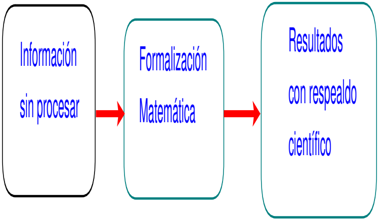
Concavidad y Optimización

Abelardo Jordán Liza

Maestría en Matemáticas Aplicadas
PUCP

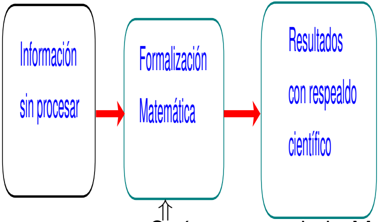
Lima, Agosto 2023

Información cuantificable se organiza para ser procesada, pasa esta etapa se generan resultados con respaldo científico.



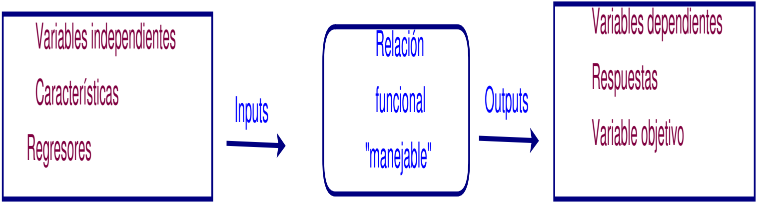
Las matemáticas como herramientas en el trabajo científico

Información cuantificable se organiza para ser procesada, pasa esta etapa se generan resultados con respaldo científico.

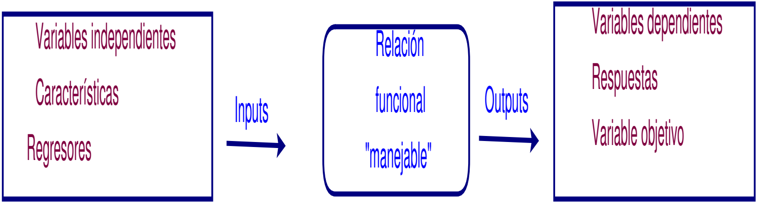


¿Qué aspectos de la Matemática?

Hacia donde se orienta este curso : ¿Los elementos esenciales?



Hacia donde se orienta este curso : ¿Los elementos esenciales?



NOTACIÓN:
Para identificar una característica cuantificable, se usan las notaciones literales.
Es usual el empleo de letras como x para referirse a los valores que tomará una variable independiente.
En el caso de m variables independientes: $x_1, x_2, \cdots, x_m$ donde x_i se refiere a los valores que toma la característica i.

EN el caso de las variables dependientes, podemos encontrar letras como y para una variable dependiente(una respuesta).
En el caso de m varias variable dependientes, podemos emplear letras como

$$y_1, \cdots, y_m, \text{ donde } y_j \text{ se refiere a los valores de la j-ésima variable dependiente}$$

$$Variable(s)\ respuesta\ \textbf{depende(n)}\ de\ las\ variables\ independientes$$

- Una variable independiente y una variable dependiente : $y = f(x)$.

$$Variable(s)\ respuesta\ \textbf{depende(n)}\ de\ las\ variables\ independientes$$

- Una variable independiente y una variable dependiente : $y = f(x)$.
- Varias variables independientes y una variable dependiente $y = f(x_1, \cdots, x_n)$.

$$Variable(s)\ respuesta\ \textbf{depende(n)}\ de\ las\ variables\ independientes$$

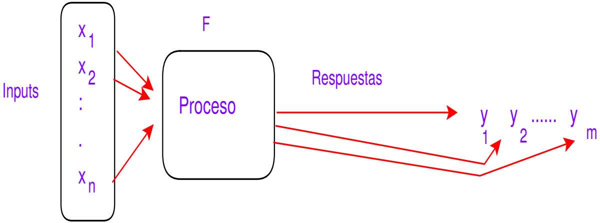
- Una variable independiente y una variable dependiente : $y = f(x)$.
- Varias variables independientes y una variable dependiente $y = f(x_1, \cdots, x_n)$.
- Varias variables independientes y varias variables dependients:

$$\begin{aligned} y_1 &= f_1(x_1, \cdots, x_n) \\ y_2 &= f_2(x_1, \cdots, x_n) \\ \vdots &\quad \vdots \\ y_m &= f_m(x_1, \cdots, x_n). \end{aligned}$$

$$Variable(s)\ respuesta\ \textbf{depende(n)}\ de\ las\ variables\ independientes$$

- Una variable independiente y una variable dependiente : $y = f(x)$.
- Varias variables independientes y una variable dependiente $y = f(x_1, \cdots, x_n)$.
- Varias variables independientes y varias variables dependients:

$$\begin{aligned} y_1 &= f_1(x_1, \cdots, x_n) \\ y_2 &= f_2(x_1, \cdots, x_n) \\ \vdots & \quad \vdots \\ y_m &= f_m(x_1, \cdots, x_n). \end{aligned}$$



En esta parte consideremos una variable dependiente y como función de las variables independientes(de valor real) $x_1, x_2, \cdots, x_n$, que podemos escribir en la forma $y = f(x_1, x_2, \cdots, x_n)$.

En esta parte consideremos una variable dependiente y como función de las variables independientes(de valor real) $x_1, x_2, \cdots, x_n$, que podemos escribir en la forma $y = f(x_1, x_2, \cdots, x_n)$.En primer lugar, se debe establecer el rango de valores que toman las variables independientes.

- Para una variable independiente: Se dice que x toma valores en un conjunto C, donde $C \subset \mathbb{R}$.
  C puede ser un conjunto discreto por ejemplo $C = \mathbb{Z}$.
  C puede ser un intervalo como por ejemplo $[a, b$ de $\mathbb{R}$.
  C puede ser un intervalo no limitado como $[0, +\infty)$.
  C en un caso extremo puede ser $\mathbb{R}$. En este caso, se dice que la variable es "libre".

Rango de las variables

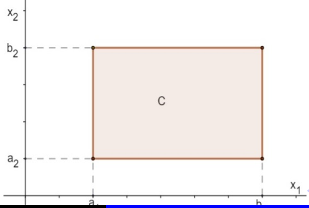
En esta parte consideremos una variable dependiente y como función de las variables independientes(de valor real) $x_1, x_2, \cdots, x_n$, que podemos escribir en la forma $y = f(x_1, x_2, \cdots, x_n)$.En primer lugar, se debe establecer el rango de valores que toman las variables independientes.

