \{u_1, u_2, \ldots, u_k\}$ de vectores no nulos de $I\!R^n$ es ortogonal entonces es l.i. (linealmente independiente).*

Demostración: Se debe mostrar que si $a_1u_1 + a_2u_2 + \cdots + a_ku_k = 0_n$ entonces cada $a_i = 0$, $i = 1, \ldots, k$. Observe que $\forall j \in \{1, 2, \ldots, k\}$

$$\begin{array}{rcl} 0 = u_j \cdot 0_n & = & u_j \cdot (a_1u_1 + a_2u_2 + \cdots + a_ku_k) \\ & = & a_1(u_j \cdot u_1) + a_2(u_j \cdot u_2) + \cdots + a_k(u_j \cdot u_k) \\ & = & a_1(0) + a_2(0) + \cdots + a_j(u_j \cdot u_j) + \cdots + a_k(0) \end{array}$$

Porque $u_j \cdot u_i = 0$ siempre que $j \neq i$, dado que T es ortogonal. Luego $0 = a_j(u_j \cdot u_j) = a_j \|u_j\|^2$ y como u_j es un vector no nulo, su norma no es cero por lo que, necesariamente, $a_j = 0$.

∎

Corolario 8.3 *Sea $\mathcal{B} = \{u_1, u_2, \ldots, u_k\}$ un conjunto de vectores no nulos ortogonales, de un subespacio S de dimensión k, entonces $\mathcal{B}$ es una base de S.*

Demostración: Como $\mathcal{B}$ es ortogonal entonces es un conjunto l.i. y como son k vectores de un subespacio de dimensión k, entonces deben generarlo y por lo tanto es una base de S.

■■■

8.2 Bases ortonormales

Definición 8.4 (Bases ortonormales)

Sea $S \subseteq I\!R^n$ un subespacio vectorial de dimensión k. Una base $\mathcal{B} = \{v_1, v_2, \ldots, v_k\}$ de S es ortonormal si

$$\begin{array}{rl} \textit{(i)} & v_i \cdot v_j = 0 \ \forall i, j = 1, \ldots, k, \quad i \neq j \\ \textit{(ii)} & \|v_i\| = 1 \ \forall i = 1, \ldots, k \end{array}$$

Es decir, si es un conjunto ortogonal de vectores unitarios de S.

Observe que la base canónica de $I\!R^n$ es una base ortonormal.

Ejemplo 8.2 El conjunto

$$\mathcal{B} = \{v_1, v_2, v_3\} = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix} \right\}$$

es una base ortonormal de $I\!R^3$ porque:

$$\begin{array}{rclcl} v_1 \cdot v_2 & = & (1/\sqrt{2})0 + (0)1 + (-1/\sqrt{2})0 & = & 0 \\ v_1 \cdot v_3 & = & (1/\sqrt{2})1/\sqrt{2} + 0(0) + (-1/\sqrt{2})1/\sqrt{2} & = & 0 \\ v_2 \cdot v_3 & = & 0(1/\sqrt{2}) + 1(0) + 0(1/\sqrt{2}) & = & 0 \\ v_1 \cdot v_1 & = & (1/\sqrt{2})^2 + (0)^2 + (-1/\sqrt{2})^2 & = & 1 \\ v_2 \cdot v_2 & = & 0^2 + 1^2 + 0^2 & = & 1 \\ v_3 \cdot v_3 & = & (1/\sqrt{2})^2 + (0)^2 + (1/\sqrt{2})^2 & = & 1 \end{array}$$

Teorema 8.5 *Sean S un subespacio de $I\!R^n$, $\mathcal{B} = \{v_1, v_2, \ldots, v_k\}$ una base ortonormal de S y B la matriz cuyas columnas son los vectores $v_1, v_2, \ldots, v_k$, entonces:*

(a) $B^t B = I_k$

(b) $[v]_{\mathcal{B}} = (v \cdot v_1, v \cdot v_2, \ldots, v \cdot v_k)^t \quad \forall v \in S$

En particular si $k = n$ entonces B es una matriz ortogonal.

Demostración: Ejercicio.

■■■

Ejemplo 8.3 Sea $\mathcal{B} = \{v_1, v_2, v_3\}$ la base ortonormal de $I\!R^3$ propuesta en el ejemplo 8.2. Un atributo apreciable de las bases ortonormales es que facilitan el cómputo de las coordenadas de un vector cualquiera. Si $x = (x_1, x_2, x_3)^t$

$$[x]_{\mathcal{B}} = (\alpha_1, \alpha_2, \alpha_3)^t \Longleftrightarrow x = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$$

y observe que para cualquier $i = 1, 2, 3$

$$\begin{array}{rcl} x \cdot v_i & = & (\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3) \cdot v_i \\ & = & \alpha_i v_i \cdot v_i \\ & = & \alpha_i \end{array}$$

Entonces, por ejemplo, las coordenadas de $x = (1, 2, 3)^t$ en la base $\mathcal{B}$ son dadas por:

$$\begin{array}{rcccccl} \alpha_1 & = & x \cdot v_1 & = & (1,2,3)^t \cdot (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}) & = & \frac{-2}{\sqrt{2}} \\ \alpha_2 & = & x \cdot v_2 & = & (1,2,3)^t \cdot (0,1,0) & = & 2 \\ \alpha_3 & = & x \cdot v_3 & = & (1,2,3)^t \cdot (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) & = & \frac{4}{\sqrt{2}} \end{array}$$

Esto es

$$[x]_{\mathcal{B}} = (\frac{-2}{\sqrt{2}}, 2, \frac{4}{\sqrt{2}})^t$$

8.3 Subespacios ortogonales

Definición 8.6 (Subespacios Ortogonales)

Dos subespacios S y T de $\mathbb{R}^n$ son ortogonales si

$$x \cdot y = 0 \quad \forall x \in S \text{ y } \forall y \in T$$

Si S y T son ortogonales se escribe $S \perp T$.

Ejemplo 8.4 Los subespacios S y T siguientes, son ortogonales:

$$\begin{array}{rcl} S & = & \{(x,y,z)^t \in \mathbb{R}^3 / 2x - 3y + z = 0\} \\ T & = & \{(x,y,z)^t \in \mathbb{R}^3 / (x,y,z)^t = t(2,-3,1)^t \text{ con } t \in \mathbb{R}\} \end{array}$$

Observe que si $x = (x_1, x_2, x_3)^t \in S$ entonces

$$2x_1 - 3x_2 + x_3 = 0$$

además si $y = (y_1, y_2, y_3)^t \in T$ existe $s \in \mathbb{R}$ tal que

$$(y_1, y_2, y_3)^t = s(2,-3,1)^t = (2s, -3s, s)^t$$

de manera que:

$$\begin{array}{rcl} x \cdot y & = & (x_1, x_2, x_3)^t \cdot (2s, -3s, s)^t \\ & = & 2sx_1 - 3sx_2 + sx_3 \\ & = & s(2x_1 - 3x_2 + x_3) \\ & = & s0 \\ & = & 0 \end{array}$$

Luego $x \cdot y = 0\, \forall x \in S$ y $\forall y \in T$. Geométricamente, S corresponde al plano que contiene el origen y es ortogonal al vector $(2,-3,1)$, y T es la recta que contiene el origen con la dirección del vector $(2,-3,1)$.

Teorema 8.7 *Si S y T son subespacios ortogonales de $\mathbb{R}^n$ entonces $S \cap T = \{0_n\}$.*

Demostración: Suponga que $x \in S \cap T$ y $x = (x_1, x_2, \ldots, x_n)^t$. Como $S \perp T$ entonces $\forall\ x \in S$ y $\forall\ y \in T$

$$x \cdot y = 0$$

Luego si $x \in S$ y $x \in T$ se tiene que

$$x \cdot x = 0 \Longrightarrow x_1^2 + x_2^2 + \cdots + x_n^2 = 0 \Longrightarrow x_i = 0 \quad \forall i$$

entonces $x = (0, 0, \ldots, 0)^t = 0_n$.

∎

Teorema 8.8 *Sean $\mathcal{B}_1 = \{u_1, u_2, \ldots, u_k\}$ y $\mathcal{B}_2 = \{v_1, v_2, \ldots, v_r\}$ bases de S y T, subespacios de $I\!R^n$, entonces*

$$S \perp T \Longleftrightarrow u_i \perp v_j \quad \forall\ i = 1, \ldots, k \quad y \quad \forall\ j = 1, \ldots, r.$$

Demostración: De la definición de subespacios ortogonales se tiene en forma directa que si $S \perp T$ entonces $u_i \perp v_j\ \forall\ i$ y $\forall\ j$. Por otra parte, si $u_i \perp v_j\ \forall\ i$ y $\forall\ j$, entonces $\forall\ x \in S$ y $\forall\ y \in T$:

$$x = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k \quad \text{y} \quad y = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_r v_r$$

Luego

$$\begin{aligned} x \cdot y &= (\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k) \cdot (\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_r v_r) \\ &= \alpha_1\beta_1 u_1 \cdot v_1 + \alpha_1\beta_2 u_1 \cdot v_2 + \cdots + \alpha_1\beta_r u_1 \cdot v_r + \\ &\quad\ \alpha_2\beta_1 u_2 \cdot v_1 + \alpha_2\beta_2 u_2 \cdot v_2 + \cdots + \alpha_2\beta_r u_2 \cdot v_r + \\ &\quad\ \cdots + \alpha_k\beta_1 u_k \cdot v_1 + \alpha_k\beta_2 u_k \cdot v_2 + \cdots + \alpha_k\beta_r u_k \cdot v_r \\ &= \alpha_1\beta_1(0) + \cdots + \alpha_k\beta_r(0) \\ &= 0 \end{aligned}$$

∎

Teorema 8.9 *Si S es un subespacio de $I\!R^n$ entonces el conjunto*

$$S^\perp = \{x \in I\!R^n / x \perp y\ \forall y \in S\}$$

es un subespacio vectorial.

Demostración: Ejercicio.

∎

Definición 8.10 (Complemento ortogonal de S, $S^\perp$)

Sea S un subespacio vectorial de $I\!R^n$, se llama complemento ortogonal de S al subespacio de todos los vectores de $I\!R^n$ que son ortogonales a cualquier vector de S:

$$S^\perp = \{x \in I\!R^n | \; x \perp y \; \forall \; y \in S\}$$

Ejemplo 8.5 Sea $S = \mathcal{C}\ell\{(1,2,3)^t\}$ la recta por el origen con vector director $(1,2,3)^t$, entonces el complemento ortogonal de S es el plano

$$S^\perp = \{(x,y,z)^t \in I\!R^3 | \; x + 2y + 3z = 0\}$$

Veámoslo:

$$\begin{array}{rcl} S^\perp & = & \{x \in I\!R^3 | \; x \perp y \; \forall \; y \in S\} \\ & = & \{(x,y,z)^t \in I\!R^3 | \; (x,y,z)^t \perp (t, 2t, 3t) \; \forall t \in I\!R\} \\ & = & \{(x,y,z)^t \in I\!R^3 | \; (x,y,z)^t \perp (1,2,3)\} \\ & = & \{(x,y,z)^t \in I\!R^3 | \; x + 2y + 3z = 0\} \end{array}$$

En general se puede observar que si $S = \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$ es un subespacio de $I\!R^n$ entonces:

$$\begin{array}{rcl} S^\perp & = & \{x \in I\!R^n | \; x \perp y \; \forall \; y \in S\} \\ & = & \{x \in I\!R^n | \; x \perp v_i \; \forall \; i = 1, 2, \ldots, k\} \\ & = & \{x \in I\!R^n | \; v_i^t x = 0 \; \forall \; i = 1, 2, \ldots, k\} \\ & = & \{x \in I\!R^n | \; Ax = 0_k\} \end{array}$$

donde A es la matriz $k \times n$ cuyas filas son los vectores $v_1^t, v_2^t, \ldots, v_k^t$.

Teorema 8.11 *Sea $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_k\}$, $k < n$, una base de un subespacio S de $I\!R^n$ entonces*

1. $\dim(S^\perp) = n - k$.
2. *Si $\mathcal{B}_2 = \{u_1, u_2, \ldots, u_{n-k}\}$ es una base de $S^\perp$ entonces $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ es una base de $I\!R^n$.*

Demostración: 1) si A es la matriz $k \times n$ formada con filas iguales a los vectores $v_1^t, v_2^t, \ldots, v_k^t$ entonces

$$S^\perp = \{x \in I\!R^n | Ax = 0\}$$

y como $\operatorname{Rng} A = k$ entonces $\dim(S^{\perp}) = n - k$. Vea también los ejercicios 1 y 4.

2) Ejercicio.

■

8.4 Proyección ortogonal sobre un subespacio

Para cada $x \in I\!R^n$ el teorema 8.12 muestra la existencia de un vector $u \in W$ con dos propiedades que generalizan el concepto de proyección ortogonal sobre un vector (o recta) visto en el capítulo 5.

Teorema 8.12 *Si $W \subseteq I\!R^n$ es un subespacio de $I\!R^n$, para todo $x \in I\!R^n$ existe un único vector $u \in W$ tal que*

(1) $(x - u) \perp w \quad \forall\, w \in W$.

(2) $\|x - u\| \leq \|x - w\| \quad \forall\, w \in W$.

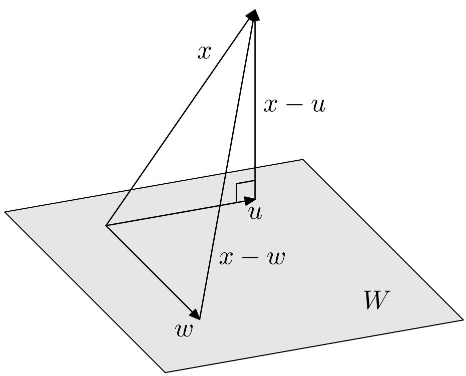


Figura 8.1: $\|x - u\| \leq \|x - w\| \quad \forall\, w \in W$: $u = \operatorname{Proy}_W x$.

Demostración: Sea $\mathcal{B} = \{w_1, w_2, \ldots, w_k\}$ una base ortonormal de W y

$$u = (x \cdot w_1)w_1 + (x \cdot w_2)w_2 + \cdots + (x \cdot w_k)w_k.$$

Claramente $u \in W$. Para demostrar la parte (1) es suficiente demostrar que

$$(x - u) \perp w_j, \quad \forall j = 1, 2, \ldots, k :$$

$$\begin{aligned}
& (x-u) \cdot w_j \\
= \; & [x - ((x \cdot w_1)w_1 + (x \cdot w_2)w_2 + \cdots + (x \cdot w_k)w_k)] \cdot w_j \\
= \; & x \cdot w_j - (x \cdot w_1)w_1 \cdot w_j - (x \cdot w_2)w_2 \cdot w_j - \cdots \\
& \qquad \qquad \qquad \cdots - (x \cdot w_k)w_k \cdot w_j \\
= \; & x \cdot w_j - (x \cdot w_j)w_j \cdot w_j \\
= \; & x \cdot w_j - (x \cdot w_j) \\
= \; & 0.
\end{aligned}$$

Dado que para $i \neq j$: $(x \cdot w_i)w_i \cdot w_j = (x \cdot w_i)0 = 0$ y que $v_j \cdot v_j = 1$.

La unicidad de u se deriva de la parte (2) y se deja como ejercicio.

(2) Ver ejercicio 6.

■

Este teorema permite definir el concepto de proyección ortogonal sobre subespacios.

Definición 8.13 (Proyección ortog. sobre un subespacio)

Sea $W \subseteq \mathbb{R}^n$ un subespacio y $x \in \mathbb{R}^n$ se llama vector proyección ortogonal de x sobre el subespacio W, o simplemente proyección ortogonal de x sobre W, al único vector $u \in W$ tal que

$$(x - u) \perp w \quad \forall \; w \in W$$

y se escribe $u = Proy_W x$. El vector $x - u$ se denomina componente de x ortogonal a W.

La demostración del teorema 8.12 provee la forma de calcular la proyección ortogonal de un vector sobre un subespacio, dada una base ortonormal de este.

Teorema 8.14 *Si* W *es un subespacio de* $\mathbb{R}^n$ *y* $\mathcal{B} = \{v_1, v_2, \ldots, v_k\}$ *una base ortonormal de* W *entonces*

1. $Proy_W x = (x \cdot v_1)v_1 + (x \cdot v_2)v_2 + \cdots + (x \cdot v_k)v_k$

2. $Proy_W x = BB^t x$

donde B *es la matriz cuyas columnas son los vectores de la base* $\mathcal{B}$.

Demostración: La parte 1) corresponde a la demostración del teorema 8.12.

Parte 2): $B = (v_1, v_2, \ldots, v_k)$ es una matriz $n \times k$ y

$$\begin{aligned} BB^t x &= (v_1, v_2, \ldots, v_k)(v_1, v_2, \ldots, v_k)^t x \\ &= (v_1, v_2, \ldots, v_k) \begin{pmatrix} v_1^t \\ v_2^t \\ \vdots \\ v_k^t \end{pmatrix} x \\ &= (v_1, v_2, \ldots, v_k) \begin{pmatrix} v_1 \cdot x \\ v_2 \cdot x \\ \vdots \\ v_k \cdot x \end{pmatrix} \end{aligned}$$

Esta última expresión es el producto de la matriz B $n \times k$ por un vector columna de k componentes, que se puede escribir como una combinación lineal de las columnas de B:

$$\begin{aligned} BB^t x &= (v_1 \cdot x)v_1 + (v_2 \cdot x)v_2 + \cdots + (v_k \cdot x)v_k \\ &= \text{Proy}_S x \end{aligned}$$

■

Observe que:

1. En el caso particular de que el subespacio W sea de dimensión 1, $W = \mathcal{C}\ell\{v\}$, la proyección ortogonal del vector x sobre W coincide con la proyección ortogonal del vector x sobre v, en este caso si v es un vector unitario se tiene que:

$$\text{Proy}_{\mathcal{C}\ell\{v\}} x = (x \cdot v)v = \text{Proy}_v x$$

dado que $\{v\}$ constituye una base ortonormal de $\mathcal{C}\ell\{v\}$.

2. Y si W es un plano que contiene el origen, $W = \mathcal{C}\ell\{v_1, v_2\}$, y $\{w_1, w_2\}$ es una base ortonormal de W entonces

$$\text{Proy}_W x = (x \cdot w_1)w_1 + (x \cdot w_2)w_2$$

Ejemplo 8.6 Sea $S = \{(x_1, x_2, x_3, x_4)^t \in I\!R^4 / x_3 = 0\}$, Determine $\text{Proy}_S x$ y la componente de x ortogonal a S.

Solución: Observe que para todo $x = (x_1, x_2, x_3, x_4)^t \in I\!R^4$ el vector $u = (x_1, x_2, 0, x_4)^t \in S$, además, $x - u = (0, 0, x_3, 0)^t$ es tal que $(x - u) \perp s$ para todo $s \in S$. Luego

$$u = x - (0, 0, x_3, 0)^t = (x_1, x_2, 0, x_4)^t = \text{Proy}_S x$$

y la componente de x ortogonal a S es $x - u = (0, 0, x_3, 0)^t$.

Por otra parte,

$$\mathcal{B} = \{v_1, v_2, v_3\} = \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{-1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ 1/\sqrt{2} \end{pmatrix} \right\}$$

es una base ortonormal de S (cada uno de sus elementos está en S y es un conjunto ortonormal). Entonces por el teorema 8.14, se tiene que

$$\begin{array}{rcl} \text{Proy}_S x & = & (x \cdot v_1)v_1 + (x \cdot v_2)v_2 + (x \cdot v_3)v_3 \\ & = & \frac{x_1 - x_4}{\sqrt{2}} v_1 + x_2 v_2 + \frac{x_1 + x_4}{\sqrt{2}} v_3 \\ & = & (x_1, x_2, 0, x_4)^t \end{array}$$

lo cual confirma el resultado ya obtenido. Verifique que también obtiene el mismo resultado si utiliza la siguiente base de S

$$\mathcal{B}_2 = \{u_1, u_2, u_3\} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

■

8.5 Construcción de bases ortonormales

Si $\{u_1, u_2, \ldots, u_k\}$ es una base cualquiera de un subespacio vectorial S_k de $\mathbb{R}^n$, es posible mediante un proceso recursivo construir una nueva base $\{v_1, v_2, \ldots, v_k\}$ de S_k ortonormal y tal que cada nuevo vector $v_i \in S_i = \mathcal{C}\ell\{u_1, u_2, \ldots, u_i\}$, $i = 1, 2, \ldots, k$. Este método utiliza fuertemente las proyecciones ortogonales sobre los subespacios S_i y se conoce como proceso de ortonormalización de Gram-Schmidt.

8.5.1 Ortonormalización de Gram-Schmidt

Paso 1: Defina $v_1 = \dfrac{u_1}{\|u_1\|}$, un vector unitario en la dirección del primer vector de base. Considere el subespacio $S_1 = \mathcal{C}\ell\{v_1\}$ y observe que $\{v_1\}$ es una base ortonormal de S_1.

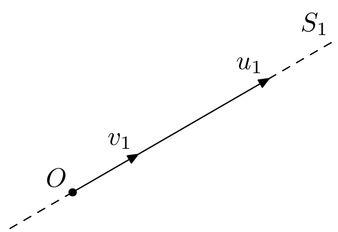


Figura 8.2: $v_1 = u_1/\|u_1\|$ y $S_1 = \mathcal{C}\ell\{v_1\}$.

Paso 2: Defina v_2 como:

$$v_2 = \frac{u_2 - \text{Proy}_{S_1} u_2}{\left\|u_2 - \text{Proy}_{S_1} u_2\right\|}$$

un vector unitario en la dirección de la componente de u_2 ortogonal a S_1. Observe que $v_2 \in S_2 = \mathcal{C}\ell\{v_1, v_2\}$, $v_2 \perp S_1$ y $\{v_1, v_2\}$ es una base ortonormal de S_2.

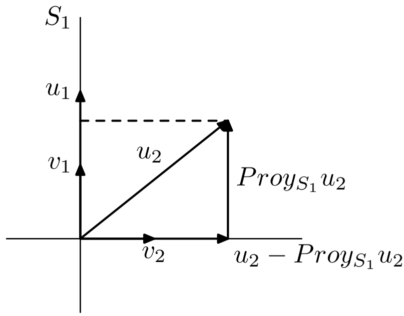


Figura 8.3: v_2: vector unitario en la dirección de $u_2 - \text{Proy}_{S_1} u_2$ y $S_2 = \mathcal{C}\ell\{v_1, v_2\}$.

Paso 3: Defina v_3 como:

$$v_3 = \frac{u_3 - \text{Proy}_{S_2} u_3}{\left\| u_3 - \text{Proy}_{S_2} u_3 \right\|}$$

un vector unitario en la dirección de la componente de u_3 ortogonal a S_2. Nuevamente $v_3 \in S_3 = \mathcal{C}\ell\{v_1, v_2, v_3\}$, $v_3 \perp S_2$ y $\{v_1, v_2, v_3\}$ es una base ortonormal de S_3.

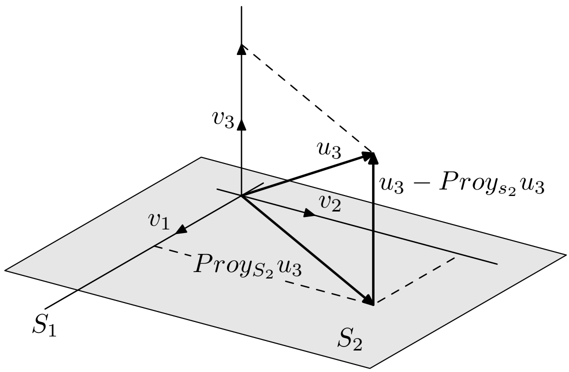


Figura 8.4: v_3: vector unitario en la dirección de $u_3 - \text{Proy}_{S_2} u_3$ y $S_3 = \mathcal{C}\ell\{v_1, v_2, v_3\}$.

El proceso continúa utilizando la misma estrategia hasta definir el vector v_k:

Paso k:

$$v_k = \frac{u_k - \text{Proy}_{S_{k-1}} u_k}{\left\| u_k - \text{Proy}_{S_{k-1}} u_k \right\|}$$

de manera que $v_k \in S_k = \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$, $v_k \perp S_{k-1}$ y $\{v_1, v_2, \ldots, v_k\}$ es una base ortonormal de S_k.

Observe que en cada etapa i, se obtiene una base ortonormal $\{v_1, v_2, \ldots, v_i\}$ de S_i lo cual permite en el siguiente paso, calcular

$$\text{Proy}_{S_i} u_{i+1} = (u_{i+1} \cdot v_1)v_1 + (u_{i+1} \cdot v_2)v_2 + \cdots + (u_{i+1} \cdot v_i)v_i$$

Ejemplo 8.7 Considere el siguiente subespacio vectorial de $I\!R^4$:

$$S = \mathcal{C}\ell\{u_1, u_2, u_3\} = \mathcal{C}\ell\{(0,1,0,1)^t, (0,1,1,0)^t, (-1,1,0,1)^t\}$$

a) Obtenga una base ortonormal para S.

b) Sea $w = (0,1,0,0)^t$, calcule $w_1 = \text{Proy}_S w$.

c) Determine $w_2 \in S^\perp$ tal que $w = w_1 + w_2$.

d) Proponga una base ortonormal para $I\!R^4$, completando la base obtenida en a) para S.

Solución:

a) Base ortonormal para S. Aplicando el procedimiento de Gram-Schmidt, a la base $\{u_1, u_2, u_3\}$ de S, se obtiene:

1) Primer vector unitario de la nueva base:

$$v_1 = u_1 / \|u_1\| = (0, 1/\sqrt{2}, 0, 1/\sqrt{2})^t$$

2) Sea $S_1 = \mathcal{C}\ell\{v_1\}$.
$\text{Proy}_{S_1} u_2 = (u_2 \cdot v_1)v_1 = (0, 1/2, 0, 1/2)^t$. Luego el segundo vector unitario de la nueva base es:

$$\begin{aligned} v_2 &= \frac{u_2 - \text{Proy}_{S_1} u_2}{\left\| u_2 - \text{Proy}_{S_1} u_2 \right\|} \\ &= \frac{(0, 1/2, 1, -1/2)^t}{\left\| (0, 1/2, 1, -1/2)^t \right\|} \\ &= (1/\sqrt{6})(0, 1, 2, -1)^t \end{aligned}$$

3) Sea $S_2 = \mathcal{C}\ell\{v_1, v_2\}$.

$$\begin{aligned} \text{Proy}_{S_2} u_3 &= (u_3 \cdot v_1)v_1 + (u_3 \cdot v_2)v_2 \\ &= \tfrac{2}{\sqrt{2}}(0, \tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}})^t + 0(\tfrac{1}{\sqrt{6}}(0,1,2,-1)^t) \\ &= (0,1,0,1)^t \end{aligned}$$

Entonces el tercer vector unitario de la nueva base es:

$$\begin{aligned} v_3 &= \frac{u_3 - \text{Proy}_{S_2} u_3}{\left\| u_3 - \text{Proy}_{S_2} u_3 \right\|} \\ &= \frac{(-1,0,0,0)^t}{\left\|(-1,0,0,0)^t\right\|} \\ &= (-1,0,0,0)^t \end{aligned}$$

Luego, el conjunto

$$\{v_1, v_2, v_3\} = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

es una base ortonormal para S.

b) Cálculo de $\mathbf{Proy_S w}$: como $\{v_1, v_2, v_3\}$ es una base ortonormal de S, entonces:

$$\begin{aligned} w_1 &= \text{Proy}_S w \\ &= (w \cdot v_1)v_1 + (w \cdot v_2)v_2 + (w \cdot v_3)v_3 \\ &= \tfrac{1}{\sqrt{2}}(0, \tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}})^t + \tfrac{1}{\sqrt{6}}(\tfrac{1}{\sqrt{6}}(0,1,2,-1)^t) + 0(-1,0,0,0)^t \\ &= (0, 2/3, 1/3, 1/3)^t \end{aligned}$$

c) Determinar $\mathbf{w_2 \in S^\perp}$ tal que $\mathbf{w = w_1 + w_2}$:

$$\begin{aligned} w_2 = w - w_1 &= (0,1,0,0)^t - (0,2/3,1/3,1/3)^t \\ &= (0, 1/3, -1/3, -1/3)^t \end{aligned}$$

el cual es un vector en $S^\perp$, por ser ortogonal a todo vector en S. En la definición de $\text{Proy}_S w$ se le llamó componente de w ortogonal a S.

d) Base ortonormal para $I\!R^4$: como w_2 es ortogonal a v_1, v_2, y v_3, entonces una base ortonormal de $I\!R^4$ es:

$$\{v_1, v_2, v_3, w_2 / \left\|w_2\right\|\}$$

específicamente:

$$\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix}0\\1\\2\\-1\end{pmatrix}, \begin{pmatrix}-1\\0\\0\\0\end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix}0\\1\\-1\\-1\end{pmatrix} \right\}$$

■

8.6 Ejercicios

1. Sean $A = \begin{pmatrix} 1 & 1 & -1 & 1 \\ -2 & 3 & 7 & -2 \\ 1 & -4 & -6 & 1 \end{pmatrix}$,
F el espacio generado por las filas de A y
$N = \{x \in I\!R^4 / Ax = 0\}$, el Núcleo de A.

(a) Determine una base para F.

(b) Determine una base para N.

(c) Demuestre que $F^\perp = N$.

(d) Determine bases ortonormales $\mathcal{B}_1$ para F y $\mathcal{B}_2$ para N.

(e) Sea $\vec{x} = (1, 2, 3, 4)$. Calcule dos vectores $\vec{a}$ y $\vec{b}$ ortogonales y tales que $\vec{b} \in N$ y $\vec{x} = \vec{a} + \vec{b}$.

(f) Muestre que para todo $\vec{x} = (x, y, z, w) \in I\!R^4$ existen dos vectores $\vec{a} \in F$ y $\vec{b} \in N$ tales que $\vec{x} = \vec{a} + \vec{b}$.

2. Sea $e_3 = (0, 0, 1)$ y $W = \mathcal{C}\ell\{e_3\}$. Determine $W^\perp$. Haga una representación gráfica de W y $W^\perp$.

3. Sea $W = \{(x, x - y + z, y + 2z, z) / x, y, z \in I\!R\}$, encuentre una base ortonormal para W y determine $W^\perp$, el complemento ortogonal de W.

4. Sean $A \in M(m, n, I\!R)$ y

$F = \{b \in I\!R^n / b = A^t y$ para algún $y \in I\!R^m\}$ el subespacio generado por las filas de A.

$N = \{x \in I\!R^n / Ax = 0\}$ el Núcleo de A.

Demuestre que $N = F^{\perp}$ y justifique que

$$\text{Rng}\,(A) + \text{Nulidad}\,(A) = n$$

5. La distancia del vector $x \in I\!R^n$ al subespacio W de $I\!R^n$ se define como $\|x - \text{Proy}_W x\|$. Calcule la distancia de $(1,1,0,1)$ al subespacio $W = \mathcal{C}\ell\{(1,1,0,0),(0,1,0,0),(0,0,1,1)\}$.

6. Sea W subespacio de $I\!R^n$ y $u = \text{Proy}_W x$ la proyección ortogonal de x sobre W. Demuestre que

$$\|x - u\| = \min_{w \in W} \|x - w\|$$

Sugerencia: Justifique que $(x-u)\cdot(w-u) = 0 \;\forall\; w \in W$, demuestre que $\|x-w\|^2 = \|x-u\|^2 + \|w-u\|^2 \;\; \forall\; w \in W$ y concluya el resultado.

7. Considere las rectas $L_1 = \{(1-t, 0, 1+t) | t \in I\!R\}$ y $L_2 = \{(t, 2+t, 0) | t \in I\!R\}$.

a) Establezca si estas rectas se intersecan o no, y si tienen un punto de intersección determine dicho punto.

b) Determine la ecuación vectorial del plano Π, que contiene el origen y no interseca las rectas L_1 y L_2.

c) Si W es el conjunto de puntos del plano Π (observe que W es un subespacio vectorial de $I\!R^3$), determine una base para $W^{\perp}$.

d) Haga un esquema para ilustrar gráficamente Π, L_1 y $W^{\perp}$ y muestre la forma de calcular la distancia de L_1 a Π. Calcule dicha distancia.

8. Sea $W = \mathcal{C}\ell\{(1,-1,1,1),(1,1,1,-1)\}$

1. Obtenga una base $\mathcal{B}$ ortonormal para $W^{\perp}$.
2. Complete $\mathcal{B}$ a una base $\mathcal{D}$ ortonormal de $I\!R^4$.
3. Calcule la proyección ortogonal de $v = (2,-1,0,0)$ sobre $W^{\perp}$.
4. Si $\mathcal{C}$ es una base canónica de $I\!R^4$. Calcule las coordenadas de $\text{Proy}_{W^{\perp}} v$ en la base $\mathcal{B}, \mathcal{D}$ y $\mathcal{C}$.

9. Sean $u = (1,-1,2), v = (2,3,0)$ vectores de $I\!R^3$ y $W = \mathcal{C}\ell\{u,v\}$.

(a) Construya una base de $\mathcal{B}$ ortonormal para W.

(b) Si $x = (1, 1, 2)$ encuentre $w \in W$ y $p \in W^\perp$ tal que $x = w + p$. Haga el dibujo correspondiente.

10. Sean $v_1 = (0, 1, 1, 0)^t$, $v_2 = (1, 1, 0, 1)^t$ y $W = \mathcal{C}\ell\{v_1, v_2\}$.

a) Determine una base para $W^\perp$.

b) Si $x = (a, a, a, 0)$, calcule $\text{Proy}_W x$.

c) ¿$y - \text{Proy}_W y$ es un vector de $W^\perp$, para todo $y \in I\!R^4$? Justifique su respuesta.

11. Sea $W = \mathcal{C}\ell\{(1, 0, 0, 1), (0, 1, 1, 0), (0, 0, 1, 1)\}$. Observe que W es un hiperplano de $I\!R^4$ y por lo tanto puede ser descrito en términos de una ecuación, en la forma:

$$W = \{(x_1, x_2, x_3, x_4)^t \in I\!R^4 / a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = 0\}$$

Determine (a_1, a_2, a_3, a_4).

12. Sea $W = \mathcal{C}\ell\{(1, 0, 0, 1), (0, 1, 1, 0)\}$. Observe que W y $W^\perp$ son planos de $I\!R^4$.

a) Determine bases ortonormales para W y $W^\perp$.

b) Si $\mathcal{B}_2 = \{a, b\}$ es la base obtenida para $W^\perp$. justifique que $W = \{x \in I\!R^4 / x \cdot a = 0 \text{ y } x \cdot b = 0\}$.

c) Determine un sistema de ecuaciones lineales, cuyo conjunto solución sea W.

d) Muestre que si $W = \mathcal{C}\ell\{v_1, v_2, \ldots, v_k\}$ es un subespacio de $I\!R^n$ de dimensión k, entonces existe un sistema homogéneo de $n - k$ ecuaciones en n variables, con conjunto solución igual a W.

e) Concluya (explique) que todo subespacio vectorial de $I\!R^n$ puede ser descrito en dos formas:

- Como el conjunto de combinaciones lineales de un conjunto de vectores dado.
- O, como el espacio solución de un cierto sistema de ecuaciones lineales homogéneo.

Capítulo 9

Regresión Lineal

Con frecuencia nuestro interés se centra en buscar relaciones entre dos o más variables, con el propósito de explicar una de ellas. El problema se plantea en términos de encontrar una función del tipo:

$$y = L(x_1, x_2, .., x_n)$$

donde $x_1, x_2, .., x_n$ representan las variables *explicativas o independientes*, y y la *variable dependiente o variable a explicar.* Una vez conocida la función L es posible utilizarla para elaborar estimaciones o predicciones de y dados los valores de las variables $x_1, x_2, .., x_n$.

El método denominado **Regresión Lineal** consiste en asumir que la función L es de tipo lineal, agregando a esto el supuesto probabilístico de normalidad en los residuos con lo que se consigue averiguar explícitamente la función L. Nosotros nos ocuparemos del problema de estimar los parámetros que caracterizan la función L, asumiendo que es lineal, sin entrar en el campo de las probabilidades, así las técnicas a desarrollar corresponden al ajuste de curvas por **Mínimos Cuadrados**.

9.1 El caso de dos variables

Analizaremos primero el caso más simple, que es el de dos variables, una a explicar y y una explicativa x, con el supuesto que existe entre ellas, aproximadamente una relación lineal de la forma $y = a_0 + a_1 x$.

El término constante a_0, lo podemos considerar como el coeficiente de una variable artificial x_0 cuyo valor es siempre 1. Una relación como la descrita la tenemos en el siguiente ejemplo.

9.1.1 Planteo del modelo $y = a_0 + a_1 x + \epsilon$

Se estudiará este modelo con la presentación de un ejemplo concreto, lo que permitirá fijar mejor las ideas y la notación que se involucra.

Ejemplo 9.1 Supongamos que en un país, el ingreso promedio anual (en miles de dólares), varía aproximadamente de acuerdo con el número de años de escolaridad, según la relación:

$$y = a_0 + a_1 x + \epsilon$$

donde las variables x, y representan:

- x : años de escolaridad,
- y : ingreso anual (miles de dólares)
- ϵ : el error.

Se desea estimar a_0 y a_1 a partir de los siguientes datos observados:

x	5	10	13	17
y	8	12	15	19

Geométricamente, el objetivo planteado es determinar la recta $y = a_0 + a_1 x$ (parámetros a_0 y a_1) que mejor ajusta los datos observados, según se muestra en el siguiente gráfico.

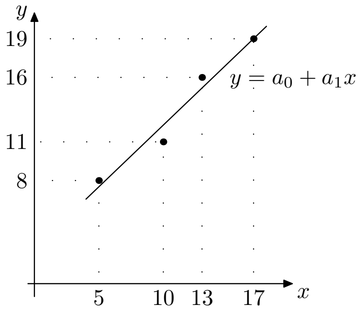


Figura 9.1: Puntos (x_i, y_i) y recta de mejor ajuste.

Tradicionalmente del gráfico 9.1, luego de precisar la idea de mejor ajuste, se derivan las condiciones necesarias para estimar a_0 y b_0, haciendo uso de las herramientas del cálculo diferencial. Nosotros sin embargo, acudiremos a una nueva representación geométrica del problema en $\mathbb{R}^4$ (en $\mathbb{R}^n$ en general) que nos dará la solución de una manera mucho más directa y fácil.

Denotemos con $x = (x_1, x_2, x_3, x_4)^t$ y con $1_4 = (1, 1, 1, 1)^t$ los vectores de $\mathbb{R}^4$, de valores observados correspondientes a las variables explicativas y con $y = (y_1, y_2, y_3, y_4)^t$ el vector de valores observados para la variable a explicar [1]. Denotamos con e_i los residuos que dependen de los datos, para diferenciarlos de los teóricos ϵ_i, que son independientes de los datos.

De acuerdo a nuestro supuesto, para los datos observados, debe verificarse que

$$y_i = a_0 + a_1 x_i + e_i, \qquad i = 1, \ldots, 4$$

lo que da lugar al siguiente sistema de ecuaciones

$$\begin{array}{lllcr} a_0 + & a_1 & + e_1 & = & 8 \\ a_0 + & a_1 10 & + e_2 & = & 12 \\ a_0 + & a_1 13 & + e_3 & = & 15 \\ a_0 + & a_1 17 & + e_4 & = & 19. \end{array}$$

[1] Observe que se usa los mismos símbolos x y y para denotar las variables del modelo y los vectores de valores observados de estas variables y a estos vectores se les llamó indistintamente variables, variables observadas o simplemente vectores

Hay $n = 4$ ecuaciones para determinar $p = 2$ coeficientes desconocidos (a_0 y a_1) y otras $n = 4$ incógnitas que son los residuos e_i. El interés fundamental es determinar los coeficientes a_0 y a_1, para lo que se dispone de un sistema con 2+4 incógnitas y 4 ecuaciones; el sistema admite, por lo tanto infinitas soluciones. En el conjunto de las infinitas soluciones queremos encontrar la **mejor** de estas soluciones, lo cual se precisa en la siguiente forma:

> Estimar (a_0, a_1) de modo que se minimicen los errores e_i, de acuerdo con algún criterio de calidad (minimización).

Entre los posibles criterios de minimización para los e_i se tienen los siguientes:

- $\min(\sum e_i^2)$: que la suma de los errores al cuadrado sea mínima.
- $\min(\sum |e_i|)$: que la suma de los valores absolutos de los errores sea mínima.
- $\min(\max e_i)$: minimizar el máximo de los errores.

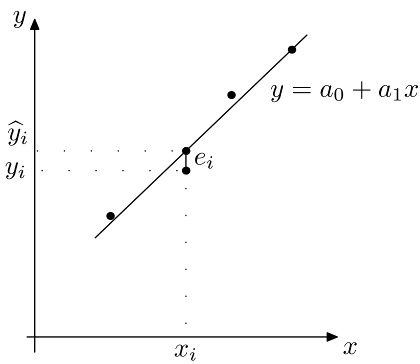


Figura 9.2: e_i: distancia vertical entre (x_i, y_i) y la recta de mejor ajuste.

