Capítulo 12

Apendice

Resumen

Este apéndice tiene como objetivo refrescar algunos resultados de los espacios vectoriales, los espacios normados y los espacios con producto interior. Al final se provee de una amplia bibliografía sobre estos tópicos.

12.1. Espacios vectoriales reales.

12.1.1. Definición de espacio vectorial. Ejemplos.

Se denota con $\mathbb{R}$ al cuerpo de los números reales. Nos limitamos en definir los espacios vectoriales reales.

Definición 1 *Un espacio vectorial V sobre $\mathbb{R}$ consiste en un conjunto no vacío V en el que se ha definido dos operaciones: adición "+" en V que a cada par de elementos x, y de V le asocia un único elemento $x+y$ de V, y, producto de números reales por elementos de V dicha también producto por escalares que a cada $\alpha \in \mathbb{R}$ y $x \in V$ le asocia un único elemento αx de V; y, estas operaciones satisfacen las propiedades siguientes:*

i. *Conmutativa: para todo $x, y \in V$, $\quad x+y=y+x$.*

ii. *Asociativa: para todo $x, y, z \in V$, $\quad (x+y)+z=x+(y+z)$.*

iii. *Existencia de elemento neutro: existe $0 \in V$ tal que para todo $x \in V$, $\quad x+0=0+x=x$.*

iv. *Existencia de opuestos aditivos: para cada $x \in V$, existe $y \in V$ tal que $\quad x+y=0$.*

v. *Para todo $\alpha \in \mathbb{R}$, $x, y \in \mathbb{R}$, $\quad \alpha(x+y)=\alpha x+\alpha y$.*

vi. *Para todo $x \in V$, $\quad \alpha, \beta \in \mathbb{R}$, $\quad (\alpha+\beta) x=\alpha x+\beta x$.*

vii. *Para todo $x \in V$, $\quad \alpha, \beta \in \mathbb{R}$, $\quad \alpha(\beta x)=(\alpha \beta) x$.*

viii. *Para todo $x \in V$, $\quad 1 \cdot x=x$.*

Los elementos de V se llaman vectores y los elementos de $\mathbb{R}$ se llaman escalares. El espacio vectorial V sobre $\mathbb{R}$ se dirá simplemente espacio vectorial real. El conjunto V con la operación adición "+" que satisface las propiedades i) a iv) se dice grupo conmutativo que se nota $(V,+)$.

El elemento $0 \in V$ de iii) es único y se denomina elemento nulo. El elemento y de iv) se escribe $-x$, además es único. La propiedad iv) se expresa como sigue: $\forall x \in V,\ \exists -x \in V$ tal que $x+(-x)=0$.

Para todo $x, y, z \in V$, se escribe $x+y+z$ en vez de $(x+y)+z$ o de $x+(y+z)$.

En todo espacio vectorial real se verifican las propiedades siguientes cuyas demostraciones son inmediatas y se dejan como ejercicio.

i. Para todo $x \in V$, $0x = 0$.

ii. Para todo $\alpha \in \mathbb{R}$, $\alpha 0 = 0$.

iii. Para todo $\alpha \in \mathbb{R}$, $x \in V$, $(-\alpha)\, x = -\alpha x$.

iv. $\alpha x = 0 \Leftrightarrow \alpha = 0$ o $x = 0$.

Ejemplos

1. El espacio vectorial $\mathbb{R}^n$. Sea $n \in \mathbb{Z}^+$. Se denota con $\mathbb{R}^n$ al conjunto $\{(x_1, ..., x_n) \mid x_i \in \mathbb{R},\ i = 1, ..., n\}$, esto es $\mathbb{R}^n = \{(x_1, ..., x_n) \mid x_i \in \mathbb{R},\ i = 1, ..., n\}$. A los elementos de $\mathbb{R}^n$ los notamos como $\overrightarrow{x}$, $\overrightarrow{y}$, etc. También escribiremos $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ y los denominaremos vectores de $\mathbb{R}^n$. El elemento nulo de $\mathbb{R}^n$ se escribe $\overrightarrow{0} = (0, ..., 0)$.

   En $\mathbb{R}^n$ se define la igualdad, adición y producto de escalares por elementos de $\mathbb{R}^n$ como sigue. Sean $\overrightarrow{x} = (x_1, ..., x_n)$, $\overrightarrow{y} = (y_1, ..., y_n)$ dos elementos de $\mathbb{R}^n$, $\alpha \in \mathbb{R}$.

   Igualdad: diremos $\overrightarrow{x} = \overrightarrow{y}$ si y solo si $x_i = y_i$, $i = 1, ..., n$.

   Adición: $\overrightarrow{x} + \overrightarrow{y} = (x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n)$.

   Producto por escalares: $\alpha \overrightarrow{x} = \alpha\,(x_1, ..., x_n) = (\alpha x_1, ..., \alpha x_n)$.