- Para una variable independiente: Se dice que x toma valores en un conjunto C, donde $C \subset \mathbb{R}$.
  C puede ser un conjunto discreto por ejemplo $C = \mathbb{Z}$.
  C puede ser un intervalo como por ejemplo $[a, b$ de $\mathbb{R}$.
  C puede ser un intervalo no limitado como $[0, +\infty)$.
  C en un caso extremo puede ser $\mathbb{R}$. En este caso, se dice que la variable es "libre".En los anteriores es de valores restringidos.

Rango de las variables

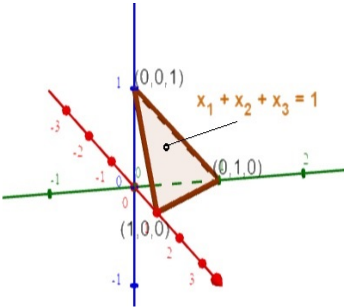
En esta parte consideremos una variable dependiente y como función de las variables independientes(de valor real) $x_1, x_2, \cdots, x_n$, que podemos escribir en la forma $y = f(x_1, x_2, \cdots, x_n)$.En primer lugar, se debe establecer el rango de valores que toman las variables independientes.

- Para una variable independiente: Se dice que x toma valores en un conjunto C, donde $C \subset \mathbb{R}$.
  C puede ser un conjunto discreto por ejemplo $C = \mathbb{Z}$.
  C puede ser un intervalo como por ejemplo $[a, b$ de $\mathbb{R}$.
  C puede ser un intervalo no limitado como $[0, +\infty)$.
  C en un caso extremo puede ser $\mathbb{R}$. En este caso, se dice que la variable es "libre".En los anteriores es de valores restringidos.
- Para varias variables independientes:
  (i) Las variables $x_1, x_2, \cdots, x_n$ son tales que $a_1 \leq x_1 \leq b_1$, $a_2 \leq x_2 \leq b_2$, $\cdots, a_n \leq x_n \leq b_n$. En tal caso C tiene la forma de una caja :
  $C = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$.
  En el ambiente bidimensional, $C = [a_1, b_1] \times [a_2, b_2]$.



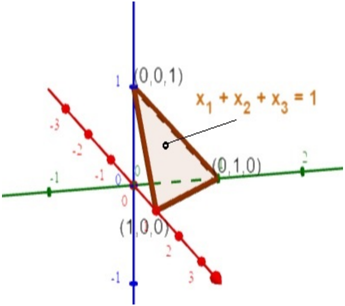
(ii) Las variables independientes están relacionadas entre si, por una relación funcional(o varias) del tipo

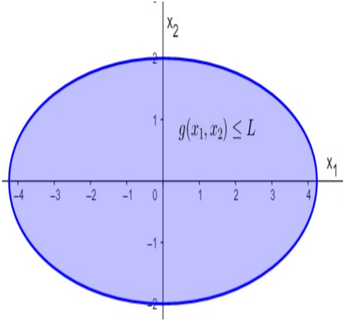
$$g(x_1, x_2, \cdots, x_n) = (\leq)L.$$



(ii) Las variables independientes están relacionadas entre si, por una relación funcional(o varias) del tipo

$$g(x_1, x_2, \cdots, x_n) = (\leq) L.$$



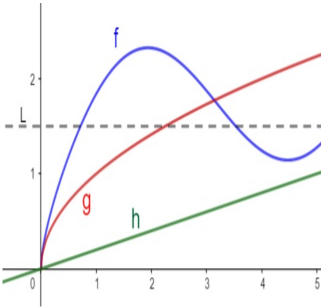


Nivel de respuesta

Los conjuntos C antes descritos se denominan " conjuntos de factibilidad para la toma de decisiones".

Los conjuntos C antes descritos se denominan " conjuntos de factibilidad para la toma de decisiones".
(Para funciones de valor real)Dado un número real L, para la función $y = f(x_1, \cdots, x_n)$,se dice que L es un nivel alcanzado en C por y si

$$\text{Existe } \overline{x} \in C \text{ tal que } f(\overline{x}) = L.$$



Para la función f, existen dos puntos x en $[0, 5]$ tales que $f(x) = L$.
Para la función g, existe un único punto x en $[0, 5]$ tal que $g(x) = L$. Para la función h, la ecuación $h(x) = L$ no tiene solución en $[0, 5]$.

- Para una función $y = f(x)$ con $x \in C$. Dado un número real m, el conjunto

$$L_m(f) := \{x \in C : f(x) = m\}$$

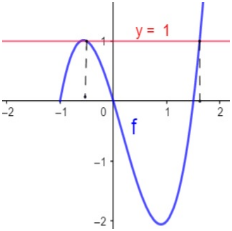
se llama conjunto de nivel m para f en C. Si f alcanza dicho nivel, entonces $L_m(f)$ es no vacío.

Conjunto de nivel o contorno

- Para una función $y = f(x)$ con $x \in C$. Dado un número real m, el conjunto
$$L_m(f) := \{x \in C : f(x) = m\}$$
se llama conjunto de nivel m para f en C. Si f alcanza dicho nivel, entonces $L_m(f)$ es no vacío.
- Para una función de varias variables $y = f(x_1, \cdots, x_n)$ con $(x_1, \cdots, x_n) \in C$. Dado un número real m, el conjunto
$$L_m(f) := \{x = (x_1, \cdots, x_n) \in C : f(x) = m\}$$
se llama conjunto(curva) "contorno" de nivel m para f en C.

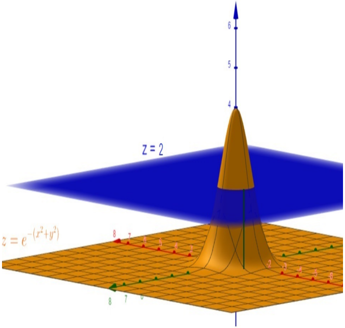
- Para una función $y = f(x)$ con $x \in C$. Dado un número real m, el conjunto
$$L_m(f) := \{x \in C : f(x) = m\}$$
se llama conjunto de nivel m para f en C. Si f alcanza dicho nivel, entonces $L_m(f)$ es no vacío.

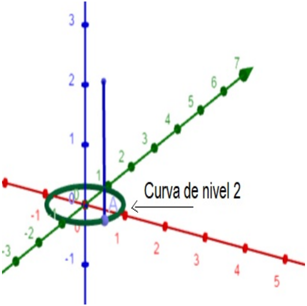
- Para una función de varias variables $y = f(x_1, \cdots, x_n)$ con $(x_1, \cdots, x_n) \in C$. Dado un número real m, el conjunto
$$L_m(f) := \{x = (x_1, \cdots, x_n) \in C : f(x) = m\}$$
se llama conjunto(curva) "contorno" de nivel m para f en C.