Nosotros utilizaremos el primer criterio, llamado *criterio de los mínimos cuadrados*, el cual conduce a cálculos simples que pueden ponerse en el contexto del álgebra vectorial y tiene una interpretación geométrica clara: minimizar la suma de los cuadrados de las distancias verticales entre los puntos (x_i, y_i) y la recta de mejor ajuste, como se ilustra en el gráfico 9.2, o equivalentemente, minimizar $\|e\|^2 = \|(e_1, e_2, e_3, e_4)\|^2 = \sum_{i=1}^4 e_i^2$ lo que se ilustra en la figura 9.3.

9.1.2 Solución: mínimos cuadrados

Las ecuaciones

$$y_i = a_0 + a_1 x_i + e_i \qquad (i = 1, \ldots, 4)$$

se expresan, en términos vectoriales, como $y = a_0 1_4 + a_1 x + e$, es decir, si se elige denotar $a = (a_0, a_1)^t$ y $X = (1_4, x)$ se tiene que:

$$y = Xa + e,$$

lo que en concreto corresponde a:

$$\begin{pmatrix} 8 \\ 12 \\ 15 \\ 19 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 1 & 10 \\ 1 & 15 \\ 1 & 19 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix}.$$

En la búsqueda de los "mejores" valores a_0 y a_1, el criterio de los mínimos cuadrado se expresa como:

$$\min_a \sum_{i=1}^{4} e_i^2 = \min_a \|e\|^2. \tag{9.1}$$

Denotemos como

$$\widehat{a} = \begin{pmatrix} \widehat{a}_0 \\ \widehat{a}_1 \end{pmatrix}$$

el vector a que satisface la relación 9.1 al cual llamaremos "a estimado" y como $\widehat{y} = \widehat{a}_0 1_4 + \widehat{a}_1 x = X\widehat{a}$ al vector "y estimado". Con esta notación $y = \widehat{y} + e$ y 9.1 se expresa como

$$\min_a \|e\|^2 = \|y - \widehat{y}\|^2 . \tag{9.2}$$

Para encontrar el vector a que satisface la relación 9.2 se usará un procedimiento geométrico basado en los resultados sobre proyecciones del capítulo anterior. Definamos W como el subespacio de $\mathbb{R}^n$ generado por las variables 1_n y x, esto es

$$W = \mathcal{C}\ell\{1_n, x\} = \mathcal{C}\ell\{1_n, x_c\} \quad \text{con} \quad x_c = x - \overline{x}1_n.$$

La variable $x_c = x - \overline{x}1_n$ es conocida como x centrada.

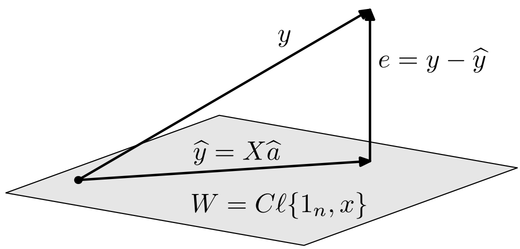


Figura 9.3: $\widehat{y} = X\widehat{a} = \text{Proy}_W y$.

Note que:

1. $\widehat{y} = \widehat{a}_0 1_n + \widehat{a}_1 x$ es un vector de W.

2. El conjunto $\{1_n, x_c\}$ formado por la variable 1_n y la variable x centrada es una base ortogonal para W:

$$1_n \cdot x_c = 1_n \cdot (x - \overline{x}1_n) = n\overline{x} - \overline{x}n = 0.$$

3. Encontrar el vector a que minimiza 9.2 es equivalente a encontrar el vector $\widehat{y}$ de W más próximo al vector y.

4. El vector $\widehat{y}$ más cercano al vector y es

$$\widehat{y} = \text{Proy}_W y = \text{Proy}_{1_n} y + \text{Proy}_{x_c} y \tag{9.3}$$

pues $\{1_n, x_c\}$ es una base ortogonal de W.

De estas observaciones, para encontrar las mejores estimaciones de a_0 y a_1, según el criterio dado, se tiene por una parte que

$$\widehat{y} = \widehat{a}_0 1_n + \widehat{a}_1 x = \text{Proy}_W y. \tag{9.4}$$

Y por otra

$$\widehat{y} = \text{Proy}_{1_n} y + \text{Proy}_{x_c} y = \frac{y \cdot 1_n}{1_n \cdot 1_n} 1_n + \frac{y \cdot x_c}{x_c \cdot x_c} x_c. \tag{9.5}$$

Además si se denomina $b_0 = \frac{y \cdot 1_n}{1_n \cdot 1_n}$ y $b_1 = \frac{y \cdot x_c}{x_c \cdot x_c}$, (9.5) se escribe también como:

$$\widehat{y} = \widehat{b}_0 1_n + \widehat{b}_1 (x - \overline{x} 1_n) = (\widehat{b}_0 - \widehat{b}_1 \overline{x}) 1_n + \widehat{b}_1 x \tag{9.6}$$

Finalmente, como las coordenadas de un vector son únicas, tenemos de (9.4) y (9.6) que

$$\begin{aligned} \widehat{a}_1 &= \widehat{b}_1 = \frac{y \cdot x_c}{x_c \cdot x_c} \\ \widehat{a}_0 &= \widehat{b}_0 - \widehat{b}_1 \overline{x} = \overline{y} - \widehat{b}_1 \overline{x} \end{aligned} \tag{9.7}$$

dado que $\widehat{b}_0 = \frac{y \cdot 1_n}{1_n \cdot 1_n} = \overline{y}$.

Las relaciones (9.7) proveen la solución buscada.

9.1.3 Aplicación al ejemplo 9.1

Ahora en términos de los datos del ejemplo se tiene

$$\overline{x} = \frac{1}{4} \sum x_i = 11.25, \qquad \overline{y} = \frac{1}{4} \sum y_i = 13.5$$

$$\text{y} \qquad x_c = x - \overline{x} 1_4 = (-6.25, -1.25, 1.75, 5.75)^t.$$

Luego las estimaciones para los coeficientes $\widehat{a}_0$, $\widehat{a}_1$ de nuestro modelo son:

$$\begin{aligned} \widehat{a}_1 &= \tfrac{y \cdot x_c}{x_c \cdot x_c} = \tfrac{73.5}{76.75} = 0.9576 \\ \widehat{a}_0 &= \overline{y} - \widehat{b}_1 \overline{x} = 13.5 - 0.9576 * 11.25 = 2.726. \end{aligned}$$

El gráfico 9.4 muestra los datos iniciales (x_i, y_i) y la recta de mejor ajuste $y = \widehat{a}_0 + \widehat{a}_1 x$.

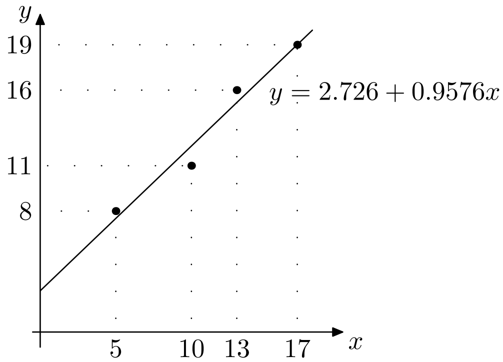


Figura 9.4: Recta $y = \widehat{a}_0 + \widehat{a}_1 x$ y puntos (x_i, y_i).

9.1.4 Calidad de las estimaciones

Para cuantificar la calidad de de las estimaciones dadas en (9.7), utilizamos el coeficiente de correlación (ver subsección 5.4.1) entre los vectores y, $\widehat{y}$, esto es

$$r(y, \widehat{y}) = \frac{\langle \widehat{y}_c, y_c \rangle_D}{\|\widehat{y}_c\|_D \, \|y_c\|_D}.$$

El subíndice c indica que las variables son centradas. Naturalmente esta medida surge de estudiar el vector de residuos e en la relación $y = \widehat{y} + e$. En la Figura 9.5 se ilustra esta relación, haciendo ver que los vectores e y $\widehat{y}$ son catetos de un triángulo rectángulo de hipotenusa y, y que el proceso de centraje reduce esta relación a $y_c = \widehat{y}_c + e$ involucrando el mismo vector de residuos e, lo que define un nuevo triángulo rectángulo de catetos $\widehat{y}_c$ y e e hipotenusa y_c. El teorema 9.1 precisa estas ideas.

De la visualización del teorema 9.1 en la Figura 9.5, se reconoce que el coeficiente $r(y, \widehat{y})$ corresponde al coseno del ángulo θ entre los vectores y_c y $\widehat{y}_c$. Por lo tanto $-1 \leq r(y, \widehat{y}) \leq 1$, y además si $r(y, \widehat{y})$ es cercano a 1 o -1 es porque θ es cercano a 0 o π, lo que significa que la magnitud del vector de residuos e es pequeña y que $\widehat{a}_0 + \widehat{b}_0 x$ explica bien la variable y. Contrariamente, si $r(y, \widehat{y})$

se aleja de 1 o -1 (se acerca a 0) los errores se hacen mayores y el modelo falla en explicar la variable y.

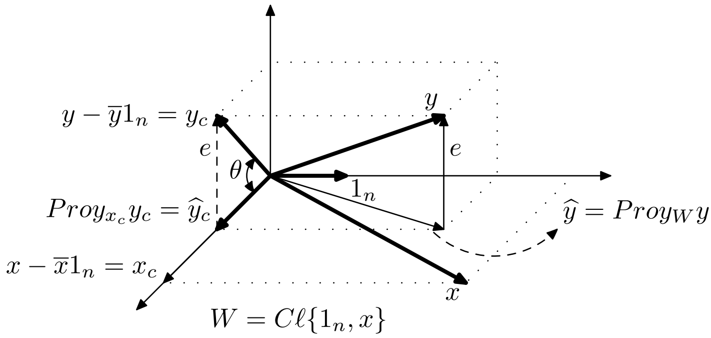


Figura 9.5: $e = y - \widehat{y} = y_c - \widehat{y}_c$.

Se deja como ejercicio verificar que el valor de r no cambia si se utiliza el producto punto en lugar del producto con la matriz de pesos $D = \frac{1}{n} I_n$, además dos vectores son D-ortogonales si y sólo si son I_n-ortogonales. Se tiene entonces

$$r(y, \widehat{y}) = \frac{y_c \cdot \widehat{y}_c}{\|y_c\| \, \|\widehat{y}_c\|}.$$

Teorema 9.1 *Sean* $y, x, 1_n \in I\!R^n$, $W = \mathcal{C}\ell\{1_n, x_c\}$, $\widehat{y} = Proy_W y$, $e = y - \widehat{y}$. *Entonces*

a) $\overline{e} = 0$ *(La media del vector de residuos es cero).*

b) $\overline{\widehat{y}} = \overline{y}$ *(y, $\widehat{y}$ tienen la misma media).*

c) $\widehat{y_c} = (\widehat{y})_c$ *(Centrar y y luego estimar es igual a estimar y y luego centrar).*

d) $e = y_c - \widehat{y}_c$.

e) $r(y, \widehat{y}) = \frac{\|\widehat{y}_c\|}{\|y_c\|}$.

Demostración:

a) De $\widehat{y} = \text{Proy}_W y$ se sigue $e = y - \widehat{y} \in W^{\perp}$ y como $1_n \in W$ entonces
$$\overline{e} = \langle e, 1_n \rangle_D = 0.$$

b) $y = \widehat{y} + e$ luego $\overline{y} = \overline{\widehat{y}} + \overline{e} = \overline{\widehat{y}}$.

c) $\text{Proy}_W y_c = \text{Proy}_{x_c} y_c + \text{Proy}_{1_n} y_c$ como $\text{Proy}_{1_n} y_c = 0$ se tiene que $\widehat{y_c} = \text{Proy}_{x_c} y_c$. Por otro lado
$$\begin{aligned}\text{Proy}_W y_c &= \text{Proy}_W y - \overline{y} 1_n \\ &= \text{Proy}_W y - \overline{y}\text{Proy}_W 1_n \\ &= \widehat{y} - \overline{y} 1_n \\ &= \widehat{y}_c\end{aligned}$$

d) $$\begin{aligned}y_c - \widehat{y_c} &= y_c - \widehat{y}_c \\ &= y_c - (\widehat{y} - \overline{\widehat{y}}) \\ &= y - \overline{y}1_n - \widehat{y} + \overline{y}1_n \\ &= y - \widehat{y} = e.\end{aligned}$$

e) $\widehat{y_c} = \widehat{y} - \overline{\widehat{y}} 1_n \in W$ y $e \in W^{\perp}$ luego
$$\langle e, \widehat{y_c} \rangle_D = \langle y_c - \widehat{y_c}, \widehat{y_c} \rangle_D = 0$$
de donde $\langle y_c, \widehat{y_c} \rangle_D = \|\widehat{y_c}\|_D^2$. Finalmente de esta última igualdad se tiene
$$r(y, \widehat{y}) = \frac{\langle y_c, \widehat{y_c} \rangle_D}{\|y_c\|_D \|\widehat{y_c}\|_D} = \frac{\|\widehat{y_c}\|_D}{\|y_c\|_D} = \frac{\|\widehat{y_c}\|}{\|y_c\|}.$$

■

Cálculos según los datos del ejemplo 9.1

En este ejemplo vimos que $x_c = (-6.25, -1.25, 1.75, 5.75)$, $\widehat{a}_1 = 0.9576$ y $y_c = (-5.5, -2.5, 2.5, 5.5)$, por lo tanto
$$\widehat{y_c} = \widehat{a}_1 x_c = (-5.98534, -1.19707, 1.6759, 5.50651).$$

Luego el valor del coeficiente de correlación es:
$$r(y, \widehat{y}) = \frac{\|\widehat{y_c}\|}{\|y_c\|} = 0.981944.$$

9.2 Regresión Lineal Múltiple

Analicemos ahora el caso de varias variables independientes (explicativas) y una variable dependiente a explicar.

Suponga que tenemos p variables $x_1, \ldots, x_p$, que explican la variable y. Y más específicamente que, salvo por un cierto error ϵ, y se puede explicar por las variables $x_1, \ldots, x_p$, mediante la siguiente relación:

$$y = b_1 x_1 + \cdots + b_p x_p + \epsilon. \tag{9.8}$$

Nuevamente, la exposición de resultados y su justificación se apoyarán en un ejemplo.

9.2.1 Planteo del modelo

Ejemplo 9.2 Consideremos que los principales factores que inciden en el rendimiento del cultivo del trigo son: a) Potasio y fósforo (kg/Ha), b) Nitrógeno (kg/Ha), c) Agua de lluvia promedio (cm^3), d) Acidez del suelo (pH) y e) Temperatura promedio ($^\circ$). Y que las respectivas variables explicativas se denotarán como:

x_1: Cantidad de potasio y fósforo (mezcla) por hectárea.

x_2: Cantidad de nitrógeno por hectárea.

x_3: Acidez del suelo (pH).

x_4: Cantidad promedio de lluvia caída (cm^3).

x_5: Temperatura promedio.

y: Rendimiento medido en quintales/Ha. (variable dependiente a explicar)

Nuestro objetivo es medir de alguna forma la incidencia en la variable y de los factores explicativos. Asumimos que el efecto de

	x_1	x_2	x_3	x_4	x_5	y
	Pot	Nit	A.lluv	PH	tem	rend
1	1100	300	6	5	10	30
2	1000	200	4	7	8	20
3	1200	350	6.7	8	10	40
4	1000	300	5	6	8	25
5	1100	300	5.5	7	9	35
6	1200	350	8	6	11	45
7	900	300	4	5	8	30
8	700	400	3.5	3	7	25
9	1200	350	6	7	7	35
10	1300	350	7	6.5	10	40

Tabla 9.1: Rendimiento de trigo, y, en 10 parcelas según los factores x_i.

cada factor sobre el rendimiento es aditivo de manera que y se expresa en términos de los factores explicativos como:

$$y = b_0x_0 + b_1x_1 + b_2x_2 + \cdots + b_5x_5 + \epsilon \tag{9.9}$$

donde se ha incluido un factor explicativo x_0 (variable cuyo valor es constante e igual a 1) y que tiene el efecto de agregar un término constante b_0 en la relación 9.9.

Al conocer los valores de los b_i de este modelo, se explica el poder de incidencia de cada factor x_i en la variable y, lo que permite estimar el rendimiento del cultivo de trigo y, para diferentes valores de los factores explicativos x_i (variables independientes), entre otros objetivos.

A fin de disponer de los datos necesarios para estimar los valores $b_0, \dots, b_5$ se realizó el siguiente experimento: cultivar trigo en 10 campos diferentes sometido a distintos tratamientos de las variables explicativas. Las mediciones resultantes de las 6 variables, en cada caso, se resumen en la tabla 9.1.

Si convenimos en que los valores observados de cada variable se denotan con los vectores $x_0, x_1, ...x_5,\ y \in I\!R^{10}$, $(x_0 = 1_n, n = 10)$, –observe que por comodidad los vectores de datos observados y las respectivas variables usan el mismo nombre– la relación 9.9 se expresa vectorialmente como:

$$\begin{pmatrix} 30\\20\\40\\25\\35\\45\\30\\25\\35\\40 \end{pmatrix} = b_0 \begin{pmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1 \end{pmatrix} + b_1 \begin{pmatrix} 1100\\1000\\1200\\1000\\1100\\1200\\900\\700\\1200\\1300 \end{pmatrix} + b_2 \begin{pmatrix} 300\\200\\350\\300\\300\\350\\300\\400\\350\\350 \end{pmatrix} + b_3 \begin{pmatrix} 6\\4\\6.7\\5\\5.5\\8\\4\\3.5\\6\\7 \end{pmatrix}$$

$$+ b_4 \begin{pmatrix} 5\\7\\8\\6\\7\\6\\5\\3\\7\\6.5 \end{pmatrix} + b_5 \begin{pmatrix} 10\\8\\10\\8\\9\\11\\8\\7\\7\\10 \end{pmatrix} + \begin{pmatrix} e_1\\e_2\\e_3\\e_4\\e_5\\e_6\\e_7\\e_8\\e_9\\e_{10} \end{pmatrix}$$

o lo que es lo mismo

$$\begin{pmatrix} 30\\20\\40\\25\\35\\45\\30\\25\\35\\40 \end{pmatrix} = \begin{pmatrix} 1 & 1100 & 300 & 6 & 5 & 10\\ 1 & 1000 & 200 & 4 & 7 & 8\\ 1 & 1200 & 350 & 6.7 & 8 & 10\\ 1 & 1000 & 300 & 5 & 6 & 8\\ 1 & 1100 & 300 & 5.5 & 7 & 9\\ 1 & 1200 & 350 & 8 & 6 & 11\\ 1 & 900 & 300 & 4 & 5 & 8\\ 1 & 700 & 400 & 3.5 & 3 & 7\\ 1 & 1200 & 350 & 6 & 7 & 7\\ 1 & 1300 & 350 & 7 & 6.5 & 10 \end{pmatrix} \begin{pmatrix} b_1\\b_2\\b3\\b_4\\b_5 \end{pmatrix} + \begin{pmatrix} e_1\\e_2\\e_3\\e_4\\e_5\\e_6\\e_7\\e_8\\e_9\\e_{10} \end{pmatrix}$$

Así el modelo se escribe como

$$y = Xb + e$$

donde y es el vector cuyas entradas son los valores que asume la variable y en los 10 campos, $X = (x_0, x_1, x_2, x_3, x_4, x_5)$ la matriz de tamaño 10×6, cuyas entradas son los valores de las variables independientes y $b = (b_0, b_1, \cdots, b_5)^t$ el vector de parámetros a estimar.

El problema es determinar el vector de parámetros $\widehat{b}$ tal que $\widehat{y} = X\widehat{b}$ sea los más cercano a y, o lo que es lo mismo, que minimice la magnitud del vector de residuos $\|y - X\widehat{b}\|$.

9.2.2 Solución geométrica

Geométricamente, el problema corresponde a encontrar la combinación lineal de los vectores $x_0, x_1, ...x_5$:

$$\widehat{y} = X\widehat{b} = \widehat{b}_0 x_0 + \widehat{b}_1 x_1 + \widehat{b}_2 x_2 + \widehat{b}_3 x_3 + \widehat{b}_4 x_4 + \widehat{b}_5 x_5$$

más próxima a y. Si denotamos con $W = \mathcal{C}\ell\{x_0, x_1, ...x_5\}$, se debe encontrar $\widehat{b} \in I\!R^6$ y $\widehat{y} = X\widehat{b} \in W$ tal que $\|y - \widehat{y}\|$ sea mínima.

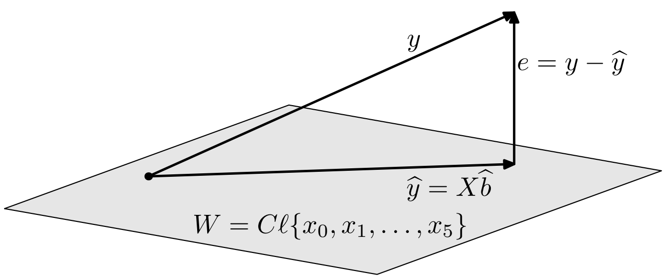


Figura 9.6: $\widehat{y} = X\widehat{b} = \text{Proy}_W y$.

Sabemos que el vector en W más cercano a y es $\widehat{y} = X\widehat{b} = \text{Proy}_W y$, sin embargo para calcular $\widehat{b}$ utilizando proyecciones sobre subespacios, necesitamos una base ortonormal de W, que no tenemos. Y ortonormalizar la base $\{x_0, x_1, ...x_5\}$ (si suponemos que son l.i.) demandaría más esfuerzo que el siguiente procedimiento alternativo.

El vector de residuos $e = y - \widehat{y}$ es ortogonal a W de manera que $(y - \widehat{y}) \cdot x_j = 0$, o lo que es lo mismo, $x_j^t (y - \widehat{y}) = 0$, para cada

$j = 0, 1, \ldots, 5$. Entonces

$$\begin{array}{rrcl} & X^t(y - \widehat{y}) & = & 0 \\ \Longrightarrow & X^t(y - X\widehat{b}) & = & 0 \\ \Longrightarrow & X^t y - X^t X\widehat{b} & = & 0 \\ \Longrightarrow & X^t X\widehat{b} & = & X^t y. \end{array}$$

Suponiendo que X^tX sea una matriz invertible, lo cual se tiene cuando los vectores columna $x_0, x_1, \ldots, x_5$ son linealmente independientes, se obtiene finalmente que:

$$\widehat{b} = (X^tX)^{-1}X^ty. \tag{9.10}$$

9.2.3 Indice de calidad

Análogamente al caso de dos variables, medimos la calidad de la estimación con el coeficiente de correlación entre los vectores y y $\widehat{y}$:

$$r(y, \widehat{y}) = \frac{\|\widehat{y}_c\|}{\|y_c\|} = \cos(\theta) \tag{9.11}$$

que se llama índice de calidad de la regresión (θ es el ángulo entre el vector y_c y su proyección ortogonal $\widehat{y}_c$, sobre $W = \mathcal{C}\ell\{x_{1c}, \ldots, x_{5c}\}$, donde x_{ic} es la variable x_i centrada).

9.2.4 Ejemplo 9.2: estimaciones con datos no centrados

A continuación realizaremos los cálculos correspondientes para estimar los vectores $\widehat{b}$ y $\widehat{y}$ de acuerdo con la información que disponemos:

$$X^tX = \begin{pmatrix} 10 & 10700 & 3200 & 55.7 & 60.5 & 88 \\ 10700 & 11730000 & 3425000 & 61640 & 66450 & 95600 \\ 3200 & 3425000 & 1050000 & 18045 & 19125 & 28200 \\ 55.7 & 61640 & 18045 & 329.39 & 346.1 & 505 \\ 60.5 & 66450 & 19125 & 346.1 & 384.25 & 538 \\ 88 & 95600 & 28200 & 505 & 538 & 792 \end{pmatrix}$$

Calculando $(X^tX)^{-1}X^ty = \widehat{b}$ obtenemos

$\widehat{b}^t = (-0.26.0822, 0.00276874, 0.0714974, 1.80473, 1.38543, 1.62571)$

con lo cual

$$\widehat{y} = X\widehat{b} =$$

$$\begin{pmatrix} 1 & 1100 & 300 & 6 & 5 & 10 \\ 1 & 1000 & 200 & 4 & 7 & 8 \\ 1 & 1200 & 350 & 6.7 & 8 & 10 \\ 1 & 1000 & 300 & 5 & 6 & 8 \\ 1 & 1100 & 300 & 5.5 & 7 & 9 \\ 1 & 1200 & 350 & 8 & 6 & 11 \\ 1 & 900 & 300 & 4 & 5 & 8 \\ 1 & 700 & 400 & 3.5 & 3 & 7 \\ 1 & 1200 & 350 & 6 & 7 & 7 \\ 1 & 1300 & 350 & 7 & 6.5 & 10 \end{pmatrix} \begin{pmatrix} -26.082 \\ 0.00277 \\ 0.07150 \\ 1.80473 \\ 1.38543 \\ 1.62571 \end{pmatrix} = \begin{pmatrix} 32.425 \\ 20.908 \\ 41.696 \\ 28.477 \\ 32.668 \\ 42.897 \\ 25.010 \\ 26.307 \\ 34.170 \\ 40.436 \end{pmatrix}$$

La calidad de los parámetros estimados se mide con el coeficiente de correlación entre y y $\widehat{y}$, como en el caso de dos variables:

$$r(y, \widehat{y}) = \frac{||\widehat{y}_c||}{||y_c||} = \frac{7.09563}{7.5} = 0.946084.$$

El siguiente gráfico muestra los valores experimentales (variable y) y los valores estimados por el modelo (variable $\widehat{y}$), o sea, los puntos $(y_i, \widehat{y}_i)$. También se agrega la recta identidad para visualizar más fácilmente los puntos $(y_i, \widehat{y}_i)$ muy próximos a la recta lo cual indica que $\widehat{y}_i$ es muy próximo a y_i.

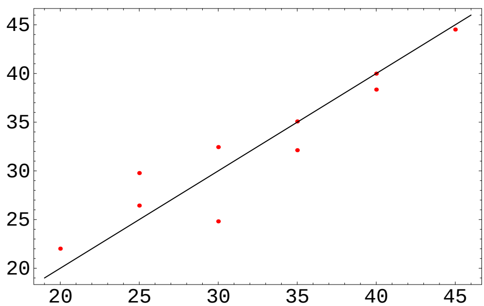


Figura 9.7: Puntos $(y_i, \widehat{y}_i)$ y recta $\widehat{y} = y$.

Resumimos los resultados obtenidos al estudiar el ejemplo anterior en la siguiente forma.

9.2.5 Resumen

Modelo de regresión lineal múltiple: cuando una variable y se explica por otras $x_1, \ldots, x_p$ variables mediante la relación

$$y = b_0 + b_1 x_1 + \ldots + b_p x_p + \epsilon,$$

los valores estimados para los parámetros $b_0, b_1, \ldots, b_p$, usando el criterio de calidad de los mínimos cuadrados, se obtienen según el teorema 9.2.

Teorema 9.2 (Mínimos cuadrados, datos no centrados) *Sean $y,\ x_0, \ldots, x_p \in I\!R^n$, vectores de valores observados de las variables del modelo de regresión lineal múltiple, con $x_0 = 1_n$ una variable de valor constante 1. En estos términos la relación asumida es*

$$y = b_0 x_0 + b_1 x_1 + \ldots + b_p x_p + e = Xb + e$$

donde $X = (x_0, x_1, \ldots, x_p)_{n \times (p+1)}$ y $b = (b_0, \ldots, b_p)^t$. Además suponga que $\{x_0, x_1, \ldots, x_p\}$ es linealmente independiente. Entonces al estimar el vector b usando mínimos cuadrados se obtiene que

$$\widehat{b} = Proy_{W_X} y = (X^t X)^{-1} X^t y$$

con $W_X = \mathcal{C}\ell\{x_0, x_1, \ldots, x_p\}$. Y el coeficiente de correlación lineal entre y y $\widehat{y}$, $r(y, \widehat{y}) = \frac{||\widehat{y_c}||}{||y_c||}$ es una medida de calidad del ajuste, donde $\widehat{y} = X\widehat{b}$.

Al calcular $\widehat{b}$, con el criterio de los mínimos cuadrados, el vector de residuos e resulta ortogonal a 1_n, de manera que la media de estos residuos es cero. Esto conduce a la igualdad entre las medias de y y $\widehat{y}$:

$$\overline{y} = \widehat{b}_0 \overline{x}_0 + \widehat{b}_1 \overline{x}_1 + \ldots + \widehat{b}_p \overline{x}_p,$$

y entonces

$$y - \overline{y} 1_n = \widehat{b}_1 (x_1 - \overline{x}_1 1_n) + \ldots + \widehat{b}_p (x_p - \overline{x}_p 1_n) = Z\widehat{a},$$

donde $\widehat{a} = (\widehat{b}_1, \ldots, \widehat{b}_p)^t$, $Z = (z_1, \ldots, z_p)_{n \times p}$ y $z_i = x_i - \overline{x}_i 1_n$ para $i = 1, \ldots, p$.

Estas variables centradas y la matriz Z pueden ser usadas en el cómputo del vector $\widehat{b}$, como lo muestra el siguiente teorema.

Teorema 9.3 (Variables centradas) *Sean* $y,\ x_0, \ldots, x_p \in I\!R^n$ *como en el teorema 9.2. Y sean* $z_i = x_i - \overline{x}_i 1_n$, $i = 1, \ldots p$ *los respectivos vectores centrados,* $Z = (z_1, \ldots, z_p)_{n \times p}$ *y* $W_Z = \mathcal{C}\ell\{z_1, \ldots, z_p\}$. *Entonces al estimar los parámetros* $b = (b_0, \ldots, b_p)^t$, *del modelo de regresión lineal múltiple*

$$y = b_0 x_0 + b_1 x_1 + \ldots + b_p x_p + e$$

usando mínimos cuadrados con los datos centrados, se tiene que

$$\begin{aligned} (\widehat{b_1}, \ldots, \widehat{b}_p)^t = & \quad Proy_{W_Z} y = (Z^t Z)^{-1} Z^t y \\ \widehat{b}_0 = & \quad \overline{y} - \textstyle\sum_{i=1}^p \widehat{b}_i \overline{x}_i . \end{aligned}$$

Demostración: Sea $H = \mathcal{C}\ell\{1_n\}$. Como $z_i \cdot 1_n = 0$ se tiene que H, W_Z son subespacios ortogonales, además $W_X = W_Z + H$. Se sigue entonces que

$$\begin{aligned} \mathrm{Proy}_{W_X} y &= \mathrm{Proy}_{W_Z} y + \mathrm{Proy}_H y \\ \mathrm{Proy}_{W_X} y &= \textstyle\sum_{i=0}^p \widehat{b}_i x_i \\ \mathrm{Proy}_{W_Z} y &= \textstyle\sum_{i=1}^p \widehat{a}_i z_i \\ \mathrm{Proy}_H y &= \overline{y} 1_n \end{aligned}$$

luego

$$\begin{aligned} \textstyle\sum_{i=0}^p \widehat{b}_i x_i = & \quad \textstyle\sum_{i=1}^p \widehat{a}_i (x_i - \overline{x}_i 1_n) + \overline{y} 1_n \\ = & \quad \textstyle\sum_{i=1}^p \widehat{a}_i x_i + (\overline{y} - \textstyle\sum_{i=1}^p \widehat{a}_i \overline{x}_i) 1_n . \end{aligned}$$

Como las coordenadas en una base son únicas, se tiene

$$\begin{aligned}\widehat{b}_i &= \widehat{a}_i, \quad i = 1, \ldots p \\ \widehat{b}_0 &= \overline{y} - \textstyle\sum_{i=1}^{p} \widehat{b}_i \overline{x}_i.\end{aligned}$$

Finalmente

$$\begin{aligned}(\widehat{b}_1, \ldots, \widehat{b}_p)^t &= \text{Proy}_{W_Z} y = (Z^t Z)^{-1} Z^t y \\ \widehat{b}_0 &= \overline{y} - \textstyle\sum_{i=1}^{p} \widehat{b}_i \overline{x}_i.\end{aligned}$$

■■■

9.2.6 Ejemplo 9.2: estimaciones usando datos centrados

Considere la matriz $X = (x_1, \ldots, x_5)$, que a diferencia del caso con datos no centrados, no incluye la primera columna de unos. Entonces centrando las variables la matriz $Z = (I_n - \frac{1}{n} 1_n 1_n^t) X$, $n = 10$ es:

$$Z = \begin{pmatrix} 30 & -20 & 0.43 & -1.05 & 1.2 \\ -70 & -120 & -1.57 & 0.95 & -0.8 \\ 130 & 30 & 1.13 & 1.95 & 1.2 \\ -70 & -20 & -0.57 & -0.05 & -0.8 \\ 30 & -20 & -0.07 & 0.95 & 0.2 \\ 130 & 30 & 2.43 & -0.05 & 2.2 \\ -170 & -20 & -1.57 & -1.05 & -0.8 \\ -370 & 80 & -2.07 & -3.05 & -1.8 \\ 130 & 30 & 0.43 & 0. - 95 & -1.8 \\ 230 & 30 & 1.43 & 0.45 & 1.2 \end{pmatrix}$$

Estimación de los parámetros b_i:

$$(\widehat{b}_1, \ldots, \widehat{b}_5) = (Z^t Z)^{-1} Z^t y = (0.0028, 0.0715, 1.805, 1.385, 1.626)$$

$$\widehat{b}_0 = \overline{y} - \textstyle\sum_{i=1}^{5} \widehat{b}_i \overline{x}_i = -26.08$$

Cálculo del vector de valores estimados para y, sin centrar y centrado, según el modelo.

$$\begin{aligned}\widehat{y} &= X(b_1, \ldots b_5)^t + b_0 1_5 \\ &= (32.43, 20.91, 41.7, 28.48, 32.67, 42.9, \\ &\qquad\qquad 25.01, 26.31, 34.17, 40.44)^t\end{aligned}$$

$$\begin{aligned}\widehat{y}_c &= (-0.07469, -11.59, 9.197, -4.022, 0.1681, \\ &\qquad 10.4, -7.489, -6.192, 1.671, 7.937)^t\end{aligned}$$

Indice de calidad, calculado con los datos anteriores:

$$r(y, \widehat{y}) = \frac{||\widehat{y}_c||}{||y_c||} = 0.946084.$$

Observación: si usamos en la definición del producto interno la matriz de pesos D, en lugar de la identidad I_n, el teorema 9.2 no se altera en lo fundamental y además

$$(\widehat{b}_1, \ldots, \widehat{b}_p) = (Z^t D Z)^{-1} Z^t y = V^{-1} Z^t D y$$

donde V es la matriz de varianza covarianza de las variables x_i.

9.3 Ejercicios

1. En un espectrofotómetro se mide absorvancia para muestras de diferentes concentraciones de una sustancia. Los datos obtenidos son los siguientes

C_n	0.25	0.50	1.00	1.50	2.00	0.00
A	0.03	0.02	0.05	0.07	0.09	0.00

Sabemos que la relación entre absorvancia y concentración es lineal

$$A = a_0 + a_1 C_n + \epsilon$$

Estime la concentración de la sustancia para una absorvancia de 0.04.

2. Use regresión lineal múltiple para estimar los valores de a, b, c, y d del polinomio cúbico $y = a + bx + cx^2 + dx^3$ que mejor ajusta los datos de la siguiente tabla.

x	3	0	-1	2	1
y	-2	3	4	-2	2

Calcule el índice de calidad de la regresión (coeficiente de correlación) y discútalo. Indicación, defina $x_1 = x$, $x_2 = x^2$, $x_3 = x^3$.

3. Un fabricante compra grandes cantidades de ciertos repuestos para máquinas. Encuentra que su costo depende del número de cajas de piezas que compre de una sola vez y que el costo por unidad disminuye al aumentar el número de unidades compradas. Supone que el costo es una función cuadrática del volumen de compra. De experiencias anteriores obtuvo la siguiente tabla.

número cajas	10	30	50	100	175
costo total \$	150	260	325	500	570

Encuentre la función de costo total.

4. El siguiente problema busca determinar la cantidad de alimento que se le debe suministrar a una vaca de acuerdo con su producción de leche. Los datos corresponden a la Finca del señor Ricardo Madriz Arias, el cual posee 50 vacas en producción. La Cooperativa de Productores de leche Dos pinos usa la fórmula : **Cantidad alim= 0.4 Producción.**

Corrobore esa ecuación con los datos de Don Ricardo, obtenidos de un muestreo de diez vacas. **Observación.** Note que en el modelo se asume que no hay término constante.

P(lb / día)	49.5	38	34.5	32.0	27.5
C.A (lb)	20	15	14	13	11

P(lb / día)	25.0	21.0	17.0	13.5	7.0
C.A (lb)	10	9	7	5	3

5. Los siguientes datos son el caudal promedio mensual obtenidos en la estación de Belén del Río Macho (datos del ICE).

Mes	may	jun	jul	ag	set	oc
Cau (m^3/s)	1.88	3.5	3.84	4.61	5.2	4.89

mes	no	di	en	feb	mar	ab
Cau $(m^3/s$	3.89	3.38	2.02	1.35	1.02	1.14

Usando una función cuadrática estime el caudal en función del mes y use esa información para estimar la fecha de máximo caudal y el valor de este (Esta estación aporta un 12% del caudal que necesita la planta Hidroeléctrica de Río Macho para funcionar).

6. Suponga que la variable y se explica aproximadamente en términos de la variable x, mediante la relación

$$y = a + bx + \epsilon.$$

Utilice regresión lineal a fin de estimar los parámetros a y b, del anterior modelo, para la siguiente tabla de datos:

x	-1	2	3	4
y	5	4	1	2

Detalle el procedimiento que lleva a su resultado.

7. Una cierta variable y se explica, aproximadamente, por la variable x, en términos del siguiente modelo:

$$y = a + bx^2 + \epsilon$$

donde ϵ es un error desconocido y a, y b son parámetros que se desean estimar. Si dispone de los siguientes datos observados:

x	1	2	3	4	5
y	85	60	36	45	75

a) Establezca el o los cambios de variable necesarios para resolver el problema utilizando un modelo de regresión lineal.

b) Determine la tabla de datos que requiere el modelo de regresión lineal, para estimar dichos parámetros.

c) Calcule las estimaciones de a y b según el modelo propuesto, y el coeficiente de correlación correspondiente.

8. Igual que en el problema anterior, una cierta variable y se explica, aproximadamente, por la variable x, en términos del siguiente modelo:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \epsilon$$

donde ϵ es un error desconocido y a_0, a_1, a_2 y a_3 son parámetros que se desean estimar. Si dispone de los siguientes datos observados:

x	-3	-2	-1	0	1	2	3
y	-3	-1	1	1	1	0	2

Responda las mismas preguntas que en el ejercicio anterior.

9. Suponga que la variable z se explica aproximadamente en términos de las variables x y y, mediante la relación

$$z = ax + by + \epsilon.$$

Utilice regresión lineal múltiple a fin de estimar los parámetros a y b, del anterior modelo, según la siguiente tabla de datos:

x	1	0	-1	1
y	-2	-1	0	2
z	5	3	1	2

Detalle las operaciones que conducen a su resultado y estime el valor de z cuando $x = 1$ y $y = 1$.

10. Se ha observado que en una granja, la producción de huevos que denominamos y depende de la cantidad suministrada de dos tipos de alimentos: x_1 y x_2. Suponga que la relación entre estas variables se explica apropiadamente por el modelo:

$$y = ax_1 + bx_2 + \epsilon,$$

y que para la tabla de datos siguiente (que se muestra parcialmente: faltan los valores observados para y),

x_1	1	0	1	2	1
x_2	0	1	1	1	2
y					

se estimaron los parámetros a, b utilizando regresión lineal múltiple, y se calculó $\hat{y}$ (y estimado) y $r(y, \widehat{y})$.

a) Haga una ilustración gráfica de $W = \mathcal{C}\ell\{x_1, x_2\}$, y y $\hat{y}$.

b) Si el vector y estimado resultó ser $\hat{y} = (2, 2, 4, 6, 6)^t$, calcule los parámetros a, b que se obtuvieron.

c) Si con estos datos, el coeficiente de correlación es $r = 0.902$. ¿Cuál es la norma del vector y_c ? (y centrado)

d) Estime la producción de huevos si las cantidades de los dos tipos de alimentos son: $x_1 = 1$ y $x_2 = 2$.

11. El rendimiento y de la alfalfa (toneladas por acre) depende linealmente de la cantidad x de agua aplicada (pulgadas por acre) de acuerdo con la relación: $y = 5 + kx$, que se puede expresar como $z = kx$, con $z = y - 5$.