   De la definición de adición en $\mathbb{R}^n$, se tiene $\overrightarrow{x}, \overrightarrow{y} \in \mathbb{R}^n \Rightarrow \overrightarrow{x} + \overrightarrow{y} \in \mathbb{R}^n$, y, de la definición de producto por escalares $\alpha \in \mathbb{R}$, $\overrightarrow{x} \in \mathbb{R}^n \Rightarrow \alpha \overrightarrow{x} \in \mathbb{R}^n$. Mas aún, se prueba fácilmente que $\mathbb{R}^n$ es un espacio vectorial real.

   El opuesto aditivo de $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ es $-\overrightarrow{x} = (-x_1, ..., -x_n) \in \mathbb{R}^n$.

   Para $n = 1$, se tiene que $\mathbb{R}$ es un espacio vectorial sobre si mismo.

   Para $n = 2$, $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$. Note que si $\overrightarrow{x} = (a, b)$, $\overrightarrow{y} = (c, d) \in \mathbb{R}^2$, $\alpha \in \mathbb{R}$ se tiene

   Igualdad: $\overrightarrow{x} = \overrightarrow{y} \Leftrightarrow a = c$ y $b = d$,

   Adición: $\overrightarrow{x} + \overrightarrow{y} = (a, b) + (c, d) = (a + c, b + d)$.

   Producto por escalares: $\alpha \overrightarrow{x} = \alpha\,(a, b) = (\alpha a, \alpha b)$.

   Para $n = 3$, $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$. Sean $\overrightarrow{x} = (x_1, y_1, z_1)$, $\overrightarrow{y} = (x_2, y_2, z_2) \in \mathbb{R}^3$, $\alpha \in \mathbb{R}$. Se tiene

$$
\begin{aligned}
\overrightarrow{x} &= \overrightarrow{y} \Leftrightarrow x_1 = x_2,\ \ y_1 = y_2,\ \ z_1 = z_2, \\
\overrightarrow{x} + \overrightarrow{y} &= (x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2), \\
\alpha \overrightarrow{x} &= \alpha\,(x_1, y_1, z_1) = (\alpha x_1, \alpha y_1, \alpha z_1).
\end{aligned}
$$

   Los elementos de $\mathbb{R}^n$ se escribirán también como vectores columna, así: $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

2. Espacio de matrices $M_{m\times n}[\mathbb{R}]$. Sean $m, n \in \mathbb{Z}^+$. Una matriz de $m \times n$ con valores en $\mathbb{R}$ es un arreglo rectangular de la forma:

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

   donde $a_{i,j} \in \mathbb{R}$, $i = 1, ..., m$, $j = 1, ..., n$.

Los números naturales $i = 1, ..., m, \quad j = 1, ..., n$ se llaman índices. Cuando i es fijo, los elementos $a_{i1}, a_{i2}, ..., a_{in}$ forman el i-ésimo renglón de la matriz y se puede considerar como un vector de $\mathbb{R}^n$, esto es, $(a_{i1}, a_{i2}, ..., a_{in}) \in \mathbb{R}^n$. Para j fijo, los elementos forman la j-ésima columna de la matriz. Esta puede considerarse como un vector columna de $\mathbb{R}^m$, es decir $\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in \mathbb{R}^m$.

A una matriz de $m \times n$ la representaremos abreviadamente como $(a_{ij})_{m\times n}$. También escribiremos $A = (a_{ij})_{m\times n}$ y simplemente (a_{ij}) si no hay peligro de confusión. Se nota con $M_{m\times n}[\mathbb{R}]$ al conjunto de todas la matrices de $m \times n$ con valores en $\mathbb{R}$.

Cuando $m = n$, los elementos de $M_{n\times n}[\mathbb{R}]$ los denominamos matrices cuadradas. Si $A = (a_{ij})_{n\times n}$ es una matriz cuadrada, si no hay peligro de confusión, escribiremos simplemente $A = (a_{ij})$.

Sean $A = (a_{ij})_{m\times n}, \ B = (b_{ij})_{m\times n} \in M_{m\times n}[\mathbb{R}]$ y $\alpha \in \mathbb{R}$. Definimos la igualdad, adición de matrices y producto de escalares por matrices, como sigue:

Igualdad: $A = B \Leftrightarrow a_{ij} = b_{ij}, \ \ i = 1, ..., m, \ \ j = 1, ..., n.$

Adición: $A + B = (a_{ij})_{m\times n} + (b_{ij})_{m\times n} = (a_{ij} + b_{ij})_{m\times n}$.

Producto por escalares: $\alpha A = \alpha (a_{ij})_{m\times n} = (\alpha a_{ij})_{m\times n}$.

Por la definición de adición en $M_{m\times n}[\mathbb{R}]$, tenemos $A, B \in M_{m\times n}[\mathbb{R}] \Rightarrow A + B \in M_{m\times n}[\mathbb{R}]$, y por la definición de producto de escalares por matrices tenemos $\alpha \in \mathbb{R}, \ A \in M_{m\times n}[\mathbb{R}] \Rightarrow \alpha A \in M_{m\times n}[\mathbb{R}]$.

Se demuestra fácilmente que $M_{m\times n}[\mathbb{R}]$ es un espacio vectorial real.