En la figura, la función $f(x) = 2x^3 - x^2 - 3x, -2 \leq x \leq 2$
El conjunto de nivel 1 resulta de resolver la ecuación
$2x^3 - x^2 - 3x = 1, -2 \leq x \leq 2.$
Este conjunto es $\{-0,618; -0{,}50; 1,6218\}.$

Para la función $z = f(x,y) = 4e^{-(x^2+y^2)}$ y el contorno de nivel $\{(x,y) : f(x,y) = 2\} = \{(x,y) : x^2 + y^2 = ln(2).\}$





Dados dos conjuntos arbitrarios A y B y una función $f : A \to B$. Dado $m \in B$, el contorno de f de nivel m está dado por

$$L_m(f) := \{x \in A : f(x) = m.\}$$

¿De qué naturaleza son A y B?

Dados dos conjuntos arbitrarios A y B y una función $f : A \to B$. Dado $m \in B$, el contorno de f de nivel m está dado por

$$L_m(f) := \{x \in A : f(x) = m.\}$$

¿De qué naturaleza son A y B?

Ejemplo:

Sea la función $f : \mathbb{R}^2 \to \mathbb{Z}^2$ definido por $f(x, y) := (i, j)$ siempre que $(x, y) \in [i - 1/2, i + 1/2[\times[j - 1/2, j + 1/2[$.

En especial se cumple $f(x, y) = (x, y)$ si y solamente si $(x, y) \in \mathbb{Z}^2$.

Dados dos conjuntos arbitrarios A y B y una función $f : A \to B$. Dado $m \in B$, el contorno de f de nivel m está dado por

$$L_m(f) := \{x \in A : f(x) = m.\}$$

¿De qué naturaleza son A y B?

Ejemplo:

Sea la función $f : \mathbb{R}^2 \to \mathbb{Z}^2$ definido por $f(x, y) := (i, j)$ siempre que $(x, y) \in [i - 1/2, i + 1/2[\times[j - 1/2, j + 1/2[$.

En especial se cumple $f(x, y) = (x, y)$ si y solamente si $(x, y) \in \mathbb{Z}^2$.

Determine el conjunto $L_{(i,j)}(f)$.

Dados dos conjuntos arbitrarios A y B y una función $f : A \to B$. Dado $m \in B$, el contorno de f de nivel m está dado por

$$L_m(f) := \{x \in A : f(x) = m.\}$$

¿De qué naturaleza son A y B?

Ejemplo:

Sea la función $f : \mathbb{R}^2 \to \mathbb{Z}^2$ definido por $f(x, y) := (i, j)$ siempre que $(x, y) \in [i - 1/2, i + 1/2[\times[j - 1/2, j + 1/2[$.

En especial se cumple $f(x, y) = (x, y)$ si y solamente si $(x, y) \in \mathbb{Z}^2$.

Determine el conjunto $L_{(i,j)}(f)$.

Ejemplo:

Sean $\mathcal{S}_2$ el conjunto de matrices simétricas reales de orden 2. Se define la función

$f : \mathcal{S}_2 \to \mathbb{R}^2$ por

$f(M) = (\lambda_1, \lambda_2)$ donde $\lambda_1 \leq \lambda_2$ son los valores propios de M.

Determine el conjunto $L_{(0,1)}(f)$.

Sea V un conjunto no vacío y $+$ una operación binaria en V. El complejo $(V, +)$ es un **grupo** abeliano, si las siguientes se satisfacen:

(i) $(x + y) + z = x + (y + z), \quad \forall x, y, z \in V.$

(ii) $\exists 0 \in V$ tal que $x + 0 = x + 0 = x, \quad \forall x \in V.$

(iii) Para cada $x \in V$, existe $-x \in V$ tal que $x + (-x) = -x + x = 0.$

(iv) $x + y = y + x$, $\forall x, y \in V.$

Ejemplos

(a) $(\mathbb{R}^n, +)$, *donde la operación* $+$ *está definida de la siguiente manera:*
Si $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n)$ *son elementos de* $\mathbb{R}^n$*, se define*

$$x + y := (x_1 + y_1, \cdots, x_n + y_n)$$

(b) *Si* $\mathcal{M}_{m\times n}$ *es el conjunto de matrices reales de orden* $m \times n$ *con la operación* $+$ *definida por :*
Sean $A = [a_{ij}]$ *y* $B = [b_{ij}]$ *elementos de* $\mathcal{M}_{m\times n}$ *, se define*

$$A + B := [a_{ij} + b_{ij}], \quad \textit{para } i = 1, \cdots, m; j = 1, \cdots, n.$$

Sea $(V, +)$ un grupo abeliano y $\cdot$ una aplicación definida en $\mathbb{R} \times V$ con valores en V tal que a cada $(\alpha, x) \in \mathbb{R} \times V$ le asigna un elemento $\alpha \cdot x \in V$. Este último elemento se denotará por αx.

Definición

El complejo $(V, +, \cdot)$ *es un espacio vectorial (real), si* $(V, +)$ *es un grupo abeliano, y se satisfacen:*

(i) $\alpha(\beta x) = (\alpha\beta)x \quad \forall \alpha, \beta \in \mathbb{R}, x \in V.$

(ii) $(\alpha + \beta)x = \alpha x + \beta x$ *y* $\alpha(x + y) = \alpha x + \alpha y, \quad \forall \alpha, \beta \in \mathbb{R}, x, y \in V.$

(iii) $1x = x, \ \forall x \in V.$

Los elementos de V se denominan vectores y particularmente el elemento $0 \in V$ se denomina el vector cero o vector nulo.

Ejemplos

(a) $(\mathbb{R}, +, \cdot)$ *con las operaciones usuales de adición y multiplicación de números reales, constituye un espacio vectorial.*

(b) *Para* $n \in \mathbb{N}$, $(\mathbb{R}^n, +, \cdot)$ *es un espacio vectorial real, donde el producto por escalar* "$\cdot$" *se define por*
$\alpha \in \mathbb{R}, x = (x_1, \cdots, x_n) \in \mathbb{R}^n : \; \alpha x := (\alpha x_1, \cdots, \alpha x_n)$.

(c) $(\mathcal{M}_{m\times n}, +, \cdot)$ *donde* $+$ *es la operación usual de adición de matrices y* $\cdot$ *es la multiplicación usual de una matriz por un escalar.*

Existen otros ejemplos de espacios vectoriales que son de relevancia para un curso de Optimización, como por ejemplo si $V = \mathbb{R}^{[0,1]}$ es el conjunto de funciones de valor real definidas en el intervalo $[0,1]$. En V se definen la adición y multiplicación por escalar, mediante:

(i) $f, g \in V : f + g$ es la función definida en $[0,1]$ por $(f+g)(x) = f(x) + g(x)$, para cada $x \in [0,1]$.