Estimar el valor de k mediante regresión lineal, asumiendo que se hicieron 8 mediciones y que los datos expresados en los vectores $\vec{x}, \vec{y}$, $\vec{1}$ ($\vec{1}$ es el vector con sus 8 entradas iguales a 1) verifican:

$\vec{x} \cdot \vec{1} = 210$ $\quad\quad$ $\vec{x} \cdot \vec{x} = 7308$ $\quad\quad$ $\vec{x} \cdot \vec{y} = 1590.48$

Indicación: $\vec{z} = (\vec{y} - 5\vec{1})$

12. Cuando se extrae agua de un pozo, este recupera parcialmente su nivel con la captación de agua de sus alrededores (acuífero).

Si en un pozo cilíndrico se extrae agua mediante un bombeo de caudal constante de $225m/h$. El descenso en el nivel del agua se da en la siguiente tabla.

t (min)	200	280	300	350	400	500	600
Des (cm)	31	33	33.5	34	35	34	35

t (min)	700	900	1100	1200	1300	1440
Des (cm)	36	38	38.5	39.5	39.5	40

Fuente:Custopio E: Ensayo de bombeo en el pozo Cornellá-13 (Barcelona 1968)

Use una función lineal para estimar el descenso después de 1000 minutos de bombeo.

13. Una cierta variable y se explica, aproximadamente, por la variable x, en términos del siguiente modelo exponencial:

$$y = \alpha e^{-\beta x + \epsilon}$$

donde ϵ es un error desconocido y α, β son los parámetros que se desean estimar.

Si dispone de los siguientes datos observados:

x	2	4	6	12	18	24
y	1.07	1.88	2.26	2.78	2.97	2.99

a) Formule el problema como un problema de regresión lineal múltiple.

b) Calcule las estimaciones de los parámetros α y β.

c) Según este modelo, cuál es el valor estimado para y, cuando $x = 10$.

14. Sea $B = \{1_n, x_1, \ldots, x_p\}$, $p < n$, subconjunto de $I\!R^n$. $D = \frac{1}{n} I_n$ la matriz de pesos. Demuestre que:

a) Si B es linealmente independiente y $z_i = x_i - \overline{x}_i 1_n$ entonces $B_c = \{z_1, \ldots, z_p\}$ es l.i.

b) Si $W = \mathcal{C}\ell\{B\}$, $W_c = \mathcal{C}\ell\{B_c\}$, $H = \mathcal{C}\ell\{1_n\}$, H es el complemento ortogonal de W_c respecto de W.

c) Si $y \in I\!R^n$, y_c es el vector centrado correspondiente, entonces $\text{Proy}_W y = \text{Proy}_W y_c$.

15. Sea $H = I_n - \frac{1}{n} 1_n 1_n^t$ matriz de $n \times n$. Demuestre que

a) H es simétrica e idempotente ($H^2 = H$.)

b) $X_c = HX$. Al multiplicar X por H, por la izquierda, produce como resultado una matriz con las correspondientes columnas de X centradas.

16. Sean $x, y \in I\!R^n$ con $\widehat{y}$ la proyección de y a lo largo de x. Demuestre que si θ es el ángulo de y con $\widehat{y}$ entonces:

$$|\cos \theta| = \frac{||\widehat{y}||}{||y||}$$

17. Si el producto interno esta definido por una matriz de pesos D_p. Demuestre que:

1. $r^2(y, \widehat{y}) = \dfrac{\text{Var}\,(\widehat{y})}{\text{Var}\,(y)}$

2. $r^2(y,\widehat{y}) = \frac{\text{Cov}^2(y,\widehat{y})}{\text{Var}(y)\,\text{Var}(\widehat{y})}$. donde r es el coeficiente de correlación, $\text{Var}(y)$ es la varianza de y, $\text{Cov}(y,\widehat{y})$ es la covarianza entre y y $\widehat{y}$.

18. En el modelo $y = a + bx + \epsilon$. Demuestre que:

$$\widehat{b} = \frac{\sum_i^n x_i y_i - n\overline{xy}}{\sum_i^n x_i^2 - n\overline{x}^2} = \frac{\text{Cov}(x,y)}{\text{Var}(x)}$$

$$\widehat{a} = \overline{y} - \widehat{b}\overline{x}. \text{ donde } n \text{ es el número de datos.}$$

19. Sean x, y vectores en $I\!R^n$ de valores observados para las variables x y y y $1_n = (1,\ldots,1)^t$. Considere los dos modelos dados en términos de los vectores de datos observados:

$$y = a_1 x + a_0 1_n + \varepsilon \tag{9.12}$$
$$y - \overline{y}1_n = a_1(x - \overline{x}1_n) + \varepsilon \tag{9.13}$$

donde $\overline{x}$ y $\overline{y}$ son las medias aritméticas de x y y respectivamente ($\overline{x} = \sum_{i=1}^n x_i/n$). Observe que el modelo centrado 9.13 se obtiene del no centrado 9.12, por la relación entre las medias de x y y:

$$\overline{y} = a_1\overline{x} + a_0$$

a) Muestre (e ilustre gráficamente) que $\text{Proy}_{1_n} x = \overline{x}1_n$ y $(x - \overline{x}1_n) \perp 1_n$.

b) Justifique que $\|a_1 x + a_0 1_n\|^2 = \|a_1(x - \overline{x}1_n) + (a_1\overline{x} + a_0)1_n\|^2 = \|a_1(x - \overline{x}1_n)\|^2 + \|\overline{y}1_n\|^2$.

c) Muestre que $\|y\|^2 = \|y - \overline{y}1_n\|^2 + \|\overline{y}1_n\|^2$.

d) Finalmente pruebe que si definimos la calidad de la regresión utilizando el coseno cuadrado del ángulo entre los vectores $y, \widehat{y}$ $\;(R^2 = \frac{\|\widehat{y}\|^2}{\|y\|^2})$, en lugar del coseno cuadrado del ángulo entre los mismos vectores, pero centrados $(r^2 = \frac{\|\widehat{y_c}\|^2}{\|y_c\|^2})$, el resultado es diferente.

$$R^2 = \frac{\|\widehat{y}\|^2}{\|y\|^2} = \frac{\|a_1 x + a_0 1_n\|^2}{\|y\|^2} = \frac{\|a_1(x - \overline{x}1_n)\|^2 + n\overline{y}^2}{\|(y - \overline{y}1_n)\|^2 + n\overline{y}^2}$$

donde $\frac{\|a_1(x-\overline{x}1_n)\|^2}{\|(y-\overline{y}1_n)\|^2} = \frac{\|\widehat{y_c}\|^2}{\|y_c\|^2} = r(y,\widehat{y})^2$, es el cuadrado del coeficiente de correlación entre y, $\widehat{y}$.

Capítulo 10

Transformaciones Lineales

EL tema central de este capítulo es el estudio de una clase de funciones especiales, llamadas transformaciones lineales.

Una de las características importantes de las transformaciones lineales es que están totalmente determinadas por sus valores en una base cualquiera del espacio. Además en los espacios de dimensión finita (como es el caso que nos ocupa) toda transformación lineal puede ser representada por una matriz y recíprocamente a toda matriz se le puede asociar una transformación lineal.

Por ejemplo cuando se estudian sistemas de ecuaciones lineales donde el recurso fundamental es la teoría de matrices, se puede establecer una conexión inmediata con las transformaciones lineales para ver que la solución de un sistema homogéneo es el núcleo de una transformación lineal, y que la solución de un sistema no homogéneo son las preimágenes bajo una transformación lineal de un cierto vector fijo.

Otras veces un enfoque basado en el lenguaje de las transformaciones lineales nos permite deducir fácil y elegantemente, propiedades relativas a las matrices.

Aparte de lo anteriormente expuesto, el estudio de las transformaciones lineales cobra mayor interés en razón de que cuando una transformación entre espacios de dimensión finita no es lineal, se acostumbra bajo ciertas hipótesis, aproximarlas por la suma de una transformación lineal más una constante. Este procedimiento es típico en numerosos problemas tanto de matemática como de otras ciencias.

10.1 Concepto de transformación lineal

Definición 10.1 (Transformación lineal)

Sean V y W e.v. sobre $I\!R$; se llama transformación lineal de V a W (que abreviamos con t.l. de V a W), a toda función

$$T : V \longrightarrow W$$

que satisface para todo $v, u \in V$ y $\alpha \in I\!R$ lo siguiente:

(a) $T(\alpha v) = \alpha T(v)$ *(decimos que T "saca" el escalar α)*

(b) $T(v + u) = T(v) + T(u)$ *(T preserva las operaciones suma de los espacios vectoriales)*

Notación: Se denota al conjunto de las transformaciones lineales de V a W por $L(V, W)$ y si $V = W$ se escribe $L(V)$ en lugar de $L(V, V)$.

Ejemplo 10.1 Sea $T : I\!R^2 \longrightarrow I\!R^3$ una función definida por:

$$T(x, y) = (x + y, x - y, y)$$

Comprobar que $T \in L(I\!R^2, I\!R^3)$.

Solución

T saca el escalar:

$$\begin{array}{rcl} T(\alpha(x, y)) & = & T(\alpha x, \alpha y) \\ & = & (\alpha x + \alpha y, \alpha x - \alpha y, \alpha y) \\ & = & \alpha(x + y, x - y, y) \\ & = & \alpha T(x, y) \end{array}$$

T preserva la suma:

$$\begin{array}{rl} & T((x_1, y_1) + (x_2, y_2)) \\ = & T(x_1 + x_2, y_1 + y_2) \\ = & (x_1 + x_2 + (y_1 + y_2), x_1 + x_2 - (y_1 + y_2), y_1 + y_2) \\ = & (x_1 + y_1, x_1 - y_1, y_1) + (x_2 + y_2, x_2 - y_2, y_2) \\ = & T(x_1, y_1) + T(x_2, y_2) \end{array}$$

Por lo tanto: $T((x_1, y_1) + (x_2, y_2)) = T(x_1, y_1) + T(x_2, y_2)$.

Observaciones

(i) Una consecuencia directa de la propiedad (b) de la definición de t.l. es que T hace corresponder el cero de V con el cero de W.

En efecto, $T(0_v) = T(0_v + 0_v) = T(0_v) + T(0_v)$. Luego $T(0_v) = 0_w$. Nótese que usamos el símbolo 0_v para referirnos al cero de V y 0_w para el de W.

(ii) Las propiedades (a) y (b) definitorias de una t.l., se pueden resumir en una sola:

$$T \in L(V, W) \iff \begin{array}{l} T(\alpha v + u) = \alpha T(v) + T(u) \\ \forall\, v, u \in V \;\; \text{y} \; \forall\, \alpha \in I\!R. \end{array}$$

10.1.1 Imágenes de los vectores de una base determinan la t.l.

Una transformación lineal queda determinada por sus valores en los vectores de una base del espacio V.

Supongamos que para la transformación $T : V \longrightarrow W$, se conocen los valores $T(v_1), \ldots, T(v_n)$, donde $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base de V, entonces para cualquier $v \in V$ se tiene que:

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

y es posible determinar $T(v)$ porque:

$$T(v) = T(\sum_{i=1}^{n} \alpha_i v_i) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n).$$

Los siguientes ejemplos ilustran esta idea.

Ejemplo 10.2 Sea la transformación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ cuyos valores en la base canónica de $\mathbb{R}^3$ son: $T(e_1) = (1, 1)$, $T(e_2) = (0, 1)$, $T(e_3) = (1, -1)$. Obtenga la expresión general de $T(x, y, z)$.

Solución

Sea $(x, y, z) \in \mathbb{R}^3$, es decir $(x, y, z) = xe_1 + ye_2 + ze_3$. Puesto que T es una t.l. se tiene que:

$$\begin{aligned} T(x, y, z) &= xT(e_1) + yT(e_2) + zT(e_3) \\ &= x(1, 1) + y(0, 1) + z(1, -1) \\ &= (x + z, x + y - z) \end{aligned}$$

Ejemplo 10.3 Encontrar la transformación lineal
$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ que a todo vector de $\mathbb{R}^2$ lo rota en un ángulo θ en la dirección positiva, sin alterar su norma.

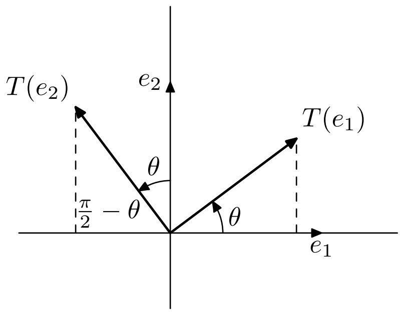


Figura 10.1: T rota vectores en un ángulo θ.

Como T queda determinada por sus valores en los vectores de una base, para determinar $T(x, y)$ es suficiente calcular $T(e_1)$ y $T(e_2)$. De la gráfica se observa que:

$$T(e_1) = (\cos\theta, \operatorname{sen}\theta) \qquad T(e_2) = (-\operatorname{sen}\theta, \cos\theta)$$

Luego

$$\begin{aligned} T(x, y) = T(xe_1 + ye_2) &= xT(e_1) + yT(e_2) \\ &= x(\cos\theta, \operatorname{sen}\theta) + y(-\operatorname{sen}\theta, \cos\theta) \\ &= (x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta) \end{aligned}$$

10.2 Relación entre transformaciones y matrices

Cada matriz en $M(m,n,\mathbb{R})$ determina una única transformación en $L(\mathbb{R}^n,\mathbb{R}^m)$, e inversamente, cada transformación lineal en $L(\mathbb{R}^n,\mathbb{R}^m)$ se puede asociar a una matriz $A \in M(m,n,\mathbb{R})$, única si fijamos una base para cada espacio, lo cual se explicará más adelante.

Esta asociación entre transformaciones y matrices se expresa por medio de una función biyectiva que preserva ciertas operaciones (esto es, el isomorfismo entre t.l. y matrices), lo cual permite afirmar que que transformaciones lineales y matrices son objetos matemáticos idénticos.

En primer término observemos que a toda matriz define una transformación lineal.

10.2.1 Toda matriz define una transformación lineal

Sea A una matriz de tamaño $m \times n$. A la matriz A le asociamos la transformación lineal:

$$T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m \text{ tal que } T_A(x) = Ax \ \ \forall\, x \in \mathbb{R}^n.$$

Claramente T_A es lineal. En efecto:

$$T_A(\alpha x + y) = A(\alpha x + y) = \alpha Ax + Ay = \alpha T_A(x) + T_A(y)$$

De lo anterior podemos concluir que toda matriz de tamaño $m\times n$ puede verse como una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$.

En estos casos conviene representar los vectores de los espacios $\mathbb{R}^n$ como vectores columna.

Ejemplo 10.4 La transformación $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, obtenida en el ejemplo 10.2 es de la forma "$T_A(x) = Ax$", porque puede ser escrita como:

$$T\begin{pmatrix} x\\ y\\ z\end{pmatrix} = \begin{pmatrix} x+z\\ x+y-z\end{pmatrix} = \begin{pmatrix} 1 & 0 & 1\\ 1 & 1 & -1\end{pmatrix}\begin{pmatrix} x\\ y\\ z\end{pmatrix}$$

Y además, se puede observar que las columnas de la matriz que se asocia a la transformación son las imágenes de los vectores de la base canónica de $I\!R^3$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix} = \left(T\begin{pmatrix}1\\0\\0\end{pmatrix} T\begin{pmatrix}0\\1\\0\end{pmatrix} T\begin{pmatrix}0\\0\\1\end{pmatrix} \right) = (T(e_1)T(e_2)T(e_3))$$

En el ejemplo 10.3 tenemos lo mismo:

$$T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\cos\theta - y\operatorname{sen}\theta\\ x\operatorname{sen}\theta + y\cos\theta\end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

y nuevamente las columnas de la matriz que asocia a la rotación T son las imágenes de los vectores canónicos:

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} = (T(e_1)\ T(e_2)).$$

10.2.2 Asociación de matrices a las transformaciones

En el caso particular de que se elijan las bases canónicas, para los espacios $I\!R^n$ y $I\!R^m$, se puede asociar a la transformación lineal $T : I\!R^n \longrightarrow I\!R^m$ una matriz A de la siguiente manera:

Para todo $x \in I\!R^n$, $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$, luego

$$\begin{array}{rcl} T(x) & = & T(\sum_{i=1}^n x_ie_i) \\ & = & x_1T(e_1) + x_2T(e_2) + \cdots + x_nT(e_n) \\ & = & (T(e_1)T(e_2)\cdots T(e_n))x \\ & = & Ax \end{array}$$

donde $A = (T(e_1)T(e_2)\cdots T(e_n))$ es una matriz $m \times n$, que se asocia a la transformación T y la denominamos matriz canónica de T.

La asociación de una matriz a una transformación, vista anteriormente, no es única, puesto que para cada pareja de bases $\mathcal{B}_1 \in I\!R^n$ y $\mathcal{B}_2 \in I\!R^n$, se puede pensar en una asociación de este tipo.

Sea $T : V \longrightarrow W$, una transformación lineal y $\mathcal{B}_1 = \{v_1, \cdots, v_n\}$ y $\mathcal{B}_2 = \{u_1, \cdots, u_m\}$ bases de V y W respectivamente. Para $i = 1, \cdots, n$, $T(v_i)$ se puede expresar de manera única como combinación lineal de $u_1, \cdots, u_m$:

$$\begin{array}{ccccc} T(v_1) & = & a_{11}u_1 & + \cdots + & a_{m1}u_m \\ T(v_2) & = & a_{12}u_1 & + \cdots + & a_{m2}u_m \\ \vdots & \vdots & \vdots & & \vdots \\ T(v_n) & = & a_{1n}u_1 & + \cdots + & a_{mn}u_m \end{array} \tag{10.1}$$

lo que también expresamos como:

$$[T(v_1)]_{\mathcal{B}_2} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \; [T(v_2)]_{\mathcal{B}_2} = \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} \cdots \; [T(v_n)]_{\mathcal{B}_2} = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Como los valores $T(v_1), \ldots, T(v_n)$ definen la transformación T, dada $\mathcal{B}_2$, los vectores $[T(v_1)]_{\mathcal{B}_2}, \ldots, [T(v_n)]_{\mathcal{B}_2}$ también determinan a T, y convenimos en la siguiente definición.

Definición 10.2 (Matriz de una transformación lineal)

La matriz de T (o representación matricial) en el par de bases $\mathcal{B}_1$ y $\mathcal{B}_2$ se define por:

$$[T]_{\mathcal{B}_1}^{\mathcal{B}_2} = ([T(v_1)]_{\mathcal{B}_2} \; \cdots \; [T(v_n)]_{\mathcal{B}_2}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Si $V = W$ y $\mathcal{B}_1 = \mathcal{B}_2$ escribimos $[T]_{\mathcal{B}_1}$ en lugar de $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$.

La matriz $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$ juega para T un papel similar que la matriz A para T_A.

Como $\mathcal{B}_1$ es base de V, todo $v \in V$ se escribe:

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

$$\Longrightarrow\ [v]_{\mathcal{B}_1} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Y dada la linealidad de T y las identidades en 10.1, se deduce que:

$$\begin{aligned}
T(v) &= \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) \\
&= \alpha_1(a_{11}u_1 + \cdots + a_{m1}u_m) \\
&\quad +\alpha_2(a_{12}u_1 + \cdots + a_{m2}u_m) \\
&\qquad \ddots \\
&\qquad\quad +\alpha_n(a_{1n}u_1 + \cdots + a_{mn}u_m) \\
&= (\alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n})u_1 \\
&\quad +(\alpha_1 a_{21} + \alpha_2 a_{22} + \cdots + \alpha_n a_{2n})u_2 \\
&\qquad \ddots \\
&\qquad\quad +(\alpha_1 a_{m1} + \alpha_2 a_{m2} + \cdots + \alpha_n a_{mn})u_m.
\end{aligned}$$

De lo cual se puede observar que:

$$\begin{aligned}
[T(v)]_{\mathcal{B}_2} &= \begin{pmatrix} \alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n} \\ \alpha_1 a_{21} + \alpha_2 a_{22} + \cdots + \alpha_n a_{2n} \\ \vdots \\ \alpha_1 a_{m1} + \alpha_2 a_{m2} + \cdots + \alpha_n a_{mn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \\
&= [T]_{\mathcal{B}_1}^{\mathcal{B}_2} [v]_{\mathcal{B}_1}.
\end{aligned}$$

Resultado que resumimos en el siguiente teorema:

Teorema 10.3 *Sea $T \in L(V,W)$, $\mathcal{B}_1 = \{v_1, \cdots, v_n\}$ una base de V, y $\mathcal{B}_2 = \{u_1, \cdots, u_m\}$ una base de W. Si $A = (a_{ij})_{m\times n} = [T]_{\mathcal{B}_1}^{\mathcal{B}_2}$ es la matriz de T en las bases $\mathcal{B}_1$ y $\mathcal{B}_2$, entonces*

$$[T(x)]_{\mathcal{B}_2} = [T]_{\mathcal{B}_1}^{\mathcal{B}_2}[x]_{\mathcal{B}_1} \quad \forall x \in V.$$

Ejemplo 10.5 Considere la t.l. $T : I\!R^3 \longrightarrow I\!R^2$ tal que

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ 3x-z \end{pmatrix}.$$

Y sean $\mathcal{C}_1 = \{e_1, e_2, e_3\}$, $\mathcal{C}_2 = \{f_1, f_2\}$ las bases canónicas de $I\!R^3$ y $I\!R^2$ respectivamente. Además $\mathcal{B}_1 = \{(1,1,2)^t, (-3,0,1)^t, (2,4,3)^t\}$ una base de $I\!R^3$ y $\mathcal{B}_2 = \{(4,1),(3,1)\}$ base de $I\!R^2$. Encuentre las matrices: $[T]_{\mathcal{C}_1}^{\mathcal{C}_2}$ y $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$.

Solución: En el primer caso resulta fácil ver que:

$$\begin{aligned}
T(e_1) &= \begin{pmatrix} 1 \\ 3 \end{pmatrix} = f_1 + 3f_2 \Longrightarrow [T(e_1)]_{\mathcal{C}_2} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \\
T(e_2) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} = f_1 + 0f_2 \Longrightarrow [T(e_2)]_{\mathcal{C}_2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\
T(e_3) &= \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0f_1 - f_2 \Longrightarrow [T(e_2)]_{\mathcal{C}_2} = \begin{pmatrix} 0 \\ -1 \end{pmatrix},
\end{aligned}$$

luego

$$[T]_{\mathcal{C}_1}^{\mathcal{C}_2} = ([T(e_1)]_{\mathcal{C}_2}[T(e_2)]_{\mathcal{C}_2}[T(e_3)]_{\mathcal{C}_2}) = \begin{pmatrix} 1 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix}.$$

Para obtener $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$ procedemos análogamente, sin embargo, no conocemos los coeficientes a_{ij} de las identidades 10.1 que corresponden a este caso, por lo que debemos plantear los respectivos sistemas de ecuaciones:

$$T\begin{pmatrix}1\\1\\2\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix} = a_{11}\begin{pmatrix}4\\1\end{pmatrix}+a_{21}\begin{pmatrix}3\\1\end{pmatrix}$$

$$T\begin{pmatrix}-3\\0\\1\end{pmatrix} = \begin{pmatrix}-3\\-10\end{pmatrix} = a_{12}\begin{pmatrix}4\\1\end{pmatrix}+a_{22}\begin{pmatrix}3\\1\end{pmatrix}$$

$$T\begin{pmatrix}2\\4\\3\end{pmatrix} = \begin{pmatrix}6\\3\end{pmatrix} = a_{13}\begin{pmatrix}4\\1\end{pmatrix}+a_{23}\begin{pmatrix}3\\1\end{pmatrix}$$

Observe que se trata de tres sistemas de ecuaciones lineales, escritos en su forma columnar, todos con la misma matriz del sistema:

$$\begin{pmatrix}4&3\\1&1\end{pmatrix}\begin{pmatrix}a_{11}\\a_{21}\end{pmatrix}=\begin{pmatrix}2\\1\end{pmatrix},$$
$$\begin{pmatrix}4&3\\1&1\end{pmatrix}\begin{pmatrix}a_{12}\\a_{22}\end{pmatrix}=\begin{pmatrix}-3\\-10\end{pmatrix},$$
$$\begin{pmatrix}4&3\\1&1\end{pmatrix}\begin{pmatrix}a_{13}\\a_{23}\end{pmatrix}=\begin{pmatrix}6\\3\end{pmatrix}.$$

Naturalmente, como en el caso del cómputo de una matriz inversa, se pueden resolver todos simultáneamente, mediante el siguiente proceso de cómputo:

$$\left(\begin{array}{cc|ccc}4&3&2&-3&6\\1&1&1&-10&3\end{array}\right)\xrightarrow{f_1,f_2}\left(\begin{array}{cc|ccc}1&1&1&-10&3\\4&3&2&-3&6\end{array}\right)$$
$$\xrightarrow{-4f_1+f_2}\left(\begin{array}{cc|ccc}1&1&1&-10&3\\0&-1&-2&37&-6\end{array}\right)$$
$$\xrightarrow[\ ]{\substack{f_2+f_1\\-f_2}}\left(\begin{array}{cc|ccc}1&0&-1&27&-3\\0&1&2&-37&6\end{array}\right).$$

Luego

$$[T]_{\mathcal{B}_1}^{\mathcal{B}_2} = \left(\left[T\begin{pmatrix}1\\1\\2\end{pmatrix}\right]_{\mathcal{B}_2}\left[T\begin{pmatrix}-3\\0\\1\end{pmatrix}\right]_{\mathcal{B}_2}\left[T\begin{pmatrix}2\\4\\3\end{pmatrix}\right]_{\mathcal{B}_2}\right)$$
$$= \begin{pmatrix}-1&27&-3\\2&-37&6\end{pmatrix}.$$

Este ejemplo muestra que si $T : I\!R^n \longrightarrow I\!R^m$ es una t.l., $\mathcal{B}_1 = \{v_1, \cdots, v_n\}$ una base de $I\!R^n$, y $\mathcal{B}_2 = \{u_1, \cdots, u_m\}$ una base de $I\!R^m$, el problema de calcular $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$ se resuelve aplicando operaciones elementales a la matriz

$$(u_1 u_2 \ldots u_m | T(v_1) T(v_2) \ldots T(v_n))$$

hasta obtener

$$(e_1 e_2 \ldots e_m | [T(v_1)]_{\mathcal{B}_2} [T(v_2)]_{\mathcal{B}_2} \ldots [T(v_n)]_{\mathcal{B}_2}) = (I_m | [T]_{\mathcal{B}_1}^{\mathcal{B}_2}).$$

Ejemplo 10.6 Usando la matriz $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$ del ejemplo anterior, calcular $T(0, 5, 6)^t$.

Solución

Sabemos que $[T(x)]_{\mathcal{B}_2} = [T]_{\mathcal{B}_1}^{\mathcal{B}_2} \; [x]_{\mathcal{B}_1}$ y como

$$\begin{pmatrix} 0 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}$$

se tiene que $\left[\begin{pmatrix} 0 \\ 5 \\ 6 \end{pmatrix} \right]_{\mathcal{B}_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Luego

$$\left[T \begin{pmatrix} 0 \\ 5 \\ 6 \end{pmatrix} \right]_{\mathcal{B}_2} = \begin{pmatrix} -1 & 27 & -3 \\ 2 & -37 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 23 \\ -29 \end{pmatrix}.$$

Finalmente

$$T \begin{pmatrix} 0 \\ 5 \\ 6 \end{pmatrix} = 23 \begin{pmatrix} 4 \\ 1 \end{pmatrix} - 29 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ -6 \end{pmatrix}.$$

Podemos verificar nuestros cálculos evaluando directamente en la trasformación lineal $T(0, 5, 6) = (0 + 5, 3(0) - 6) = (5, -6)$.

10.2.3 Matrices de cambio de base

Cuando consideramos la t.l. identidad de $I\!R^n$, que denotamos como I —no confundir con la matriz identidad— y un par de bases para $\mathbb{R}^n$: $\mathcal{B}_1 = \{u_1, \cdots, u_n\}$ y $\mathcal{B}_2 = \{v_1, \cdots, v_n\}$, la matriz de I en las bases $\mathcal{B}_1$ y $\mathcal{B}_2$, $[I]_{\mathcal{B}_1}^{\mathcal{B}_2}$, tiene una interpretación especialmente importante. Observe que para todo $x \in \mathbb{R}^n$, por el teorema 10.3:

$$[x]_{\mathcal{B}_2} = [I(x)]_{\mathcal{B}_2} = [I]_{\mathcal{B}_1}^{\mathcal{B}_2}[x]_{\mathcal{B}_1}$$

luego $[I]_{\mathcal{B}_1}^{\mathcal{B}_2}$ tiene el efecto de cambiar las coordenadas de x de la base $\mathcal{B}_1$ a la base $\mathcal{B}_2$.

Definición 10.4 (Matriz de cambio de base)

Si $\mathcal{B}_1$ y $\mathcal{B}_2$ son dos bases de un mismo espacio V e I la transformación identidad de V, se llama matriz de cambio de base de $\mathcal{B}_1$ a $\mathcal{B}_2$ a la matriz:

$$[I]_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Naturalmente, el cómputo de matrices de cambio de base, resulta en un caso particular de proceso para determinar una matriz $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$, ya visto en el ejemplo 10.5.

Ejemplo 10.7 Considere las bases de $\mathbb{R}^3$:

$$\mathcal{B}_1 = \{u_1, u_2, u_3\} = \{(1,0,1)^t, (0,1,1)^t, (-1,0,1)^t\},$$

$$\mathcal{B}_2 = \{v_1, v_2, v_3\} = \{(0,-1,1)^t, (1,0,-1)^t, (1,0,1)^t\}$$

y determine la matriz de cambio de base de $\mathcal{B}_1$ a $\mathcal{B}_2$.

Solución:

$$[I]_{\mathcal{B}_1}^{\mathcal{B}_2} = ([I(u_1)]_{\mathcal{B}_2}[I(u_2)]_{\mathcal{B}_2}[I(u_3)]_{\mathcal{B}_2}) = ([u_1]_{\mathcal{B}_2}[u_2]_{\mathcal{B}_2}[u_3]_{\mathcal{B}_2}).$$

Por otra parte,

$$[u_j]_{\mathcal{B}_2} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{pmatrix} \iff u_j = a_{1j}v_1 + a_{2j}v_2 + a_{3j}v_3$$

$$\iff u_j = (v_1 v_2 v_3) \begin{pmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{pmatrix}.$$

Luego deben resolverse simultáneamente estos tres sistemas de ecuaciones lineales, todos com matriz del sistema igual a $(v_1 v_2 v_3)$:

$$(v_1 v_2 v_3 | u_1 u_2 u_3) = \left(\begin{array}{rrr|rrr} 0 & 1 & 1 & 1 & 0 & -1 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 1 & 1 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{\cdots} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{array}\right) = (I_3|[I]_{\mathcal{B}_1}^{\mathcal{B}_2}).$$

De donde se obtiene la matriz de cambio de base buscada:

$$[I]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Observe que $[(0,1,3)^t]_{\mathcal{B}_1} = (1,1,1)^t$ y que el producto:

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

produce el vector de coordenadas de $(0,1,3)^t$ en la base $\mathcal{B}_2$.

10.2.4 Composición de t.l. y producto matricial

Con la representación matricial de las transformaciones lineales se obtiene una correspondencia directa entre la composición de transformaciones lineales y la multiplicación matricial

Definición 10.5 (Composición de transf. lineales)

Dados los espacios vectoriales V, W y U y dos t.l. $S \in L(V,W)$ y $T \in L(W,U)$:

$$V \xrightarrow{S} W \xrightarrow{T} U.$$

Se define la función $T \circ S : V \longrightarrow U$ por

$$(T \circ S)(x) = T(S(x)) \quad \forall x \in V.$$

Se comprueba fácilmente que $T \circ S$ es lineal y además que la composición de transformaciones es asociativa, esto es:

$$\text{si } V_1 \xrightarrow{T_1} V_2 \xrightarrow{T_2} V_3 \xrightarrow{T_3} V_4 \quad \text{entonces}$$

$$(T_3 \circ T_2) \circ T_1 = T_3 \circ (T_2 \circ T_1).$$

La composición de transformaciones corresponde con la multiplicación matricial en los términos que establece el siguiente teorema:

Teorema 10.6 *Considere las t.l. S y T como en el esquema:*

$$\begin{array}{ccccc} V & \xrightarrow{S} & W & \xrightarrow{T} & U \\ \mathcal{B}_1 & \longrightarrow & \mathcal{B}_2 & \longrightarrow & \mathcal{B}_3 \end{array}$$

Donde $\mathcal{B}_1 = \{v_1, \cdots, v_n\}$ *es una base de* V,
$\mathcal{B}_2 = \{w_1, \cdots, w_m\}$ *es una base de* W *y*
$\mathcal{B}_3 = \{u_1, \cdots, u_p\}$ *es una base de* U.

$$\textit{Entonces} \quad [T \circ S]_{\mathcal{B}_1}^{\mathcal{B}_3} = [T]_{\mathcal{B}_2}^{\mathcal{B}_3}[S]_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Demostración: Sea $[T]_{\mathcal{B}_2}^{\mathcal{B}_3} = (a_{ij})_{p\times m}$ y $[S]_{\mathcal{B}_1}^{\mathcal{B}_2} = (b_{ij})_{m\times n}$. Probaremos que la i-ésima columna de $[T \circ S]_{\mathcal{B}_1}^{\mathcal{B}_3}$ es igual a $[T]_{\mathcal{B}_2}^{\mathcal{B}_3} b_i$. donde b_i es la columna i-ésima de $[S]_{\mathcal{B}_1}^{\mathcal{B}_2}$.

$$\begin{array}{rcl} (T \circ S)(v_i) = T(S(v_i)) & = & T(\sum_{s=1}^m b_{si} w_s) \\ & = & \sum_{si}^m b_{si} T(w_s) \\ & = & \sum_{s=1}^m b_{si} \left(\sum_{j=1}^p a_{js} u_j\right) \\ & = & \sum_{j=1}^p \left(\sum_{s=1}^m a_{js} b_{si}\right) u_j. \end{array}$$

Por lo tanto, la columna i-ésima de

$[T \circ S]_{\mathcal{B}_1}^{\mathcal{B}_2}$ es

$$\left(\sum_{s=1}^m a_{is} b_{si}, \cdots, \sum_{s=1}^m a_{ps} b_{si}\right)^t$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} \begin{pmatrix} b_{1i} \\ \vdots \\ b_{pi} \end{pmatrix}$$

$$= \quad [T]_{\mathcal{B}_2}^{\mathcal{B}_3} b_i.$$

■■■

Si en particular, en el teorema anterior todos los espacios vectoriales son el mismo: $V = W = U$, ambas transformaciones son la identidad: $T = S = I$ y $\mathcal{B}_1 = \mathcal{B}_3$, obtenemos:

$$[I]_{\mathcal{B}_1} = [I \circ I]_{\mathcal{B}_1}^{\mathcal{B}_1} = [I]_{\mathcal{B}_2}^{\mathcal{B}_1}[I]_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Por otra parte $[I]_{\mathcal{B}_1} = I_n$, puesto que si $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_n\}$, entonces $[v_i]_{\mathcal{B}_1} = e_i$ es el i-ésimo vector canónico, luego

$$[I]_{\mathcal{B}_1} = ([v_1]_{\mathcal{B}_1}[v_2]_{\mathcal{B}_1} \ldots [v_n]_{\mathcal{B}_1}) = (e_1 e_2 \ldots e_n) = I_n.$$

Por tanto

$$I_n = [I]_{\mathcal{B}_2}^{\mathcal{B}_1}[I]_{\mathcal{B}_1}^{\mathcal{B}_2} \Longrightarrow ([I]_{\mathcal{B}_1}^{\mathcal{B}_2})^{-1} = [I]_{\mathcal{B}_2}^{\mathcal{B}_1}.$$

Ejemplo 10.8 Utilizando la matriz del ejemplo 10.5.

$$[T]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{pmatrix} -1 & 27 & -3 \\ 2 & -37 & 6 \end{pmatrix}.$$

Calcule $[T]_{\mathcal{B}_1}^{\mathcal{E}}$, donde $\mathcal{E} = \{(-1, 1)^t, (4, -5)^t\}$.

Solución: Como $I \circ T = T$ utilizando el teorema 10.6:

$$[T]_{\mathcal{B}_1}^{\mathcal{E}} = [I \circ T]_{\mathcal{B}_1}^{\mathcal{E}} = [I]_{\mathcal{B}_2}^{\mathcal{E}}\ [T]_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Luego es suficiente calcular $[I]_{\mathcal{B}_2}^{\mathcal{E}}$:

$I((4, 1)) = (4, 1) = -24(-1, 1) - 5(4, -5)$,

$I((3, 1)) = (3, 1) = -19(-1, 1) - 4(4, -5)$, entonces

$$[I]_{\mathcal{B}_2}^{\mathcal{E}} = \begin{pmatrix} -24 & -19 \\ -5 & -4 \end{pmatrix}.$$

Finalmente

$$\begin{aligned} [T]_{\mathcal{B}_1}^{\mathcal{E}} &= \begin{pmatrix} -24 & -19 \\ -5 & -4 \end{pmatrix}\begin{pmatrix} -1 & 27 & -3 \\ 2 & -37 & 6 \end{pmatrix} \\ &= \begin{pmatrix} -14 & 55 & -42 \\ -3 & 13 & -9 \end{pmatrix}. \end{aligned}$$

10.2.5 Matrices asociadas a una misma transformación

Supongamos que $T : V \longrightarrow W$ es una t.l. y que $\mathcal{B}_1$ y $\mathcal{B}_3$ son bases del espacio vectorial V, en tanto que $\mathcal{B}_2$ y $\mathcal{B}_4$ son bases de W. El teorema 10.6, permite también relacionar las matrices, $[T]_{\mathcal{B}_3}^{\mathcal{B}_4}$ y $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$ asociadas a una misma transformación lineal T, en bases distintas.

Sean I_v e I_w las transformaciones identidad de V y W respectivamente y consideremos la siguiente composición de transformaciones:

$$V \xrightarrow{I_v} V \xrightarrow{T} W \xrightarrow{I_w} W$$

Podemos observar que $T = I_w \circ T \circ I_v$, de manera que:

$$[T]_{\mathcal{B}_3}^{\mathcal{B}_4} = [I_w]_{\mathcal{B}_2}^{\mathcal{B}_4}[T]_{\mathcal{B}_1}^{\mathcal{B}_2}[I_v]_{\mathcal{B}_3}^{\mathcal{B}_1}.$$

Ejemplo 10.9 Sea L la recta de $I\!R^3$ engendrada por el vector $v = (1,1,1)$ y que contiene el origen. Queremos determinar la transformación $T : I\!R^3 \longrightarrow I\!R^3$, que a todo vector (x,y,z) lo rota un ángulo θ sobre el eje que determina la recta L.

Como en otras situaciones, para determinar la transformación T es suficiente con definir la imágenes por T de los vectores de una base de $I\!R^3$. Pero en este caso, debemos elegir una base apropiada para facilitar el proceso: consideremos el vector $v_3 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^t$ sobre el eje L y completemos con otros dos vectores hasta obtener una base ortonormal, por ejemplo:

$$v_1 = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)^t \quad \text{y} \quad v_2 = (\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}})^t.$$

Sea $\mathcal{B}$ la base obtenida: $\mathcal{B} = \{v_1, v_2, v_3\}$. Como v_1 y v_2 son ortogonales al eje de rotación, los vectores del plano que engendran rotan como en el caso de la rotación de ángulo θ en $I\!R^2$ (ejemplo 10.3):

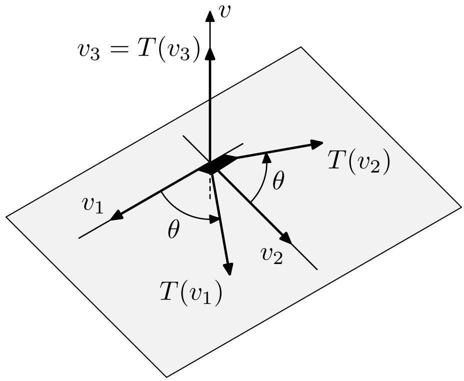


Figura 10.2: Rotación sobre el eje L de ángulo θ.