El elemento neutro de $M_{m\times n}[\mathbb{R}]$ es la matriz nula o matriz cero y la representamos con $0 = (0)_{m\times n}$, es decir que la matriz nula 0 es aquella que sus elementos son $0 \in \mathbb{R}$. El opuesto aditivo de $A = (a_{ij})_{m\times n}$ es la matriz notada $-A = (-a_{ij})_{m\times n}$.

Por otro lado, si $n = 1$ las matrices de $M_{m\times 1}[\mathbb{R}]$ coinciden con los vectores columna de $\mathbb{R}^m$ y si $m = 1$, las matrices de $M_{1\times n}[\mathbb{R}]$ coinciden con los vectores fila de $\mathbb{R}^n$.

3. Los espacios $C([a, b])$ y $C^1([a, b])$.

Revisemos brevemente algunos conceptos sobre funciones reales, operaciones con funciones reales, límites, continuidad, derivación e integración que son tratados en el curso de Análisis Matemático.

Sea $A \subset \mathbb{R}$, $A \neq \emptyset$. Una función real f definida en el conjunto A es un subconjunto F del producto cartesiano $A \times \mathbb{R}$ que satisface con las dos propiedades siguientes:

i. Para cada $x \in A$, existe un único $y \in \mathbb{R}$ tal que $y = f(x)$, o bien $(x, y) \in F$.
ii. Si $(x, y_1), \ (x, y_2) \in F$, entonces $y_1 = y_2$.

Sea f una función real definida en A. Escribiremos $f : \begin{cases} A & \rightarrow & \mathbb{R} \\ x & \rightarrow & f(x), \end{cases}$ que se lee f es la función de A en $\mathbb{R}$ que a cada $x \in A$ le asocia un único elemento $f(x)$ en $\mathbb{R}$. Se dirá también f es la función real de A en $\mathbb{R}$ que a cada $x \in A$ le asocia o le corresponde $f(x) \in \mathbb{R}$. El conjunto A se llama dominio de f y se designa con $\mathrm{Dom}(f)$. El conjunto $\mathbb{R}$ se llama conjunto de llegada de f y el conjunto $\mathrm{Rec}(f) = \{f(x) \mid x \in A\}$ se llama recorrido de f. Claramente $\mathrm{Rec}(f) \subset \mathbb{R}$ y $\mathrm{Rec}(f) \neq \emptyset$. Se designa con $\mathcal{F}(A)$ al conjunto de todas las funciones definidas en A.

La función nula $0 \in \mathcal{F}(A)$ se define como $0(x) = 0 \quad \forall x \in A$, y la función unidad $\mathbf{1} \in \mathcal{F}(A)$ está definida como $\mathbf{1}(x) = 1 \quad \forall x \in A$.

En $\mathcal{F}(A)$ se define la igualdad, adición y producto por escalares o producto de números reales por funciones como sigue:

Igualdad: Sean $f, g \in \mathcal{F}(A)$, $f = g$ si y solo si $f(x) = g(x) \quad \forall x \in A$.

Adición: Sean $f, g \in \mathcal{F}(A)$. Se define $f + g \in \mathcal{F}(A)$ como $(f + g)(x) = f(x) + g(x) \quad \forall x \in A$.

Producto por escalares: Sean $\alpha \in \mathbb{R}, \ \ f \in \mathcal{F}(A)$. Se define $\alpha f \in \mathcal{F}(A)$ como $(\alpha f)(x) = \alpha f(x) \quad \forall x \in A$.

Se demuestra fácilmente que $\mathcal{F}(A)$ es un espacio vectorial real denominado espacio de funciones definidas en A.

Funciones continuas

Sea A un intervalo de $\mathbb{R}$, $f \in \mathcal{F}(A)$, $x_0 \in A$ y $L \in \mathbb{R}$. Se dice que $f(x)$ tiende a L cuando x tiende a x_0 que se escribe $f(x) \underset{x \to x_0}{\to} L$, si y solo si se satisface la siguiente condición:

$$\forall \varepsilon > 0, \ \exists \delta > 0 \text{ tal que } \forall x \in A \ \text{ con } \ 0 < |x - x_0| < \delta \Rightarrow |f(x) - L| < \varepsilon.$$

Escribiremos también $\lim\limits_{x \to x_0} f(x) = L$ que se lee límite de $f(x)$ cuando x tiende a x_0 es igual a L.

Sea $A \subset \mathbb{R}$, $A \neq \emptyset$ y $f \in \mathcal{F}(A)$. Se dice que f es continua en $x_0 \in A$ si y solo si se satisfacen las dos condiciones siguientes:

i. $f(x_0)$ está bién definido.

ii. $\lim\limits_{x \to x_0} f(x) = f(x_0)$.

Se dice f continua en A si y solo si f es continua en todo punto $x_0 \in A$.

Se designa con $C(A)$ al conjunto de todas las funciones continuas en A. En el curso de análisis matemático se prueba que la suma de dos funciones continuas es continua, y que el producto de un número real λ por una función continua f es también una función continua, esto es,

$$f, g \in C(A) \Rightarrow f + g \in C(A), \qquad \lambda \in \mathbb{R}, \ f \in C(A) \Rightarrow \lambda f \in C(A).$$

Se prueba además que $C(A)$ es un espacio vectorial real.