(ii) Para $f \in V, \alpha \in \mathbb{R}$, se define la función αf en $[0,1]$ por $(\alpha f)(x) = \alpha f(x)$ para cada $x \in [0,1]$.

En tal caso, $(\mathbb{R}^{[0,1]}, +, \cdot)$ constituye un espacio vectorial.

COMBINACIÓN LINEAL: Dada una colección $x_1, x_2, \cdots, x_p$ de vectores de V y números reales $\alpha_1, \alpha_2, \cdots, \alpha_p$; el vector

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p \tag{1}$$

se llama una combinación lineal de los vectores $x_1, x_2, \cdots$ y x_p.
Si z es el vector (1), también se dice que z es combinación lineal del conjunto $\{x_1, x_2, \cdots, x_p\}$.

Ejemplos:

(a) En $\mathbb{R}$, $z = 4$ es combinación lineal de los vectores 3 y 5.

(b) En $\mathbb{R}^2, z = (3, 8)$ es combinación lineal de $(1, 3)$ y $(2, 4)$

(c) En $\mathbb{R}^3, z = (1, -1, 0)$ es combinación lineal de los vectores $(2, 1, 1), (0, 1, 0)$ y $(0, 0, -1)$.

(d) En $\mathcal{M}_{2\times3}$, el vector $A = \begin{bmatrix} 4 & -8 & 9 \\ -4 & -3 & -1 \end{bmatrix}$ es combinación lineal de los vectores $\begin{bmatrix} 2 & 1 & -3 \\ 1 & 0 & 1 \end{bmatrix}$ y $\begin{bmatrix} 0 & 2 & -1 \\ 2 & 1 & 1 \end{bmatrix}$.

Se dice que un conjunto finito de vectores $\mathbb{X} = \{x_1, x_2, \cdots, x_p\} \subset V$ es linealmente dependiente, si al menos uno de sus elementos se puede escribir como una combinación lineal de los otros elementos de $\mathbb{X}$. Equivalentemente, $\mathbb{X}$ es linealmente dependiente, si existen escalares $\alpha_1, \alpha_2, \cdots, \alpha_p$ no todos nulos tales que

$$\sum_{i=1}^{p} \alpha_i x_i = 0$$

Se dice que un conjunto finito de vectores $\mathbb{X} = \{x_1, x_2, \cdots, x_p\} \subset V$ es linealmente dependiente, si al menos uno de sus elementos se puede escribir como una combinación lineal de los otros elementos de $\mathbb{X}$. Equivalentemente, $\mathbb{X}$ es linealmente dependiente, si existen escalares $\alpha_1, \alpha_2, \cdots, \alpha_p$ no todos nulos tales que

$$\sum_{i=1}^{p} \alpha_i x_i = 0$$

Nota: Si $\mathbb{X}$ es linealmente dependiente, entonces cualquier colección finita $\mathbb{Y}$ tal que $\mathbb{X} \subset \mathbb{Y}$, será también linealmente dependiente.

Se dice que un conjunto finito de vectores $\mathbb{X} = \{x_1, x_2, \cdots, x_p\} \subset V$ es linealmente dependiente, si al menos uno de sus elementos se puede escribir como una combinación lineal de los otros elementos de $\mathbb{X}$. Equivalentemente, $\mathbb{X}$ es linealmente dependiente, si existen escalares $\alpha_1, \alpha_2, \cdots, \alpha_p$ no todos nulos tales que

$$\sum_{i=1}^{p} \alpha_i x_i = 0$$

Nota: Si $\mathbb{X}$ es linealmente dependiente, entonces cualquier colección finita $\mathbb{Y}$ tal que $\mathbb{X} \subset \mathbb{Y}$, será también linealmente dependiente.
Para $\mathbb{X}$ un conjunto infinito de elementos de V, se dice que $\mathbb{X}$ es linealmente dependiente si existe una subcolección finita de $\mathbb{X}$ que es linealmente dependiente.

Se dice que un conjunto finito de vectores $\mathbb{X} = \{x_1, x_2, \cdots, x_p\} \subset V$ es linealmente dependiente, si al menos uno de sus elementos se puede escribir como una combinación lineal de los otros elementos de $\mathbb{X}$. Equivalentemente, $\mathbb{X}$ es linealmente dependiente, si existen escalares $\alpha_1, \alpha_2, \cdots, \alpha_p$ no todos nulos tales que

$$\sum_{i=1}^{p} \alpha_i x_i = 0$$

Nota: Si $\mathbb{X}$ es linealmente dependiente, entonces cualquier colección finita $\mathbb{Y}$ tal que $\mathbb{X} \subset \mathbb{Y}$, será también linealmente dependiente.
Para $\mathbb{X}$ un conjunto infinito de elementos de V, se dice que $\mathbb{X}$ es linealmente dependiente si existe una subcolección finita de $\mathbb{X}$ que es linealmente dependiente.

Una familia de vectores $\mathbb{X}$ es **linealmente independiente** si no es linealmente dependiente, esto significa que para cualquier subfamilia finita de $\mathbb{X}$, $\{x_i \in \mathbb{X} : i \in I\}$ (I un conjunto finito), tenemos que

$$\sum_{i \in I} \alpha_i x_i = 0 \Rightarrow \alpha_i = 0, \ \forall i \in I.$$

Se dice que un conjunto finito de vectores $\mathbb{X} = \{x_1, x_2, \cdots, x_p\} \subset V$ es linealmente dependiente, si al menos uno de sus elementos se puede escribir como una combinación lineal de los otros elementos de $\mathbb{X}$. Equivalentemente, $\mathbb{X}$ es linealmente dependiente, si existen escalares $\alpha_1, \alpha_2, \cdots, \alpha_p$ no todos nulos tales que

$$\sum_{i=1}^{p} \alpha_i x_i = 0$$

Nota: Si $\mathbb{X}$ es linealmente dependiente, entonces cualquier colección finita $\mathbb{Y}$ tal que $\mathbb{X} \subset \mathbb{Y}$, será también linealmente dependiente.
Para $\mathbb{X}$ un conjunto infinito de elementos de V, se dice que $\mathbb{X}$ es linealmente dependiente si existe una subcolección finita de $\mathbb{X}$ que es linealmente dependiente.

Una familia de vectores $\mathbb{X}$ es **linealmente independiente** si no es linealmente dependiente, esto significa que para cualquier subfamilia finita de $\mathbb{X}$, $\{x_i \in \mathbb{X} : i \in I\}$ (I un conjunto finito), tenemos que

$$\sum_{i \in I} \alpha_i x_i = 0 \Rightarrow \alpha_i = 0, \ \forall i \in I.$$

" Si la colección es $\{x_1, x_2\}$" entonces la independencia lineal de esta colección indica que cada vector no es múltiplo del otro.
Si en una colección $\mathbb{X}$ está presente el elemento cero, entonces $\mathbb{X}$ es linealmente dependiente.