Como en $\mathbb{R}^2$, de las relaciones de trigonométricas seno y coseno se tiene que:

$$T(v_1) = \cos\theta v_1 + \operatorname{sen}\theta v_2$$

$$\text{y} \quad T(v_2) = -\operatorname{sen}\theta v_1 + \cos\theta v_2.$$

Además, como v_3 está en el eje de rotación no se modifica al aplicarle T:

$$T(v_3) = v_3.$$

Luego

$$[T(v_1)]_{\mathcal{B}} = \begin{pmatrix} \cos\theta \\ \operatorname{sen}\theta \\ 0 \end{pmatrix}, \quad [T(v_2)]_{\mathcal{B}} = \begin{pmatrix} -\operatorname{sen}\theta \\ \cos\theta \\ 0 \end{pmatrix}, \quad [T(v_3)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Con lo que disponemos de la matriz de T en la base $\mathcal{B}$:

$$[T]_{\mathcal{B}} = ([T(v_1)]_{\mathcal{B}}[T(v_2)]_{\mathcal{B}}[T(v_3)]_{\mathcal{B}}) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Por otra parte, si $\mathcal{C} = \{e_1, e_2, e_3\}$ es la base canónica de $\mathbb{R}^3$, las matrices de cambio de base $[I]_{\mathcal{B}}^{\mathcal{C}}$ y $[I]_{\mathcal{C}}^{\mathcal{B}}$ se obtienen fácilmente:

$$[I]_{\mathcal{B}}^{\mathcal{C}} = ([v_1]_{\mathcal{C}}\ [v_2]_{\mathcal{C}}\ [v_3]_{\mathcal{C}}) = (v_1, v_2, v_3) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$[I]_{\mathcal{C}}^{\mathcal{B}} = ([I]_{\mathcal{B}}^{\mathcal{C}})^{-1} = ([I]_{\mathcal{B}}^{\mathcal{C}})^t = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Finalmente podemos obtener la matriz de T en la base canónica:

$$[T]_{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}\ [T]_{\mathcal{B}}\ [I]_{\mathcal{C}}^{\mathcal{B}} =$$

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Si denotamos $\alpha = 1 + 2\cos\theta$, $\beta = 1 - \cos\theta - \sqrt{3}\operatorname{sen}\theta$ y $\gamma = 1 - \cos\theta + \sqrt{3}\operatorname{sen}\theta$, el producto de estas matrices es:

$$[T]_{\mathcal{C}} = \frac{1}{3}\begin{pmatrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{pmatrix}.$$

De manera que

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{3}\begin{pmatrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

10.3 Núcleo e Imagen

Esta sección comprende el estudio de dos subespacios asociados a toda transformación lineal: el núcleo y la imagen. En particular nos interesa destacar los métodos de cálculo, algunos resultados sobre dimensión y la relación del núcleo y la imagen con la inyectividad y sobreyectividad.

10.3.1 Definición de núcleo e imagen

Definición 10.7 (Núcleo e imagen)

Sea $T \in L(V, W)$.

a) El conjunto $Nuc(T) = \{v \in V | T(v) = 0_w\}$ se llama núcleo de T.

b) El conjunto $Img(T) = \{T(x) | x \in V\}$ se llama imagen de V bajo T, o simplemente imagen de T.

Teorema 10.8 *Sea $T \in L(V, W)$. El $Nuc(T)$ y la $Img(T)$ son subespacios de V y W respectivamente*

Demostración: $\text{Nuc}(T) \neq \emptyset$:

$$T(0_v) = T(0_v + 0_v) = T(0_v) + T(0_v) \Longrightarrow T(0_v) = 0_w,$$

luego $0_v \in \text{Nuc}(T)$.

Sean $v, u \in \text{Nuc}(T)$ y $\alpha \in I\!R$. Vamos a probar que $\alpha v + u \in \text{Nuc}(T)$.

$$T(\alpha v + u) = \alpha T(v) + T(u) = \alpha 0_w + 0_w = 0_w.$$

Por lo tanto $\alpha v + u \in \text{Nuc}(T)$. De manera igualmente fácil se deduce que $\text{Img}(T)$ es un subespacio de W.

■

Ejemplo 10.10 Sea $A = \begin{pmatrix} 2 & -2 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}$ y $T \in L(I\!R^2, I\!R^3)$ definida por $T(x, y) = A \begin{pmatrix} x \\ y \end{pmatrix}$. Calcular una base de $\text{Nuc}(T)$.

Solución

Por definición:

$$(x, y) \in \text{Nuc}(T) \Longleftrightarrow T(x, y) = (0, 0, 0).$$

Por lo tanto tienen lugar las siguientes equivalencias:

$$\begin{aligned}(x,y) \in \text{ Nuc}\,(T) &\iff (2x-2y, -x+y, x-y) = (0,0,0)\\ &\iff \left\{\begin{array}{lcl} 2x-2y &=& 0\\ -x+y &=& 0\\ x-y &=& 0\end{array}\right.\end{aligned}$$

Luego $x = y$ y por lo tanto: $\text{Nuc}\,(T) = \mathcal{C}\ell\{(1,1)\}$.

En este ejemplo hemos visto que encontrar una base del núcleo es lo mismo que encontrar una base del espacio solución de un sistema homogéneo. El teorema 10.3, permite generalizar este resultado a una t.l. definida entre cualesquiera espacios de dimensión finita.

En efecto, si A es la matriz de $T \in L(V,W)$ en un par de bases $\mathcal{B}$ y $\mathcal{C}$ de V y W respectivamente, entonces la solución del sistema $Ax = 0$, donde $x = [v]_{\mathcal{B}}$ caracteriza las coordenadas de los vectores v del núcleo de T.

Ejemplo 10.11 Sea $T \in L(I\!R^4, I\!R^3)$ tal que:

$$[T]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & 2 & 1 & -1\\ 0 & -1 & 0 & 1\\ 1 & 3 & 2 & 2\end{pmatrix} = A.$$

1. Obtenga una base del $\text{Nuc}\,(T)$.

2. Obtenga la expresión general de $T(z)$ en términos de las entradas de $z \in I\!R^4$.

   Las bases $\mathcal{C}$ y $\mathcal{B}$ se definen así: $\mathcal{C}$ es la base canónica de $I\!R^3$ y

   $$\mathcal{B} = \{(1,1,0,0), (1,0,1,0), (1,0,0,1), (0,0,0,1)\}$$

Solución

(1) Sea $x = (a,b,c,d)$, entonces

$$[T(x)]_{\mathcal{C}} = A[x]_{\mathcal{B}} = 0 \iff x \in \text{ Nuc}\,(T).$$

Luego hay que resolver el sistema

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & -1 & 0 & 1 \\ 1 & 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} r \\ s \\ t \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

donde $[x]_{\mathcal{B}} = (r,\ s,\ t,\ u)^t$.

Realizando operaciones elementales fila sobre A se obtiene el sistema equivalente:

$$\begin{pmatrix} 1 & 0 & 0 & -3 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} r \\ s \\ t \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cuya solución es $r = 3u$, $s = u$, $t = -4u$. Por tanto

$$\begin{array}{l} x \in \operatorname{Nuc}(T) \iff \\ x = 3u(1,1,0,0) + u(1,0,1,0) - 4u(1,0,0,1) + u(0,0,0,1) \end{array}$$

Luego x es igual a:

$$x = u(0,3,1,-3).$$

Por tanto una base del Nuc (T) está formada por el único elemento $(0,3,1,-3)$.

(2) Sea $z = (a,b,c,d)$, $[T(z)]_{\mathcal{C}} = A[z]_{\mathcal{B}}$.

Es suficiente calcular las coordenadas de z en $\mathcal{B}$, en términos de a,b,c,d:

$$\begin{array}{rcl} z & = & l(1,1,0,0) + h(1,0,1,0) + k(1,0,0,1) + r(0,0,0,1) \\ & = & (l+h+k, l, h, k+r). \end{array}$$

Entonces

$$\left\{ \begin{array}{rcl} h & = & c \\ l & = & b \\ k & = & a-l-h \\ r & = & d-k \end{array} \right.$$

Por lo tanto $k = a - b - c$ y $r = d - a + b + c$.

$$\begin{aligned}[T(z)]_{\mathcal{C}} &= \begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & -1 & 0 & 1 \\ 1 & 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} b \\ c \\ a-b-c \\ d-a+b+c \end{pmatrix}. \\ &= \begin{pmatrix} 2a-b-d \\ d+b-a \\ b+3c+2d \end{pmatrix}\end{aligned}$$

Luego

$$T(z) = (2a - b - d, d + b - a, b + 3c - a + 2d).$$

10.3.2 Inyectividad y sobreyectividad

La secuencia de teoremas que siguen (10.10,...,10.13) conforman una sola unidad en cuanto que establecen las relaciones básicas entre las propiedades de inyectividad y sobreyectividad con el núcleo y las dimensiones de los espacios involucrados. Como veremos, estos resultados nos permitirán comprender mejor la naturaleza de las transformaciones lineales.

Definición 10.9 (Inyectividad y Sobreyectividad)

Sea $f : A \longrightarrow B$, una función del conjunto A al conjunto B.

a) Se dice que f es inyectiva si todo elemento $z \in B$ tiene a lo más una preimagen $x \in A$, $z = f(x)$, o equivalentemente si:

$$f(x) = f(y) \Longrightarrow x = y$$

b) f es sobreyectiva si todo elemento $z \in B$ es imagen de algún $x \in A$, o sea, $Img(f) = B$. Equivalentemente, f es sobreyectiva si la ecuación en la variable x:

$$f(x) = z \quad \textit{tiene solución} \ \forall\ z \in B$$

c) Cuando f es inyectiva y sobreyectiva se dice que f es biyectiva.

Ejemplo 10.12 Sea $T \in L(\mathbb{R}^3, \mathbb{R}^3)$ definida por $T(x, y, z) = (x - y, -x + y + z, -y + z)$. ¿Es T sobreyectiva?

Solución: T es sobreyectiva si $T(x, y, z) = (a, b, c)$ tiene solución $\forall\ (a, b, c) \in \mathbb{R}^3$.

$$
\begin{array}{ll}
 & T(x,y,z) = (a,b,c) \\
\iff & (x-y, -x+y+z, -y+z) = (a,b,c) \\
\iff & \left\{ \begin{array}{lcl} x - y & = & a \\ -x + y + z & = & b \\ -y + z & = & c \end{array} \right.
\end{array}
$$

Resolviendo el sistema se tiene que $(x, y, z) = (2a + b - c, a + b - c, a + b)$. Luego para cada $(a, b, c) \in \mathbb{R}^3$ el sistema tiene solución por lo tanto es sobreyectiva.

Verificar la inyectividad de una transformación lineal a partir de la definición, aunque no es en extremo difícil, es más difícil que si se emplea el teorema siguiente.

Teorema 10.10 *Sea $T \in L(V, W)$. Se tiene que:*

$$T \text{ es inyectiva} \iff Nuc\,(T) = \{0\}.$$

Demostración: "$\Rightarrow$": Supongamos que T es inyectiva. Sea $x \in \mathrm{Nuc}\,(T)$ entonces $T(x) = 0 = T(0)$. Luego $x = 0$ (por la inyectividad). De donde $\mathrm{Nuc}\,(T) = \{0\}$.

"$\Leftarrow$": Supongamos que $\mathrm{Nuc}\,(T) = \{0\}$ y que $T(x) = T(y)$. Hay que probar que $x = y$. Como T es lineal:

$$
\begin{array}{ll}
 & T(x) = T(y) \\
\Longrightarrow & T(x) - T(y) = 0 \\
\Longrightarrow & T(x - y) = 0 \\
\Longrightarrow & x - y \in \mathrm{Nuc}\,(T) \\
\Longrightarrow & x - y = 0 \\
\Longrightarrow & x = y.
\end{array}
$$

■■■

Teorema 10.11 *Sea $T \in L(V, W)$*

(a) Si $V = \mathcal{C}\ell\{v_1, \cdots, v_m\}$ entonces

$$\text{Img}\,(T) = \mathcal{C}\ell\{T(v_1), \cdots, T(v_m)\}$$

(b) Si T es inyectiva y $v_1, \cdots, v_m$ son l.i. entonces $T(v_1), \cdots, T(v_m)$ son l.i.

$$\textit{En particular:} \quad \dim V = \dim(\,\text{Img}\,(T)).$$

Demostración:

(a) La prueba se propone como ejercicio.

(b) Veamos que: $T(v_1), \cdots, T(v_m)$ son l.i. :
Sea $d_1, \cdots, d_m$ escalares en $I\!R$ tales que:

$$\sum_{i=1}^{m} d_i T(v_i) = 0$$

entonces $\sum_{i=1}^{m} d_i T(v_i) = T\left(\sum_{i=1}^{m} d_i v_i\right) = 0$,
de donde $\sum_{i=1}^{m} d_1 v_1 \in \text{Nuc}\,(T)$. Pero como T es inyectiva entonces $\text{Nuc}\,(T) = \{0\}$. Por tanto $\displaystyle\sum_{i=1}^{m} d_i v_i = 0$ y se deduce que $d_1 = d_2 = \cdots = d_m = 0$, porque $\{v_1, \ldots, v_m\}$ es l.i.

■■■

Ejemplo 10.13 Sea la transformación lineal $T : I\!R^3 \longrightarrow I\!R^4$ definida por:

$$T(x, y, z) = (x + z, x + y + 2z, y + z, x + z).$$

Obtenga una base de $\text{Img}\,(T)$ y del $\text{núcleo}(T)$.

Solución

Cálculo de una base de $\text{Img}\,(T)$: consideramos la base canónica de $I\!R^3$, $\{e_1, e_2, e_3\}$. Por el teorema 10.11 se concluye que $T(e_1)$,

$T(e_2)$, $T(e_3)$ generan a $\mathrm{Img}\,(T)$.

$$\mathrm{Img}\,(T) = \mathcal{C}\ell\{T(e_1), T(e_2), T(e_3)\}.$$

Pero por otra parte:

$$T(e_1) = (1,1,0,1) \qquad T(e_2) = (0,1,1,0) \qquad T(e_3) = (1,2,1,1).$$

Y se observa que $T(e_3) = T(e_1) + T(e_2)$. Luego una base para $\mathrm{Img}\,(T)$ es:

$$\mathcal{B}_1 = \{(1,1,0,1), (0.1,1,0)\}.$$

Ahora calculemos una base para $\mathrm{Nuc}\,(T)$:

$$\begin{aligned} & (x,y,z) \in \mathrm{Nuc}\,(T) \\ \Longleftrightarrow \quad & T(x,y,z) = (0,0,0,0) \\ \Longleftrightarrow \quad & (x+z, x+y+2z, y+z, x+z) = (0,0,0,0). \end{aligned}$$

Y resolviendo se tiene que $x+z=0$, $y+z=0$, luego

$$\mathrm{Nuc}\,(T) = \{(-z,-z,z) | z \in I\!R\} = \mathcal{C}\ell\{(1,1,-1)\}$$

de donde $\mathcal{B}_2 = \{(1,1,-1)\}$ es una base para $\mathrm{Nuc}\,(T)$.

Teorema 10.12 *Sea $T \in L(V,W)$ y $\dim V = \dim W = n$. Entonces*

$$T \textit{ es inyectiva} \quad \Longleftrightarrow \quad T \textit{ es sobreyectiva.}$$

Demostración: "$\Rightarrow$": Sea $\{v_1, \cdots, v_n\}$ una base de V, por teorema 10.11(b) se tiene que $\{T(v_1), \cdots, T(v_n)\}$ es una base de $\mathrm{Img}\,(T)$. Luego $\dim \mathrm{Img}\,(T) = \dim W$ y consecuentemente $\mathrm{Img}\,(T) = W$.

"$\Leftarrow$": sea $\{z_1, \cdots, z_n\}$ una base de W. Como $\mathrm{Img}\,(T) = W$, existe

$$v_i \in V \text{ tal que } T(v_i) = z_i, \ \ i = 1, 2, \cdots, n.$$

Vamos a probar que $\{v_1, \cdots, v_n\}$ es una base de V y luego que $\mathrm{Nuc}\,(T) = \{0\}$ con lo cual la prueba sería completa. Sean $d_1, \cdots, d_n$ elementos de $I\!R$ tales que $\displaystyle\sum_{i=1}^{n} d_i v_i = 0$. Por tanto

$$\sum_{i=1}^{n} d_i T(v_i) = \sum_{i=1}^{n} d_i z_i = 0$$

luego $d_1 = \cdots = d_n = 0$. Es decir $v_1, \cdots, v_n$ son l.i. y constituyen una base de V, puesto que $\dim V = n$.

Sea ahora $x \in \operatorname{Nuc}(T)$, existen $\alpha_1, \cdots, \alpha_n \in I\!R$ tales que

$$x = \sum_{i=1}^{n} \alpha_i v_i.$$

Entonces $0 = T(x) = \displaystyle\sum_{i=1}^{n} \alpha_i T(v_i) = \sum_{i=1}^{n} \alpha_i z_i$. De donde

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

Por lo tanto $x = 0$, $\operatorname{Nuc}(T) = \{0\}$ y T es inyectiva.

∎

Observación

El resultado anterior no es válido si $\dim V \neq \dim W$.

Teorema 10.13 *Sea $T \in L(V, W)$ y V de dimensión finita, entonces*

$$\dim V = \dim(\, Nuc\,(T)) + \dim(\, Img\,(T))$$

Demostración: Para los casos especiales $T = 0$ y T inyectiva la validez del resultado es evidente.

Supongamos que $T \neq 0$ y T no es inyectiva. Sea $\{v_1, \cdots, v_k\}$ una base del $\operatorname{Nuc}(T)$, es decir $\dim(\operatorname{Nuc}(T)) = k$. Por el teorema de completación de la base existen $u_1, \cdots, u_r \in V$ tales que $\{v_1, \cdots, v_k, u_1, \cdots, u_r\}$ es una base de V.

Vamos a probar que $\{T(u_1), T(u_2), \cdots, T(u_r)\}$ es una base de $\operatorname{Img}(T)$, con lo cual se tendría que $\dim \operatorname{Img}(T) = r$ y por lo tanto que $\dim V = r + k = \dim(\operatorname{Img}(T)) + \dim(\operatorname{Nuc}(T))$.

- $T(u_1), \cdots, T(u_r)$ son l.i:

sean $d_i \in K$, tales que $\sum_{i=1}^{r} d_i T(u_i) = 0$.

Luego $T(\sum_{i=1}^{r} d_i u_i) = 0$. De donde $\sum_{i=1}^{r} d_i u_i \in \operatorname{Nuc}(T)$.

Como $\{v_1, \cdots, v_k\}$ es una base del $\operatorname{Nuc}(T)$, entonces existen $\alpha_i \in \mathbb{R}$ tales que

$$\sum_{i=1}^{r} d_i u_i = \sum_{j=1}^{k} \alpha_j v_j.$$

Como $v_1, \cdots, v_k, u_1, \cdots, u_r$ son l.i., se concluye que

$$d_1 = \cdots, = d_r = \alpha_1 = \cdots, = \alpha_k = 0.$$

De donde $T(u_1), \cdots, T(u_r)$ son l.i.

- $T(u_1), \cdots, T(u_r)$ generan a $\operatorname{Img}(T)$:

  sea $y \in \operatorname{Img}(T)$ entonces existe $x \in V$ tal que $T(x) = y$.

  Sea $x = \sum_{i=1}^{k} d_i v_i + \sum_{j=1}^{r} \beta_j u_j$; luego

$$T(x) = \sum_{i=1}^{k} d_i T(v_i) + \sum_{j=1}^{r} \beta_j T(u_j) = \sum_{j=1}^{r} \beta_j T(u_j).$$

  Por lo tanto $T(u_1), \cdots, T(u_r)$ generan a $\operatorname{Img}(T)$.

■

Ejemplo 10.14 En el ejemplo 10.13 en la página 312, observe que $T : R^3 \longrightarrow \mathbb{R}^4$, $\dim(\operatorname{Nuc}(T)) = 1$ y $\dim(Img(T)) = 3$, así se verifica que:

$$\dim(\mathbb{R}^3) = \dim(\operatorname{Nuc}(T)) + \dim(Img(T))$$

esto es $3 = 1 + 2$.

10.3.3 Transformaciones invertibles

Definición 10.14 (Inversa de una t. lineal)

Sea $T \in L(V, W)$. T se llama invertible si existe una transformación lineal $T^{-1} : W \longrightarrow V$ tal que $T \circ T^{-1} = I_w$ y $T^{-1} \circ T = I_v$, donde I_v e I_w son las identidades de V y W respectivamente.

Teorema 10.15 *(a) Sea $T \in L(V, W)$.*

$$T \text{ es invertible } \Longleftrightarrow T \text{ es biyectiva}$$

(b) Sea $T \in L(V, W)$, $\dim V = n = \dim W$, B_1 y B_2 bases de V y W respectivamente. Entonces T es invertible si y sólo si $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$ lo es. Y si este es el caso, entonces

$$([T]_{\mathcal{B}_1}^{\mathcal{B}_2})^{-1} = [T^{-1}]_{\mathcal{B}_2}^{\mathcal{B}_1}.$$

Demostración:

(a) Ejercicio.

(b) "$\Rightarrow$": Sea T invertible, existe $T^{-1} \in L(W, V)$ tal que

$$T \circ T^{-1} = I_w \text{ y } T^{-1} \circ T = I_v.$$

Por el teorema 10.6 en la página 302 se tiene:

$$[I_w]_{\mathcal{B}_2}^{\mathcal{B}_2} = [T \circ T^{-1}]_{\mathcal{B}_2}^{\mathcal{B}_2} = [T]_{\mathcal{B}_1}^{\mathcal{B}_2}[T^{-1}]_{\mathcal{B}_2}^{\mathcal{B}_1}.$$

Por otra parte, es claro que

$$[I_w]_{\mathcal{B}_2}^{\mathcal{B}_2} = I_n.$$

De manera similar se comprueba que:

$$I_n = [T^{-1}]_{\mathcal{B}_2}^{\mathcal{B}_1}[T]_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Se concluye que $A^{-1} = [T^{-1}]_{\mathcal{B}_2}^{\mathcal{B}_1}$.

"$\Leftarrow$": Si A es invertible entonces el sistema $Ax = 0$ tiene solución $S = \{0\}$. Por el teorema 10.3 se tiene:

$$T(v) = 0 \Longleftrightarrow [T(v)]_{\mathcal{B}_2} = A[v]_{\mathcal{B}_1} = 0.$$

Por la tanto: $\text{Nuc}\,(T) = \{0\}$. Luego por el teorema 10.12 T es biyectiva y por la parte (a) sigue que T es invertible.

■■■

Ejemplo 10.15 Sea $T \in L(I\!R^3)$ definida por

$$[T]_{\mathcal{C}} = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 1 & 1 \end{pmatrix} = A$$

donde $\mathcal{C}$ es la base canónica de $I\!R^3$.

Compruebe que T es invertible, calcule $T^{-1}(x, y, z)$ y la matriz de T^{-1} en la base $\mathcal{C}$.

Solución: Para calcular la matriz inversa se aplican las operaciones elementales por filas, a la matriz A.

$$\left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ -1 & 3 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{f_1 + f_2} \left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 5 & 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{-5f_3 + f_2} \left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 1 & 1 & -5 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[(1/3)f_3]{f_2, f_3} \left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} \end{array}\right)$$

$$\xrightarrow{-1f_3 + f_2} \left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} \end{array}\right)$$

$$\xrightarrow{-2f_2 + f_1} \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{1}{3} & -\frac{2}{3} & \frac{4}{3} \\ 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} \end{array}\right)$$

Luego $[T^{-1}]_{\mathcal{C}} = A^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{4}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} \end{pmatrix}$.

Por otra parte, dado que

$$[T^{-1}(x)]_{\mathcal{C}} = A^{-1}[x]_{\mathcal{C}} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{4}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

se obtiene el resultado siguiente:

$$[T^{-1}(x_1, x_2, x_3)]_{\mathcal{C}} = T^{-1}(x_1, x_2, x_3)$$

$$= \frac{1}{3}(x_1 - 2x_2 + 4x_3, x_1 + x_2 - 2x_3, -x_1 - x_2 + 5x_3).$$

∎

Observación: otra forma de resolver el problema anterior es:

$$\begin{aligned} [T(x)]_{\mathcal{C}} &= \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 1 & 1 \end{pmatrix} [x]_{\mathcal{C}} . \\ T(x, y, z) &= (x_1 + 2x_2, -x_1 + 3x_2 + 2x_3, x_2 + x_3)^t \end{aligned}$$

Luego se calcula directamente de $T(x, y, z)$, la expresión general de $T^{-1}(x, y, z)$.

10.4 Ejercicios

1. Sea $T : I\!R^3 \to I\!R^2$ una transformación lineal tal que:

$$T(1,2,0) = (1,1), \;\; T(-1,1,2) = (1,1), \;\; T(0,3,3) = (-1,0)$$

Para $(x,y,z) \in I\!R^3$ determine $T(x,y,z)$.

2. Defina una transformación lineal $T : I\!R^2 \to I\!R^2$, que asocie los puntos del cuadrado de vértices $(0,0)$, $(1,0)$, $(1,1)$ $(0,1)$ (puntos de los lados e interiores) con puntos del paralelogramo de vértices $(0,0)$, $(3,2)$, $(4,6)$ $(1,4)$, (puntos de los lados e interiores).

3. Determine una transformación lineal $T : I\!R^3 \to I\!R^3$, que asigne el cero como imagen de todos los puntos de la recta $L : \; (t, 2t, 0)$ y tal que el conjunto de imágenes de los puntos del plano $\Pi_1 : \; 2x + y - z = 0$ sea el plano $\Pi_2 : \; y + z = 0$, (en algunos casos se escribirá: $T(L) = 0$ y $T(\Pi_1) = \Pi_2$).

4. Determine una transformación lineal $T : I\!R^3 \longrightarrow I\!R^3$ tal que $T(P_1) = P_2$, donde

$$P_1 = \{(x,y,z) \in I\!R^3 | (x,y,z) = (1+t, 1+s, -2t+s), \; t,s \in I\!R\}$$

$$\text{y} \quad P_2 = \{(x,y,z) \in I\!R^3 | x - 2y - z = 0\}$$

Es decir, el conjunto de imágenes de los vectores del plano P_1 es P_2.

5. Considere la transformación $T : I\!R^3 \longrightarrow I\!R^3$ correspondiente a una reflexión sobre el plano $\pi : x+z = 0$. Observe que T deja invariantes a los vectores del plano π e invierte la dirección de aquellos que son ortogonales a π.

i) Determine $T(x,y,z)$.

ii) Calcule bases para $\operatorname{Nuc}(T)$ e $\operatorname{Img}(T)$.

6. Construya una t.l. $T : I\!R^4 \to I\!R^3$ tal que

$\operatorname{Nuc}(T) = \mathcal{C}\ell\{(1,1,1,0), (2,-1,1,1)\}$ y

$\operatorname{Img}(T) = \mathcal{C}\ell\{(1,1,1), (0,0,1)\}$

7. Considere la transformación lineal identidad I en $I\!R^3$ y $\mathcal{B} = \{u_1, u_2, u_3\}$ una base de $I\!R^3$.

1. Calcule $[I]_{\mathcal{B}}$ y $[I]_{\mathcal{B}}^{\mathcal{C}}$ con $\mathcal{C} = \{e_1, e_2, e_3\}$ la base canónica de $I\!R^3$.
2. Si $u_1 + u_2 - u_3 = e_1, \quad u_1 + u_2 = e_2, \quad au_1 + bu_3 = e_3$. Encuentre $[I]_{\mathcal{C}}^{\mathcal{B}}$.

8. Sea $T : I\!R^3 \to I\!R^2$ lineal tal que:

$$[T]_{\mathcal{D}}^{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

a) Si $[v]_{\mathcal{D}} = (2, 1, 0)^t$. Encuentre $[T(v)]_{\mathcal{B}}$.

b) Si $\mathcal{B} = \{(1, -1), (0, 1)\}$. Calcule las coordenadas de $T(v)$ en la base canónica $\mathcal{C}$ de $I\!R^2$.

c) Encuentre $[T]_{\mathcal{D}}^{\mathcal{C}}$.

9. Sea $T : I\!R^3 \to I\!R^3$ la transformación lineal dada por:

$T(x, y, z) = (x, y + z, x + z)$. Considere las bases de $I\!R^3$:

$$\mathcal{B} = \{(1, 0, 1), (1, 1, 0), (1, 1, 1)\}$$

$$\mathcal{B}_1 = \{(1, 0, 0,), (0, 1, 1), (1, 1, 0)\}$$

1. Hallar la matriz de T respecto a la base $\mathcal{B}$ $([T]_{\mathcal{B}})$.
2. Hallar la matriz de transición de $\mathcal{B}_1$ a $\mathcal{B}$ $([I]_{\mathcal{B}_1}^{\mathcal{B}})$.
3. Usar 1. y 2. para encontrar la matriz de T respecto a $\mathcal{B}_1$ y $\mathcal{B}$ (es decir la matriz $[T]_{\mathcal{B}_1}^{\mathcal{B}}$).

10. Sea $\mathcal{B} = \{(1, 1, 0)^t, (0, 1, 1)^t, (1, 1, 1)^t\}$ una base y la transformación lineal $T : I\!R^3 \longrightarrow I\!R^3$, definida por:

$$T[(1, 1, 0)^t] = (0, 1, 1)^t, \qquad T[(0, 1, 1)^t] = (0, 0, 0)^t,$$

$$\text{y} \qquad T[(1, 1, 1)^t] = (2, 2, 2)^t.$$

a) Sin hacer cálculos, dé una base para $\mathrm{Nuc}\,(T)$ (Justifique).

b) Determine $[I]_{\mathcal{B}}^{\mathcal{C}}$ y $[I]_{\mathcal{C}}^{\mathcal{B}}$ donde $\mathcal{C}$ es la base canónica de $I\!R^3$ e I la transformación identidad de $I\!R^3$ a $I\!R^3$.

c) Calcule $[T]_{\mathcal{C}}$ y dé una fórmula general para $T[(x, y, z)^t]$.

11. Sea $T : I\!R^n \to I\!R^m$ lineal. Si $\mathcal{B} = \{u_1, \ldots, u_n\}$ es base de $I\!R^n$.

1. Demuestre que $\text{Img}\,(T) = \mathcal{C}\ell\{T(u_1), \ldots, T(u_n)\}$.
2. Si T es inyectiva entonces $\{T(u_1), \ldots, T(u_n)\}$ es l.i.
3. Si A es la matriz de T en las bases canónicas respectivas, demuestre que las columnas de A generan $\text{Img}\,(T)$.
4. Si

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

Encuentre bases para $\text{Img}\,(T)$, $\text{Nuc}\,(T)$.

12. Sea $\mathcal{B} = \{(1,1,0)^t, (0,-1,1)^t, (1,0,-1)^t\}$ una base y la transformación lineal $T : I\!R^3 \longrightarrow I\!R^3$, definida por:

$$T[(1,1,0)^t] = (2,0,1)^t, \qquad T[(0,-1,1)^t] = (0,-1,1)^t,$$

$$T[(1,0,-1)^t] = (4,-1,3)^t.$$

i) Determine $[T]_{\mathcal{B}}^{\mathcal{C}}$, donde $\mathcal{C}$ es la básica canónica de $I\!R^3$.
ii) Use i) para determinar la fórmula general de $T[(x,y,z)^t]$,
iii) Calcule una base para el núcleo de T. ¿T es sobreyectiva? Justifique.

13. Sean $\mathcal{B} = \{v_1, v_2\}$ y $\mathcal{D} = \{w_1, w_2\}$ bases de $I\!R^2$ tales que $w_1 = v_1 - v_2$ y $w_2 = 3v_1$, y considere las transformaciones lineales:

$$I : I\!R^2 \longrightarrow I\!R^2 \quad \text{la transformación identidad}$$
$$T : I\!R^2 \longrightarrow I\!R^2 \quad \text{tal que } [T]_{\mathcal{B}}^{\mathcal{D}} = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}.$$

a) Calcule $[I]_{\mathcal{B}}^{\mathcal{D}}$.

b) Encuentre $[T]_{\mathcal{B}}$.

c) Calcule $[T(2v_1 - v_2)]_{\mathcal{D}}$.

d) Si $w_1 = (1,-2)^t$ y $w_2 = (0,1)^t$, determine $T(2v_1 - v_2)$.

14. Sean V, W subespacios de $I\!R^n$ de igual dimensión y $T : V \to W$, lineal. Demuestre que:

T es inyectiva si y sólo si T es sobreyectiva.

15. Sea W subespacio de $I\!R^n$, $T : I\!R^n \to W$ tal que

$$T(v) = \text{Proy}_W v.$$

1. Demuestre que T es lineal.
2. Si $n = 3$ y $W = \{(x, y, z) | x + y = 0\}$, encuentre una base $\mathcal{B}$ de W y calcule $[T]_{\mathcal{C}}^{\mathcal{B}}$, donde $\mathcal{C}$ es la base canónica de $I\!R^3$.
3. Use 2) para determinar $T(x, y, z)$.
4. Calcule la distancia de $(-1, 2, 5)$ a W, utilizando la transformación T y el resultado obtenido en 3).

16. Sean $T : I\!R^3 \to I\!R^3$ definida por

$$[T]_{\mathcal{B}} = A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

a) Compruebe que T es invertible.

b) Calcule $\left[T^{-1}\right]_{\mathcal{B}}$ y $T^{-1}(x, y, z)$
donde $\mathcal{B} = \{(1, 0, 0,), (0, 1, 1), (1, 1, 0)\}$.

17. Sea $T : I\!R^3 \to I\!R^4$,

$$T(x, y, z) = (x + y, x - y + z, y - z, x + z)$$

a) Demuestre que T es lineal.

b) Obtenga $\mathcal{D}$ base de $\text{Img}(T)$.

c) Sea $\mathcal{B} = \{(1, 0, 0), (1, 1, 0), (1, 0, 1)\}$ base de $I\!R^3$.

Se define $S : I\!R^3 \to \text{Img}(T)$, $S(x) = T(x) \ \forall \ x \in I\!R^3$.

Calcule $[S]_{\mathcal{B}}^{\mathcal{D}}$ y S^{-1}

18. Sea $T : I\!R^3 \to I\!R^4$ definida por:

$$T(x, y, z) = (x - y, x - z, 0, x - y)$$

a) Encuentre bases para el $\text{Nuc}(T)$ e $\text{Img}(T)$.

b) Decida si T es inyectiva. ¿Es T sobreyectiva?

c) Si $\mathcal{C}$ es la base canónica de $I\!R^4$ encuentre $\mathcal{B}$ base de $I\!R^3$ tal que $[T]_{\mathcal{B}}^{\mathcal{C}}$ tenga la segunda columna nula.

19. Sea $W = \{(x,y,z) \in \mathbb{R}^3 | 4x = 3y\}$.

a) Justifique por qué W es un subespacio de $\mathbb{R}^3$.

b) Verifique que $\mathcal{B} = \{(3,4,0),(-6,-8,2)\}$ es una base de W.

c) Construya a partir de $\mathcal{B}$ una base ortonormal para W.

d) Dé una base ortonormal para $W^\perp$.

e) Considere la transformación lineal:

$$T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 \text{ definida por } T(v) = \text{Proy}_W v$$

i) Sin más cálculos, establezca: $\text{Img}\, T$, $\text{Nuc}\, T$.

ii) Indique el rango de T y la dimensión del núcleo de T.

iii) ¿Es T Inyectiva, sobreyectiva? (Justifique.)

20. Sean W un subespacio de $\mathbb{R}^3$ y $T : \mathbb{R}^3 \to W$ definida por:

$$T(x) = \text{Proy}_W x$$

1. Demuestre que T es una transformación lineal.
2. Si W es el plano de ecuación $2x - 3y + z = 0$ encuentre:
   (a) Una fórmula para $T(x,y,z)$.
   (b) El núcleo de T.
   (c) El conjunto $T(\mathbb{R}^3)$.
   (d) La distancia del punto $(2,-1,0)$ al subespacio W.

21. Sea A una matriz de 2×2, y $T_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, tal que $T_A(x) = Ax \;\; \forall x \in \mathbb{R}^2$

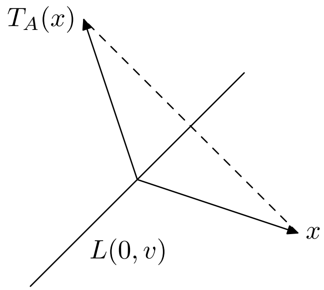
1. Demuestre que la matriz asociada a T_A en la base canónica de $\mathbb{R}^2$ es A.
2. Demuestre que si A es una matriz ortogonal ($A^t A = I$) entonces T_A preserva norma y ángulo, esto es que para cualquier $x, y \in \mathbb{R}^2$ se tiene:

$$||T_A(x)|| = ||x|| \;\; y \;\; \text{ang}(T_A(x), T_A(y)) = \text{ang}(x,y)$$

22. Sea $v = (v_1, v_2) \in \mathbb{R}^2$, $\|v\| = 1$ y L la recta que pasa por el origen en la dirección de v.

Considere la transformación lineal[1] T_A tal que $T_A(x)$ es el punto simétrico de x respecto a la recta L.

Nos referiremos a T_A como la reflexión respecto a la recta L.



1. Justifique que $\forall\ x \in I\!R^2$,

$$T_A(x) = 2\text{Proy}_v x - x$$

   y muestre que

   i) $T_A(e_1) = (2v_1^2 - 1, 2v_1v_2)$ y $T_A(e_2) = (2v_1v_2, 2v_2^2 - 1)$

   ii) $||T_A(e_1)|| = 1$.

2. Calcule A y $\det(A)$.
3. Muestre que T_A preserva norma.
4. Muestre que si $e_1 + T_A(e_1)$ es no nulo es un vector paralelo a v. Similarmente, si $e_2 + T_A(e_2) \neq 0$ entonces es paralelo a v.

23. En cada caso, encuentre A tal que T_A sea una reflexión respecto:

i) al eje x,

ii) a la recta $y = x$,

(Ver pie de página en página 324). Verifique en ambos casos que A es una matriz ortogonal.

24. Encuentre A para que T_A sea una rotación en 30 grados. (Vea pie de página en página 324 y ejemplo 10.3). Encuentre explícitamente $T_A(x, y)$

[1] En este y los problemas que siguen $A \in M(2, I\!R)$ y T_A denota la t.l. definida por $T_A : I\!R^2 \longrightarrow I\!R^2$, tal que $T_A(x) = Ax\ \ \forall x \in I\!R^2$

25. Considere las transformaciones lineales T_R y T_S definidas por:

$$R = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ y } S = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$$

a) Que tipo de movimientos realizan T_R y T_S.

b) Encuentre C tal que T_C realiza un movimiento equivalente a realizar T_S y luego T_R.

c) Muestre que T_C preserva norma.

26. Sea la transformación lineal $T_A(x, y) = (x + ky, y)$ (llamada deslizamiento de factor k en la dirección x)

1. Haga un gráfico que muestre el efecto de T_A en la base canónica.
2. Encuentre A
3. Si T_B es la t.l. que rota en un ángulo θ. Encuentre la t.l. T_C que rota y luego desliza.

27. Sea T_A tal que su imagen sobre la recta L_i generada por e_i que pasa por el origen, es la recta $T_A(L_i)$ rotada 30 grados con respecto a L_i en el sentido del movimiento de las manecillas del reloj, $i = 1, 2$.

1. Encuentre la matriz A. Note que hay varias respuestas que dependen de dos parámetros.
2. Muestre que las únicas transformaciones $T \in L(I\!R^2)$ que transforman rectas L que pasan por cero en rectas $T(L)$, rotadas 30 grados respecto a L, en el sentido del movimiento de las manecillas del reloj, son las determinadas en 1.

28. En cada caso, determine la transformación $T : I\!R^3 \longrightarrow I\!R^3$, es decir, determine la matriz A tal que $T(x) = Ax$:

1. T es una rotación derecha[2] de ángulo θ sobre el eje y.
2. T es una reflexión sobre el plano $y = 0$.
3. T es una rotación derecha de ángulo θ sobre el eje generado por $v = (1, 0, 1)$.

[2]Una rotación sobre el eje determinado por el vector v se dice que es derecha si corresponde al giro normal de atornillar e izquierda al de desatornillar.

4. T es una reflexión sobre el plano $x + z = 0$.
5. T es una reflexión sobre el eje plano $2x - y + z = 0$.
6. T es una rotación derecha de ángulo θ sobre el eje generado por $v = (2, -1, 1)$. Este caso, puede requerir de computadora para efectuar y simplificar las operaciones involucradas.

Capítulo 11

Vectores y Valores Propios

Las ideas de vector y valor propio constituyen conceptos centrales del álgebra lineal y resultan una valiosa herramienta en la solución de numerosos problemas de la matemática. Aquí se presentan las definiciones y resultados principales que permiten determinarlos, y se aplican al problema de reconocer las cónicas y superficies cuádricas, a partir de su ecuación.

11.1 Concepto de valor y vector propio

Definición 11.1 (Vector y valor propio)

Sea A una matriz $n \times n$, decimos que un número real λ es un **valor propio** *de A si existe una columna x de $I\!R^n$, $x \neq 0$, tal que $Ax = \lambda x$. El vector x se llama* **vector propio** *de A asociado a λ.*

También se utiliza el término valor característico y autovalor para nombrar un valor propio y, correspondientemente, vector característico y autovector, para el vector propio.

Ejemplo 11.1 Sea $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ y $x = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Como x es un vector no nulo en el $\text{Nuc}\,(A)$, verifica que $Ax = 0_3$, o lo que es lo mismo:

$$Ax = 0x.$$

Luego $x = (1, -1, 1)^t$ es un vector propio de A asociado al valor propio $\lambda = 0$. Por otra parte, $\forall\ a \in I\!R$, $a \neq 0$, $x = (a, -a, a)^t$ también es vector propio de A asociado a 0.