En particular, si $A = [a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ es un intervalo cerrado y acotado de $\mathbb{R}$, el conjunto $C(A)$ se denota $C([a, b])$ y se le denomina espacio de funciones continuas en $[a, b]$.

Funciones derivables

Sean $A \subset \mathbb{R}$, $A \neq \emptyset$, f una función real definida en A y $x_0 \in A$. Si $\lim\limits_{h \to 0} \dfrac{f(x_0 + h) - f(x_0)}{h}$ existe, este se denomina derivada de f en x_0 que se escribe $f'(x_0)$ o también $\dfrac{df}{dx}(x_0)$; esto es,

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Se dice que f es derivable en A si $f'(x_0)$ existe en todo punto $x_0 \in A$ y se define una nueva función f' llamada función derivada de f.

Se designa con $C^1(A)$ al conjunto de todas las funciones f tales que f' es continua en A, y diremos que f es de clase C^1 en A. Particularmente si $A = [a, b]$, escribiremos $C^1([a, b])$ y diremos espacio de funciones de clase C^1en $[a, b]$.

Funciones integrables

En esta parte proponemos algunos resultados importantes de la teoría de la integración de funciones reales acotadas.

Definición 2 *Sean $a, b \in \mathbb{R}$ con $a < b$, $n \in \mathbb{Z}^+$. Una subdivisión o partición del intervalo $[a, b]$ se nota con $\Pi(n)$ y se define como el conjunto $\{x_0, x_1, ..., x_n\}$, donde $x_0 = a$, $x_n = b$, $x_i < x_{i+1}$, $i = 0, ..., n-1$.*

Si $\Pi(n)$ es una subdivisión de $[a, b]$, $[x_{i-1}, x_i]$, $i = 1, ..., n$ designa el i-ésimo subintervalo de $[a, b]$. Se pone $h_i = x_i - x_{i-1}$ la longitud del intervalo $[x_{i-1}, x_i]$, $i = 1, ..., n$, y $h = \max\limits_{i=0,1,...,n-1} h_i$.

Definición 3 *Sean $m, n \in \mathbb{Z}^+$. Se dice que $\Pi(m)$ es una subdivisión mas fina que $\Pi(n)$ si se verifica que $\Pi(n) \subset \Pi(m)$.*

Particularmente, si $\Pi(m)$, $\Pi(n)$ son dos subdivisiones de $[a,b]$, $\Pi(m)\cup\Pi(n)$ es una subdivisión mas fina que $\Pi(m)$ y $\Pi(n)$.

Definición 4 *Sea f una función real definida en $[a,b]$. Se dice que f es acotada en $[a,b]$ si y solo si $f([a,b])=\{f(x)\mid x\in[a,b]\}$ es acotado, es decir, existe $\beta>0$ tal que $|f(x)|\leq\beta \quad \forall x\in[a,b]$.*

Sean $n\in\mathbb{Z}^+$, $\Pi(n)$ una subdivisión de $[a,b]$ y f una función acotada en $[a,b]$. Se pone

$$\begin{aligned}\alpha_i &= \operatorname*{Inf}_{x\in[x_{i-1},x_i]} f(x), \quad \beta_i=\operatorname*{Sup}_{x\in[x_{i-1},x_i]} f(x), \quad i=1,...,n,\\ \alpha &= \operatorname*{Min}_{i=1,...,n}\alpha_i, \quad \beta=\operatorname*{Max}_{i=1,...,n}\beta_i.\end{aligned}$$

Definición 5 *La oscilación de f en $[x_{i-1},x_i]$, $i=1,...,n$ se define como*

$$\omega_i=\operatorname*{Sup}_{x\in[x_{i-1},x_i]} f(x)-\operatorname*{Inf}_{x\in[x_{i-1},x_i]} f(x), \quad i=1,...,n.$$

Definición 6 *Sea f una función real definida en $[a,b]$. Se dice que f es una función escalonada si y solo si existe una subdivisión $\Pi(n)=\{x_0=a,x_1,...,x_n=b\}$ y $c_1,...,c_n\in\mathbb{R}$ tales que*

$$f(x)=c_i, \quad x\in\,]x_{i-1},x_i[\,, \quad i=1,...,n.$$

La subdivisión $\Pi(n)$ se dice asociada a f.

Note que f es una función definida en todo $[a,b]$ y en cada subintervalo abierto $]x_{i-1},x_i[\,$, $i=1,...,n$, f es constante.

Definición 7 *Sea s una función escalonada en $[a,b]$ con $\Pi(n)=\{a=x_0,x_1,...,x_n=b\}$ la partición asociada a s y $s(x)=c_i$, $x\in\,]x_{i-1},x_i[\,$, $i=1,...,n$. La integral de s sobre el intervalo $[a,b]$ se nota $\int_a^b s(x)\,dx$ y se define como $\int_a^b s(x)\,dx=\sum\limits_{i=1}^{n}c_ih_i$, donde $h_i=x_i-x_{i-1}$, $i=1,...,n$.*
La notación $\int_a^b s(x)\,dx$ se lee integral de la función s con respecto de x en el intervalo $[a,b]$. El número real a es extremo inferior de integración, y el número real b el extremo superior de integración.