Sea V un espacio vectorial, se dice que un subconjunto W de V, genera V si todo elemento de V se puede escribir como una combinación lineal de una colección finita de elementos de W. Esto significa

$$\forall x \in V, \exists \text{ escalares } \alpha_1, \alpha_2, \cdots, \alpha_k, \text{ y vectores } x_1, x_2, \cdots, x_k \in W : x = \sum_{i=1}^{k} \alpha_i x_i.$$

EJEMPLOS:

(i) Si $V = \mathbb{R}$ es el conjunto de los números reales, entonces $W = \{1\}$ genera V (también $W = \{-1, 1\}$ genera V.)

Sea V un espacio vectorial, se dice que un subconjunto W de V, genera V si todo elemento de V se puede escribir como una combinación lineal de una colección finita de elementos de W. Esto significa

$$\forall x \in V, \exists \text{ escalares } \alpha_1, \alpha_2, \cdots, \alpha_k, \text{ y vectores } x_1, x_2, \cdots, x_k \in W : x = \sum_{i=1}^{k} \alpha_i x_i.$$

EJEMPLOS:

(i) Si $V = \mathbb{R}$ es el conjunto de los números reales, entonces $W = \{1\}$ genera V (también $W = \{-1, 1\}$ genera V.)

(ii) $V = \mathbb{R}^2$ es generado por $W = \{(1,0); (0,-1)\}$ dado que cualquier elemento $x = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ puede escribirse como $x = \alpha_1(1,0) + (-\alpha_2)(0,-1)$.

Sea V un espacio vectorial, se dice que un subconjunto W de V, genera V si todo elemento de V se puede escribir como una combinación lineal de una colección finita de elementos de W. Esto significa

$$\forall x \in V, \exists \text{ escalares } \alpha_1, \alpha_2, \cdots, \alpha_k, \text{ y vectores } x_1, x_2, \cdots, x_k \in W : x = \sum_{i=1}^{k} \alpha_i x_i.$$

EJEMPLOS:

(i) Si $V = \mathbb{R}$ es el conjunto de los números reales, entonces $W = \{1\}$ genera V (también $W = \{-1, 1\}$ genera V.)

(ii) $V = \mathbb{R}^2$ es generado por $W = \{(1,0); (0,-1)\}$ dado que cualquier elemento $x = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ puede escribirse como $x = \alpha_1(1,0) + (-\alpha_2)(0,-1)$.

(iii) $\mathbb{R}^2$ también es generado por $\{(0,1), (3,1), (1,1)\}$.

Sea V un espacio vectorial, se dice que un subconjunto W de V, genera V si todo elemento de V se puede escribir como una combinación lineal de una colección finita de elementos de W. Esto significa

$$\forall x \in V, \exists \text{ escalares } \alpha_1, \alpha_2, \cdots, \alpha_k, \text{ y vectores } x_1, x_2, \cdots, x_k \in W : x = \sum_{i=1}^{k} \alpha_i x_i.$$

EJEMPLOS:

(i) Si $V = \mathbb{R}$ es el conjunto de los números reales, entonces $W = \{1\}$ genera V (también $W = \{-1, 1\}$ genera V.)

(ii) $V = \mathbb{R}^2$ es generado por $W = \{(1,0); (0,-1)\}$ dado que cualquier elemento $x = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ puede escribirse como $x = \alpha_1(1,0) + (-\alpha_2)(0,-1)$.

(iii) $\mathbb{R}^2$ también es generado por $\{(0,1), (3,1), (1,1)\}$.

(iv) $V = \{(\alpha, 0) : \alpha \in \mathbb{R}\}$ es generado por $\cdots$?

Sea V un espacio vectorial, se dice que un subconjunto W de V, genera V si todo elemento de V se puede escribir como una combinación lineal de una colección finita de elementos de W. Esto significa

$$\forall x \in V, \exists \text{ escalares } \alpha_1, \alpha_2, \cdots, \alpha_k, \text{ y vectores } x_1, x_2, \cdots, x_k \in W : x = \sum_{i=1}^{k} \alpha_i x_i.$$

EJEMPLOS:

(i) Si $V = \mathbb{R}$ es el conjunto de los números reales, entonces $W = \{1\}$ genera V (también $W = \{-1, 1\}$ genera V.)

(ii) $V = \mathbb{R}^2$ es generado por $W = \{(1,0); (0,-1)\}$ dado que cualquier elemento $x = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ puede escribirse como $x = \alpha_1(1,0) + (-\alpha_2)(0,-1)$.

(iii) $\mathbb{R}^2$ también es generado por $\{(0,1), (3,1), (1,1)\}$.

(iv) $V = \{(\alpha, 0) : \alpha \in \mathbb{R}\}$ es generado por $\cdots$?

(v) $V = \{(0, x, -x) : x \in \mathbb{R}\}$ es generado por {(0,1,-1)}.

Sea V un espacio vectorial, se dice que un subconjunto W de V, genera V si todo elemento de V se puede escribir como una combinación lineal de una colección finita de elementos de W. Esto significa

$$\forall x \in V, \exists \text{ escalares } \alpha_1, \alpha_2, \cdots, \alpha_k, \text{ y vectores } x_1, x_2, \cdots, x_k \in W : x = \sum_{i=1}^{k} \alpha_i x_i.$$

EJEMPLOS:

(i) Si $V = \mathbb{R}$ es el conjunto de los números reales, entonces $W = \{1\}$ genera V (también $W = \{-1, 1\}$ genera V.)

(ii) $V = \mathbb{R}^2$ es generado por $W = \{(1,0); (0,-1)\}$ dado que cualquier elemento $x = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ puede escribirse como $x = \alpha_1(1,0) + (-\alpha_2)(0,-1)$.

(iii) $\mathbb{R}^2$ también es generado por $\{(0,1), (3,1), (1,1)\}$.

(iv) $V = \{(\alpha, 0) : \alpha \in \mathbb{R}\}$ es generado por $\cdots$?

(v) $V = \{(0, x, -x) : x \in \mathbb{R}\}$ es generado por {(0,1,-1)}.

(vi) $V = \{(x, 0, y) : x, y \in \mathbb{R}\}$ es generado por $\cdots$?

Definición

Dado un espacio vectorial V. Un subconjunto $\mathcal{B}$ de V se denomina una base de V, si $\mathcal{B}$ es un conjunto linealmente independiente y genera V.

EJEMPLOS:

(i) $\{1\}$ es una base de $\mathbb{R}$.