Este ejemplo permite ver que en general si $\text{Nuc}\,(A) \neq \{0\}$, entonces los vectores no nulos en $\text{Nuc}\,(A)$ son vectores propios de A asociados al valor propio 0.

Ejemplo 11.2 Sea $B = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ y $x = \begin{pmatrix} 0 \\ t \\ 0 \end{pmatrix}$ $t \in I\!R$. Si se efectúa el producto matricial se verifica que

$$Bx = 1x$$

luego $\forall t \in I\!R, t \neq 0$, $x = (0, t, 0)^t$ es un vector propio de B asociado al valor propio $\lambda = 1$.

En los ejemplos anteriores se observa que para cada valor propio existe una infinitud de vectores propios. De hecho es fácil ver que si x es un vector propio de A asociado al valor propio λ, $Ax = \lambda x$, entonces $A(\alpha x) = \lambda(\alpha x)$, luego αx también es vector propio de A asociado a λ. Además, si y es otro vector propio de A también asociado al valor propio λ, $Ay = \lambda y$, entonces

$$A(x + y) = Ax + Ay = \lambda x + \lambda y = \lambda(x + y)$$

luego $x + y$ es otro vector propio de A asociado a λ. De esta manera, para una matriz A $n \times n$, si λ es un valor propio de A el conjunto de vectores propios de A asociados a λ, incluyendo el vector cero que no es vector propio, es un subespacio vectorial de $I\!R^n$.

Definición 11.2 (Subespacio propio o característico)

Sea λ un valor propio de A, el conjunto $V_\lambda = \{x\,|Ax = \lambda x\}$ se llama ***subespacio propio*** *o espacio característico de A asociado a λ. Y la dimensión de V_λ se denomina multiplicidad geométrica de λ.*

Observe que:

$$\begin{aligned} V_\lambda &= \{x|Ax = \lambda x\} \\ &= \{x|Ax - \lambda x = 0\} \\ &= \{x|(A - \lambda I)x = 0, \} \end{aligned}$$

de manera que el subespacio propio de A asociado al valor propio λ corresponde al núcleo de la matriz $A - \lambda I$. Equivalente se muestra que también es el núcleo de $\lambda I - A$.

Teorema 11.3 *Si $\lambda_1, \ldots, \lambda_k$, son k valores propios de A, diferentes entre sí y asociados respectivamente a los vectores propios $v_1, \ldots, v_k$, entonces $v_1, \ldots, v_k$ son linealmente independientes.*

Demostración: La prueba se hace por inducción sobre el número de valores propios de A. Sea k este número:

Para $k = 2$; se debe probar que

$$\alpha_1 v_1 + \alpha_2 v_2 = 0 \Longrightarrow \alpha_1 = \alpha_2 = 0. \tag{11.1}$$

De la hipótesis en 11.1 se tiene que

$$0 = A\left(\alpha_1 v_1 + \alpha_2 v_2\right) = \alpha_1 A v_1 + \alpha_2 A v_2$$

Luego

$$\lambda_1 \alpha_1 v_1 + \lambda_2 \alpha_2 v_2 = 0 \tag{11.2}$$

Sumando $-\lambda_1(\alpha_1 v_1 + \alpha_2 v_2 = 0)$ a ecuación (11.2) se obtiene:

$$-\lambda_1 \alpha_2 v_2 + \lambda_2 \alpha_2 v_2 = 0 \quad \text{o sea} \quad \left(-\lambda_1 + \lambda_2\right) \alpha_2 v_2 = 0$$

Como $-\lambda_1 + \lambda_2 \neq 0$ entonces $\alpha_2 = 0$. Sustituyendo $\alpha_2 = 0$ en $\alpha_1 v_1 + \alpha_2 v_2 = 0$ se llega a $\alpha_1 = 0$.

Supongamos ahora que el resultado vale para $k-1$ y que

$$\alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} + \alpha_k v_k = 0.$$

Procediendo de manera similar al caso $k = 2$, se llega a la ecuación:

$$\begin{aligned}(-\lambda_k + \lambda_1)\,\alpha_1 v_1 + (-\lambda_k + \lambda_2)\,\alpha_2 v_2 + \cdots \\ + (-\lambda_k + \lambda_{k-1})\,\alpha_{k-1} v_{k-1} = 0.\end{aligned}$$

Como por hipótesis de inducción $v_1, \ldots, v_{k-1}$ son l.i. entonces $(-\lambda_k + \lambda_i)\,\alpha_i = 0$ para $i = 1, \ldots, k-1$. De esto se sigue fácilmente que $\alpha_1 = \cdots = \alpha_k = 0$.

■

11.1.1 Cálculo de valores y vectores propios

La búsqueda y cálculo de valores propios para una matriz A se apoya en las equivalencias siguientes:

$$\begin{aligned} & \lambda \text{ es un valor propio de } A \\ \Longleftrightarrow\ & Ax = \lambda x \quad \text{tiene soluciones no nulas} \\ \Longleftrightarrow\ & (A - \lambda I)x = 0 \quad \text{tiene infinitas soluciones} \\ \Longleftrightarrow\ & \det(A - \lambda I) = 0. \end{aligned}$$

Con lo que obtenemos el siguiente teorema:

Teorema 11.4 *Las proposiciones que siguen son equivalentes:*

(a) λ *es un valor propio de* A.

(b) *det*$(A - \lambda I) = 0$.

Ejemplo 11.3 Considere la matriz

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

Para determinar posibles valores propios de A debemos resolver la ecuación en la variable λ: $\det(A - \lambda I) = 0$, denominada ecuación característica. Lo cual requiere calcular y factorizar el polinomio $P(\lambda) = \det(A - \lambda I)$:

$$\begin{aligned}
P(\lambda) &= \det(A - \lambda I) \\
&= \begin{vmatrix} 5-\lambda & 6 & -6 \\ -3 & -4-\lambda & 6 \\ 0 & 0 & 2-\lambda \end{vmatrix} \\
&= (2-\lambda)[(5-\lambda)(-4-\lambda) - (-3)6] \\
&= (2-\lambda)[\lambda^2 - \lambda - 2] \\
&= (2-\lambda)[(\lambda-2)(\lambda+1)] \\
&= -(\lambda-2)^2(\lambda+1).
\end{aligned}$$

De esta manera $\det(A - \lambda I) = 0 \iff -(\lambda-2)^2(\lambda+1) = 0 \iff \lambda = 2$ o $\lambda = -1$. Y los únicos valores propios de A son -1 y 2.

Definición 11.5 (Polinomio característico)

Si $A \in M(n, I\!R)$ se llama **polinomio característico** *de A y se denota P_A, al polinomio de grado n:*

$$P_A(\lambda) = \det(A - \lambda I).$$

Si $P_A(\lambda)$ se factoriza en factores lineales y, eventualmente, algunos factores irreducibles de grado mayor igual que 2 cuyo producto denotamos $Q(\lambda)$:

$$P_A(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \dots (\lambda - \lambda_r)^{n_r} Q(\lambda),$$

se dice que n_i es la **multiplicidad algebraica** *del valor propio λ_i.*

Ejemplo 11.4 El polinomio característico de la matriz A dada en el ejemplo 11.3, es:

$$P_A(\lambda) = -(\lambda-2)^2(\lambda+1) = -\lambda^3 + 3\lambda^2 - 4$$

Luego los valores propios son: $\lambda = 2$ y $\lambda = -1$ con multiplicidad algebraica 2 y 1 respectivamente.

El hecho de que los valores propios de A, en el ejemplo 11.3, resultaran enteros es debido a una selección especial —con fines didácticos— de la matriz A. En general si A es una matriz $n \times n$, con entradas reales se puede demostrar que $\det(A - \lambda I)$ es un polinomio en λ, de grado n con coeficientes reales, que tiene a lo sumo n raíces reales. El cómputo o aproximación de estas raíces reales en general demanda la utilización de métodos numéricos. En este material, se evita dicha dificultad eligiendo matrices cuya ecuación característica pueda ser resuelta por métodos de factorización.

Observaciones:

- Una matriz A, de orden n, tiene a lo sumo n valores propios ya que el polinomio característico a lo sumo tiene n ceros.
- Los valores propios de una matriz triangular son los elementos en la diagonal. Nótese que si A es triangular entonces $A - \lambda I$ también lo es, de donde se sigue el resultado.
- Las matrices A y $C^{-1}AC$ tienen igual polinomio característico y por tanto los mismos valores propios. En efecto:

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

El procedimiento de cálculo de valores y vectores propios se resume así :

Paso 1: Calcular y factorizar el polinomio $\det(A - \lambda I)$, a efecto de obtener los valores propios de A. Es decir, resolver la ecuación característica: $\det(A - \lambda I) = 0$.

Paso 2: Para cada valor propio λ, hallar una base de V_λ, resolviendo el sistema $(A - \lambda I)\, x = 0$.

Ejemplo 11.5 Para la matriz A del ejemplo 11.3, ya conocemos el resultado del paso 1:

$$|A - \lambda I| = -(\lambda - 2)^2(\lambda + 1) = 0 \Longleftrightarrow \lambda = 2 \ \text{ ó } \ \lambda = -1.$$

Ahora determinamos los espacios característicos asociados a estos valores propios:

$V_{\lambda=2}$: $Ax = 2x \Longleftrightarrow (A - 2I)x = 0$ y resolviendo este sistema homogéneo:

$$\left(\begin{array}{rrr|r} 3 & 6 & -6 & 0 \\ -3 & -6 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow[\frac{1}{3}f_1]{f_1+f_2} \left(\begin{array}{rrr|r} 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Así $x_1 = -2x_2 + 2x_3$ y $V_{\lambda=2} = \mathcal{C}\ell\{(-2,1,0)^t, (2,0,1)^t\}$.

$V_{\lambda=-1}$: $Ax = -1x \Longleftrightarrow (A + I)x = 0$ y resolviendo:

$$\left(\begin{array}{rrr|r} 6 & 6 & -6 & 0 \\ -3 & -3 & 6 & 0 \\ 0 & 0 & 0 & 3 \end{array}\right) \xrightarrow{\cdots} \left(\begin{array}{rrr|r} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Así $x_1 = -x_2$ y $x_3 = 0$ luego $V_{\lambda=-1} = \mathcal{C}\ell\{(-1,1,0)^t\}$.

11.2 Diagonalización de matrices

Algunas matrices A permiten una factorización en términos de una matriz invertible y una matriz diagonal, en la que se utilizan sus vectores y valores propios. Para obtener esta factorización se realiza un proceso denominado diagonalización de matrices.

11.2.1 Caracterización de matrices diagonalizables

Definición 11.6 (Matriz diagonalizable)

Una matriz A es diagonalizable si existe una matriz C invertible y una matriz D diagonal tales que

$$C^{-1}AC = D.$$

Esto también significa que A se factoriza como

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

Para determinar las matrices C y D en esta factorización, observe que

$$C^{-1}AC = D \Longleftrightarrow AC = CD.$$

Y escribiendo apropiadamente la igualdad $AC = CD$ se reconoce en ella la presencia de los vectores y valores propios de A: si

$$C = (C_1, C_2, \ldots, C_n),$$

donde $C_1, C_2, \ldots, C_n$ son las columnas de C y

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

entonces, por una parte

$$AC = (AC_1, AC_2, \ldots, AC_n)$$

y por otra, si se recuerda el efecto sobre las columnas de una matriz al multiplicarla por matriz diagonal (por la derecha), se tiene que

$$CD = (\lambda_1 C_1, \lambda_2 C_2, \ldots, \lambda_n C_n).$$

Finalmente

$$\begin{aligned} & AC = DC \\ \iff \quad & (AC_1, AC_2, \ldots, AC_n) = (\lambda_1 C_1, \lambda_2 C_2, \ldots, \lambda_n C_n) \\ \iff \quad & AC_i = \lambda_i C_i \qquad \forall\, i = 1, 2, \ldots n. \end{aligned}$$

De lo cual observamos que los elementos en la diagonal de la matriz D, buscada, son valores propios de A y las columnas de C son los respectivos vectores propios. Además, como C es una matriz invertible sus columnas deben ser l.i.

Entonces

A es diagonalizable $\iff$ A tiene n vectores propios l.i.

o equivalentemente, como se propone en el siguiente teorema.

Teorema 11.7 *Sea $A \in M(n, \mathbb{R})$ tal que el polinomio característico se puede factorizar como:*

$$P_A(\lambda) = \det(A - \lambda I) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_r)^{n_r}$$

donde $n_1 + n_2 + \cdots + n_r = n$, $\lambda_1, \ldots, \lambda_r$ son todos los valores propios distintos de A, y $V_{\lambda_1}, \ldots, V_{\lambda_r}$ los espacios propios correspondientes. Entonces las siguientes proposiciones son equivalentes:

(a) *La matriz A es diagonalizable.*

(b) *Existe una base $\mathcal{B} = \{v_1, \ldots, v_n\}$, para $\mathbb{R}^n$, de vectores propios de A.*

(c) *Para cada λ_i, $i = 1, \ldots, r$, su multiplicidad geométrica es igual a su multiplicidad algebraica. Es decir, $dim(V_{\lambda_i}) = n_i$, para todo $i = 1, \ldots, r$.*

(d) $dim(V_{\lambda_1}) + \cdots + dim(V_{\lambda_r}) = n$.

(e) *Todo vector $x \in \mathbb{R}^n$ se puede escribir de manera única en la forma $x = x_1 + \cdots + x_r$, con $x_i \in V_{\lambda_i}$.*

Una demostración de este teorema se puede consultar en [7].

Ejemplo 11.6 Considere la matriz

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

Su polinomio característico es:

$$\begin{aligned} P_A(\lambda) &= \det(A - \lambda I) \\ &= \begin{vmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 1 & 2-\lambda \end{vmatrix} \\ &= (2-\lambda)[(1-\lambda)(1-\lambda) - 1] \\ &= -\lambda(\lambda - 2)^2. \end{aligned}$$

Así los valores propios son 2 y 0, y determinando los correspondientes espacios propios (o característicos) V_2 y V_0 tenemos:

V_2: $Ax = 2x \Longleftrightarrow (A - 2I)x = 0$ y resolviendo:

$$\left(\begin{array}{rrr|r} -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right) \xrightarrow{\cdots} \left(\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Luego $x_1 = 0$ y $x_2 = 0$, de manera que $V_2 = \mathcal{C}\ell\{(0, 0, 1)^t\}$.

V_0: $Ax = 0x \Longleftrightarrow Ax = 0$ y resolviendo:

$$\left(\begin{array}{rrr|r} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \xrightarrow{\cdots} \left(\begin{array}{rrr|r} 1 & 0 & -2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Entonces $x_1 = 2x_3$ y $x_2 = -2x_3$, luego $V_0 = \mathcal{C}\ell\{(2, -2, 1)^t\}$.

De esto tenemos que $\dim(V_2) = 1$ y $\dim(V_0) = 1$, luego por la parte (c) del teorema 11.7, la matriz A no es diagonalizable.

Observe que para efectos de determinar si la matriz A es diagonalizable o no, en el ejemplo anterior, el cómputo de V_0 es un paso innecesario. Porque al determinar que

$$\dim(V_2) = 1 \neq 2 = \text{multiplicidad algebraica de} \lambda = 2,$$

se tiene que A no es diagonalizable.

Ejemplo 11.7 Para la matriz A del ejemplo 11.3, en página 330,

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

se obtuvo que

$$P_A(\lambda) = -(\lambda - 2)^2(\lambda + 1).$$

Además, en el ejemplo 11.5 se calculó:

$$V_{\lambda=2} = \mathcal{C}\ell\{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}\} \quad \text{y} \quad V_{\lambda=-1} = \mathcal{C}\ell\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\}.$$

Luego es posible elegir una base $\mathcal{B}$ para $I\!R^3$ de vectores propios de A. Específicamente:

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

Así se tiene que A es diagonalizable, es decir, que $A = CDC^{-1}$ donde

$$C = \begin{pmatrix} -2 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Observe que el orden de las columnas C se eligió arbitrariamente. Si se cambia este orden, otra elección para C y D puede ser:

$$C = \begin{pmatrix} -2 & -1 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

En ambos casos se comprueba que $A = CDC^{-1}$ verificando que $AC = CD$.

El teorema 11.7 anterior, presenta un caso particular importante cuando $r = n$. Esto es, si A tiene n valores propios distintos entonces es clara la existencia de n vectores propios l.i. y por lo tanto la matriz A es diagonalizable.

Corolario 11.8 *Si una matriz A, $n \times n$, tiene n valores propios distintos entonces es diagonalizable.*

Ejemplo 11.8 La matriz $A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ tiene como polinomio característico a:

$$P_A(\lambda) = \det \begin{pmatrix} -\lambda & 0 \\ 1 & 1-\lambda \end{pmatrix} = \lambda(\lambda - 1).$$

Por lo tanto A tiene dos valores propios distintos cuyos espacios característicos tienen dimensión 1, necesariamente. Y por lo tanto A es diagonalizable.

11.2.2 Matrices ortogonalmente diagonalizables

El proceso de diagonalización de una matriz A presenta una situación especial cuando la matriz C, cuyas columnas son una base de vectores propios de A, resulta ser ortogonal. Es decir cuando la base de vectores propios de A es una base ortonormal.

Definición 11.9 (Diagonalización ortogonal)

Una matriz A, $n \times n$, es ortogonalmente diagonalizable si existe una matriz C ortogonal y una matriz D diagonal, tales que

$$C^t AC = D$$

Recordemos que C es ortogonal si $C^tC = I$, es decir, si $C^{-1} = C^t$. Así la diagonalización ortogonal es un caso particular de diagonalización, en el cual la base de vectores propios que forman las columnas de C es una base ortonormal.

Es fácil ver que las matrices ortogonalmente diagonalizables son simétricas:

$$\begin{array}{ll} & A \text{ es ortogonalmente diagonalizable} \\ \iff & \text{existe } C \text{ ortogonal y } D \text{ diagonal} \\ & \text{ tales que} \quad C^tAC = D \\ \Longrightarrow & A = CDC^t \\ \Longrightarrow & A^t = (CDC^t)^t = CDC^t = A. \end{array}$$

El resultado recíproco también es cierto y se basa en los siguientes teoremas:

Teorema 11.10 *Si $A \in M(n, I\!R)$ es simétrica, su polinomio característico sólo tiene raíces reales.*

Además, cuando la matriz es simétrica los vectores propios asociados a valores propios distintos, no solo son l.i. como en el caso general, si no que también son ortogonales.

Teorema 11.11 *Sea $A \in M(n, I\!\!R)$ simétrica y λ_1, λ_2 valores propios distintos de A con vectores propios asociados v y u respectivamente, entonces v y u son ortogonales.*

Demostración: Observe que

$$Av \cdot u = (Av)^t u = v^t A^t u = v^t (Au) = v \cdot Au$$

De esto se deduce que $\lambda_1(v \cdot u) = \lambda_2(v \cdot u)$ luego $(\lambda_1 - \lambda_2)(v \cdot u) = 0$, de donde $(v \cdot u) = 0$ dado que $\lambda_1 \neq \lambda_2$.

■■■

Y finalmente el siguiente resultado garantiza la diagonalización ortogonal de las matrices simétricas.

Teorema 11.12 *Sea A una matriz simétrica, es decir $A^t = A$. Entonces existe una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de $I\!\!R^n$, ortonormal, formada por vectores propios de A.*

Así, se obtiene la siguiente equivalencia.

Teorema 11.13 *A es simétrica si y solo si A es ortogonalmente diagonalizable.*

Demostración: Supongamos que A es simétrica y sea C la matriz cuyas columnas son los vectores propios $v_1, \ldots, v_n$, de A, ortonormales. Es decir $C = (v_1 \ldots v_n)$. Por multiplicación matricial se comprueba que:

(a)

$$C^t C = \begin{pmatrix} v_1^t \\ \vdots \\ v_n^t \end{pmatrix} (v_1 \ldots v_n) = \begin{pmatrix} v_1^t \cdot v_1 & \cdots & v_1^t \cdot v_n \\ \vdots & & \vdots \\ v_n^t \cdot v_1 & \cdots & v_n^t \cdot v_n \end{pmatrix} = I_n.$$

(b) C^tAC es diagonal. En efecto

$$\begin{aligned} C^tAC &= C^tA\,[v_1 \ldots v_n] \\ &= C^t\,[Av_1 \ldots Av_n] \\ &= C^t\,[\lambda_1 v_1 \ldots \lambda_n v_n] \\ &= C^t\,[v_1 \ldots v_n]\,D_\lambda \\ &= C^tCD_\lambda \\ &= D_\lambda. \end{aligned}$$

donde D_λ es la matriz diagonal que tiene a $\lambda_1, \ldots \lambda_n$ como entradas en la diagonal, en ese orden.

Luego A es ortogonalmente diagonalizable. El recíproco fue demostrado al inicio de esta sección.

∎

Ejemplo 11.9 Sea $A = \begin{pmatrix} 1 & 2 & -4 \\ 2 & 4 & 2 \\ -4 & 2 & 1 \end{pmatrix}$. Como A es simétrica, conocemos que es diagonalizable ortogonalmente. El proceso para determinar la matriz C ortogonal y la matriz D diagonal tales que $A = CDC^t$, es el mismo que en la diagonalización general, pero ahora con el cuidado de elegir las columnas de C como una base ortonormal de vectores propios de A:

Polinomio característico de A:

$$\begin{aligned} P_A(\lambda) &= \det(A - \lambda I) \\ &= \begin{vmatrix} 1-\lambda & 2 & -4 \\ 2 & 4-\lambda & 2 \\ -4 & 2 & 1-\lambda \end{vmatrix} \\ &= -(\lambda+4)(\lambda-5)^2 \end{aligned}$$

Los espacios propios V_{-4} y V_5:

V_{-4}**:** $Ax = -4x \Longleftrightarrow (A + 4I)x = 0$ y resolviendo:

$$\left(\begin{array}{rrr|r} 5 & 2 & -4 & 0 \\ 2 & 8 & 2 & 0 \\ -4 & 2 & 5 & 0 \end{array}\right) \xrightarrow{\quad\cdots\quad} \left(\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Entonces $V_{-4} = \mathcal{C}\ell\{(2, -1, 2)^t\}$.

V_5**:** Resolviendo $(A - 5I)x = 0$:

$$\left(\begin{array}{rrr|r} -4 & 2 & -4 & 0 \\ 2 & -1 & 2 & 0 \\ -4 & 2 & -4 & 0 \end{array}\right) \xrightarrow{\quad \cdots \quad} \left(\begin{array}{rrr|r} 1 & -1/2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Así $V_5 = \mathcal{C}\ell\{(-1,0,1)^t, (1,2,0)^t\}$.

Para la elección de una base de vectores propios ortonormales como columnas de la matriz C que se busca, se requiere ortonormalizar las bases obtenidas para cada espacio característico. El teorema 11.11 garantiza que la unión de las bases resultantes es una base ortonormal.

Así, una base ortonormal para V_{-4} es $\mathcal{B}_1 = \{(2/3, -1/3, 2/3)^t\}$ y para V_5 es $\mathcal{B}_2 = \{(-1/\sqrt{2}, 0, 1/\sqrt{2})^t, (\frac{1}{3\sqrt{2}}, \frac{2\sqrt{2}}{3}, \frac{1}{3\sqrt{2}})^t\}$ con lo que obtenemos las siguientes posibles elecciones para C y D:

$$C = \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{2\sqrt{2}}{3} & \frac{-1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

Fácilmente se verifica que $A = CDC^t$ comprobando que $AC = CD$.

11.3 Valores y vectores propios de operadores

En esta sección se denominan operadores a las transformaciones lineales de un espacio vectorial a él mismo. Y como se recordará, dada una base del espacio vectorial, hay una asociación directa entre matrices y transformaciones lineales (operadores en este caso), la que permite trasladar los conceptos de valor propio, vector propio, y diagonalización a los operadores.

Si T es un operador en $\mathbb{R}^n$, es decir, una transformación lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, y consideramos la matriz A de T, en la base canónica, entonces

$$Av = \lambda v \Longleftrightarrow T(v) = \lambda v$$

Definición 11.14 (Valor y Vector propio)

Sea T un operador en $\mathbb{R}^n$, λ es un valor propio de T si existe $v \in \mathbb{R}^n$ $v \neq 0_n$, tal que $T(v) = \lambda v$. Y v es el vector propio de T asociado a λ.

La definición anterior puede generalizarse a un operador de cualquier espacio vectorial con dimensión finita utilizando la relación

$$[T\,(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}\,[v]_{\mathcal{B}}$$

para "pasar" el concepto de valor y vector propio de matrices a los operadores de estos espacios.

Teorema 11.15 *λ es un valor propio del* **operador** *T asociado a v si y solo si λ es valor propio de $A = [T]_{\mathcal{B}}$ asociado a $x = [v]_{\mathcal{B}}$.*

Demostración:

$$\begin{array}{ll} & \lambda \text{ es un valor propio de } A \text{ asociado a } x \\ \Longleftrightarrow & Ax = \lambda x \\ \Longleftrightarrow & [T]_{\mathcal{B}}[v]_{\mathcal{B}} = \lambda[v]_{\mathcal{B}} \\ \Longleftrightarrow & [T(v)]_{\mathcal{B}} = [\lambda v]_{\mathcal{B}} \\ \Longleftrightarrow & T(v) = \lambda v. \end{array}$$

■

Ejemplo 11.10 Sea $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ el operador definido por

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

observe que

$$T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Luego $(0, 1, 0)^t$ es un vector propio de T asociado al valor propio $\lambda = 1$.

11.3.1 Diagonalización de operadores

Sea $A \in M(n, I\!R)$ una matriz diagonalizable y $T(x) = Ax$ el respectivo operador, es decir, si $\mathcal{C}$ es la base canónica de $I\!R^n$ entonces

$$A = [T]_{\mathcal{C}}.$$

Como A es diagonalizable existe una base de vectores propios de A, $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, que naturalmente también es una base de vectores propios de T puesto que

$$Av_i = \lambda_i v_i \Longleftrightarrow T(v_i) = \lambda_i v_i.$$

Luego, si definimos $C = (v_1, v_2, \ldots, v_n)$ la matriz cuyas columnas son los vectores de la base $\mathcal{B}$ conocemos que:

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

donde $D = \mathrm{diag}(\lambda_i)$ es una matriz diagonal. En esta situación observe que:

- La matriz de T en la base es $\mathcal{B}$ es diagonal:

$$\begin{array}{rcl} [T]_{\mathcal{B}} & = & ([T(v_1)]_{\mathcal{B}}, [T(v_2)]_{\mathcal{B}}, \ldots, [T(v_n)]_{\mathcal{B}}) \\ & = & ([\lambda_1 v_1]_{\mathcal{B}}, [\lambda_2 v_2]_{\mathcal{B}}, \ldots, [\lambda_n v_n]_{\mathcal{B}}) \\ & = & (\lambda_1 [v_1]_{\mathcal{B}}, \lambda_2 [v_2]_{\mathcal{B}}, \ldots, \lambda_n [v_n]_{\mathcal{B}}) \\ & = & (\lambda_1 e_1, \lambda_2 e_2, \ldots, \lambda_n e_n) \\ & = & \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & 0 & & \lambda_n \end{pmatrix} = D. \end{array}$$

- La matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es:

$$\begin{array}{rcl} [Id]_{\mathcal{B}}^{\mathcal{C}} & = & ([Id(v_1)]_{\mathcal{C}}, [Id(v_2)]_{\mathcal{C}}, \ldots, [Id(v_n)]_{\mathcal{C}}) \\ & = & (v_1, v_2, \ldots, v_n) \\ & = & C, \end{array}$$

  y consecuentemente $[Id]_{\mathcal{C}}^{\mathcal{B}} = C^{-1}$. Observe que en este caso se denotó la transformación identidad como Id.

Los dos resultados anteriores hacen que la relación

$$A = CDC^{-1}$$

también se escriba como:

$$[T]_{\mathcal{C}} = [Id]^{\mathcal{C}}_{\mathcal{B}}[T]_{\mathcal{B}}[Id]^{\mathcal{B}}_{\mathcal{C}}.$$

Lo que permite la siguiente definición del concepto de diagonalización para operadores:

Definición 11.16 (Operadores diagonalizables)

Se dice que el operador $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es diagonalizable si existe una base $\mathcal{B}$ de $\mathbb{R}^n$ tal que $[T]_{\mathcal{B}}$ es diagonal.

O equivalentemente:

Teorema 11.17 *T es diagonalizable si y solo si existen matrices $D = diag(\lambda_i)_{n\times n}$ diagonal, y P invertible cuyas columnas son una base $\mathcal{B}$ de $\mathbb{R}^n$, tales que*

$$P^{-1}[T]_{\mathcal{C}}P = [T]_{\mathcal{B}} = D,$$

donde $\mathcal{C}$ es la base canónica de $\mathbb{R}^n$.

Demostración: Si T es diagonalizable, sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de $\mathbb{R}^n$ tal que $[T]_{\mathcal{B}} = \mathrm{diag}(\lambda_i)_{n\times n} = D$ y $P = [v_1 \ldots v_n] = [\mathrm{Id}]^{\mathcal{C}}_{\mathcal{B}}$. Entonces $D = [T]_{\mathcal{B}} = [\mathrm{Id}]^{\mathcal{B}}_{\mathcal{C}} [T]_{\mathcal{C}} [\mathrm{Id}]^{\mathcal{C}}_{\mathcal{B}} = P^{-1} [T]_{\mathcal{C}} P$.

Por otra parte si $P^{-1} [T]_{\mathcal{C}} P = D$ entonces $[T]_{\mathcal{C}} P = PD$. Definiendo $P = [v_1 \ldots v_n]$ y $\mathcal{B} = \{v_1, \ldots, v_n\}$ se tiene $[T]_{\mathcal{C}} v_i = [\lambda_i v_i]_{\mathcal{C}}$, es decir $[T(v_i)]_{\mathcal{C}} = [\lambda_i v_i]_{\mathcal{C}}$. Por lo que $T(v_i) = \lambda_i v_i$, o lo que es equivalente $[T]_{\mathcal{B}} = \mathrm{diag}(\lambda_i)_{n\times n}$.

■■■

Definición 11.18 (Op. ortogonalmente diagonalizable)

Un operador $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es ortogonalmente diagonalizable si existe una base ortonormal $\mathcal{B}$ de $\mathbb{R}^n$ tal que $[T]_{\mathcal{B}}$ es diagonal.

Se puede deducir que si existe una base $\mathcal{B}$ de $\mathbb{R}^n$ tal que $[T]_{\mathcal{B}}$ es ortogonalmente diagonalizable entonces T es ortogonalmente diagonalizable.

En efecto, sea H ortogonal tal que $H^t\,[T]_{\mathcal{B}}\,H = \text{diag}(\lambda_i)$ con $[S]_{\mathcal{B}} = H$ y $\mathcal{B} = \{v_1, \ldots, v_n$. Luego $[S^{-1}TS]_{\mathcal{B}} = \text{diag}(\lambda_i)$ por lo que $T(Sv_i) = \lambda_i Sv_i$.

Ejemplo 11.11 Sea $\mathcal{C}$ la base canónica de $I\!R^3$.
El operador $T : I\!R^3 \longrightarrow I\!R^3$ definido por la fórmula

$$T\,(x, y, z) = (x - z, y + z, -x + y + 2z)$$

es ortogonalmente diagonalizable puesto que

$$[T]_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

es simétrica.

Para calcular la matriz ortogonal C tal que $C^t AC = D_\lambda$ se debe calcular una base ortonormal de $I\!R^3$, o lo que es equivalente, una base ortonormal de cada espacio propio. El polinomio característico de T es $p(\lambda) = \lambda(1 - \lambda)(\lambda - 3)$. Las bases ortonormales de los espacios propios V_0, V_1 y V_3 son, respectivamente, $\left\{\frac{1}{\sqrt{3}}\,(1, -1, 1)\right\}$, $\left\{\frac{1}{\sqrt{2}}\,(1, 1, 0)\right\}$ y $\left\{\frac{1}{\sqrt{6}}\,(-1, 1, 2)\right\}$. Por lo tanto

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

11.4 Diagonalización de formas cuadráticas

La diagonalización de matrices produce, mediante un cambio de varible, una simplificación de las formas cuadráticas con importantes aplicaciones.

Definición 11.19 (Forma cuadrática)

Una función $f : I\!R^n \longrightarrow I\!R$ tal que $f\,(x) = x^t Ax$ se llama forma cuadrática, donde A es una matriz simétrica denominada matriz de la forma cuadrática.

Ejemplo 11.12 Consideremos la forma cuadrática

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_3 + 2x_2x_3.$$

La matriz (simétrica) de f es $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$. Lo cual se verifica, efectuando el producto matricial:

$$\begin{aligned} f(x) &= x^tAx \\ f[(x_1, x_2, x_3)^t] &= (x_1, x_2, x_3)\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= (x_1, x_2, x_3)\begin{pmatrix} x_1 - x_3 \\ x_2 + x_3 \\ -x_1 + x_2 + x_3 \end{pmatrix} \\ &= x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_3 + 2x_2x_3. \end{aligned}$$

Observe que una forma cuadrática es un polinomio de segundo grado en el que todos sus términos son cuadráticos. Por ejemplo, para el caso de tres variables tenemos:

$$f(x_1, x_2, x_3) = ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_1x_3 + gx_2x_3.$$

La matriz A de esta forma cuadrática, se obtiene al reconocer que:

i) si $i \neq j$, el coeficiente del término x_ix_j del polinomio define las entradas a_{ij} y a_{ji},

$$a_{ij} = a_{ji} = \frac{\text{coeficiente de } x_{ij}}{2},$$

ii) cuando $i = j$, el coeficiente de x_i^2 es la i-ésima entrada en la diagonal de A.

Específicamente:

$$A = \begin{pmatrix} a & d/2 & e/2 \\ d/2 & b & g/2 \\ e/2 & g/2 & c \end{pmatrix}.$$

En general, si $x^t = (x_1, \ldots, x_n)$ y $A = (a_{ij})$, la forma cuadrática $f(x)$ se escribe como:

$$f(x) = x^t Ax = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j$$

o

$$f(x) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j.$$

La idea principal en esta sección es que con un cambio de variable apropiado, es posible eliminar de una forma cuadrática la expresión que corresponde a la segunda sumatoria. Es decir, remover todos los términos de la forma $a_{ij} x_i x_j$ con $i < j$, que en lo sucesivo denominaremos términos cruzados.

Ejemplo 11.13 La matriz A de la forma cuadrática en el ejemplo 11.12, se puede factorizar como:

$$A = CDC^t$$

donde,

$$C = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Esto se obtuvo en el ejemplo 11.11, donde se observa que los valores propios de A son: $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 3$ y los respectivos vectores propios definen las columnas de C, de manera que $C^t AC = D$.

Luego, con el cambio de variable $y = C^t x$, la forma cuadrática se transforma en:

$$\begin{aligned} x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_3 + 2x_2x_3 &= x^t Ax \\ &= x^t CDC^t x \\ &= (C^t x)^t D(C^t x) \\ &= y^t Dy \\ &= y_2^2 + 3y_3^2 \end{aligned}$$

la cual no incluye términos cruzados.

En general tenemos el siguiente resultado:

Teorema 11.20 *Sea* $C = (v_1 \dots v_n)$ *cuyas columnas son una base de vectores propios, ortonormalizados, de la matriz simétrica* A. *El cambio de variable*

$$y = C^t x$$

transforma la forma cuadrática $f(x) = x^t A x$ *en*

$$f(x) = y^t D y = \sum_{i=1}^{n} \lambda_i y_i^2$$

donde $\lambda_1, \dots, \lambda_n$ *son los valores propios de* A, *asociados a* $v_1, \dots, v_n$ *respectivamente.*

Demostración: Por el teorema 11.13, $C^t A C = D_\lambda$. Entonces $f(x) = f(Cy) = (Cy)^t ACy = y^t (C^t AC) y = y^t D_\lambda y = \sum_{i=1}^{n} \lambda_i y_i^2$.

■

Esta reducción de las formas cuadráticas tiene aplicaciones importantes en la Estadística y otras disciplinas. Nosotros la usaremos para reconocer curvas y superficies cuadráticas.

11.5 Rotación de cónicas y superficies cuadráticas

Las curvas y superficies cuadráticas se asocian a casos particulares, $n = 2$ y $n = 3$, de ecuaciones de la forma

$$x^t A x + Bx + d = 0$$

cuyo primer término es una forma cuadrática. En el caso de las cónicas, el polinomio de segundo grado, en dos variables, de la ecuación que la define:

$$ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0$$

puede ser escrito matricialmente como

$$ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = x^tAx + Bx + f$$

donde,

$$A = \begin{pmatrix} a & c/2 \\ c/2 & b \end{pmatrix} \text{ y } B = (d,\ e).$$

Similarmente, un polinomio cuadrático en tres variables se escribe matricialmente:

$$\begin{aligned} & ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_1x_3 + fx_2x_3 + gx_1 + hx_2 + kx_3 + l \\ = \quad & (x_1, x_2, x_3) \begin{pmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + (g, h, k) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + l \\ = \quad & x^tAx + Bx + l. \end{aligned}$$

Las ecuaciones que permiten reconocer una cónica o superficie cuadrática, se caracterizan porque no tienen términos cruzados. Sin embargo, las que tienen términos cruzados se pueden reducir a las primeras, con la aplicación del teorema 11.20 y el cambio de variable involucrado introduce un nuevo sistema de ejes (o de coordenadas), en términos del cual se simplifica la ecuación y puede ser identificada.

Pero antes de hacer las reducciones mencionadas, debemos elaborar un "inventario" de las curvas que reconocemos a partir de su ecuación canónica.

11.5.1 Cónicas y sus ecuaciones canónicas

Las elipses, hipérbolas y parábolas son las curvas cuadráticas que denominamos cónicas —se definen por la intersección entre un plano y un cono— y aparecen con mucha frecuencia en los cursos de matemática y sus aplicaciones.

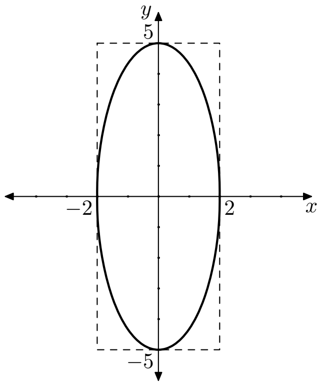


Figura 11.2: Elipse de ecuación $\frac{x^2}{2^2} + \frac{y^2}{5^2} = 1,\ \alpha > \beta$.

Elipse

La **elipse** no rotada y centrada en (x_0, y_0) es una curva cuadrática que se puede caracterizar por la ecuación

$$\frac{(x - x_0)^2}{\alpha^2} + \frac{(y - y_0)^2}{\beta^2} = 1.$$

De esta ecuación se dice que está escrita en la forma canónica.

Véase a continuación la forma de los gráficos de la elipse centrada en el origen, en los casos $\alpha > \beta$ y $\alpha < \beta$ respectivamente.

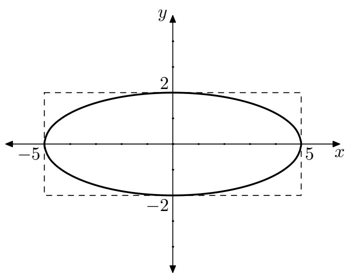


Figura 11.1: Elipse de ecuación $\frac{x^2}{5^2} + \frac{y^2}{2^2} = 1,\ \alpha > \beta$.

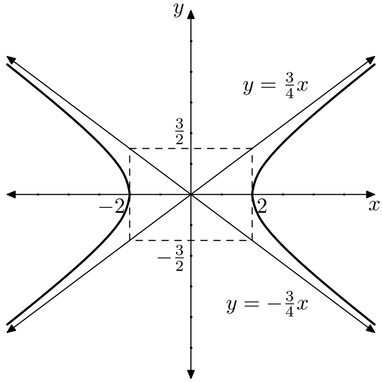


Figura 11.3: Hipérbola de ecuación $\frac{x^2}{2^2} - \frac{y^2}{1.5^2} = 1,\ \alpha > \beta$.

Hipérbola

Análogamente la **hipérbola** centrada en (x_0, y_0), no rotada, es una curva cuadrática cuya ecuación canónica tiene la forma

$$\pm\frac{(x-x_0)^2}{\alpha^2} \mp \frac{(y-y_0)^2}{\beta^2} = 1.$$

Las rectas que pasan por el punto (x_0, y_0), en la dirección de los vectores (α, β) y $(-\alpha, \beta)$:

$$(x, y) = (x_0, y_0) \mp t(\alpha, \beta),$$

se llaman asíntotas de la hipérbola. También pueden ser escritas como

$$y = \pm\frac{\beta}{\alpha}(x - x_0) + y_0.$$

Parábola

Las formas de las ecuaciones, que convenimos en denominar canónicas, para las **parábolas** son:

$$\begin{aligned}(x - x_0) &= a(y - y_0)^2 \\ (y - y_0) &= a(x - x_0)^2\end{aligned}$$

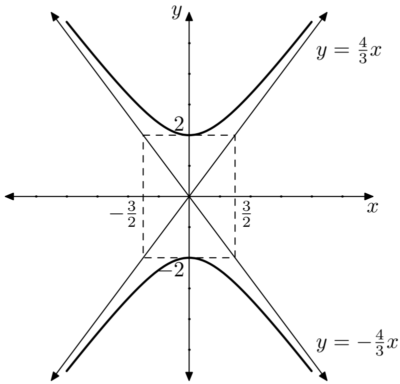


Figura 11.4: Hipérbola de ecuación $-\frac{x^2}{1.5^2} + \frac{y^2}{2^2} = 1,\ \alpha > \beta.$

donde el punto (x_0, y_0) corresponde con el vértice de la parábola y cuando la ecuación es del primer tipo, la parábola abre en la dirección del eje X, hacia la derecha de x_0 si $a > 0$ y hacia la izquierda de x_0 si $a < 0$. La segunda ecuación corresponde a una parábola que abre en la dirección del eje Y, hacia arriba de y_0 si $a > 0$ y hacia abajo de y_0 si $a < 0$.