Sean s, t dos funciones escalonadas en $[a,b]$ y f una función acotada en $[a,b]$ tales que

$$s(x)\leq f(x)\leq t(x) \quad \forall x\in[a,b],$$

la función s se llama función escalonada inferior a f, y t se llama función escalonada superior a f. Particularmente, sea $n\in\mathbb{Z}^+$ y $\Pi(n)$ una subdivisión de $[a,b]$. Se definen las funciones escalonadas s_n y t_n como sigue:

$$\begin{aligned}s_n(x) &= \alpha_i=\operatorname*{Inf}_{x\in[x_{i-1},x_i]} f(x), \quad x\in\,]x_{i-1},x_i[\,, \quad i=1,...,n,\\ t_n(x) &= \beta_i=\operatorname*{Sup}_{x\in[x_{i-1},x_i]} f(x), \quad x\in\,]x_{i-1},x_i[\,, \quad i=1,...,n.\end{aligned}$$

Estas funciones satisfacen la siguiente desigualdad:

$$s_n(x)\leq f(x)\leq t_n(x) \quad \forall x\in\,]x_{i-1},x_i[\,, \quad i=1,...,n.$$

Se tiene que s_n es una función escalonada inferior a f, t_n es una función escalonada superior a f. Las integrales de estas funciones se definen como:

$$S_n(f)=\int_a^b s_n(x)\,dx=\sum_{i=1}^{n}\alpha_ih_i, \quad T_n(f)=\int_a^b t_n(x)\,dx=\sum_{i=1}^{n}\beta_ih_i,$$

donde $h_i = x_i - x_{i-1}, \quad i = 1, ..., n$. Se verifica

$$\alpha (b - a) \leq S_n (f) \leq T_n (f) \leq \beta (b - a)$$

y

$$0 \leq T_n (f) - S_n (f) \leq \sum_{i=1}^{n} \omega_i h_i \leq (\beta - \alpha) (b - a),$$

donde ω_i es la oscilación de f, $\alpha = \operatorname*{Inf}_{x \in [a,b]} f(x), \quad \beta = \operatorname*{Sup}_{x \in [a,b]} f(x)$.

Definición 8 *Sea f una función real, acotada en $[a, b]$.*

i. La integral inferior de f se designa con $\underline{I}(f)$ y se define como $\underline{I}(f) = \operatorname*{Sup}_{n \in \mathbb{Z}^+} S_n (f) = \operatorname*{Sup}_{n \in \mathbb{Z}^+} \int_a^b s_n (x)\, dx$, donde s_n es escalonada inferior a f.

ii. La integral superior de f se designa con $\overline{I}(f)$ y se define como $\overline{I}(f) = \operatorname*{Inf}_{n \in \mathbb{Z}^+} T_n (f) = \operatorname*{Inf}_{n \in \mathbb{Z}^+} \int_a^b t_n (x)\, dx$, donde t_n es escalonada superior a f.

Se verifica inmediatamente que si $s_n \leq f \leq t_n, \quad \int_a^b s_n (x)\, dx \leq \underline{I}(f) \leq \overline{I}(f) \leq \int_a^b t_n (x)\, dx$.

Definición 9 *Sea f una función real, acotada en $[a, b]$. Se dice que f es integrable en $[a, b]$ si y solo si $\underline{I}(f) = \overline{I}(f)$. En tal caso, escribimos $I(f) = \int_a^b f(x)\, dx$ y al número real $I(f)$ lo denominamos la integral de la función f en el intervalo $[a, b]$.*

Las funciones monótonas en $[a, b]$ (crecientes, decrecientes), las funciones continuas en $[a, b]$ son ejemplos de funciones integrables en $[a, b]$.

Se denota con $\mathcal{I}([a, b])$ al conjunto de todas las funciones integrables en $[a, b]$. Con las operaciones habituales de adición " + " de funciones y producto de escalares por funciones, $\mathcal{I}([a, b])$ es un espacio vectorial real denominado espacio de funciones integrables en $[a, b]$. Se tiene

i. $f, g \in \mathcal{I}([a, b]) \Rightarrow f + g \in \mathcal{I}([a, b])$, e $\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx$,

ii. $\lambda \in \mathbb{R}, f \in \mathcal{I}([a, b]) \Rightarrow \lambda f \in \mathcal{I}([a, b])$, e $\int_a^b \lambda f(x)\, dx = \lambda \int_a^b f(x)\, dx$.
Si $f \in \mathcal{I}([a, b])$ y $\alpha \in [a, b]$ se define $\int_\alpha^\alpha f(x)\, dx = 0$.

Sea $f \in \mathcal{I}([a, b])$. Se verifican las siguientes propiedades:

i. Si $c \in [a, b]$, $\int_a^b f(x)\, dx = \int_a^c f(x) + \int_c^b f(x)\, dx$.

ii. $\int_a^b f(x)\, dx = \int_{a+c}^{b+c} f(x - c)\, dx \quad \forall c \in \mathbb{R}$.

iii. Para $\alpha \neq 0$, $\int_a^b f(x)\, dx = \frac{1}{\alpha} \int_{\alpha a}^{\alpha b} f\left(\frac{x}{\alpha}\right) dx$.

iv. Si $f(x) \geq 0 \quad \forall x \in [a, b]$, $\int_a^b f(x)\, dx \geq 0$.

v. $\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx$.

vi. Si $[c, d] \subset [a, b]$ y $f(x) \geq 0 \quad \forall x \in [a, b]$, $\int_c^d f(x)\, dx \leq \int_a^b f(x)\, dx$.