(ii) $\mathcal{B} = \{(1,0),(0,1)\}$ es una base de $\mathbb{R}^2$, $\mathcal{C} = \{(1,2),(0,-1)\}$ es también una base de $\mathbb{R}^2$

(iii) $\mathcal{B} = \{\underbrace{(1,0,\cdots,0)}_{e_1}, \underbrace{(0,1,\cdots,0)}_{e_2}, \cdots, \underbrace{(0,0,\cdots,1)}_{e_n}\}$ es una base de $\mathbb{R}^n$, llamada la base canónica.

(iv) En el espacio vectorial $\mathcal{M}_{m\times n}$ (de las matrices reales de orden $m \times n$), si A_{ij} es la matriz cuyo ij-elemento es uno y los otros son cero, entonces la colección de matrices $\mathcal{B} = \{A_{ij} : i = 1,\cdots,m; j = 1,\cdots,n\}$ es una base de $\mathcal{M}_{m\times n}$.

IMPORTANTE!:

(i) Si un espacio vectorial V tiene una base con un número finito k de elementos, entonces toda base también tendrá k elementos. En tal caso, se dice que la dimensión de V es k y se escribe $dim(V) = k$.

(ii) Una colección de n vectores de $\mathbb{R}^n$ será una base de $\mathbb{R}^n$, si esta colección es linealmente independiente.

IMPORTANTE!:

(i) Si un espacio vectorial V tiene una base con un número finito k de elementos, entonces toda base también tendrá k elementos. En tal caso, se dice que la dimensión de V es k y se escribe $dim(V) = k$.

(ii) Una colección de n vectores de $\mathbb{R}^n$ será una base de $\mathbb{R}^n$, si esta colección es linealmente independiente.

$$\boxed{\mathbb{R}^n \textbf{ es un espacio vectorial n-dimensional}}$$

(iii) Si $\mathcal{B} = \{v^1, \cdots, v^n\}$ es una base de $\mathbb{R}^n$ y A es una matriz invertible de orden n, entonces el conjunto $\{Av^1, \cdots, Av^n\}$ es también una base de $\mathbb{R}^n$.

(iv) $\mathcal{M}_{m\times n}$ es un espacio vectorial de dimensión $m \times n$.

Definición

Dado un espacio vectorial real V, se dice que la aplicación $\langle\ \cdot\ ,\ \cdot \rangle$ definida en $V \times V$ con valores en $\mathbb{R}$, es un producto interno en V, si satisface las siguientes condiciones:

(i) $\langle x, x\rangle > 0, \forall\ x \in V \setminus \{0\}$ $(\langle x, x\rangle = 0,$ *si y solo si* $, x = 0)$

(ii) $\langle x, y\rangle = \langle y, x\rangle,\ \forall x, y \in V.$

(iii) $\langle \alpha x + y, z\ \rangle = \alpha\langle x, z\rangle + \langle y, z\rangle,\ \forall x, y, z \in V, \alpha \in \mathbb{R}.$

También se suele denotar al producto interno de los vectores $\mathbf{x}$, $\mathbf{y}$ por $\mathbf{x}.\mathbf{y}$
Si V es un espacio vectorial dotado de un producto interno $\langle\ \rangle$, entonces se dice que V es un espacio producto interno.

Definición

Dado un espacio vectorial real V, se dice que la aplicación $\langle \cdot \, , \, \cdot \rangle$ definida en $V \times V$ con valores en $\mathbb{R}$, es un producto interno en V, si satisface las siguientes condiciones:

(i) $\langle x, x \rangle > 0, \forall\, x \in V \setminus \{0\}$ $(\langle x, x \rangle = 0,$ *si y solo si* $, x = 0)$

(ii) $\langle x, y \rangle = \langle y, x \rangle, \ \forall x, y \in V.$

(iii) $\langle \alpha x + y, z \, \rangle = \alpha \langle x, z \rangle + \langle y, z \rangle, \ \forall x, y, z \in V, \alpha \in \mathbb{R}.$

También se suele denotar al producto interno de los vectores $\mathbf{x}$, $\mathbf{y}$ por $\mathbf{x.y}$
Si V es un espacio vectorial dotado de un producto interno $\langle \ \rangle$, entonces se dice que V es un espacio producto interno.

Propiedades básicas:

(1) Para todo $x \in V$, se cumple $\langle x, 0 \rangle = 0$.

(2) Si para todo $x \in V$ se satisface $\langle x, y \rangle = 0$, entonces $y = 0$.

Producto interno canónico en $\mathbb{R}^n$

Considerando el espacio vectorial $\mathbb{R}^n$ con su base canónica, se define el producto interno (canónico) $\langle\, , \,\rangle$ en $\mathbb{R}^n$ mediante:

$$x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{R}^n : \langle x, y\rangle = \sum_{i=1}^{n} x_i y_i \tag{2}$$

Se verifican (i)-(iii) de la definición de producto interno, particularmente $\langle x, x\rangle = \sum_{i=1}^{n} x_i^2 \geq 0, \forall x \in \mathbb{R}^n$, mientras que $\langle x, x\rangle = 0$, sii $\sum_{i=1}^{n} x_i^2 = 0$ sii $x_i = 0$ para cada $i = 1, \cdots n$, sii $x = 0$.

Nota

(a) *Si los elementos de $\mathbb{R}^n$ son considerados como vectores columna de n componentes, entonces el producto interno canónico de x con y suele denotarse por $x^t y$.*

(b) *Si A es una matriz simétrica definida positiva de orden n, entonces la aplicación $\langle\, , \,\rangle_A$ definida en $\mathbb{R}^n \times \mathbb{R}^n$ por*

$$x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{R}^n : \langle x, y\rangle_A := x^t A y$$

resulta ser un producto interno en $\mathbb{R}^n$.

(c) *Un espacio vectorial de dimensión finita y dotado de un producto interno, se denominará un ESPACIO EUCLIDIANO. $\mathbb{R}^n$ como espacio euclidiano, se asumirá que está dotado del producto interno canónico.*

Se dice que un espacio vectorial V es normado, si existe una aplicación $\begin{array}{rl} \|\ \|: V & \to \mathbb{R} \\ x & \mapsto \|x\| \end{array}$ llamada norma en V que satisface las siguientes condiciones:

(i) $\|x\| \geq 0,\ \forall x \in V$ $(\|x\| = 0 \Leftrightarrow x = 0)$.

(ii) $\|\alpha x\| = |\alpha|\|x\|,\ \forall \alpha \in \mathbb{R},\ x \in V$.

(iii) $\|x + y\| \leq \|x\| + \|y\|,\ \forall x, y \in V$.