Cónicas degeneradas

En general, la ecuación $ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0$ corresponde a alguna de las cónicas vistas. Sin embargo, en ciertos casos, el conjunto de puntos que satisfacen la ecuación es vacío, o representan rectas o puntos aislados. Por ejemplo, la ecuación $x^2 + y^2 + 1 = 0$ no tiene soluciones, por otra parte, la ecuación

$$x_1^2 - x_2^2 = 0$$

la satisfacen sólo los puntos de las rectas $x_2 = \pm x_1$. Y la ecuación $x_1^2 - 2x_1 + 1 + y_1^2 = 0$, tiene sólo el punto $(1, 0)$ como solución.

En casos como estos dos últimos se suele decir que las cónicas son degeneradas.

11.5.2 Ejes principales, ángulo de rotación

Para identificar la cónica que corresponde a una ecuación

$$ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0$$

se debe eliminar el término cruzado cx_1x_2 y hacer transformaciones en la ecuación resultante, hasta obtener una de las formas canónicas vistas.

Lo cual supone, en primer término, escribir la ecuación en su forma matricial:

$$x^tAx + Bx + f = 0$$

donde $A = \begin{pmatrix} a & \frac{c}{2} \\ \frac{c}{2} & b \end{pmatrix}$, $B = (d\ ,\ e)$.

Luego se determina, $\{v_1, v_2\}$, una base ortonormada de $I\!R^2$ formada de vectores propios de A, asociados respectivamente a λ_1 y λ_2. Así, con $C = (v_1, v_2)$ y $D = \text{diag}(\lambda_1, \lambda_2)$ se tiene que $A = CDC^t$, y el cambio de variable $y = C^tx$ conduce a:

$$\begin{aligned}
& ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0 \\
\Longleftrightarrow\quad & x^tAx + Bx + f = 0 \\
\Longleftrightarrow\quad & x^tCDC^tx + Bx + f = 0 \\
\Longleftrightarrow\quad & (C^tx)^tD(C^tx) + BCy + f = 0 \\
\Longleftrightarrow\quad & y^tDy + BCy + f = 4 \\
\Longleftrightarrow\quad & (y_1, y_2)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + (r, s)\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + f = 0 \\
\Longleftrightarrow\quad & \lambda_1y_1^2 + \lambda_2y_2^2 + ry_1 + sy_2 + f = 0
\end{aligned}$$

donde $(r, s) = BC$.

En esta última ecuación, si λ_1 y r no son cero, (similarmente para λ_2 y s), se agrupan los términos $\lambda_1y_1^2 + ry_1$ y $\lambda_2y_2^2 + sy_2$ y completando cuadrados se les da la forma $\lambda_1(y_1 \pm h)^2 \pm t$ y $\lambda_2(y_2 \pm k)^2 \pm u$, para obtener:

$$\lambda_1(y_1 \pm h)^2 + \lambda_2(y_2 \pm k)^2 = -(f \pm t \pm u)$$

y finalmente, con más transformaciones algebraicas, se obtiene una de las ecuaciones canónicas.

De todas estas transformaciones, resulta de especial importancia la interpretación para el cambio de variable $y = C^tx$:

- La matriz $C^t = C^{-1} = [Id]_{\mathcal{C}}^{\mathcal{B}}$ es la matriz de cambio de base de la canónica a $\mathcal{B}$, esto porque:

$$C = (v_1, v_2) = ([Id(v_1)]_{\mathcal{C}}, [Id(v_2)]_{\mathcal{C}}) = [Id]_{\mathcal{B}}^{\mathcal{C}}.$$

  Así el cambio de variable $y = C^t x$, cambia el vector $x = (x_1, x_2)^t$, coordenadas en la base canónica de un punto, a $y = (y_1, y_2)^t$, coordenadas del mismo punto en la base $\mathcal{B} = \{v_1, v_2\}$.

- Las rectas determinadas por los vectores v_1, v_2 son los **ejes principales** de la cónica. Los vectores v_1, v_2 siempre se pueden escoger de modo que $\det(C) = \det[(v_1, v_2)] = 1$, para que cambio de variable $y = C^t x$ pueda ser interpretado como una rotación de ejes.

- Si $\det(C) = 1$, el **ángulo de rotación** θ es el ángulo entre el vector canónico $e_1 = (1, 0)$ y el vector v_1.

Ejemplo 11.14 Calcular el ángulo de rotación, los ejes principales y la ecuación canónica respecto de estos ejes, de la cónica definida por la ecuación $x_1^2 - 4x_1x_2 + x_2^2 - 4 = 0$.

Solución:

$$\begin{aligned} & x_1^2 - 4x_1x_2 + x_2^2 = 4 \\ \iff & (x_1, x_2)\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 4 \\ \iff & x^t A x = 4. \end{aligned}$$

Los valores propios de la matriz simétrica $A = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$ son $\lambda_1 = -1$ y $\lambda_2 = 3$, con vectores propios ortonormalizados:

$$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ y } v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Eligiendo $C = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ y $D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$ se tiene que $A = CDC^t$ y sustituyendo en la ecuación de la cónica, tene-

mos:
$$\begin{aligned}
& x_1^2 - 4x_1x_2 + x_2^2 = 4 \\
\Longleftrightarrow\quad & x^tAx = 4 \quad\Longleftrightarrow\quad x^tCDC^tx = 4 \\
\Longleftrightarrow\quad & (C^tx)^tD(C^tx) = 4 \quad\Longleftrightarrow\quad y^tDy = 4 \\
\Longleftrightarrow\quad & (y_1, y_2)\begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 4 \\
\Longleftrightarrow\quad & -y_1^2 + 3y_2^2 = 4 \\
\Longleftrightarrow\quad & -\frac{y_1^2}{2^2} + \frac{y_2^2}{(2/\sqrt{3})^2} = 1.
\end{aligned}$$

Entonces observe que el cambio de variable $y = C^tx$, es decir,

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

permitió reducir la ecuación de la cónica a la forma canónica de una hipérbola centrada en el origen. Además (y_1, y_2) es el vector de coordenadas en la base

$$\mathcal{B} = \{v_1, v_2\} = \{\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\}$$

de los puntos (x_1, x_2) en la hipérbola, dados por sus coordenadas en la base canónica. Así, los ejes principales son las rectas L_1 y L_2 generadas por los vectores v_1 y v_2 con ecuaciones vectoriales $(x_1, x_2) = t(1, 1)$ y $(x_1, x_2) = t(1-, 1)$, es decir, sus ecuaciones cartesianas son: $y = x$ y $y = -x$, respectivamente.

El cambio de variable corresponde a una rotación de ejes de ángulo $\theta = \frac{\pi}{4}$, el ángulo entre v_1 y e_1.

Las pendientes de las asíntotas son dadas por $m = \pm\frac{2/\sqrt{3}}{2} = \pm\frac{1}{\sqrt{3}}$, luego las ecuaciones de las asíntotas referidas a coordenadas de los ejes principales son: $y_2 = \frac{1}{\sqrt{3}}y_1$ y $y_2 = \frac{-1}{\sqrt{3}}y_1$. Observe que el cambio de variable $x = Cy$, permite conocer las coordenadas en términos de la base canónica, de los vectores que generan las asíntotas.

En la figura 11.5, el primer gráfico muestra los vectores v_1 y v_2 que generan los nuevos ejes (ejes principales), y en el segundo se muestra la hipérbola, trazada respecto de estos nuevos ejes, con indicación de las rectas asíntotas.

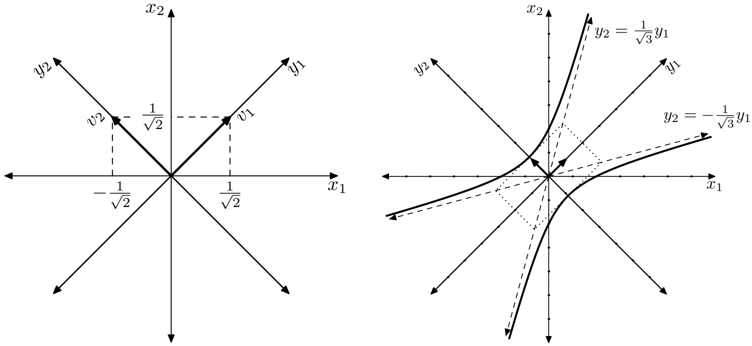


Figura 11.5: Hipérbola rotada $x_1^2 - 4x_1x_2 + x_2^2 - 4 = 0$

11.5.3 Superficies cuadráticas usuales

El problema de la rotación de superficies se resuelve en modo similar a las cónicas. Se ilustra este procedimiento en el ejemplo 11.15, luego se presentan las ecuaciones canónicas de algunas de las superficies cuadráticas más comunes y que se estudian normalmente en los cursos de Cálculo Superior.

Ejemplo 11.15 Considere la superficie determinada por la ecuación

$$-4x_1^2 - 4x_2^2 - 7x_3^2 + 4x_1x_2 + 10x_1x_3 + 10x_2x_3 = 3$$

y calcule su ecuación canónica, así como las ecuaciones de cambio de variables. Dé las ecuaciones vectoriales de los nuevos ejes.

Solución: La ecuación de la superficie escrita en forma matricial es $x^tAx = 3$, con

$$A = \begin{pmatrix} -4 & 2 & 5 \\ 2 & -4 & 5 \\ 5 & 5 & -7 \end{pmatrix} \text{ y } x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

El polinomio característico de A factorizado es

$$p(\lambda) = (3 - \lambda)(6 + \lambda)(12 + \lambda)$$

y los valores propios son $\lambda_1 = 3$, $\lambda_2 = -6$ y $\lambda_3 = -12$. Por lo tanto la ecuación de la superficie referida a los nuevos ejes es

$$3y_1^2 - 6y_2^2 - 12y_3^2 = 3.$$

La ecuación canónica será entonces:

$$y_1^2 - \frac{y_2^2}{(\frac{1}{\sqrt{2}})^2} - \frac{y_3^2}{(\frac{1}{2})^2} = 1.$$

El gráfico de esta figura es como el de figura 11.8 en la página 359, llamada hiperboloide de dos hojas.

La ecuación de cambio de variables es $y = C^t x$ con $y = (y_1, y_2, y_3)^t$ y $C = (v_1, v_2, v_3)$ la matriz cuyas columnas son los vectores propios de A ortonormados, v_1, v_2, v_3, asociados a λ_1, λ_2 y λ_3 respectivamente:

$$v_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \; v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix} \text{ y } v_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\-1\\2\end{pmatrix}.$$

Haciendo la multiplicación $C^t x$ se obtiene $y = (y_1, y_2, y_3)^t$ donde:

$$\begin{array}{rcl} y_1 & = & \frac{1}{3}\sqrt{3}\,(x_1 + x_2 + x_3) \\ y_2 & = & \frac{1}{2}\sqrt{2}\,(-x_1 + x_2) \\ y_3 & = & \frac{1}{3}\sqrt{6}\left(-\frac{1}{2}x_1 - \frac{1}{2}x_2 + x_3\right). \end{array}$$

■

Ecuaciones canónicas de algunas superficies

En las siguientes ecuaciones canónicas de superficies cuadráticas, los parámetros α, β y γ se consideran no nulos.

Elipsoide : El elipsoide centrado en el punto (x_0, y_0, z_0), no rotado, es la superficie cuadrática definida por la ecuación

$$\frac{(x - x_0)^2}{\alpha^2} + \frac{(y - y_0)^2}{\beta^2} + \frac{(z - z_0)^2}{\gamma^2} = 1.$$

llamada ecuación canónica del elipsoide. La Figura 6 ilustra el gráfico de un elipsoide.

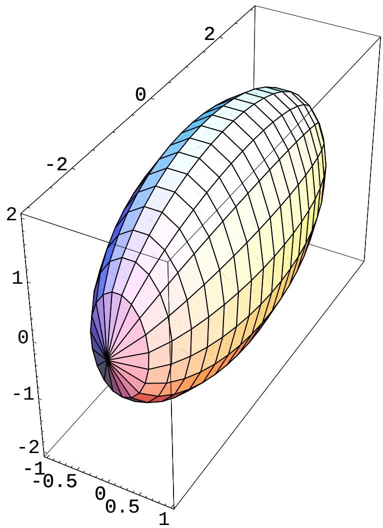


Figura 11.6: Elipsoide centrado de ecuación $x^2 + \frac{y^2}{9} + \frac{z^2}{4} = 1$.

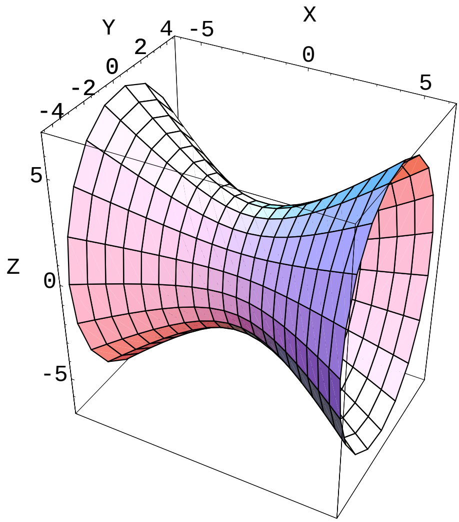


Figura 11.7: hiperboloide de una hoja: $-\frac{x^2}{9} + \frac{y^2}{4} + \frac{z^2}{9} = 1$.

Hiperboloide de una hoja : El hiperboloide de una hoja, centrado en el punto (x_0, y_0, z_0) , no rotado, es la superficie cuadrática definida por la ecuación canónica

$$\frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2} - \frac{(z-z_0)^2}{\gamma^2} = 1.$$

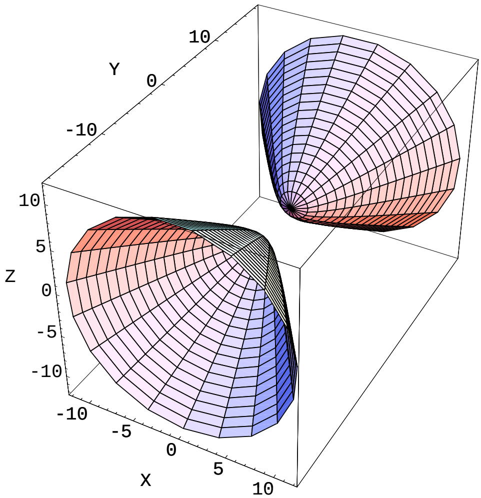


Figura 11.8: hiperboloide de dos hojas: $-\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{4} = 1.$

Un único signo menos puede afectar a cualquiera de términos de la ecuación y determina el eje sobre el cual se extiende y abre el hiperboloide.

La Figura 11.7 ilustra el gráfico de un hiperboloide de una hoja.

Hiperboloide de dos hoja : El hiperboloide de dos hojas, centrado en el punto (x_0, y_0, z_0) , no rotado, tiene una ecuación canónica como el de una hoja, pero con dos signos menos

$$\frac{(x-x_0)^2}{\alpha^2} - \frac{(y-y_0)^2}{\beta^2} - \frac{(z-z_0)^2}{\gamma^2} = 1.$$

En este caso, el único signo más (o positivo) afecta a cualquiera de términos de la ecuación y determina el eje sobre el cual se extiende y abre el hiperboloide de dos hojas.

La Figura 11.8 ilustra el gráfico de un hiperboloide de dos hojas.

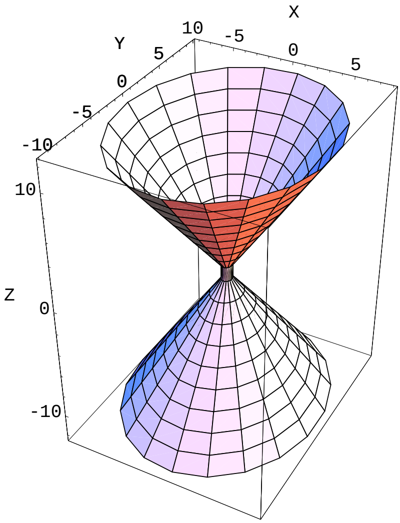


Figura 11.9: cono elíptico de ecuación $\frac{z^2}{9} = \frac{x^2}{4} + \frac{y^2}{6}$.

Cono elíptico : El cono elíptico, centrado en el punto (x_0, y_0, z_0), no rotado, es la superficie cuadrática definida por la ecuación canónica

$$\frac{(x - x_0)^2}{\alpha^2} + \frac{(y - y_0)^2}{\beta^2} - (z - z_0)^2 = 0.$$

Nótese que las superfices definidas por

$$\frac{(x - x_0)^2}{\alpha^2} + \frac{(z - z_0)^2}{\beta^2} - (y - y_0)^2 = 0$$

$$\text{y} \quad \frac{(z - z_0)^2}{\alpha^2} + \frac{(y - y_0)^2}{\beta^2} - (x - x_0)^2 = 0$$

son también conos elípticos.

La Figura 11.9 ilustra el gráfico de un cono elíptico.

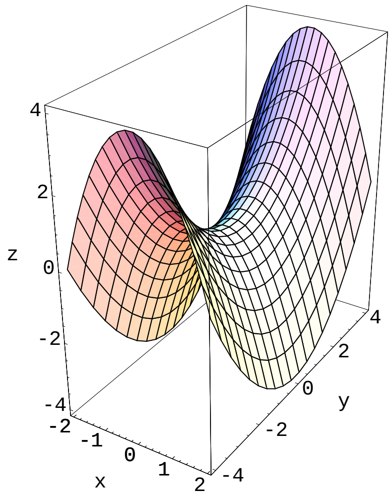


Figura 11.10: Paraboloide hiperbólico centrado: $z = -x^2 + y^2/4$.

Paraboloides hiperbólico y elíptico:

El **paraboloide hiperbólico**, centrado en el punto (x_0, y_0, z_0), no rotado, es la superficie cuadrática definida por la ecuación canónica

$$\pm\frac{(y-y_0)^2}{\alpha^2} \mp \frac{(x-x_0)^2}{\beta^2} = (z-z_0)\,.$$

La Figura 11.10 ilustra el gráfico de un paraboloide hiperbólico.

Muy similarmente, el **paraboloide elíptico**, centrado en el punto (x_0, y_0, z_0), no rotado, tiene como ecuación canónica

$$\pm\frac{(y-y_0)^2}{\alpha^2} \pm \frac{(x-x_0)^2}{\beta^2} = (z-z_0)\,.$$

La Figura 11.11 ilustra el gráfico de un paraboloide elíptico centrado.

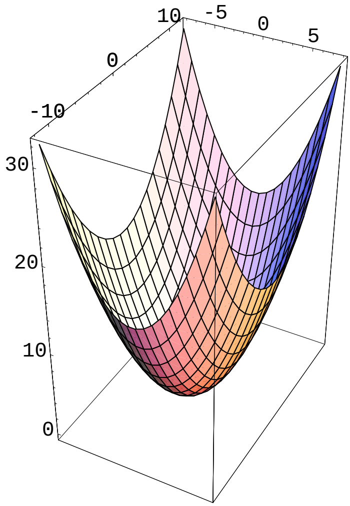


Figura 11.11: paraboloide elíptico centrado: $z = \frac{x^2}{4} + \frac{y^2}{9}$.

11.6 Ejercicios

1. Sea $A = \begin{pmatrix} 3 & -2 & 3 \\ 1 & 6 & -3 \\ 1 & 2 & 1 \end{pmatrix}$.

a) Calcule los valores propios de A.

b) Halle una base de $I\!R^3$ formada de vectores propios de A.

c) Calcule dos matrices C y D (diagonal) tales que $C^{-1}AC = D$.

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

a) Calcule una base de cada espacio propio de A.

b) Calcule una base ortonormal de $I\!R^3$ formada de vectores propios de A.

c) Obtenga una matriz ortogonal P tal P^tAP es diagonal.

3. Sea $A = \begin{pmatrix} 2 & a & 0 & 0 \\ 0 & 2 & b & 0 \\ 0 & 0 & 2 & c \\ 0 & 0 & 0 & 2 \end{pmatrix}$ con $a, b, c \in I\!R$.

a) Verifique que $\forall\ a, b, c \in I\!R$, 2 es un valor propio de A con multiplicidad algebraica 4.

b) Obtenga condiciones sobre a, b, c para que la multiplicidad geométrica de 2 sea 1, es decir, $\dim(V_2) = 1$. Repita este ejercicio para la condición $\dim(V_2) = 2$.

c) Obtenga condiciones sobre a, b, c para que A sea diagonalizable.

4. Hallar una matriz que tiene valores propios 1, -1, -1 y vectores propios respectivos $(1, 1, -1)$, $(0, 1, 1)$ y $(1, 0, -1)$.

5. Sean las matrices

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

a) Calcule el polinomio característico de A y B. Factorícelos en factores lineales y diga cual es la multiplicidad algebraica de cada valor propio.

b) Obtenga una base para cada espacio propio.

c) Defina un operador T tal que $[T]_C = A$ y $[T]_B$ es diagonal, donde C es la base canónica de $I\!R^3$ y B es una base que debe determinar a partir de (b).

6. Sea $A = \begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$ con $a \in I\!R$, $a \neq 0$.

a) Compruebe que $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^t$ es un vector propio de A. Hallar el valor propio λ asociado. Calcule una base de V_λ.

b) Compruebe que $1 - a$ es un valor propio de A y calcule una base del espacio propio correspondiente.

c) Compruebe que todo vector $x \in I\!R^3$ se puede escribir como $x = y + z$ donde $y \in V_\lambda$ y $z \in V_{1-a}$ y que $V_\lambda \cap V_{1-a} = \{0\}$.

7. Sea la matriz $C = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

a) Hacer lo mismo que en (a) y (b) del ejercicio 1.

b) Si existe una base de $I\!R^4$ formada por vectores propios de C, obténgala. Si no existe, justifíquelo.

8. Determine los valores propios de A y sus respectivos espacios propios y decida si A es o no diagonalizable. Justifique su respuesta.

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

9. Sea el operador $T : I\!R^3 \longrightarrow I\!R^3$ definido por $T(x, y, z) = (x + z, y - z, x - y)$.

a) Calcule el polinomo característico de T y factorícelo. Halle una base $\mathcal{B}$ de $I\!R^3$ tal que $[T]_\mathcal{B}$ sea diagonal. Es decir, $\mathcal{B}$ está formada por vectores propios de T.

b) Calcule una matriz H ortogonal tal que $H^t\ [T]_C\, H = D$ es diagonal, donde C es la base canónica de $I\!R^3$.

10. Sean $\lambda_1 \neq \lambda_2$ dos valores propios de un operador T de $I\!R^n$. Pruebe que

a) $V_{\lambda_1} \cap V_{\lambda_2} = \{0\}$.

b) $V_{\lambda_1} \perp V_{\lambda_2}$ siempre que $[T]_B$ sea simétrica, donde B es una base de $I\!R^n$. Ayuda: note que $(Ax)^t y = x^t(Ay)$ con A una matriz simétrica; x, y son vectores columna.

11. Escriba el polinomio característico de una matriz triangular. Haga lo mismo para una matriz diagonal y la matriz identidad en tanto que casos particulares de matrices triangulares.

12. Sea λ un valor propio de la matriz A, asociado a x. Compruebe que :

a) Para todo $r \in I\!N$, λ^r es un valor propio de A^r, asociado a x.

b) Si $p(x) = a_0 + a_1 x + \cdots + a_m x^m$ con $a_i \in R$, entonces $p(\lambda)$ es un valor propio de $p(A) = a_0 I + a_1 A + \cdots + a_m A^m$ asociado a x.

c) Si A es invertible, entonces $\frac{1}{\lambda}$ es un valor propio de A^{-1} asociado a x.

13. Sea A una matriz $n \times n$ con entradas en $I\!R$. Compruebe que

a) A y A^t tienen el mismo polinomio característico. Indicación : note que $\det(A - \lambda I)^t = \det(A - \lambda I)$.

b) $C^{-1}AC$ y A tienen igual polinomio característico.

14. Sean A y B dos matrices $n \times n$, con entradas en $I\!R$. Pruebe que si λ es un valor propio de AB también lo es de BA. Ayuda: distinga los casos $\lambda \neq 0$ y $\lambda = 0$. Al tratar este último, considere el polinomio característico de AB

15. Sea A una matriz tal que $A^2 = A$. Pruebe que si λ es un valor propio de A entonces λ^res valor propio de A, para todo $r \in I\!N$. Deduzca entonces que $\lambda = 0$ o $\lambda = 1$. Indicación: recuerde que la matriz A siendo $n \times n$, no puede tener más de n valores propios.

16. Sea $T : I\!R^3 \longrightarrow I\!R^3$ definido por

$$T(x, y, z) = (5x + 4y + z, 4x + 8y - 4z, x - 4y + 5z).$$

a) Determine si T es ortogonalmente diagonalizable. Justifique.

b) Calcule una base del núcleo de T.

c) Sin hacer más cálculos diga si $\lambda = 0$ es un valor propio de T. Justifique su respuesta.

17. Sea $\{v_1,\ v_2,\ v_3\}$ una base ortonormal de $I\!R^3$ y $S = \mathcal{C}\ell\{v_1, v_2\}$. Considere la transformación lineal $T : I\!R^3 \longrightarrow I\!R^3$ tal que:

$$T(x) = 2\text{Proy}_S x - x,$$

a) Muestre que 1 y -1 son valores propios de T.

b) Determine bases para los espacios característicos $V_{\lambda=1}$ y $V_{\lambda=-1}$.

c) ¿Es T diagonalizable ortogonalmente? Justifique su respuesta y si es afirmativa, encuentre una base $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ sea diagonal.

18. Sea $A = (v_1 v_2 v_3 v_4)$ una matriz en $M(4, I\!R)$, cuyas columnas son los vectores $v_1,\ v_2,\ v_3$ y v_4. Y suponga que

$$\text{Nuc}\,(A) = \{(t, -s, 0, 2s)^t | t, s \in I\!R\}.$$

Especifique si las siguientes afirmaciones son ciertas o falsas, o la información que se da no es suficiente para decidir. En cada caso justifique su respuesta.

a) Las columnas de A forman un conjunto de vectores l.i.

b) $(0, 0, 0, 0)^t$ es un vector propio de A.

c) La dimensión de espacio característico asociado a $\lambda = 0$ es dos.

d) Dado que $(1, 0, 0, 0)^t \in \text{Nuc}\,(A)$ entonces $v_1 = (0, 0, 0, 0)^t$.

e) La transformación lineal $T(x) = Ax$ es sobreyectiva.

19. En cada uno de los siguientes casos, obtenga las ecuaciones vectoriales de los ejes principales de la sección cónica y su ecuación canónica. Calcule el ángulo de rotación y trace un gráfico de la sección cónica (incluya las asíntotas).

a) $2x^2 - 3xy - 2y^2 + 10 = 0$.

b) $7x_1^2 + 32x_1x_2 - 17x_2^2 - 50 = 0$.

c) $x_1^2 - 2x_2^2 + 2x_2 = 0$.

d) $16x^2 - 24xy + 9y^2 = 25$.

e) $3x_1^2 - 2x_1x_2 + 3x_2^2 + 4\sqrt{2}x_1 = 6$.

f) $2x^2 + 4xy + 5y^2 - \frac{28}{\sqrt{5}}x - \frac{46}{\sqrt{5}}y = 11$.

g) $x^2 - 2xy - y^2 = 10$.

h) $6\sqrt{3}x_1x_2 - 6x_1^2 + 3x_1 - 3\sqrt{3}x_2 = 24$

20. Después del cambio de variable (una rotación) la ecuación de una cónica es $\frac{-y_1^2}{2} + y_2^2 + y_2 = 0$.

(a) Obtenga la ecuación canónica e identifique la cónica.

(b) Si el ángulo de rotación es $\frac{\pi}{3}$ calcule una ecuación vectorial de cada uno de los ejes principales.

(c) Trace la gráfica de la curva.

(d) Escriba la ecuación de la cónica con término x_1x_2.

21. Considere la cónica de ecuación

$$3x_1^2 - 8x_1x_2 - 3x_2^2 - 4\sqrt{5}x_1 + 2\sqrt{5}x_2 = 15.$$

Si $x = (x_1, x_2)^t$, y $y = (y_1, y_2)^t$:

a) Diagonalice la matriz A apropiada y dé la matriz P del cambio de variable $y = P^t x$, que elimina los términos cruzados en la ecuación anterior y además corresponde a una rotación.

b) Transforme la ecuación dada hasta obtener la ecuación canónica de la respectiva cónica.

c) En el sistema de coordenadas de las variables x_1 y x_2, trace el gráfico aproximado de la cónica, señalando las coordenadas de los vectores que generan los ejes y_1 y y_2, y mostrando las asíntotas (si es el caso).

22. Considere la superficie de ecuación

$$-x^2 + 5y^2 - z^2 - 2xy + 10xz - 2yz = 12.$$

Si A es la matriz simétrica, correspondiente a la forma cuadrática de esta ecuación, se conoce que:

$$|A - \lambda I| = -(\lambda - 3)(\lambda - 6)(\lambda + 6)$$

Además, $V_{\lambda=3} = \mathcal{C}\ell\{(1,1,1)^t\}$ y $V_{\lambda=6} = \mathcal{C}\ell\{(1,-2,1)^t\}$

a) Determine la matriz A y $V_{\lambda=-6}$.

b) Dé una matriz P para el cambio de variable $\mathbf{y} = P^t\mathbf{x}$, que elimina los términos cruzados en la ecuación anterior.

c) Obtenga la ecuación canónica de esta superficie.

d) Si las nuevas variables se denominan x', y', y z', determine el coseno del ángulo entre los ejes correspondientes a las variables x y x'.

23. Sea la superficie

$$x_1^2+x_2^2+x_3^2+2ax_1x_2+2bx_2x_3+2x_1x_3+\sqrt{3}x_1-\sqrt{3}x_2+\sqrt{3}x_3 = 0.$$

(a) Escriba la ecuación en la forma $x^tAx + Bx + d = 0$. ¿Si $P = (1, 1, 0)^t$, para cuáles valores de a y b, P es un punto de la superficie?.

(b) Sea $\mathcal{B} = \{v_1, v_2, v_3\}$ la base ortonormal de $I\!R^3$ que determina los ejes principales, y suponga que $a = b = -1$. Determine la base $\mathcal{B}$ y el cambio de variable que elimina los términos "cruzados" de la ecuación. Obtenga la ecuación sin términos cruzados.

(c) Determine $[P]_{\mathcal{B}}$. ¿$[P]_{\mathcal{B}}$ satisface la ecuación sin términos cruzados?

24. Considere las siguientes ecuaciones de superficies cuadráticas:

a) $x_1x_2 = x_3$.
b) $-3y^2 + 6xy + 6yz = 12$.

i) Diagonalice la matriz A apropiada y dé una matriz P para el cambio de variable $y = P^tx$, que elimina los términos cruzados en la ecuación.

ii) Transforme la ecuación dada hasta obtener la ecuación canónica de la respectiva superficie e identifíquela.

iii) Halle una ecuación vectorial para cada uno de los ejes rotados que denominamos L_i, $i = 1, 2, 3$.

iv) Sean $E_1,\ E_2,\ E_3$ los ejes determinados por la base canónica, encuentre el ángulo entre E_i y L_i para $i = 1, 2, 3$.

25. Considere la superficie de ecuación

$$-4x_1^2 + 2x_2^2 + 4x_3^2 + 6x_1x_3 = 20.$$

Si A es la matriz simétrica, correspondiente a la forma cuadrática de esta ecuación, y se conoce que:

$$A = PDP^t$$

donde $P = \begin{pmatrix} \frac{1}{\sqrt{10}} & 0 & \frac{-3}{\sqrt{10}} \\ 0 & 1 & 0 \\ \frac{3}{\sqrt{10}} & 0 & \frac{1}{\sqrt{10}} \end{pmatrix}$ y $D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{pmatrix}$.

a) Dé la matriz A e indique el cambio de variable que elimina el término x_1x_3 de la anterior ecuación.

b) Transforme la ecuación dada hasta obtener la respectiva forma canónica.

c) Si las nuevas variables se denominan y_1, y_2, y y_3, determine el coseno del ángulo entre los ejes correspondientes a las variables x_3 y y_3.

26. Sea A una matriz 2×2, simétrica y con valores propios λ_1 y λ_2. Compruebe que la sección cónica de ecuación $f(z) = z^t Az + d = 0$ con $d > 0$, es:

(a) Una elipse si $\lambda_1 \lambda_2 > 0$.

(b) Una hipérbola si $\lambda_1 \lambda_2 < 0$.

Analice el caso $d < 0$.

27. Halle una regla en términos de los valores propios, que caracterice el cono elíptico. Demuéstrela.

28. Sea $A = (a_{ij})_{n \times n}$ tal que $\sum_{i=1}^{n} a_{ij} = 1$ para todo j. Compruebe que 1 es valor propio de A.

Apéndice A

Exámenes

En este apéndice se presentan los exámenes aplicados en el curso MA-1004 Algebra Lineal, de la Escuela de Matemática en la Universidad de Costa Rica, correspondientes al I y II ciclo lectivo de 1996 y al I y II ciclo de 1997.

Salvo unas pocas excepciones, en cada ciclo, en el primer examen se evalúa la materia de los capítulos 1, 2, 3 y 4. El segundo parcial corresponde con los capítulos 5, 6, 7 y 8 y los restantes temas se evalúan en el tercer parcial.

A.1 Exámenes Parciales I

A.1.1 I ciclo lectivo de 1996

Universidad de Costa Rica	**I Ciclo 1996**
Escuela de Matemáticas	**MA-1004 Algebra Lineal**

PRIMER EXAMEN PARCIAL

PARTE I

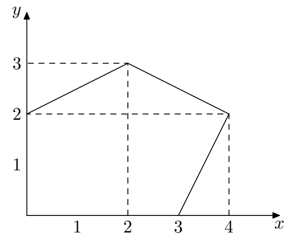
Marque con una x la única alternativa correcta

1. La región marcada es el conjunto de soluciones factibles de un programa lineal cuya función a maximizar es:

$$Z = -3x + 4y.$$

Entonces:

1. El máximo de Z es 6
2. Z no alcanza el máximo
3. El máximo de Z es 10
4. El máximo de Z es 8



2. El sistema $Ax = b$ (con b diferente de cero y A una matriz $n \times n$), tiene solución única. Entonces necesariamente se tiene:

1. b es solución del sistema
2. A es no invertible
3. La matriz aumentada del sistema es invertible
4. El determinate de A es diferente de cero.

3. Sea $a \in \mathbb{R}$ y el sistema:

$$\begin{array}{rcl} x - ay & = & 1 \\ 2x + y & = & 3 \end{array}$$

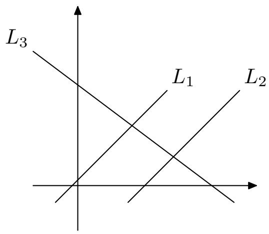
Entonces:

1. Para cualquier valor de a, el sistema no tiene solución.
2. Para cualquier valor de a, el sistema tiene solución única.
3. Si, $a \neq \frac{-1}{2}$, el sistema tiene solución única.
4. Si $a = \frac{-1}{2}$, el sistema tiene solución única.

4. Sea A una matriz $n \times n$ tal que $A^3 = A$. Entonces necesariamente se tiene que:

1. Hay sufiente información para determinar el valor A.
2. El único valor para $\det(A)$ es 1.
3. El único valor para $\det(A)$ es 0.
4. El $\det(A)$ solo puede ser 0, 1 ó -1.
5. Ninguna de las anteriores.

5. Las rectas L_1, L_2 y L_3 corresponden a un sistema no homogéneo de ecuaciones 3×2.



Si L_1 y L_2 tienen igual pendiente, entonces:

1. La solución del sistema es única.
2. El gráfico no provee suficiente información para concluir acerca de la solución del sistema.
3. El sistema no tiene solución.
4. Hay infinitas soluciones.
5. Ninguna de las anteriores

6. Sea $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & b \\ 0 & a & a+b \end{pmatrix}$ y B la matriz escalonada reducida equivalente a A. Entonces

1. Si $b \neq 0$, $B = I_3$
2. Si $a = 0$, $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$
3. Si $a \neq 0$,$B = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{pmatrix}$
4. Si $ab = 0$, B tiene al menos una fila de ceros.

7. Si $A \in M(n, I\!R)$ y tanto A como A^t son matrices triangulares superiores, entonces se cumple que:

1. A es invertible
2. A es la matriz identidad

3. A es simétrica
4. A es definida positiva

PARTE II

Los siguientes ejercicios son de solución rápida y corta, por lo que solo nos interesa su respuesta (no los cálculos), la cual debe anotar en el espacio correspondiente.

8. Sea A una matriz de 3×3. En cada caso marque:

- Con una [**S**] si A es invertible.
- Con un [**N**] si A no es invertible.
- Con una [**Q**] si no es posible concluir que A sea o no invertible.

1. [ ] A es equivalente a una matriz escalonada reducida.
2. [ ] $\det(A) = -3$
3. [ ] $A^t A = I_3$
4. [ ] $E_1 E_2 E_3 A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$ Con E_1, E_2, E_3 matrices elementales.
5. [ ] $\det(A) = 0$
6. [ ] $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix}$

9. Sea $A \in M(n, I\!R)$ con $n > 1$,A invertible. Calcule:

1. $\dfrac{|A^t A (A^t)^{-1}|}{|A|} =$ donde $|A| = \det(A)$
2. $\det(\det(A) I_n) =$ con $\det(A) = 3$

10. Sea $A = (A_1, A_2, A_3, A_4)$ una matriz de 4×4 donde A_1, A_2, A_3, A_4 son las columnas de A y $\det A = 4$ Calcule:

1. $\det(A_2, A_1, A_4, A_3)$=
2. $\det(3A_1, 2A_2 + A_3, A_3, A_4)$=

3. $\det(3A_1, A_2, 2A_2 + A_3, A_4)$=
4. Si $b = A_1 + 2A_2 + A_4$ La solución del sistema $Ax = b$ es $x =$

PARTE III

Resuelva en su cuaderno los siguientes ejercicios:

11. Para enviar un mensaje codificado de 4 letras, se remplaza cada letra por el número de su posición en el alfabeto, y cada número es colocado en una matriz M. Por ejemplo la palabra **CABO** viene dada por la matriz:

$M = \begin{pmatrix} 3 & 1 \\ 2 & 16 \end{pmatrix}$. Para evitar la detección del mensaje se envía codificado en la forma MC, donde

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

Si una persona recibe la matriz $\begin{pmatrix} 32 & 8 \\ 25 & -11 \end{pmatrix}$. Diga cuál es el mensaje.

A 1	B 2	C 3	D 4	E 5	F 6	G 7	H 8	I 9
J 10	K 11	L 12	M 13	N 14	 15	O 16	P 17	Q 18
R 19	S 20	T 21	U 22	V 23	W 24	X 25	Y 26	Z 27

12. Resuelva el problema de programación lineal siguiente:

Minimizar $z = x_1 + x_2 + 4x_3 + 7x_4$.

Sujeto a las restricciones:

$$\begin{array}{rcl} x_1 + x_2 + 5x_3 + 2x_4 & = & 8 \\ 2x_1 + x_2 + 8x_3 & = & 14 \\ x_i \geq 0 & & i = 1, 2, 3, 4. \end{array}$$

13. Si $A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$. Encuentre el conjunto solución de la ecuación $Ax = x$ con $x \in I\!R^2$.

A.1.2 II ciclo lectivo de 1996

Universidad de Costa Rica **II Ciclo 1996**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

PRIMER EXAMEN PARCIAL

PARTE I

Responda las preguntas en el espacio correspondiente. Al final de cada una hay un espacio para que, tan breve como le sea posible, ponga los argumentos que justifican su respuesta.