Por otro lado, si $f(x) \geq 0 \quad \forall x \in [a,b]$, geométricamente $\int_a^b f(x)\,dx$ se interpreta como el área de la región

$$\Omega = \left\{(x,y) \in \mathbb{R}^2 \mid 0 \leq y \leq f(x)\,,\ \ x \in [a,b]\right\}.$$

Ejercicios

1. Demuestre que el conjunto $\mathbb{R}^n$ en el que se ha definido la igualdad y las operaciones de adición y producto por escalares, es un espacio vectorial real.
2. Demuestre que el conjunto de matrices $M_{m\times n}[\mathbb{R}]$ en el que se ha definido la igualdad y las operaciones de adición y producto de números reales por matrices, es un espacio vectorial real.
3. Demuestre que el conjunto $\mathcal{F}(A)$ de funciones reales definidas en A en el que se ha definido la igualdad de funciones, y las operaciones de adición y producto de números reales por funciones, es un espacio vectorial real.
4. Pruebe que con las operaciones de adición de funciones y producto de números reales por funciones definidas en $\mathcal{F}(A)$, los siguientes conjuntos son espacios vectoriales reales.

    i. Conjunto $\mathcal{I}([a,b])$ de funciones integrables en $[a,b]$.

    ii. Conjunto $C([a,b])$ de funciones continuas en $[a,b]$.

    iii. Conjunto $C^1([a,b])$ de funciones derivables con derivada continua en $[a,b]$.

12.1.2. Subespacios vectoriales. Ejemplos.

Definición 10 *Sea V un espacio vectorial sobre $\mathbb{R}$ y W un subconjunto no vacío de V. Se dice que W es un subespacio de V si W es un espacio vectorial real con las mismas operaciones definidas en V.*

El cualquier espacio vectorial V, $W = V$ y $W = \{0\}$ con $0 \in V$ son subespacios de V llamados subespacios triviales de V.

Teorema 1 *Un subconjunto no vacío W de un espacio vectorial V es un subespacio de V si y solo si se satisfacen las tres condiciones siguientes:*

i. $0 \in V \Rightarrow 0 \in W$.

ii. $x,\, y \in W \Rightarrow x + y \in W$.

iii. $\alpha \in \mathbb{R},\ \ x \in W \Rightarrow \alpha x \in W$.

Si $W_1,\ W_2$ son dos subespacios de V, entonces $W_1 \cap W_2$ es también en subespacio de V.

Definición 11 *Sean V un espacio vectorial real, $S_1,\ S_2$ dos subconjuntos no vacíos de V. Se define $S_1 + S_2$ como sigue*

$$S_1 + S_2 = \{x + y \mid x \in S_1\ \ y \in S_2\}$$

y se denomina subconjunto suma de S_1 con S_2.

Sea $x_0 \in V$ y W un subespacio de V. El subconjunto

$$x_0 + W = \{x_0 + x \mid x \in W\}$$

se llama trasladado del subespacio W o subconjunto afín.

Teorema 2 *Si W_1, W_2 son subespacios de un espacio vectorial V, $W_1 + W_2$ es un subespacio de V. El subespacio $W_1 + W_2$ se llama suma de los subespacios W_1 y W_2.*

Definición 12 *Sean W_1, W_2 dos subespacios de V. Se dice que V es suma directa de W_1 y W_2 que se escribe $V = W_1 \oplus W_2$ si y solo si se satisfacen las dos condiciones siguientes:*

i. $V = W_1 + W_2$.

ii. $W_1 \cap W_2 = \{0\}$.

Teorema 3 *Sean W_1, W_2 dos subespacios de V. Entonces, $V = W_1 \oplus W_2$ si y solo si cada $x \in V$ se escribe de manera única en la forma $x = x_1 + x_2$, donde $x_1 \in W_1$, $x_2 \in W_2$.*

Ejemplos

1. El conjunto $W = \{(x_1, ..., x_{n-1}, 0) \mid x_i \in \mathbb{R},\ i = 1, ..., n-1\}$ es un subespacio de $\mathbb{R}^n$.

2. Sea $n = 3$, $W_1 = \{(x, y, 0) \mid x,\ y \in \mathbb{R}\}$, $W_2 = \{(2t, 3t, t) \mid t \in \mathbb{R}\}$ son subespacios de $\mathbb{R}^3$. Se verifica que $W_1 \cap W_2 = \{0\}$ y que $\mathbb{R}^3 = W_1 \oplus W_2$.