La norma en $\mathbb{R}^n$ inducida por el producto interno canónico:

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \cdots + x_n^2}, \qquad \forall x = (x_1, \cdots, x_n) \in \mathbb{R}^n.$$

se denomina *norma euclidiana* en $\mathbb{R}^n$.

Definición

Sea V un espacio normado, dados $\overline{x} \in V, r \in \mathbb{R}_{++}$, se definen los conjuntos

(a) *Bola abierta de centro $\overline{x}$ y radio r:* $\mathcal{B}_r(\overline{x}) := \{x \in V : \|x - \overline{x}\| < r\}$.

(b) *Bola cerrada de centro $\overline{x}$ y radio r:* $\overline{\mathcal{B}}_r(\overline{x}) := \{x \in V : \|x - \overline{x}\| \leq r\}$.

(c) *Esfera de centro $\overline{x}$ y radio r:* $\mathcal{S}_r(\overline{x}) := \{x \in V : \|x - \overline{x}\| = r\}$.

Dada una sucesión $\{x_k\}$ en $\mathbb{R}^n$, se dice que ésta converge a $\ell \in V$, si para cada $\epsilon > 0$, existe $k_o \in \mathbb{N}$ tal que

$$k > k_0 \Rightarrow \|x_k - \ell\| < \epsilon.(\text{ equivalentemente } k > k_0 \Rightarrow x_k \in \mathcal{B}_\epsilon(\ell))$$

En tal caso, se escribe $\lim\limits_{k\to+\infty} x_k = \ell$ y se dice que $\{x_k\}$ es una sucesión convergente.

Propiedades:

(i) Toda sucesión convergente, es acotada.

(ii) Si $\{x_k\}$ y $\{y_k\}$ son sucesiones convergentes y $\alpha \in \mathbb{R}$, entonces $\{\|x_k\|\}, \{\alpha x_k\}, \{x_k \pm y_k\}$ son sucesiones convergentes.

(iii) La sucesión $\{x_k\}$ converge a $\overline{x}$ si, y solamente si $||x_k - \overline{x}|| \to 0$.

Definición

Consideramos a $\mathbb{R}^n$ un espacio métrico cuya métrica es generada por una norma $\|\ \|$.

(a) *$C \subset \mathbb{R}^n$ es un conjunto abierto, si $\forall x \in C$, existe $r > 0$ tal que $\mathcal{B}_r(x) \subset C$.*

(b) *$D \subset \mathbb{R}^n$ es un conjunto cerrado, si su complemento es abierto.*

(c) *Se dice que $\overline{x}$ es u punto adherente de C, si para todo $r > 0$ se cumple $C \cap \mathcal{B}_r(\overline{x}) \neq \emptyset$. El conjunto de tales puntos se denomina la clausura de C y se denota por $cl(C)$ así como por $\overline{C}$.*

En la teoría de optimización, generalmente deseamos tener condiciones "favorables" para el tratamiento de los problemas relacionados. La convexidad de conjuntos, aporta un rico despliegue geométrico-algebraico que debe ser explotado tanto para los conjuntos de las variables de decisión como para otros conjuntos determinados por la función objetivo.
En esta parte, presentamos algunas definiciones y resultados en un espacio vectorial real E, si el caso lo demanda lo especificaremos en el ambiente $\mathbb{R}^n$.

Definición

Para $x, y \in E$, definimos segmentos de extremos x e y, mediante

$$[x, y] := \{x + t(y - x) : t \in [0, 1]\ \}$$

$$[x, y[:= \{x + t(y - x) : t \in [0, 1[\ \}$$

análogamente para $]x, y]$ *y* $]x, y[$.

Sea C un subconjunto de E. Se dice que:

a) C es *convexo*, si

$$\forall x, y \in C, \forall t \in [0,1] \text{ se cumple } \; tx + (1-t)y \in C \tag{3}$$

Es decir, si C contiene a todos los segmentos cerrados que conectan a dos cualesquiera de sus propios puntos . Equivalentemente, C es convexo si $[x, y] \subset C, \; \forall x, y \in C$.

b) C es **afín**, si

$$\forall x, y \in C, \forall t \in \mathrm{R} \text{ se cumple } \; tx + (1-t)y \in C \tag{4}$$

Todo conjunto afín es convexo.

Dado un espacio vectorial E,

(i) E y $\emptyset$ son subconjuntos convexos y afines a la vez de E.

(ii) Si $\{C_i\}_{i\in I}$ es una colección arbitraria de subconjuntos convexos de E, entonces $\cap_{i\in I} C_i$ es un subconjunto convexo de E.

(iii) Si $\{C_i\}_{i\in I}$ es una colección arbitraria de subconjuntos afines de E, entonces $\cap_{i\in I} C_i$ es un subconjunto afín de E.

Definición

a) *Sea* $C \subset E$*, se dice que* $z \in E$ *es combinación convexa de elementos de* C *si existen* $m \in \mathbb{N}$*,* $\{x_i\}_{i=1}^m \subset C$ *y* $\{t_i\}_{i=1}^m \subset [0,1]$ *tales que*

$$\sum_{i=1}^{m} t_i = 1 \quad y \quad z = \sum_{i=1}^{m} t_i x_i \tag{5}$$

b) *Sea* $C \subset E$*, se dice que* $z \in E$ *es combinación afín de elementos de* C *si existen* $m \in \mathbf{N}$*,* $\{x_i\}_{i=1}^m \subset C$ *y* $\{t_i\}_{i=1}^m \subset \mathrm{R}$ *tales que (5).*

Proposición

$C \subset E$ *es convexo (afín) si y solamente si, contiene a cualquier combinación convexa (afín) de sus propios elementos.*

Ejemplos

(i) *El espacio vectorial E y todo conjunto unitario de E son ejemplos de conjuntos convexos y afines a la vez.*

(ii) *Toda bola abierta o cerrada de R^n es un conjunto convexo mas no es afín.*

(iii) *Dado $p \in \mathrm{R}^n$ y $\alpha \in \mathrm{R}$, el conjunto $\{x \in \mathrm{R}^n : p^t x = \alpha\}$ es afín.*

(iv) *El simplex unitario de R^n denotado por $\triangle_n$ es un conjunto convexo y está definido por*

$$\triangle_n := \{x = (x_1, \cdots, x_n) : \sum_{i=1}^{n} x_i = 1, x_i \geq 0, \forall i = \overline{1, n}\}$$

(v) *Una matrix de $\mathcal{M}_n$ se llama estocástica si todas sus entradas son no-negativas y los elementos de cada columna suman uno. Si $\mathcal{E}_n$ es el conjunto de estas matrices, entonces $\mathcal{E}_n$ es un conjunto convexo de $\mathcal{M}_n$.*

Sea A un conjunto afín de $\mathbb{R}^n$, entonces para $a \in A$, el conjunto

$$V_a = A - \{a\} := \{x - a : x \in A\}$$

es un subespacio vectorial de $\mathbb{R}^n$. (Esto es independiente de la elección de a)
En tal caso, se dice que A es un subespacio paralelo a V_a, y se define la dimensión de A como la dimensión de V_a.