1. Sea A una matriz de 3×3, $B = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 2 & a \\ 0 & 0 & 1 \end{pmatrix}$, $\det(AB) = 4$. Entonces

$\det(A) =$ __________ $\qquad \det(A^{-1}) =$ __________

$\det(3A)^{-1} =$ __________ $\qquad \operatorname{Rng}(A) =$ __________

Justificación

2. Sea $A = \begin{pmatrix} 1 & 0 \\ 4 & 2 \end{pmatrix}$. Si E_1, E_2, E_3, E_4 son matrices elementales tales que $E_2E_1A = I$ y $E_3E_4 = A$. Entonces

$$E_1 = \begin{pmatrix} & \\ & \end{pmatrix} \qquad E_2 = \begin{pmatrix} & \\ & \end{pmatrix}$$

$$E_3 = \begin{pmatrix} & \\ & \end{pmatrix} \qquad E_4 = \begin{pmatrix} & \\ & \end{pmatrix}$$

Justificación

3. Sea A una matriz de 3×3, **escalonada, de rango 2**.

1. El sistema $Ax = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ tiene (solución: única, ninguna, infinitas) ____________
2. El sistema $Ax = 0$ tiene (solución: única, ninguna, infinitas) ____________

Justificación

4. Sea A una matriz de 3×3, cuyas columnas: A_1, A_2, A_3 son **linealmente independientes** y verifican :

$$3A_1 - 2A_2 + A_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Entonces

1. El conjunto solución del sistema $Ax = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ es $S =$ ____________
2. La primera columna de A^{-1} es $C_1 =$ ____________

Justificación

5. La función objetivo del programa lineal:
Min $z = x_1 + 2x_3 - 4$
Sujeto a las restricciones:

$$-x_1 + x_2 + x_3 = 1, \;\; 2x_1 + x_4 = 2, \;\; x_i \geq 0;$$

alcanza su valor mínimo en $x = (\quad, \quad, \quad, \quad)$ el cual es $z =$ ____________

Justificación

PARTE II: Desarrollo

6. Sea $P \in M(n, I\!R)$, P es una matriz ortogonal si $PP^t = I_n$. Demuestre que si $x \in I\!R^n,\ x \neq 0$

$$P = I_n - \frac{2}{x^t x} x x^t$$

es una matriz simétrica y ortogonal.(Justifique en detalle cada paso)

7. Sea $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

1. Encuentre los valores de λ que satisfacen la ecuación:

$$\det(\lambda I - A) = 0$$

2. Escriba el sistema $Ax = 3x$ como un sistema homogéneo. **Sin hacer cálculos** diga cuántas soluciones tiene este sistema y de los argumento que justifican su afirmación.
3. Encuentre el conjunto solución del sistema anterior.

8. Plantee y resuelva (Por cualquier método) el siguiente problema.

Una compaía distribuidora vende automóviles y camiones,por los que obtiene 10 um(Unidad monetaria equivalente a 50000 colones) y 20 um en cada camión.

El fabricante puede proveer un máximo de 30 automóviles y 26 camiones por semana.

Tanto automóviles como camiones requieren de un tiempo de preparación en el taller, para revisión e instalación de ciertos componentes. Cada automóvil requiere, en promedio, de 2 horas de taller y cada camión 3 hrs. El distribuidor dispone de 96 hrs de tiempo taller.

¿Cuántos automóviles y cuántos camiones debe ordenar semanalmente el distribuidor para maximizar su ganancia.? (no deben sobrar).

A.1.3 I ciclo lectivo 1997

Universidad de Costa Rica **I Ciclo 1997**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

PRIMER EXAMEN PARCIAL

1. Considere el programa lineal max $z = 3x_1 + 2x_2 + 20$ sujeto a

$$\begin{aligned} 6x_1 + 4x_2 &\le 120 \\ 3x_1 + 10x_2 &\le 180 \\ x_1, x_2 &\ge 0 \end{aligned}$$

a) (10 pts.) Grafique la región de soluciones factibles.

b) (5 pts.) Calcule el valor máximo de la función objetivo z.

c) (10 pts.) Dé tres puntos donde la función z alcanza su valor máximo.

2. a) (5 pts.) Exprese

$$(x, y, z)\begin{pmatrix} 1 & 1 & \alpha \\ 1 & 1 & \beta \\ \alpha & \beta & 1 \end{pmatrix} = (0, 0, 0)$$

como un sistema de ecuaciones lineales en las incógnitas x, y, z.

b) (10 pts.) Encuentre la relación entre α y β para que el sistema tenga solución única. Obtenga dicha solución.

c) (10 pts.) Si $\alpha = \beta$, encuente el rango de la matriz asociada.

3. (15 pts.) Sea $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \\ -1 & -3 & 3 \end{pmatrix}$. Exprese A como el producto de una matriz invertible y una matriz triangular.

4. Definición: Se dice que una matriz X es antisimétrica si $X^t = -X$. Sea A una matriz $n \times n$:

a) (5 pts.) Deduzca que $S = A + A^t$ es simétrica y que $R = A - A^t$ es antisimétrica.

b) (5 pts.) Compruebe que $A = \frac{S+R}{2}$ y exprese $X = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$ como la suma de una matriz simétrica y una antisimétrica.

5. Sean A y B matrices cuadradas de $n \times n$ y $\det(B) = 4$.

a) (10 pts.) Si $\det(AB) = 8$. Calcule $\det(A^{-1})$ y $\det(C^{-1}ACB)$ donde C es una matriz invertible.

b) (5 pts.) Si $S^t = -S$ y n es impar. Pruebe que S no es invertible.

c) (10 pts.) Si Q es una matriz ortogonal, $(Q^{-1} = Q^t)$, deduzca que $\det(Q) = \pm 1$ y calcule $\det\left[\left(5BQ^2\right)^{-1}\right]$

A.1.4 II ciclo lectivo de 1997

Universidad de Costa Rica **II ciclo 1997**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

Examen Parcial I

INSTRUCCIONES: Justifique cada una de sus respuestas

1. (16 Pts) Considere la matriz $A = \begin{pmatrix} 2 & 5 \\ -3 & 8 \\ 4 & -1 \end{pmatrix}$

a) Sea $u = (4, -6, 8)^t$ ¿Se puede expresar el vector u como combinación lineal de los vetores columna? .

b) Pruebe que el vector $(5, 8, -1)^t$ no se puede expresar como combinación lineal de los vectores $(2, -3, 4)^t$ y $(4, -6, 8)^t$.

c) ¿Es el conjunto de vectores $\{(2, -3, 4), (4, -6, 8), (5, 8, -1)\}$ linealmente independiente?.

d) ¿Para cuál o cuáles valores de a, el vector $v = (1, a, 2a-1)^t$ es combinación lineal de esos vectores columna de A?

2. (18 Pts) Analice para cuáles valores de α y β, el sistema:

$$\left\{ \begin{array}{rcl} 4x_1 - 6x_2 + 2x_3 & = & \alpha \\ -6x_1 + 9x_2 + \beta x_3 & = & 1 \end{array} \right.$$

(a) No tiene solución.

(b) Tiene infinitas soluciones.

(c) Tiene solución única.

3. (16 Pts) Suponga que B es la matriz 4×3 cuyas columnas son b_1, b_2, b_3. Sea A la matriz 4×4 de columnas $b_1, b_2, b_3, b_1 + b_2$.

(a) Suponiendo $\rho(B) = 3$ (rango de B), obtenga el conjunto solución del sistema

$$Ax = 0 \text{ con } x \in I\!R^4.$$

(b) Determine la nulidad de A.

4. (16 Pts) Maximice $z = 4x_1 + 2x_2 + 5x_3$ sujeto a las restricciones

$$\left\{ \begin{array}{rcl} x_1 + 2x_2 + 3x_3 & \leq & 6 \\ x_1 + x_2 + 2x_3 & \leq & 4 \\ x_i & \geq & 0 \;\; \forall\, i \end{array} \right.$$

y dar los valores de x_1, x_2, x_3 que maximizan a z.

5. (18 Pts) Si $B = P^{-1}AP$ con $A, B, P \in M(n, I\!R)$ y P invertible

(a) Muestre que $\det B = \det A$

(b) Muestre que $\det(\lambda I - B) = \det(\lambda I - A)$ con $\lambda \in I\!R$

(c) Si A es invertible y si $A^{-1} = A^t$ demuestre que $\det(2A) = 2^n$ o $\det(2A) = -2^n$.

6. (16 Pts) Si $A = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

(a) Exprese A^{-1} como producto de tres matrices elementales.

(b) Exprese A^t como producto de tres matrices elementales.

A.2 Exámenes Parciales II

A.2.1 I ciclo lectivo de 1996

Universidad de Costa Rica **I Ciclo 1996**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

EXAMEN PARCIAL II

PARTE I (7 puntos)

Marque con una x la única alternativa correcta

1. Considere las siguientes tres rectas en $I\!R^3$:

$L_1: \quad (x,y,z) = (1,0,1) + t(a,b,c)$

$L_2: \quad \dfrac{x-a}{2} = \dfrac{y-b}{-1} = \dfrac{z-c}{2}$

L_3 : definida por la intersección de los planos:

$$ax + by + cz = 6 \quad , \quad 2x - y + 2z = 0$$

Entonces independientemente de (a,b,c) se tiene que:

1. L_2 es paralela a L_3
2. L_1 y L_2 se intersecan en (a,b,c)
3. L_3 es perpendicular a L_1 y L_2
4. L_1 está contenida en L_3

2. Si S es un plano en $I\!R^3$ que contiene los puntos $(0,0,0)$ $(2,0,0)$ y $(0,0,1)$ entonces S es:

1. El plano $z = 0$
2. El plano $y = 0$
3. El plano $x = 0$
4. el plano $x - z = 0$

3. La dimensión del subespacio generado por los vectores
$(1,1,0,0)$, $(2,2,0,0)$, $(0,1,0,0)$, $(2,0,0,3)$
$(1,-2,0,8)$, $(0,0,0,0)$ es:

a) 2
b) 3
c) 4
d) 6

4. La recta $\{(0,t,t)|t \in I\!R\}$:

1. Está en el plano $6x + 4y - 5z = 0$.
2. Es paralela al plano $6x + 4y - 5z = 0$.
3. Es paralela al plano $5x - 3y + 3z = 1$.
4. Es perpendicular al plano $5x - 3y + 3z = 1$.

5. Sean $x, y \in I\!R^n$, $||x|| = ||y||$, entonces $(x + 2y) \cdot (x - 2y)$ es igual a:

1. 0
2. $5||x||^2 - 4||x||||y||$
3. $-3||x||^2$
4. $3||x||$

6. Sean el subespacio vectorial

$$U = \{(x, y, z, w)|2x - y = z + w = 0\}$$

y los vectores $u = (1, 2, -1, -1)$, $v = (0, 0, 0, 2)$
Entonces:

1. $u \in U \;\; y \;\; v \notin U$
2. $u \notin U \;\; y \;\; v \in U$
3. $(u - v) \in U$
4. $(u + v) \in U$

7. Sean los vectores $u_1 = (1, -1, 0)$, $u_2 = (2, \alpha - 3, 1 - \alpha)$ y $u_3 = (1, 1, \alpha - 1)$. Entonces

1. Para cualquier valor de α los vectores son l.i.
2. Los vectores son l.d. únicamente para $\alpha = -1$
3. Los vectores son l.d. para cualquier α distinto de 1 y -1
4. Los vectores son l.d. para $\alpha = 1$ o $\alpha = -1$

SEGUNDA PARTE (10 puntos)

Los siguientes ejercicios son de solución rápida y corta, por lo que solo nos interesa su respuesta (no los cálculos), la cual debe anotar en el espacio correspondiente.

8. Sea $\{v_1, v_2, v_3\}$ un conjunto ortonormal de vectores de $I\!R^n$

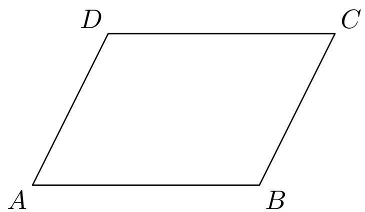
1. $||v_1 - 2v_2 - v_3||^2 =$
2. $(v_1 - v_2) \cdot (v_2 + v_3) =$

9. Sea $B = \{(1, -1, 2)\ , (1, 1, -1)\ , (0, 1, 1)\}$ una base de $I\!R^3$

1. Si las coordenadas de V en la base B son $[v]_B = (-1, 2, 1)^t$ entonces v=
2. Si $u = (1, -3, 6)$ entonces $[v]_B =$

10. Considere el paralelogramo de vértices A, B, C, D

con $A = (-1, 2, 3)$, $B = (0, 1, 5)$, $C = (7, -3, 8)$. Entonces D=



11. Sean $u, v \in I\!R^n$, $u \neq 0$, $v \neq 0$. Entonces el coseno del ángulo entre :

u y $v - \text{Proy}_u(v)$ es

y entre u y $\text{Proy}_u(v)$ es

12. El rendimiento (y) de la alfalfa (toneladas por acre) depende linealmente de la cantidad (x) de agua aplicada (pulgadas por acre) de acuerdo con la relación:

$y = 4 + kx$, que se puede expresar como $z = kx$ con $z = y - 4$.

Estimar el valor de k mediante regresión lineal, asumiendo que se hicieron 8 mediciones y que los datos expresados en los vectores $\vec{x}, \vec{y}$, $\vec{1}$ ($\vec{1}$ es el vector con sus 8 entradas iguales a 1) verifican:

$\vec{x} \cdot \vec{1} = 210$ $\quad$ $\vec{x} \cdot \vec{x} = 7308$ $\quad$ $\vec{x} \cdot \vec{y} = 1590.48$

Indicación $\vec{z} = (\vec{y} - 4\vec{1})$

III PARTE (18 puntos)

Resuelva en su cuaderno los siguientes ejercicios:

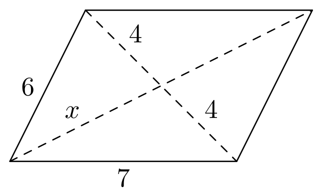
13. Sea $W = \mathcal{C}\ell\{(2,1,0),(1,1,1)\}$

1. Obtenga una base de $W^{\perp}$
2. Sea $v = (3,-7,5)$. Encuentre $a \in W$ y $b \in W^{\perp}$ tales que $v = a + b$.

14.

a) Demuestre que : $||A+B||^2 + ||A-B||^2 = 2(||A||^2 + ||B||^2)$

b) Use (a) para calcular la mediana relativa al mayor lado del triángulo de lados 6, 7, 8 cm.



15. Sea $S = \{(x,y,z)|2x + 3y - z = 0\}$ y $L = \mathcal{C}\ell\{(1,1,1)\}$

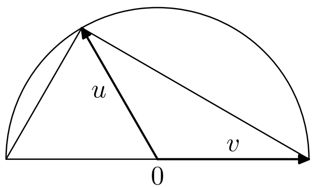
1. Demuestre que S es un subespacio de $I\!R^3$ y encuentre una base.
2. Calcule la distancia de un punto arbitrario de la recta L al plano S
3. Determine todos los puntos de la recta L que están a una distancia del plano S igual a $\sqrt{14}$.

A.2.2 II ciclo lectivo de 1996

Universidad de Costa Rica **II Ciclo 1996**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

SEGUNDO EXAMEN PARCIAL

1.
1. Sean $u, v \in I\!R^n$ tales que $||u|| = ||v||$ Demuestre que $u - v$ y $u + v$ son perpendiculares. (7 pts)
2. Demuestre que todo triángulo inscrito en un semicírculo es rectángulo. (8 pts)



2. Sea $\{v_1, v_2, v_3\}$ un conjunto **linealmente independiente** de vectores de $I\!R^3$. (15 pts)

1. Demuestre que $\mathcal{B} = \{v_1, v_1 + v_2, v_2 - v_3\}$ es una base de $I\!R^3$
2. Calcule las coordenadas de v en la base $\mathcal{B}$ para $v = v_1 + v_2 + v_3$

3. Sea Π el plano con ecuación $\quad x - 2y - z = 3 \quad$ (15 pts)

1. Decida si la recta L de ecuación: $(x, y, z) = t(0, -1, 2),\ t \in I\!R$ es paralela al plano Π
2. Obtenga la ecuación de un plano que contenga a la recta L y sea perpendicular con Π.

4. Sea la recta L definida por las siguientes ecuaciones paramétricas: (15 pts)

$$\left\{ \begin{array}{rcl} x & = & 3+2t \\ y & = & 4+t \\ z & = & 5-6t \end{array} \right. \qquad t \in I\!R$$

Encontrar el punto $Q \in L$ que esté más próximo del punto $R = (4,1,7)$

5. Sea $W = \{(x,y,z) | 2x - z = 0,\ y - 2x = 0\}$ (15 pts)

1. Demuestre que W es un subespacio de $I\!R^3$.
2. Encuentre una base para el subespacio ortogonal $W^{\perp}$
3. Trace un gráfico con W y $W^{\perp}$.
4. Encuentre $u \in W,\ v \in W^{\perp}$ tal que

$$u \ + \ v = (1,5,0)$$

6. Sean $x_1 = (1,0,1,2,1),\ \ x_2 = (0,1,1,1,2)$ vectores de $I\!R^5$. (15 pts)

y $W = \mathcal{C}\ell\{x_1,\ x_2\}$ el subespacio generado por ellos.

1. Encuentre una base $\mathcal{B}$ **ortonormal** para W
2. Sea $y = (4,5,6,5,4)$ vector de $I\!R^4$. Calcule $\widehat{y} = \text{Proy}_W(y)$ la proyección de y sobre W.

7. **Utilizando los resultados del problema 6**, resuelva el siguiente problema de regresión lineal. (15 pts)

Se ha observado que en una granja la producción de huevos y está relacionada con las cantidades de dos comidas fijas $x_1,\ x_2$ por $y = ax_1 + bx_2 + \epsilon$

1. Estime los valores de a , b de acuerdo con los siguientes datos

x_1	1	0	1	2	1
x_2	0	1	1	1	2
y	4	5	6	5	4

2. Calcule el valor de R índice de calidad de la regresión y comente.
3. Estime la producción de huevos para $x_1 = x_2 = 2$
   Nota: Ud puede resolver este problema sin utilizar los resultados del ejercicio 6, pero: Invertirá más tiempo y su puntaje máximo será 7 pts

A.2.3 I ciclo lectivo de 1997

Universidad de Costa Rica **I Ciclo 1997**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

Examen Parcial II

1. (15 Pts) Sean $A = (1, 1, 0), B = (-3, 2, 1)$ y $C = (3, -4, 4)$ los tres vértices de un triángulo:

a) Determine el vector $\overrightarrow{DA}$ que corresponde a la altura del triángulo sobre el lado $\overline{BC}$.

b) Calcule el área de este triángulo.

2. (15 Pts) Sea $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2} | a + b = 0\ ,\ c + d = 0 \right\}$

a) Muestre que W es un subespacio vectorial de $M_{2\times 2}$.

b) Determine un conjunto generador de W.

3. a) (8 Pts) Determine una base para el subespacio S de $I\!R^4$ generado por los vectores:

$$(1, 1, 2, 2),\ (3, 3, 3, 3),\ (2, 2, 1, 1) \text{ y } (1, 1, -1, -1).$$

b) (12 Pts) Calcule la distancia de $(1, 2, 0, 1)$ al subespacio S.

4. (20 Pts) Considere las rectas:

$$L_1 : \{(2 - t, 3, 1 + t) | t \in I\!R\}$$

$$L_2 : \{(5 + t, 5 + 2t, t) | t \in I\!R\}$$

a) Determine el punto R donde L_1 interseca a L_2.

b) Dé la ecuación normal de un plano que contenga a L_1 y a L_2.

c) Si $P \in L_1$ y $Q \in L_2$, $P \neq R \neq Q$. Muestre que el triángulo de vértices P, Q y R es rectángulo.

5. (15 Pts.) Sea $\mathcal{B} = \{w_1, w_2, w_3, w_4\}$ una base ortonormal de $I\!R^4$, considere los subespacios de $I\!R^4$:

$$S = \mathcal{C}\ell\{w_1, w_2\} \quad \text{y} \quad T = \mathcal{C}\ell\{w_3, w_4\}$$

a) Demuestre que S y T son subespacios ortogonales.

b) Si $x = 2w_1 - w_2 - 4w_3 + 2w_4$:

  i) Determine vectores $x_1 \in S$ y $x_2 \in T$, tales que $x = x_1 + x_2$.

  ii) Calcule $\text{Proy}_T x$.

  iii) Determine la distancia de x al subespacio S.

6. (15 Pts) Sea $B = \{v_1, v_2, v_3, v_4\}$ un conjunto ortogonal de vectores no nulos de $I\!R^5$ y sea A la matriz 5×5 cuyas filas son:

$$v_1,\ v_2,\ v_3,\ v_4,\ v_1 + v_2 + v_3.$$

Si $S = \{x \in I\!R^5 / Ax = 0\}$ y $w \in S$, entonces:

a) ¿Cuál es la dimensión de S?

b) Dé un conjunto generador de $S^\perp$.

c) Determine una base ortonormal de S.

d) Determine $\text{Proy}_S(v_1 + w)$.

e) Dé un vector ortogonal a w.

A.2.4 II ciclo lectivo de 1997

Universidad de Costa Rica **II Ciclo 1997**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

Examen Parcial II

1. (35Pts) Sean $P = (3,0,0)$ $,Q = (1,1,3)$ y $R = (0,-2,1)$ puntos en $I\!R^3$.

a) (5Pts) Determine el área del paralelogramo de lados no paralelos $\overrightarrow{PQ}$ y $\overrightarrow{PR}$.

b) (5Pts) Muestre que $\vec{n} = (1,-1,1)$ es ortogonal a $\overrightarrow{PQ}$ y $\overrightarrow{PR}$.

c) (7 Pts) Calcule el volumen del paralelepípedo de lados $\vec{n}, \overrightarrow{PQ}, \overrightarrow{PR}$.

d) (6 Pts) Sea $B = \{\overrightarrow{PQ}, \overrightarrow{PR}\}$ base del subespacio W de$I\!R^3$. Construya una base ortonormal para W.

e) (6 Pts) Calcule la proyección del vector $\vec{n}+\overrightarrow{PQ}+\overrightarrow{PR}$ sobre el subespacio W:

$$\text{Proy}_W(\vec{n} + \overrightarrow{PQ} + \overrightarrow{PR})$$

f) (6 Pts) Encuentre el subespacio $W^{\perp}$.

2. (25Pts) Sean:

$$W = \left\{(x,y,z,w) \in I\!R^4 | x = y + z - w\right\}$$

$$U = \left\{(x,y,z,w) \in I\!R^4 | w = 0, \ z = 5y - 2x\right\}$$

a) (10 Pts) Demuestre que U es subespacio de $I\!R^4$ y encuentre una base (demuestre que es base).

b) (15 Pts) Encuentre bases y las dimensiones para W y $W \cap U$. (No demuestre).

3. (25pts) Considere las rectas:

$$L_1 := \begin{cases} x = 1 - t \\ y = 2 + t \\ z = 2 - 3t \end{cases}, \ t \in I\!R \qquad L_2 := \begin{cases} x = 2 - s \\ y = 1 - s \\ z = s \end{cases}, \ s \in I\!R$$

a) (10Pts) Demuestre que L_1y L_2 no se intersecan.

b) (10Pts) Determine una ecuación del plano π que contiene a L_1 y es paralelo a L_2.

c) (5pts) Calcule la distancia entre L_1 y L_2.

4. (15Pts) Sea el subespacio de $M(2, I\!R)$

$$W = \left\{ \left(\begin{array}{cc} a & b \\ c & -a \end{array} \right) \middle| \, a, b, c \in I\!R \right\}$$

a) (5Pts) Calcule una base y la dimensión de W.

b) (5Pts) Exprese la matriz $\left(\begin{array}{cc} 1 & -2 \\ 2 & -1 \end{array} \right)$ como una combinación lineal de la base encontrada en a).

c) (5Pts) El producto punto definido en $I\!R^n$ (así como la definición de perpendiculariedad) se puede extender de manera natural al espacio de matrices como sigue:

$$\left(\begin{array}{cc} a_1 & a_2 \\ a_3 & a_4 \end{array} \right) \cdot \left(\begin{array}{cc} b_1 & b_2 \\ b_3 & b_4 \end{array} \right) = a_1 b_1 + a_2 b_2 + a_3 b_3 + a_4 b_4$$

Si $A = \left(\begin{array}{cc} 1 & -1 \\ 4 & -1 \end{array} \right)$, encuentre una matriz B perpendicular a A.

A.3 Exámenes Parciales III

A.3.1 I ciclo lectivo de 1996

Universidad de Costa Rica **I Ciclo 1996**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

EXAMEN PARCIAL III

Cada pregunta vale 20 puntos.

1. Sea S el subespacio de $I\!R^4$ definido por:

$$S = \mathcal{C}\ell\{(1,0,-1,2),(1,1,1,0),(2,0,-2,1)\}$$

1. Encuentre una base ortonormal para S
2. Sea $u = (0,3,0,6)$. Calcule la proyección ortogonal de u sobre S.
3. Encuentre vectores $s \in S$, $t \in S^{\perp}$(subespacio ortogonal) de modo que se verifique: $u = s + t$.
4. Calcule la distancia del punto u a S.

2. Sea $T : I\!R^3 \to I\!R^3$,
$T(x,y,z) = (5x+4y+z, 4x+8y-4z, x-4y+5z)$

1. Encuentre una base para el subespacio $\text{Img}\,(T)$.
2. Determine la nulidad $\eta = \dim \text{Nuc}\,(T)$
3. Diga si T es diagonalizable ortogonalmente. (Justifique).
4. Diga sin hacer más cálculos si $\lambda = 0$ es un valor propio de T. (Justifique).

3. Sea $T : I\!R^2 \to I\!R^3$ una transformacion lineal, $\mathcal{B} = \{v_{1,} v_2\}, \mathcal{E} = \{w_1, w_2, w_3\}$ bases de $I\!R^2$ y $I\!R^3$ respectivamente y

$[T]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & x \\ 2 & y \\ 3 & z \end{pmatrix}$ la matriz asociada a T en las bases $\mathcal{B}$ y $\mathcal{E}$.

1. Si $T(v_2) = w_1 + 2w_2 + w_3$. Encuentre x, y, z.
2. Si $v = 2v_1 - v_2$. Calcule $[T(v)]_{\mathcal{E}}$ (Coordenadas de $T(v)$ en la base $\mathcal{E}$).
3. Decida si T es inyectiva; sobreyectiva (Justifique)

4. Sea $T : I\!R^3 \to I\!R^3$ un operador definido por $T(x) = Ax$, $x \in I\!R^3$

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

1. Encuentre una base $\mathcal{B}$ de $I\!R^3$, formada por vectores propios de T. Escriba la matriz $[T]_{\mathcal{B}}$ (matriz de T en las bases $\mathcal{B}$, $\mathcal{B}$)
2. Encuentre una matriz P tal que :$P^{-1}AP = [T]_{\mathcal{B}}$

5. Considere la cónica cuya ecuación es : $6\sqrt{3}xy - 6x^2 = 27$

1. Realice una rotación de los ejes x, y de modo que en los nuevos ejes $x\prime, y\prime$ la ecuación de la cónica este en su forma canónica.
2. En un solo gráfico represente los sistemas de ejes cartesianos $xy, x\prime y\prime$ y trace el gráfico de la cónica.
3. Calcule el coseno del ángulo de rotación.

A.3.2 II ciclo lectivo de 1996

Universidad de Costa Rica **II Ciclo 1996**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

TERCER EXAMEN PARCIAL

1. Sea $T : I\!R^3 \to I\!R^4$ lineal, representada en las bases $\mathcal{B} = \{v_1, v_2, v_3\}$ de $I\!R^3$ y $\mathcal{C}$ canónica de $I\!R^4$ por la matriz

$$[T]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 2 & 2 & 2 \end{pmatrix}$$

1. (4p) Encuentre $T(v_2)$ y $T(u)$ para $u = v_1 - 2v_2 + v_3$.
2. (10p) Encuentre un conjunto de generadores para cada subespacio: $\mathrm{Img}\,(T)$, $\mathrm{Nuc}\,(T)$, e indique si T es: inyectiva, sobreyectiva.
3. (6p) Si $\mathcal{E} = \{(1,2,1),(1,0,0),(2,2,2)\}$ es otra base de $I\!R^3$ y $[I]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (I: identidad en $I\!R^3$). Encuentre $[I]_{\mathcal{E}}^{\mathcal{B}}$ y $[T]_{\mathcal{E}}^{\mathcal{C}}$.

2. Sea $T_A : I\!R^2 \to I\!R^2$ la transformación lineal definida por $T_A(x) = Ax$, A matriz de 2×2. Sean además las rectas en $I\!R^2$: $L_1 : \ y = -x$, $L_2 : \ y = 2x$

1. (7p) Si T_A es la reflexión respecto a la recta L_1, encuentre la matriz A y $T_A(x,y)$.
2. (6p) Encuentre en que se transforman L_1, L_2 mediante T_A, esto es equivalente a calcular los dos subconjuntos siguientes:

$$T_A(L) = \{T_A(x,y) | (x,y) \in L\}, \quad L = L_1 \ , \ L = L_2$$

3. (7p) Haga un gráfico con L_1, L_2, $\mathrm{Img}\,(T_A)$, $\mathrm{Nuc}\,(T_A)$, $T(L_1)$, y $T(L_2)$. Seale con claridad en el gráfico cada conjunto.

3. Considere las matrices:

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

1. (8p) Demuestre que la matriz A no es diagonalizable
2. (12p) Encuentre una matriz P invertible y una matriz D diagonal tal que $P^{-1}BP = D$.

4. Sea A una matriz simétrica de 2×2 con valores propios $\lambda_1 = 4$ $\lambda_2 = -2$ y vectores propios correspondientes a cada valor propio: $v_1 = (1, 1)$, $v_2 = (1, -1)$.

1. (7p) Calcule A
2. (6p)Sea la cónica con ecuación $x^t Ax = 1$ con $x^t = (x_1, x_2)$. Obtenga la ecuación referida a un nuevo sistema de coordenadas $x_1\prime, x_2\prime$ en el cual la cónica esté en su forma canónica e identifíquela.
3. (7p) Haga un dibujo de la cónica, indicando, los ejes originales x_1, x_2 y los nuevos $x_1\prime, x_2\prime$.

5. Sea $\mathcal{B}$ la base canónica de $I\!R^3$ y $[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 3 \end{pmatrix}$

1. (6p) Obtenga una base para cada subespacio propio de T
2. (2p) Obtenga una base $\mathcal{E}$ de $I\!R^3$ tal que $[T]_{\mathcal{E}}$ sea diagonal.
3. (6p) Obtenga una matriz C ortogonal tal que.

$$[T]_{\mathcal{B}} = C^t [T]_{\mathcal{E}} C$$

4. (6p) Escriba la ecuación canónica de la superficie cuadrática.

$$3x_1^2 + 4x_2^2 + 3x_3^2 + 2x_1x_3 = 8$$

e identifíquela.

A.3.3 I ciclo lectivo de 1997

Universidad de Costa Rica **I Ciclo 1997**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

Examen Parcial III

1) (20 Pts) Sean $\mathcal{B} = \{v_1, v_2\}, \mathcal{B}' = \{w_1, w_2\}$ bases de $I\!R^2$ tales que:

$$w_1 = -v_2 \qquad w_2 = v_1 + v_2$$

a) Sea $I : I\!R^2 \to I\!R^2$ la transformación identidad.
Encuentre $[I]_{\mathcal{B}}^{\mathcal{B}'}$.
b)Sea $T : I\!R^2 \to I\!R^2$ una transformación lineal tal que
$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$. Encuentre $[T]_{\mathcal{B}}^{\mathcal{B}'}$
c) Calcule $[T(2v_1 - v_2)]_{\mathcal{B}'}$ (coordenadas de $T(2v_1 - v_2)$ en la base $\mathcal{B}'$)
d) Suponga que $\mathcal{B}$ es la base canónica de $I\!R^2$. Calcule las coordenadas en base $\mathcal{B}$ de los vectores $T(u)$ y $T(v)$, si $[u]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ y $[v]_{\mathcal{B}'} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

2) (20 pts.) Sean $\mathcal{B}_1$ y $\mathcal{B}_2$ bases de $I\!R^2$ y $I\!R^3$ respectivamente. Sea $T : I\!R^2 \to I\!R^3$ una transformación lineal tal que:

$$[T]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{pmatrix} 0 & 80 \\ 1 & -7 \\ 0 & -20 \end{pmatrix}.$$

a) Encuentre $[v]_{\mathcal{B}_1}$ si $[T(v)]_{\mathcal{B}_2} = \begin{pmatrix} -240 \\ 23 \\ 60 \end{pmatrix}$
b) Encuentre la dimensión de $\mathrm{Img}\,(T)$ y $\mathrm{Nuc}\,(T)$.
c) ¿Es T inyectiva? ¿Es T sobreyectiva? Justifique su respuesta.
d) Sea $S : I\!R^3 \to I\!R^2$ una transformación lineal tal que
$[S]_{\mathcal{B}_2}^{\mathcal{B}_1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$. Calcule $[T\ o\ S]_{\mathcal{B}_2}$.

3) (15 pts)Sea $T : I\!R^3 \to I\!R^3$ definida por $T(v) = Av$, donde:

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix}$$

a) Compruebe que T es invertible. Justifique su respuesta.

b) Calcule $\left[T^{-1}\right]_{\mathcal{C}}$,donde $\mathcal{C}$ es la base canónica de $I\!R^3$.

c)Sea $\mathcal{B} = \{(1,1,0),(1,0,1),(0,1,1)\}$ base de $I\!R^3$. Calcule $\left[T^{-1}\right]_{\mathcal{B}}^{\mathcal{C}}$

4) (15 pts) Sea $A = \frac{1}{2}\begin{pmatrix} 12 & 5 \\ 5 & -12 \end{pmatrix}$

a) Calcule los valores propios de A y encuentre una base para cada subespacio propio.

b) Considere la cónica dada por la ecuación $X^tAX = 13$. Encuentre los ejes principales, la ecuación canónica y el coseno del ángulo de rotación.

c) Identifique la curva y grafíquela.

5) (20 pts.) Sea $A = \begin{pmatrix} 1 & a & 0 \\ a & 1 & a \\ 0 & a & 1 \end{pmatrix}, a \in I\!R$

a) ¿Es A ortogonalmente diagonalizable? Justifique su respuesta.

b) Compruebe que $\lambda = 1$ es un valor propio de A y encuentre un vector propio asociado.

c) Compruebe que $\left(1, \sqrt{2}, 1\right)$ y $(1, -\sqrt{2}, 1)$ son vectores propios de A y calcule sus respectivos valores propios.

d) ¿Qué condición debe cumplir el parámetro"a" para que $X^tAX = 1$ sea un elipsoide? Justifique su respuesta.

6) (10 ptos.) Se estima que para cierta planta vegetal su producción anual P es directamente proporcional a su superficie foliar S. Un agrónomo obtuvo la siguiente tabla para cuatro plantas:

S=Superficie foliar (m^2)	10	12	13	15
P=Producción anual (kg)	30	35	40	44

a) Estime la constante de proporcionalidad k para el modelo $P = kS$.

b) ¿Es esta una buena aproximación? Justifique su respuesta.

c) Si una de estas plantas produce 25 kg al año, estime su superficie foliar.

A.3.4 II ciclo lectivo de 1997

Universidad de Costa Rica **II Ciclo 1997**
Escuela de Matemáticas **MA-1004 Algebra Lineal**

Examen Parcial III

1. (20 pts.) Considere la siguiente tabla de valores de las variables x, y :

x	-2	-1	0	1	2
y	1	0	-2	-1	0

Se desea estimar los parámetros a_0, a_1 y a_2 de la función cuadrática

$$y = a_0 + a_1 x + a_2 x^2$$

que mejor ajusta los datos de la tabla.

1. (7 pts.) Utilizando el modelo de regresión lineal múltiple:

$$y = xb + e$$

Escriba y, x y b que corresponden a los datos de este problema.

2. (10 pts.) Calcule la estimación de los parámetros a_0, a_1 y a_2 que produce el modelo en (1.1). Puede utilizar que

$$\begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}^{-1} = \frac{1}{70}\begin{pmatrix} 34 & 0 & -10 \\ 0 & 7 & 0 \\ -10 & 0 & 5 \end{pmatrix}.$$

3. (3 pts.) ¿Cuál es el valor y estimado para $x = -1$?

   Nota Al resolver 1.2) y 1.3) deben anotar todas las operaciones necesarias para obtener los resultados.

2. (15 pts.) Sea

$$\mathcal{B} = \left\{ \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\}$$

una base ortonormal de $I\!R^2$ y $\mathcal{C} = \{e_1, e_2\}$ la base canónica de $I\!R^2$. Para los operadores T , I (Identidad) en $I\!R^2$ se pide:

1. (5 pts.) Calcule $[I]_{\mathcal{B}}^{\mathcal{C}}$ y $[I]_{\mathcal{C}}^{\mathcal{B}}$.
2. (5 pts.) Si T es la transformación lineal tal que

$$T\left[\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ y } T\left[\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

   Calcule $[T]_{\mathcal{C}}$.

   Sugerencia: Calcule primero $[T]_{\mathcal{B}}^{\mathcal{C}}$

3. (5 pts.) Demuestre que $\left\| T\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] \right\| = \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|$.

3. (8 pts.) Sea $T : I\!R^2 \longrightarrow I\!R^3$ definida por:

$$T(x, y) = (x - y,\ y,\ -x + y)$$

1. (4 pts.) Determine una base para $\operatorname{Img}(T)$ y sin hacer más cálculos determine el $\operatorname{Nuc}(T)$
2. (4 pts.) Si

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

   calcule $[T]_{\mathcal{B}}^{\mathcal{C}}$ donde $\mathcal{C}$ es la base canónica de $I\!R^3$.

4. (25 pts) Sea $\mathcal{B} = \{u,\ v,\ w\}$ una base ortonormal de $I\!R^3$ y considere la transformación

$$T : I\!R^3 \longrightarrow I\!R^3 \ , \qquad T(x) = \text{Proy}_u x$$

1. (6 pts) Determine Img (T) y Nuc (T).
2. (4 pts) ¿T es inyectiva? ¿T es sobreyectiva? Justifique.
3. (8 pts) Justifique que 0 y 1 son valores propios de T y determine los espacios característicos $V_{\lambda=0}$, $V_{\lambda=1}$.
4. (3 pts) ¿Es T un operador diagonalizable? (Justifique).
5. (4 pts) Si $u = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $v = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ y $w = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

   Escriba el vector $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ como una suma de dos vectores propios de T.

   Observación: No puede usar 4.5) para responder 4.1), 4.2), 4.3), 4.4).

5. (18 pts) Sea

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

1. (10 pts) Obtenga los valores propios de A y los respectivos espacios característicos.
2. (4 pts) ¿A es diagonalizable? ¿A es diagonalizable ortogonalmente? Justifique.
3. (4 pts) Si B es una matriz 3×3 tal que su polinomio característico es:

   $$P_B(\lambda) = (\lambda - 1)\left(3 - \lambda^2\right).$$

   ¿Se puede garantizar que B es diagonalizable?

6. (14 pts)

1. (7 pts) Determine la ecuación $ax^2 + by^2 + cxy = d$ de la cónica cuya forma canónica es:

$$-\frac{(x')^2}{3} + \frac{(y')^2}{2} = 1$$

donde $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

2. (7 pts) Represente gráficamente la curva en 6.1) mostrando claramente los ejes correspondientes a las variables $x,\ y,\ x'$ y y'.

Apéndice B

Respuestas a algunos ejercicios

B.1 Ejercicios 1.8 (pag. 36)

1. b) La solución del sistema es $S = \emptyset$.

4. En todos los casos S es el conjunto solución del sistema:

1. $S = \{(2 + 3t, t, 6)|t \in I\!R\}$ 2. $S = \emptyset$ 3. $S = \{(2, 1, -7)\}$.

4. $S = \emptyset$ 5. $S = \{(-2 + t, 1 - 2t, t, 3)|t \in I\!R\}$

6. $S = \{(5, 0, 3, 0)\}$.

5. En todos los casos S es el conjunto solución del sistema homogéneo:

1. $S = \{(3t, t, 0)|t \in I\!R\}$

2. $S = \{(2t, s, t)|t, s \in I\!R\}$.

3. $S = \{(t, 0, 0)|t \in I\!R\}$.

8. a) $a = 0$ y $b \neq 2$, b) $a \neq 0$ y $b \neq 2$, c.i) $a \neq 0$ y $b = 2$ c.ii) $a = 0$ y $b = 2$.

12. a,b) 1) No, 2) Sí, infinitas 3) Sí, infinitas

c) En todos los casos hay y es única.

23. a) $A = \begin{pmatrix} 1 & 0 & -2 & 1 \\ 2 & 0 & 2 & -2 \\ -1 & -3 & 1 & -2 \\ 2 & 3 & 0 & 1 \end{pmatrix}$, $l = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ y $b = \begin{pmatrix} 0 \\ 4 \\ 1 \\ 1 \end{pmatrix}$.

b) y c) La solución del sistema es

$$S = \left\{ \left(\frac{4}{3}, -\frac{5}{9}, \frac{2}{3}, 0 \right) + t\,(3, -5, 6, 9)\,|t \in I\!R \right\}.$$

Tómese $l_0 = \left(\frac{4}{3}, -\frac{5}{9}, \frac{2}{3}, 0\right)$ y $d = t\,(3, -5, 6, 9)$.