   Note que si $\overrightarrow{x} = (x, y, z) \in \mathbb{R}^3$, existen $\overrightarrow{x}_1 = (x - 2z, y - 3z, 0) \in W_1$ y $\overrightarrow{x}_2 = (2z, 3z, z) \in W_2$ tales que $\overrightarrow{x} = \overrightarrow{x}_1 + \overrightarrow{x}_2$. El vector $\overrightarrow{x} = \overrightarrow{x}_1 + \overrightarrow{x}_2$ se escribe de esta manera en forma única como $\overrightarrow{x}_1 \in W_1$, $\overrightarrow{x}_2 \in W_2$.

3. El espacio $C([a, b])$ de funciones continuas en $[a, b]$ es un subespacio de $\mathcal{I}([a, b])$.

4. Un polinomio$P$ de grado $\leq n$ con coeficientes reales se define como

$$P(x) = a_0 + a_1 x + ... + a_n x^n = \sum_{k=0}^{n} a_k x^k, \qquad x \in \mathbb{R},\ a_i \in \mathbb{R},\ i = 0, 1, ..., n.$$

   El polinomio nulo se define como $P(x) = 0 \quad \forall x \in \mathbb{R}$.

   Se designa con $\mathbb{K}_n[\mathbb{R}]$ el conjunto de todos los polinomios de grado $\leq n$. Se define la igualdad de polinomios, adición y producto de números reales por polinomios como sigue:

   Igualdad: Sean P, $Q \in \mathbb{K}_n[\mathbb{R}]$ con $P(x) = \sum\limits_{k=0}^{n} a_k x^k$, $\quad Q(x) = \sum\limits_{k=0}^{n} b_k x^k$, $x \in \mathbb{R}$.

$$P(x) = Q(x) \Leftrightarrow a_k = b_k, \quad k = 0, 1, ..., n.$$

   Adición: Sean P, $Q \in \mathbb{K}_n[\mathbb{R}]$ con $P(x) = \sum\limits_{k=0}^{n} a_k x^k$, $\quad Q(x) = \sum\limits_{k=0}^{n} b_k x^k$, $x \in \mathbb{R}$. Se define $P + Q \in \mathbb{K}_n[\mathbb{R}]$ como $(P + Q)(x) = \sum\limits_{k=0}^{n} (a_k + b_k) x^k \quad x \in \mathbb{R}$.

   Producto por escalares: Sean $\lambda \in \mathbb{R}$, $P \in \mathbb{K}_n[\mathbb{R}]$ con $P(x) = \sum\limits_{k=0}^{n} a_k x^k$, $x \in \mathbb{R}$. Se define $\lambda P \in \mathbb{K}_n[\mathbb{R}]$ como $(\lambda P)(x) = \lambda P(x) = \sum\limits_{k=0}^{n} \lambda a_k x^k$, $\quad x \in \mathbb{R}$.

   Se demuestra que $\mathbb{K}_n[\mathbb{R}]$ es un espacio vectorial real denominado espacio de polinomios de grado $\leq n$.

   Sea $V = C([a, b])$. El conjunto de todos los polinomios de grado $\leq n$ restringidos a $[a, b]$ con las operaciones de adición y producto por escalares, es un subespacio de $C([a, b])$. A este subespacio lo notaremos con $\mathbb{K}_n([a, b])$.

Bases de V

Definición 13 *Sea A un subconjunto no vacío de un espacio vectorial V. Se dice que $x \in V$ es una combinación lineal de elementos de A si existe un número finito $x_1, ..., x_n \in A$ y $\alpha_1, ..., \alpha_n \in \mathbb{R}$ tales que $x = \sum\limits_{i=1}^{n} \alpha_i x_i$.*

Particularmente, si $A = \{x_1, ..., x_n\} \subset V$, $x \in V$ es combinación lineal de elementos de A si existen $\alpha_1, ..., \alpha_n \in \mathbb{R}$ tales que $x = \sum_{i=1}^{n} \alpha_i x_i$.

Teorema 4 *Sea A un subconjunto no vacío de V. El subconjunto W de V constituído por todas las combinaciones lineales de elementos de A es un subespacio de V. Este subconjunto W de V se denomina subespacio generado por A. Escribiremos $W = L(A)$.*

Definición 14 *Sea A un subconjunto no vacío de un espacio vectorial V. Si $L(A) = V$ diremos que A genera a V o que A es un conjunto generador de V.*

Si $A \subset V$ y $L(A) = V$, entonces $x \in V$ si y solo si existen $\alpha_1, ..., \alpha_n \in \mathbb{R}$, $x_1, ..., x_n \in A$ tales que $x = \sum_{i=1}^{n} \alpha_i x_i$. Particularmente, si $A = \{x_1, ..., x_n\} \subset V$, escribiremos explícitamente al subespacio generado por A como

$$W = L(x_1, ..., x_n) = \left\{ \sum_{i=1}^{n} \alpha_i x_i \,|\alpha_i \in \mathbb{R}, \quad i = 1, ..., n \right\}.$$

Si $W = V$, escribiremos $V = L(x_1, ..., x_n)$ y diremos que $\{x_1, ..., x_n\}$ es un conjunto generador de V.