Ejemplos

(a) *Para cada $x \in \mathbb{R}^n$, el conjunto $A = \{x\}$ es afín, y es traslación de $\cdots$.*

(b) *Para dos puntos distintos x , y de $\mathbb{R}^n$, la recta $\{tx + (1-t)y : t \in \mathbb{R}\}$ es un conjunto afín, paralelo al subespacio vectorial $\{\alpha(x - y) : \alpha \in \mathbb{R}\}$.*

(c) *Sean M Una matriz de orden $m \times n$, $b \in \mathbb{R}^m$ y $A := \{x \in \mathbb{R}^n : Mx = b\}$, A es un conjunto afín de dimensión $\cdots$*

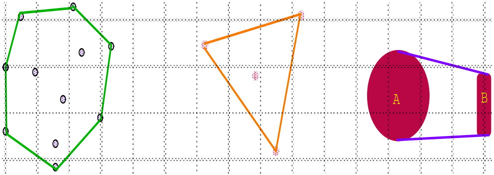
Dado un subconjunto C de un espacio vectorial E, la cápsula convexa de C denotada por $co(C)$ se define como el "menor" conjunto convexo que contiene a C. Es decir, si ζ es la colección de subconjuntos convexos de E que contienen a C, entonces

$$co(C) = \bigcap_{D \in \zeta} D$$

Dado un subconjunto C de un espacio vectorial E, la cápsula convexa de C denotada por $co(C)$ se define como el "menor" conjunto convexo que contiene a C. Es decir, si ζ es la colección de subconjuntos convexos de E que contienen a C, entonces

$$co(C) = \bigcap_{D \in \zeta} D$$

Politopo: Un conjunto C es un politopo si es la cápsula convexa de un número finito de puntos de E.



$co\{x_1, \cdots, x_m\}$ $co\{x_1, x_2, x_3, x_4\}$ $co(A \cup B)$

Dado un subconjunto C de un espacio vectorial E, la cápsula afín de C denotada por $aff(C)$ se define como el "menor" conjunto afín que contiene a C. Es decir, si ζ es la colección de subconjuntos afines de E que contienen a C, entonces

$$aff(C) = \bigcap_{D \in \zeta} D$$

Dado un subconjunto C de un espacio vectorial E, la cápsula afín de C denotada por $aff(C)$ se define como el "menor" conjunto afín que contiene a C. Es decir, si ζ es la colección de subconjuntos afines de E que contienen a C, entonces

$$aff(C) = \bigcap_{D \in \zeta} D$$

Ejemplos:

(i) Para cada $x \in E$, se cumple $aff\{x\} = \{x\}$.

(ii) Sean x , y dos vectores diferentes de E, entonces $aff\{x, y\} = \{x + t(y - x) : t \in \mathbb{R}\}$.

(iii) Si x, y, z son elementos no colineales de E, entonces $aff\{x, y, z\} = \cdots$

Dado un subconjunto C de un espacio vectorial E, la cápsula afín de C denotada por $aff(C)$ se define como el "menor" conjunto afín que contiene a C. Es decir, si ζ es la colección de subconjuntos afines de E que contienen a C, entonces

$$aff(C) = \bigcap_{D \in \zeta} D$$

Ejemplos:

(i) Para cada $x \in E$, se cumple $aff\{x\} = \{x\}$.

(ii) Sean x , y dos vectores diferentes de E, entonces $aff\{x, y\} = \{x + t(y - x) : t \in \mathbb{R}\}$.

(iii) Si x, y, z son elementos no colineales de E, entonces $aff\{x, y, z\} = \cdots$

En general, $co(C)$ consiste precisamente de todas las combinaciones convexas de elementos de C. Análogamente, $aff(C)$ consiste de todas las combinaciones afines de elementos de C.

Definición

Dados $p \in \mathbb{R}^n \setminus \{0\}$, $\alpha \in \mathbb{R}$, *el hiperplano* $H(p,\alpha)$ *se define como el conjunto*

$$H(p,\alpha) := \{x \in \mathbb{R}^n \ : \ \langle p, x\rangle = \alpha\} \tag{6}$$

Note que cuando $\alpha = 0$, $H(p,0)$ es un subespacio vectorial de $\mathbb{R}^n$ de dimensión $n-1$ y $x \in H(p,0)$ si y solo si, $x \perp p$. En tal caso, se dice que el p es un vector ortogonal al subespacio vectorial $H(0,p)$(comúnmente se dice que p es normal a $H(p,0)$.)
En general, se dice que el hiperplano $H(p,\alpha)$ es paralelo al subespacio vectorial $H(p,0)$ y que tiene vector normal p.

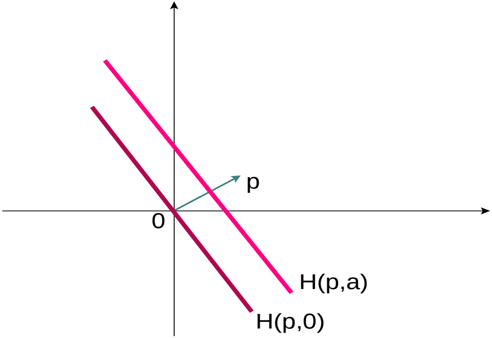


Fig1 Hiperplano $H(p,\alpha)$

Definición

Dado el hiperplano $H(p,\alpha)$, se generan los siguientes subconjuntos:

(a) $H(p,\alpha)^{\leq} := \{x \in \mathbb{R}^n : \langle p, x\rangle \leq \alpha\}$ *y* $H(p,\alpha)^{\geq} := \{x \in \mathbb{R}^n : \langle p, x\rangle \geq \alpha\}$ *que se denominan semiespacios cerrados.*

(b) $H(p,\alpha)^{<} := \{x \in \mathbb{R}^n : \langle p, x\rangle < \alpha\}$ *y* $H(p,\alpha)^{>} := \{x \in \mathbb{R}^n : \langle p, x\rangle > \alpha\}$ *que se denominan semiespacios abiertos.*

Las denominaciones cerrado y abierto, a la vez concuerdan con la naturaleza topológica de estos conjuntos.

Nota

Dado el hiperplano $H(p,\alpha)$, éste coincide con $H(tp, t\alpha)$ para cualquier $t \in \mathbb{R}$ no nulo. Particularmente, podemos exigir una representación del hiperplano con un vector normal de norma uno o también si $\alpha \neq 0$ podemos imponer que $\alpha = 1$.