14. a) La escritura matricial del sistema es:

$$\begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

b) La solución del sistema es:

$$S = \{(3, 0, 2, 0) + t\,(1, 2, 1, -1)\,|t \in I\!R\}$$

Se puede tomar $h = (3, 0, 2, 0)$ y $u = t\,(1, 2, 1, -1)$.

15. $a = \frac{1}{3}$ y $\alpha = \frac{5}{3}$. La solución del sistema es

$$S = \left\{ \left(\frac{5}{3} + t, \frac{5}{3} - 2t, t \right) |t \in I\!R \right\}$$

16. a) Para cualesquiera α y β el sistema propuesto es equivalente a un sistema no homogéneo 3×3.

El sistema propuesto es inconsistente para todo $\beta \in I\!R$ y $\alpha = 0$.

b) Sea $\alpha \neq 0$, la solución del sistema es $\left\{ \left(1, 0, \frac{1}{\alpha}\right) \right\}$.

17. a) $b = \frac{1}{2}$ y $a \in I\!R$ o bien, $a = 2$ y $b \neq \frac{1}{12}$.

b) $a \neq 2$ y $b \neq \frac{1}{2}$.

c) $a = 2$ y $b = \frac{1}{12}$.

19. a) La solución es $x = p$; $y = q$; $z = 600 - p$; $w = -p - q$; $m = 500 - q$

b) Si $p = q = 0$, entonces $x = y = w = 0$, $z = 600$, $m = 500$.

20. a) La solución del sistema es $x = -z + t + 100, y = -z + t - w + 300, s = t - w$, donde z, t, w son variables libres. b) si $z = 100, t = s = 50$, la solución particular es, $x = 50, w = 0, y = 250$.

21. a) $h = -k$
b) $x = h + w = y$, $z = 600 - h - w$. Note que x,y,z y w deben ser no negativos
c) $w = 400, h = 150$ entonces $x = y = 550$ y $z = 50$.

22. $I_1 = \frac{5}{13}$, $I_2 = \frac{6}{13}$, $I_3 = \frac{1}{13}$.

B.2 Ejercicios 2.8 (pag. 84)

1. a) $\begin{pmatrix} 4 & -1 & 0 \\ 0 & 4 & 2 \\ -1 & 0 & 4 \end{pmatrix}$ b) $\begin{pmatrix} -8 & -3 \\ -5 & 9 \\ 2 & 1 \end{pmatrix}$

c) $\begin{pmatrix} -8 & -5 & -8 \\ -8 & -5 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ d) $\begin{pmatrix} -6 & -5 & -4 \\ -8 & -6 & 1 \\ 1 & 3 & 0 \end{pmatrix}$

e) $\begin{pmatrix} 2 & 3 & -4 \\ 3 & 9 & -3 \\ -4 & -3 & 10 \end{pmatrix}$ f) $\begin{pmatrix} 11 & -4 \\ -4 & 10 \end{pmatrix}$

3. $x = -2 \pm \sqrt{2}$.

4. Si la matriz diagonal D multiplica por la izquierda (derecha) a A, cada entrada d_{ii} de D multiplica la i-ésima fila (columna) de A.

6. b) $a = 1, \quad b = 2, \quad c = 1$ y $d = 4$.

11. b) $D^p = (d_{ij}^p.)$.

12. b) $A + A^2 = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 1 & 2 & 1 & 2 \\ 2 & 2 & 0 & 2 \\ 2 & 2 & 1 & 1 \end{pmatrix}$

a_{ij} : número de maneras de comunicarse de i a j usando a lo más un intermediario.

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

17. B es triangular superior con todos los elementos de la diagonal iguales a 1.

20. $B = (A^{-1}D - C)^t = \begin{pmatrix} 3 & 0 & 1/3 \\ -1 & -4 & -4 \end{pmatrix}$.

23. $C = \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{2} & -\frac{1}{2} & 0 \\ -15 & 5 & 2 \end{pmatrix}$ y $CA = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$

24. a) $B = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} \\ 0 & 0 & 1 & -\frac{4}{5} \end{pmatrix}$

$C = E(-\frac{3}{5}f_3)E(-f_2 + f_3)E(\frac{1}{3}f_2)E(-f_1 + f_2)$

b) $B = \begin{pmatrix} 1 & 0 & 0 & \frac{11}{5} \\ 0 & 1 & 0 & \frac{6}{5} \\ 0 & 0 & 1 & -\frac{4}{5} \end{pmatrix}$

$C' = E(-\frac{2}{3}f_3 + f_1)E(-\frac{2}{3}f_3 + f_2)E(f_2 + f_1)C$

25. a) A es equivalente a $E = \begin{pmatrix} 1 & 0 & \frac{1}{4} \\ 0 & 1 & \frac{11}{6} \\ 0 & 0 & 0 \end{pmatrix}$

b) $E = E(\frac{1}{3}f_2)E(f_2+f_3)E(-6f_1+f_3)E(-2f_1+f_2)E(\frac{1}{4}f_1)A$

26. a) $\begin{pmatrix} 1 & 0 \\ 1 & 2 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 13 \end{pmatrix}$

b) $(2, 1, -1) \begin{pmatrix} -1 & 3 & 5 \\ 1 & 1 & 1 \\ 2 & 0 & 3 \end{pmatrix} = (-3, 7, 8)$

27. a) x_1 pertenece al conjunto de combinaciones lineales de las columnas de A y también x_2 para todo a, b.

b) x_1 si y x_2 no.

c) $c + a - 2b = 0$.

28. a) $(3+\alpha)\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} \forall \alpha \in I\!R$

c) Los rangos son 1 y 2, respectivamente.

29. a) $3\begin{pmatrix} 1 \\ 2 \end{pmatrix} + (-2)\begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$

b) 2.

47. b) $2a^2 + 2b^2 = 1$.

49. a) $L = \begin{pmatrix} a & d & 0 \\ b & e & 0 \\ c & f & -1 \end{pmatrix}$, $C_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ y $C_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

b) $\begin{vmatrix} a & d \\ b & e \end{vmatrix} \neq 0$ (en este caso la solución es única).

50. a)$\Longrightarrow$b): $Ax = 0 \Longrightarrow A^{-1}(Ax) = A^{-1}0 = 0 \Longrightarrow \left(A^{-1}A\right)x = 0 \Longrightarrow I_n x = x = 0$

b)$\Longrightarrow$a): Hay varias formas de probar esto; una de ellas es: si la solución de $Ax = 0$ es $\{0\}$ entonces las columnas de A son l.i. Luego el rango de A es n por lo que A es invertible.

B.3 Ejercicios 3.5 (pag.105)

1. a) $C = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -3 & -3 \\ 0 & 0 & 1 \end{pmatrix}$, $|A| = 3$.

b) $C = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & -2 & -1 & -1 \\ 0 & 0 & -7 & -2 \\ 0 & 0 & 0 & \frac{10}{7} \end{pmatrix}$, $|A| = 20$.

2. $|A| = 15$, $|B| = 36$.

4. $|A| = 48$.

7. (i) $|A| = 2$, (ii)$|A| = 2$.

9. Todo $x \in I\!R$, $x \neq \frac{1}{3}$.

10. 4, 9, -1.

11. Linealmente independiente, pues $|A| \neq 0$.

1.

Mat	Val.Propios	Vectores Propios
A	1,-1	{(-1,1)},{(1,1)}
B	no tiene	
C	-3, -4	{(1,1)},{(-5,2)}
D	-3 ,-2, 1	{(2,1,1)}, {(1,2,1)}, { (-2,1,0)}
E	1	{(0,0,1), (1,0,0),(0,0,0)}

B.4 Ejercicios 4.4 (pag. 140)

2. 1. Por el "método gráfico" se observa que el máximo de la función objetivo se alcanza en el vértice $(5, 10)$ y $z(5, 10) = 5$.

2. La solución es la misma porque la región de soluciones factibles no cambia al quitar la restricción $x_1 - x_2 \leq 3$.

3. Si se quita la restricción $x_2 \leq 10$ entonces la región de soluciones factibles es no limitada y la función objetivo es no acotada. Por ejemplo, para $\alpha > 0$ y $x = \max\left\{\alpha, 3 + \frac{\alpha}{4}\right\}$, los puntos $(x, 3x - \alpha)$ están en la región factible y $z(x, 3x - \alpha) = \alpha$, de donde $z \to +\infty$ cuando $\alpha \to +\infty$.

3. $x_1 = 50$, $x_2 = 75$. El valor óptimo de la función objetivo es 300.

6. 1. El programa lineal canónico equivalente es: min $z = -x_3 + 5x_4 + 8$ sujeto a

$$\left\{ \begin{array}{rcl} x_1 + 3x_3 - 2x_4 & = & 6 \\ x_2 + 2x_3 + 4x_4 & = & 2 \end{array} \right.$$

con x_3, $x_4 \geq 0$.

2. $x_3 = 1$, $x_1 = 3$, $x_4 = x_2 = 0$, $z\,(3, 0, 1, 0) = 7$.

7. El programa lineal propuesto es equivalente (haciendo operaciones elementales) al programa lineal canónico:
min $z = x_3 + 4x_4 + 4x_5 - 2$ sujeto a

$$\left\{ \begin{array}{rcl} x_1 + x_3 + 6x_4 + 3x_5 & = & 2 \\ x_2 + 3x_3 + 21x_4 + 10x_5 & = & 8 \end{array} \right.$$

con $x_i \geq 0$ para $i = 1, \ldots, 5$. La solución óptima es $x_1 = 2$, $x_2 = 8$, $x_3 = x_4 = x_5 = 0$, $z\,(2, 8, 0, 0, 0) = -2$.

8. La solución óptima es $x_1 = 0$, $x_2 = 25$, $x_3 = 11.67$ y $z(0, 25, 11.67) = 56.67$.

9. Sea x (respectivamente y) la cantidad de café que compra la compañía al distribuidor A (respectivamente B), en toneladas. El modelo de programación lineal cuya solución provee la cantidad de café que compra la compañía a cada distribuidor, al costo mínimo, es: min $c = 125x + 200y$ sujeto a

$$\begin{cases} 0.2x + 0.4y & \geq & 280 \\ 0.5x + 0.2y & \geq & 200 \end{cases}$$

con x, $y \geq 0$.

Solución óptima: $x = 150$, $y = 625$. El costo mínimo correspondiente es $c\,(150, 625) = 143.750$ dólares.

10. Debe plantar $x = 45$ acres del cultivo A y $y = 50$ acres del cultivo B. El ingreso máximo correspondiente es 12.450 dólares.

11. Deben fabricarse diariamente $x = 8$ unidades del producto A y $y = 2$ del producto B. La ganancia máxima es 260 dólares.

12. 1. El modelo es: max $V = 0.15x + 0.26y$ sujeto a:

$$\begin{cases} 0.1x + 0.2y & \leq 100 \\ x + y & \leq 700 \\ x & \leq 500 \\ y & \leq 400 \end{cases}$$

con $x \geq 0$, $y \geq 0$.

3. Debe vender $x = 400$ bolsas de maní y $y = 300$ bolsas de dulces para una venta máxima de 138 dólares.

B.5 Ejercicios 5.5 (Pag. 177)

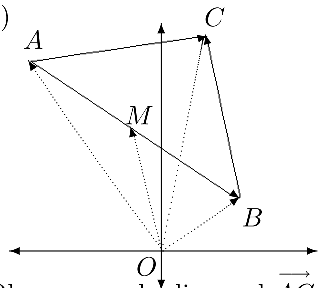
1. El cuadrilátero de vértices A, B, C y D, como se muestra en el dibujo, es un paralelogramo entonces $\overrightarrow{AB} = \overrightarrow{DC}$ (por el ejemplo 5.3), luego $B - A = C - D$ y $A - B + C - D = 0$. Por otra parte, si $A - B + C - D = 0$ entonces $C - D = B - A$ y $C - B = D - A$ por lo cual $\overrightarrow{DC} = \overrightarrow{AB}$ y $\overrightarrow{BC} = \overrightarrow{AD}$, o sea, los lados opuestos del cuadrilátero son paralelos y es por tanto un paralelogramo.

3. a) Si $A = (a_1, a_2)$ entonces $B = (a_1 + 2, a_2 - 3)$, por ejemplo: $A = (0, 2), B = (2, -1), C = (2, 0)$ y $D = (4, -3)$.

b) Si $A = (a_1, a_2)$ entonces $B = (a_1 - 1, a_2 - 1/2)$, por ejemplo $A = (1, 1), B = (0, 1/2))$.

c) Los vértices A, B, C, D deben cumplir, por ejemplo, que $\overrightarrow{AB} = \vec{u}, \overrightarrow{BC} = \vec{v}, \overrightarrow{DC} = \vec{u}$ y $\overrightarrow{AD} = \vec{v}$. Si $A = (a_1, a_2)$ entonces $B = (a_1 + 2, a_2 - 3), C = (a_1 - 2, a_2 - 1)$ y $D = (a_1 - 4, a_2 + 2)$.

5. a)



b) $M = (-1/2, 3)$.

c) Si, porque los lados $\overrightarrow{AC}$ y $\overrightarrow{BC}$ tienen la misma magnitud 82/4.
d) Si, porque el triángulo A, B, C es isósceles.

7. Observe que la diagonal $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AD} + \overrightarrow{DC}$, y dividiendo entre 2 que: $\overrightarrow{AC}/2 = \overrightarrow{AB}/2 + \overrightarrow{BC}/2 = \overrightarrow{AD}/2 + \overrightarrow{DC}/2$, luego $\overrightarrow{M_1B} + \overrightarrow{BM_2} = \overrightarrow{M_4D} + \overrightarrow{DM_3}$ y $\overrightarrow{M_1M_2} = \overrightarrow{M_4M_3}$, entonces el cuadrilátero de vértices M_1, M_2, M_3, M_4 es un paralelogramo.

8. 1) $(-3/5, 4/5)$,
3) $\frac{1}{\sqrt{30}}(-1, 2, -3, 4)$,
5) $\frac{1}{\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}}(x_1, x_2, \ldots, x_n)$.

10. 1) $\dfrac{1}{\sqrt{26}}$, 3) $\dfrac{-3}{\sqrt{60}}$, 5) $\sum_{i=1}^n x_i y_i / \sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}$.

12. $\vec{u}$ y $\vec{v}$ no nulos, son l.d. $\Longleftrightarrow$existe $\alpha \in I\!\!R$ tal que $\vec{u} = \alpha\vec{v}$ $\Longleftrightarrow\vec{u}$ y $\vec{v}$ son paralelos.

14. 1. a) No es rectángulo, b) Area: 10.

2. a) Si es rectángulo, b) Area: $9\sqrt{2}$.

3. a) No es rectángulo, b) Area: $\sqrt{34}/2$.

4. a) Es rectángulo sólo si $|\vec{u} \cdot \vec{v}| = 1$, b) Area: $\sqrt{4 - (\vec{u} \cdot \vec{v})^2}$.

16. $\left\|\vec{a} + \alpha\vec{b}\right\|^2 = (\vec{a} + \alpha\vec{b}) \cdot (\vec{a} + \alpha\vec{b}) = \vec{a} \cdot \vec{a} + 2\alpha\vec{a} \cdot \vec{b} + \alpha^2\vec{b} \cdot \vec{b}$ como $\vec{a}$ y $\vec{b}$ son perpendiculares entonces $\vec{a} \cdot \vec{b} = 0$ y se obtiene que:

$\left\|\vec{a}+\alpha\vec{b}\right\|^2 = \|\vec{a}\|^2 + \alpha^2\left\|\vec{b}\right\|^2$. Luego, como $\alpha^2\left\|\vec{b}\right\|^2 \geq 0$, para toda $\alpha \in I\!R$ entonces $\left\|\vec{a}+\alpha\vec{b}\right\|^2 \geq \|\vec{a}\|^2$ de donde se deduce el resultado.

18. Las diagonales, $\vec{a}+\vec{b}$ y $\vec{a}-\vec{b}$, de un paralelogramo son ortogonales $\Longleftrightarrow(\vec{a}+\vec{b})\cdot(\vec{a}-\vec{b}) = 0 \Longleftrightarrow\vec{a}\cdot\vec{a}+\vec{b}\cdot\vec{b} = 0$ $\Longleftrightarrow\|\vec{a}\|^2 - \left\|\vec{b}\right\|^2 = 0 \Longleftrightarrow$los lados del paralelogramo tienen la misma magnitud, o sea, es un rombo.

La demostración del otro resultado es similar.

21. 1) $\vec{a} = \frac{3}{13}(2,3)$, $\vec{b} = (\frac{-45}{13}, \frac{30}{13})$.

3) $\vec{a} = \frac{1}{3}(-1,2,3,4)$, $\vec{b} = (4/3, -11/3, 2, 2/3)$.

5) $\vec{a} = (0,\cdots,0,x_i,0,\cdots,0)$, $\vec{b} = (x_1,\cdots,x_{i-1},0,x_{i+1},\cdots,x_n)$.

23. Por hipótesis $\vec{a}\cdot\vec{v}_i = 0$ para toda $i = 1,\ldots,k$, entonces:

$$\begin{array}{rcl} \vec{a}\cdot\vec{v} & = & \vec{a}\cdot(x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k) \\ & = & x_1\vec{a}\cdot\vec{v}_1 + x_2\vec{a}\cdot\vec{v}_2 + \cdots + x_k\vec{a}\cdot\vec{v}_k \\ & = & x_1 0 + x_2 0 + \cdots + x_k 0 \\ & = & 0 \end{array}.$$

B.6 Ejercicios 6.5 (Pag. 204)

1. i) a) $\left\{\begin{array}{l} x = -3t-1 \\ y = -3t \\ z = 11t+6 \end{array}\right.$ b) $\frac{x+1}{-3} = \frac{y}{-3} = \frac{z-6}{11}$.

ii) a) $\left\{\begin{array}{l} x = -2t+2 \\ y = 3 \\ z = 3t+4 \end{array}\right.$ b) $\frac{x-2}{-2} = \frac{z-4}{3}, y = 3$.

3. a) $(x,y,z) = (2,3,-5) + t(2,1,-3)$. b) $(x,y,z) = t(0,1,0)$. c) $(x,y,z) = (-1,3,1) + t(0,0,1)$.

5. a)$2x-y-z = 1$, b) $7x+6y+8z = -4$, c)$5x+5y+2z = -13$.

7. $(28/3, 11/3, -58/3)$.

11. a) Como $X = (x,y)$ y $P = (p_1,p_2)$ son puntos de la recta, el vector $\overrightarrow{PX}$ tiene la dirección de la recta y es por lo tanto perpendicular a $\vec{a}$. Luego $0 = \overrightarrow{PX} \cdot \vec{a} = X \cdot \vec{a} - P \cdot \vec{a}$ entonces $ax + by = ap_1 + bp_2$ para todo (x, y) en la recta.

b) $ax+by = c = ax_0+by_0$ si y solo si $a(x-x_0)+b(y-y_0) = 0$ si y solo si $(a, b) \perp (x-x_0, y-y_0)$ si y solo si $(x-x_0, y-y_0) = t(v_1, v_2)$ para un vector (v_1, v_2) perpendicular a (a, b) si y solo si $(x, y) = (x_0, y_0) + t(v_1, v_2)$ si y solo si (x, y) es un punto de una recta que contiene a (x_0, y_0) en la dirección (v_1, v_2) perpendicular a (a, b).

c) Dos rectas $-m_1x + y = b_1$ y $-m_2x + y = b_2$ en $I\!R^2$ son perpendiculares si y solo si sus vectores normales son perpendiculares, esto es si $(-m_1, 1) \perp (-m_2, 1)$ lo cual ocurre si y solo si $m_1m_2 + 1 = 0$.

13. a) $d = \left\|\text{Proy}_{\vec{a}}\overrightarrow{PQ}\right\| = \dfrac{|\overrightarrow{PQ} \cdot \vec{a}|}{\|\vec{a}\|^2}\,\|\vec{a}\| = \dfrac{|(Q-P) \cdot \vec{a}|}{\|\vec{a}\|}$

b) De a) si $Q = (x_0, y_0)$, $P = (p_1, p_2)$ y $\vec{a} = (a, b)$ entonces: $d = \dfrac{|((x_0, y_0) - (p_1, p_2)) \cdot (a, b)|}{\|(a, b)\|} = \dfrac{|ax_0 + by_0 - ap_1 - bp_2|}{\sqrt{a^2 + b^2}}$ y como (p_1, p_2) es un punto de la recta se tiene que $ap_1 + bp_2 + c = 0$, o sea, $c = -ap_1 - bp_2$, entonces $d = \dfrac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$

c) Aplicando el resultado en b) con la ecuación de la recta $mx - y + b = 0$, y el punto $(x_0, y_0) = (0, 0)$ se tiene que: $d = \dfrac{|mx_0 - y_0 + b|}{\sqrt{m^2 + (-1)^2}} = \dfrac{|b|}{\sqrt{m^2 + 1}}$

15. a) Distancia$(\mathcal{P}, \ell) = \dfrac{|ap_1 + bp_2 + cp_3 + d|}{\sqrt{a^2 + b^2 + c^2}}$ donde $P = (p_1, p_2, p_3)$ es un punto de la recta ℓ que es paralela al plano $\mathcal{P} : ax + by + cz + d = 0$.

b)El vector dirección de la recta dada, $(-3, 1, 2)$, es perpendicular al vector normal del plano dado, $(-1, 5, -4)$, luego la recta es paralela al plano y la distancia entre la recta y el plano es $25/\sqrt{42}$ unidades lineales.

17. Sean $X = P + t\vec{v}$ y $X = Q + t\vec{w}$ las ecuaciones vectoriales de las dos rectas ℓ_1 y ℓ_2 en $I\!R^n$, $n > 2$. Considere los planos $\mathcal{P}_1$ con ecuación vectorial $X = P + t\vec{v} + s\vec{w}$ y $\mathcal{P}_2$ con ecuación $X = Q + t\vec{v} + s\vec{w}$. Observe que $\mathcal{P}_1$ y $\mathcal{P}_2$ son planos paralelos porque sus direcciones las determinan los mismos vectores $\vec{v}$

y $\vec{w}$. Además, claramente todo punto $X = P + t\vec{v}$ de la recta ℓ_1 está contenido en $\mathcal{P}_1$, basta elegir $s = 0$, luego $\ell_1 \subset \mathcal{P}_1$. Similarmente se observa que $\ell_2 \subset \mathcal{P}_2$.

a) $\mathcal{P}_1 : -8x + 3y + 7z = -3$, $\mathcal{P}_2 : -8x + 3y + 7z = 24$

b) $27/\sqrt{122}$ c) $(x, y, z) = t(-8, 3, 7)$.

19. a) $(3, 5, 7)$, b) $\emptyset$, c) $\ell(P, \vec{v})$.

21. $d(t) = \|Q - X(t)\|^2 = \|Q - P - t\vec{v}\|^2 = (Q - P - t\vec{v}) \cdot (Q - P - t\vec{v}) = (Q - P) \cdot (Q - P) - (2(Q - P) \cdot \vec{v})t + (\vec{v} \cdot \vec{v})t^2$ lo cual es un polinomio de grado 2 en la variable t o parábola. Como su coeficiente del término cuadrático, $\vec{v} \cdot \vec{v}$, es positivo, se trata de una parábola que abre hacia arriba y su vértice determina un único punto $(t_0, d(t_0))$ donde $d(t_0) \leq d(t)$ para todo $t \in \mathbb{R}$. Entonces $d(t)$ es mínima cuando t es la coordenada t_0 del vértice que es dada por: $t_0 = \dfrac{-b}{2a} = \dfrac{2(Q - P) \cdot \vec{v}}{2\vec{v} \cdot \vec{v}} = \dfrac{(Q - P) \cdot \vec{v}}{\vec{v} \cdot \vec{v}}$.

Por otra parte $Q - X(t_0) = (Q - P) - t_0\vec{v}$ y $(Q - X(t_0)) \cdot \vec{v} = ((Q - P) - t_0\vec{v})\vec{v} = (Q - P) \cdot \vec{v} - t_0\vec{v} \cdot \vec{v} = 0$ sustituyendo por el valor t_0 dado antes.

23. a)$(1, 1, -1)$, b)$x + y - z = 4$, c)$4/\sqrt{3}$.

25. $(x, y, z) = (2, 1, -3) + t(4, -3, 1)$.

B.7 Ejercicios 7.6 (Pag. 240)

1. a) $S = \{(1, x, y)/x, y \in \mathbb{R}\}$ no es subespacio de $\mathbb{R}^3$ porque $(0, 0, 0) \notin S$.

c) $S = \{A \in M(n, \mathbb{R}) | A$ es simétrica$\}$ es un subespacio de $M(n, \mathbb{R})$ porque si $A \in S$, $B \in S$ y $\alpha \in \mathbb{R}$ entonces $A^t = A$ y $B^t = B$, luego $(A + \alpha B)^t = A^t + \alpha B^t = A + \alpha B$ y $A + \alpha B \in S$.

e) No es subespacio, porque no contiene el cero de $\mathbb{R}^n$.

g) Es subespacio, porque si $x, y \in S$ y $\alpha \in \mathbb{R}$ entonces $x \cdot a = 0$ y $y \cdot a = 0$, de manera que $(x + \alpha y) \cdot a = x \cdot a + \alpha y \cdot a = 0$, luego $(x + \alpha y) \in S$.

i) Si A y B son matrices diagonales $A + \alpha B$ es una matriz diagonal para todo $\alpha \in \mathbb{R}$, entonces S es un subespacio de $M(n, \mathbb{R})$.

3. $\mathcal{B} = \{(1/2, 3/2, 1, 0)^t, (-1/2, -1/2, 0, 1)^t\}$.

5. a) $\mathcal{B}_1 = \{(1, 0, 1, 3), (0, 1, 1, -1)\}$

b) $(1, 1, 1, 1) \notin F$, $[2, 6, 8, 0]_{\mathcal{B}_1} = (2, 6)$

c) $\mathcal{B}_2 = \{(1, 0, 11/3), (0, 1, -4/3)\}$ o
$\mathcal{B}_2 = \{(1, 2, 1), (0, 1, -4/3)\}$ o
$\mathcal{B}_2 = \{(1, 2, 1), (1, -1, 5)\}$ son algunas elecciones posibles.

d) $[A_2]_{\mathcal{B}_2} = (1, -1)$, con $\mathcal{B}_2$ la primera base en c).

e) Si, porque el máximo número de filas l.i. es igual al máximo número de columnas l.i. (el rango) en cualquier matriz.

7. a) Si, porque generan a $I\!R^3$ y como la dimensión de este espacio es 3 entonces son una base. b) No siempre, solo si v_1, v_2, v_3 son l.i. c) Si tiene solución, porque v_2, v_3, v_4 generan $I\!R^3$ y por lo tanto v_1, v_2, v_3, v_4 también lo generan, pero la solución no es única porque estos cuatro vectores no son l.i.

9.

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

$$\left[\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix}$$

11. $a + b = 1$ o $a = 2$.

14. Base hiperplano W: $\mathcal{B} = \{(1, 1, 0, 0), (0, 1, 1, 0), (0, 2, 1, 1)\}$, $(a_1, a_2, a_3, a_4) = (-1, 1, -1, -1)$.

B.8 Ejercicios 8.6 (Pag. 260)

1. a) Una entre otras: $\mathcal{B}_F = \{(1, 1, -1, 1)^t, (0, 1, 1, 0)^t\} = \{v_1, v_2\}$.

b) $\mathcal{B}_N = \{(2, -1, 1, 0)^t, (-1, 0, 0, 1)^t\} = \{w_1, w_2\}$.

c) Ya se conoce que son subespacios de $I\!R^4$. Ahora sean $x \in F$ y $y \in N$ entonces $x = \alpha_1(1,1,-1,1) + \alpha_2(0,1,1,0) = \alpha_1 v_1 + \alpha_2 v_2$ y $y = \beta_1(2,-1,1,0) + \beta_2(-1,0,0,1) = \beta_1 w_1 + \beta_2 w_2$ de manera que $x \cdot y = \alpha_1\beta_1 v_1 \cdot w_1 + \alpha_1\beta_2 v_1 \cdot w_2 + \alpha_2\beta_1 v_2 \cdot w_1 + \alpha_2\beta_2 v_2 \cdot w_2 = \alpha_1\beta_1 0 + \alpha_1\beta_2 0 + \alpha_2\beta_1 0 + \alpha_2\beta_2 0 = 0$.

d) $\mathcal{B}_1 = \{(1/2, 1/2, -1/2, 1/2), (0, 1/\sqrt{2}, 1/\sqrt{2}, 0)\}$
$\mathcal{B}_2 = \{\frac{1}{\sqrt{6}}(2,-1,1,0), \{\frac{1}{6}(-\sqrt{3}, -\sqrt{3}, \sqrt{3}, 3\sqrt{3})$.

e) $\vec{a} = (1, 7/2, 3/2, 1)$ y $\vec{b} = (0, -3/2, 3/2, 3)$.

f) $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ es una base de $I\!R^4$ luego para todo $\vec{x} \in I\!R^4$ existen a_1, a_2, a_3, a_4 tales que:
$\vec{x} = a_1(1/2, 1/2, -1/2, 1/2) + a_2(0, 1/\sqrt{2}, 1/\sqrt{2}, 0) + a_3(2/\sqrt{6}, -1/\sqrt{6}, 1/\sqrt{6}, 0) + a_4(-\sqrt{3}/6, -\sqrt{3}/6, \sqrt{3}/6, \sqrt{3}/2)$
de manera que existen
$\vec{a} = a_1(1/2, 1/2, -1/2, 1/2) + a_2(0, 1/\sqrt{2}, 1/\sqrt{2}, 0) \in F$ y $\vec{b} = \frac{a_3}{\sqrt{6}}(2,-1,1,0) + a_4(-\sqrt{3}/6, -\sqrt{3}/6, \sqrt{3}/6, \sqrt{3}/2) \in N$,
tales que $\vec{x} = \vec{a} + \vec{b}$. Por otra parte, observe que $\vec{a} = \text{Proy}_F \vec{x}$ y $\vec{b} = \text{Proy}_N \vec{x}$.

3. $\mathcal{B} = \{\frac{1}{\sqrt{2}}(1,1,0,0), \frac{1}{\sqrt{6}}(1,-1,2,0), \frac{1}{2}(-1,1,1,1)\}$
$W^\perp = \mathcal{C}\ell\{(1,-1,-1,3)\}$.

5. $1/\sqrt{2}$.

8. 1) $\mathcal{B} = \{(-1/\sqrt{2}, 0, 1/\sqrt{2}, 0), (0, 1/\sqrt{2}, 0, 1/\sqrt{2})\}$.

2) $\mathcal{D} = \{\frac{1}{\sqrt{2}}(-1,0,1,0), \frac{1}{\sqrt{2}}(0,1,0,1), \frac{1}{2}(1,-1,1,1), \frac{1}{2}(1,1,1,-1)\}$.

3) $\text{Proy}_{W^\perp} v = (1, -1/2, -1, -1/2)$.

4) $[\text{Proy}_{W^\perp} v]_{\mathcal{B}} = \begin{pmatrix} -2/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$,

$$[\text{Proy}_{W^\perp} v]_{\mathcal{D}} = \begin{pmatrix} -2/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}, \ [\text{Proy}_{W^\perp} v]_{\mathcal{C}} = \begin{pmatrix} 1 \\ -1/2 \\ -1 \\ -1/2 \end{pmatrix}$$

11. $(a_1, a_2, a_3, a_4) = (-1, 1, -1, 1)$.

B.9 Ejercicios 9.3 (pag. 282)

1. $\widehat{C}_n = 0.7958$, ($\widehat{a}_0 = 0.006491$ y $\widehat{a}_1 = 0.04211$).

2. $\widehat{a} = \frac{25}{7}$, $\widehat{b} = \frac{-71}{42}$, $\widehat{c} = \frac{-15}{14}$, $\widehat{d} = \frac{1}{3}$.

3. γ: costo total, $\quad \gamma = 95.7114 + 5.68527x - 0.0169673x^2$.

4. A: cantidad alimento (libras/día), p: producción (libras/día). $A = kp + \epsilon$, $\quad \widehat{k} = 0.403266 \approx 0.4$.

5. γ: caudal promedio mensual (m^3/s), x : mes, mayo $= 1$, junio $= 2$, ..., abril $= 12$. $\widehat{\gamma} = 1.676 + 1.013x - 0.096x^2$.

11. $\widehat{k} = 0.0739573$.

12. Descenso $= 37.8083$.

B.10 Ejercicios 10.4 (pag.319)

1. $T(x, y, z) = (-\frac{13}{3}x + \frac{8}{3}y - 3z, -3x + 2y - 2z)$.

6. Complete una base para $I\!R^4$ y defina T como:

$T(1, 1, 1, 0) = (0, 0, 0)$, $\quad T(2, -1, 1, 1) = (0, 0, 0)$,

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

7. 1) $[I]_B = I_3$, $\; [I]_B^C = (u_1, u_2, u_3)$

2) $[I]_C^B = \begin{pmatrix} 1 & 1 & a \\ 1 & 1 & 0 \\ -1 & 0 & b \end{pmatrix}$

8. a) $[T(v)]_B = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ $\qquad [T(v)]_C = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$

c) $[T]_D^C = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 1 \end{pmatrix}$

9. 1) $[T]_B = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & -1 \\ 2 & 1 & 3 \end{pmatrix}$ $\qquad$ 2) $[I]_{B_1}^B = \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$

3) $[T]_B^{B_1} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & -1 & 0 \\ 0 & 3 & 1 \end{pmatrix}$.

11. 4) $\operatorname{Img}(T) = \mathcal{C}\ell\{(1,-1)\}$, $\operatorname{Nuc}(T) = \mathcal{C}\ell\{(-2,1,0),(-3,0,1)\}$

15. 2) $B = \left\{\frac{1,-1,0}{\sqrt{2}},(0,0,1)\right\}, [T]_C^B = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$

3) $\frac{1}{2}\sqrt{2}$.

16. A invertible, luego T lo es.

$$\left[T^{-1}\right]_B = \frac{1}{3}\begin{pmatrix} 1 & -2 & 4 \\ 1 & 1 & -2 \\ -1 & -1 & 5 \end{pmatrix}$$

$$T^{-1}(x,y,z) = \frac{1}{3}\begin{pmatrix} 0 & 9 & -12 \\ 0 & 3 & -3 \\ 1 & -3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

17. b) $D = \{(1,1,0,1),(1,-1,1,0),(0,1,-1,1)\}$

$$[S]_B^D = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \left[S^{-1}\right]_D^B = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

18. a)$\mathcal{B}_{Img} = \{(1,1,0,1),(-1,0,0,-1)\}$ $\mathcal{B}_{Nuc} = \{(1,1,1)\}$

b) No es sobreyectiva ni inyectiva.

c) $B = \{(1,0,1),(1,1,1),(0,0,1)$.

24. $\mathrm{A} = \frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \qquad T_A(x,y) = \frac{1}{2}(x-y, x+\sqrt{3}y).$

25. a) T_S : rotación en $60°$

T_R : reflexión respecto eje y.

b) $T_C = (T_r \text{ o } T_s)$, con $C = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$

B.11 Ejercicios 11.6 (pag. 362)

1 a) Los valores propios son 2 y 4.
b) Una base es $\{(-1,1,1),(3,0,1),(-2,1,0)\}$.

2 a) Los valores propios son -2 y 2. Una base de $V_{\lambda=2}$ es $\{(-\sqrt{2},-1,1)\}$ y para $V_{\lambda=-2}$ es $\{(-1/\sqrt{2},1,0),(1/\sqrt{2},0,1)\}$.

3. b) Si $a \neq 0, b \neq 0, c \neq 0$. $\dim V_2 = 1$. Si sólo uno de los parámetros a, b, c es cero, $\dim V_2 = 2$. c) $a = b = c = 0$.

5. (a) Los valores propios de A son 1, 2 y -1. Los de B son 0 y 3, de mult. alg. 2 y 1, respectivamente.

(b) Para la matriz A: $V_1 = \mathcal{C}\ell\{(3,2,1)\}$, $V_2 = \mathcal{C}\ell\{1,3,1)\}$, $V_{-1} = \mathcal{C}\ell\{(1,0,1)\}$. Para la matriz B :
$V_0 = \mathcal{C}\ell\{(-1,1,0),(-1,0,1)\}$ y $V_3 = cl\{(1,1,1)\}$.

6. a) $\lambda = 2a+1$ y una base de V_λ es $\{(1,1,1)\}$.

b) Una base de V_{1-a} es $\{(-1,1,0),(-1,0,1)\}$.

7. a) Los valores propios de A son -1 y 3, de mult. alg. 1.

b) $V_3 = \mathcal{C}\ell\{(1,1,0,0)\}$ y $V_{-1} = \mathcal{C}\ell\{(-1,1,0,0)\}$. La suma de las dimensiones de los espacios propios es menor que 4, por eso no existe una base de $\mathbb{R}^4$ formada por vectores propios de C.

9. a) $p(x = -(x-1)(x-2)(x+1)$ y
$\mathcal{B} = \{(-1,1,2),(1,-1,1),(1,1,0)\}$.

b) $H = \begin{pmatrix} \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \end{pmatrix}$

16. (a) T es ortogonalmente diagonalizable puesto que la matriz de T en la base canónica es simétrica.

(b) La nulidad de T es 1.

(c) $\lambda = 0$ es valor propio de T puesto que $\operatorname{Nuc}(T) \neq \{0\}$.

19. En cada caso los ejes principales son las rectas generadas por los vectores que se dan.

(a) $v_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ 1 \end{pmatrix}, v_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} -1 \\ 3 \end{pmatrix}$,

(b)$v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}$,

(d) La gráfica es dos rectas,

(e) $v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$,

(f) $v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$,

(g) $v_1 = \begin{pmatrix} 1 \\ \sqrt{2}+1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -\sqrt{2}+1 \end{pmatrix}$

20. (a) La ecuación canónica es:

$$-\frac{y_1^2}{\left(\frac{1}{\sqrt{2}}\right)^2} + \frac{\left(y_2+\frac{1}{2}\right)^2}{\left(\frac{1}{2}\right)2} = 1.$$

Es una hipérbola con centro $\left(0, -\frac{1}{2}\right)$.

(b) Se sabe que $v_1 = (\cos\theta, \sin\theta)$ y $v_2 = (-\sin\theta, \cos\theta)$ son las direcciones de los ejes principales. Por lo tanto las ecuaciones vectoriales son:

$L_1 : (x, y) = t\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$

$L_2 : (x, y) = t\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$; $t \in I\!R$

(d) $5x_1^2 - x_2^2 - 6\sqrt{3}x_1x_2 - 4\sqrt{3}x_1 + 4x_2 = 0.$

23. (a) $A = \begin{pmatrix} 1 & a & 1 \\ a & 1 & b \\ 1 & b & 1 \end{pmatrix}$, $B = \left(\sqrt{3}, -\sqrt{3}, \sqrt{3}\right)$, $x^t = (x_1\, x_2\, x_3)$ y $d = 0$.

Para que P sea un punto de la superficie $a = -1$ y $b \in I\!R$.

(b)$\mathcal{B} = \{(\frac{-1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})^t, (\frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}})^t, (\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^t\}$. El cambio de variable requerido es $y = P^t x$, donde P es la matriz con columnas iguales a los vectores de la base $\mathcal{B}$. La ecuación se reduce a: $3y_3^2 + 3y_3 = 0$.

(c) $[P]_{\mathcal{B}} = (\frac{-1}{\sqrt{2}}, \frac{3}{\sqrt{6}}, 0)^t$ y claramente satisface la ecuación porque $y_3 = 0$.

Bibliografía

[1] H. Anton. *Algebra Lineal.* Editorial Limusa.

[2] Apostol, Tom. *Calculus, Vol II.* Editorial Reverté, España, 1978.

[3] Barahona D., Manuel.*Investigacin de Operaciones I* (versin preliminar). Universidad de Atacama (Chile), 1995.

[4] Diaday, E., Lemaire, J., Pouget, J. et Testu, F. *Éléments d'analyse de données.* Dunod, Paris, 1989.

[5] Halmos P. *Espacios Vectoriales Finito-Dimensionales.* CECSA, México, 1965.

[6] Hamdy A., Taha.*Investigacin de Operaciones : una introduccin* (traduccin de Jos de Jess Acosta F.). Representaciones y Servicios de Ingeniera, S.A. Mxico, 1981.

[7] Hoffman, K. y Kunze, R.*Algebra Lineal* (traduccin de Hugo A. Finsterbusc). Ediciones Zacatenco, Mxico, 1987.

[8] Lang, Serge. *Algebra Lineal.* Fondo Educativo Interamericano, Bogotá, 1974.

[9] Larson, E. *Introduccin al Algebra Lineal.* Limusa, Mxico, 1995.

[10] Schrage, Linus. *Linear and Quadratic Programming with LINDO.* The Scientific Press, Palo Alto, California.

[11] Thie, Paul R.*An Introduction to Linear Programming and Game Theory* (Second Edition). Jhon Wiley & Sons, N.Y., 1988.