Definición 15 *Sea A un subconjunto no vacío de un espacio vectorial V.*

i. Se dice que A es linealmente independiente si para todo $x_1, .., x_n \in A$, $\sum_{i=1}^{n} \alpha_i x_i = 0 \Rightarrow \alpha_i = 0, \quad i = 1, ..., n$.

ii. Se dice que A es linealmente dependiente si A no es linealmente independiente.

De la definición de dependencia lineal se sigue que A es linealmente dependiente si existen $x_1, ..., x_n \in A$ y $\alpha_1, ..., \alpha_n \in \mathbb{R}$ no todos nulos tales que $\sum_{i=1}^{n} \alpha_i x_i = 0$.

Sean A, B dos subconjuntos no vacíos de un espacio vectorial V tales que $A \subset B$. Entonces, si A es linealmente dependiente, B también lo es; y , si B es linealemente independiente, A también lo es.

Definición 16 *Se dice que un subconjunto $\mathcal{B}$ de un espacio vectorial V es una base de V si y solo si se satisfacen las dos condiciones siguientes:*

i. $\mathcal{B}$ es linealmente independiente.

ii. $\mathcal{B}$ genera a V.

Definición 17

i. Un espacio vectorial real V es de dimensión finita n si toda base $\mathcal{B}$ de V está constituida por exactamente n elementos. Al único número natural n se le llama dimensión de V y se le denota $\dim V$, esto es $\dim V = n$.

ii. Se dice que un espacio vectorial real V es de dimensión infinita si cualquier base $\mathcal{B}$ de V tiene un número infinito o numerable de elementos.

Ejemplos

1. El espacio $\mathbb{R}^n$ es un espacio vectorial de dimensión finita n. La base $\mathcal{B} = \{\overrightarrow{e}_1, ..., \overrightarrow{e}_n\}$ se conoce como base canónica de $\mathbb{R}^n$, donde

$$\overrightarrow{e}_1 = (1, 0, ..., 0), \cdots, \overrightarrow{e}_n = (0, ..., 0, 1).$$

Sea $\overrightarrow{x}=(x_1,...,x_n)\in\mathbb{R}^n$. Se tiene $\overrightarrow{x}=\sum_{i=1}^n x_i\overrightarrow{e}_i=x_1\overrightarrow{e}_1+...+x_n\overrightarrow{e}_n$.

Para $n=2$, el conjunto $\mathcal{B}=\left\{\overrightarrow{i},\overrightarrow{j}\right\}$ con $\overrightarrow{i}=(1,0)$, $\overrightarrow{j}=(0,1)$, es la base canónica de $\mathbb{R}^2$. Note que $\overrightarrow{x}=(a,b)\in\mathbb{R}^2$ se escribe en la forma $\overrightarrow{x}=a\overrightarrow{i}+b\overrightarrow{j}$.

Para $n=3$, el conjunto $\mathcal{B}=\left\{\overrightarrow{i},\overrightarrow{j},\overrightarrow{k}\right\}$ con $\overrightarrow{i}=(1,0,0)$, $\overrightarrow{j}=(0,1,0)$, $\overrightarrow{k}=(0,0,1)$, es la base canónica de $\mathbb{R}^3$. Si $\overrightarrow{x}=(a,b,c)\in\mathbb{R}^3$ entonces $\overrightarrow{x}=a\overrightarrow{i}+b\overrightarrow{j}+c\overrightarrow{k}$.

2. Sea $V=C([0,2\pi])$. Consideremos las funciones $\varphi_0,\varphi_1,...,\varphi_n$ definidas en $[0,2\pi]$ como sigue:

$$\varphi_0(x)=1,\ \ \varphi_1(x)=\operatorname{sen}x,\ \ \varphi_2(x)=\operatorname{sen}(2x),...,\varphi_n(x)=\operatorname{sen}(nx).$$

El conjunto $\mathcal{B}=\{\varphi_0,\varphi_1,...,\varphi_n\}$ es linealmente independiente y genera un espacio W constituido por todas las combinaciones lineales de $\varphi_0,\varphi_1,..,\varphi_n$, esto es

$$\begin{aligned} W &= \left\{\sum_{i=1}^n \alpha_i\varphi_i\,|\alpha_i\in\mathbb{R},\ \ i=0,1,...,n\right\} \\ f &\in W\Leftrightarrow f(x)=\alpha_0+\alpha_1\operatorname{sen}(x)+...+\alpha_n\operatorname{sen}(nx),\quad x\in[0,2\pi]\end{aligned}$$

donde $\alpha_0,...,\alpha_n\in\mathbb{R}$ son elegidos apropiadamente.

3. El espacio vectorial de matrices de $M_{m\times n}[\mathbb{R}]$ es de dimensión finita $m\times n$. La base canónica $\mathcal{B}$ de $M_{m\times n}[\mathbb{R}]$ está formada por las matrices $A_1=\left(a_{ij}^{(1)}\right),...,A_{m\times n}=\left(a_{ij}^{(m\times n)}\right)$ donde

$$a_{ij}^{(1,1)}=\begin{cases}1, & \text{si }\ i=1,\,j=1,\\ 0, & \text{si } 1<i\le m,\ \ 1<j\le n,\end{cases}\qquad\cdots\quad a_{ij}^{(m\times n)}=\begin{cases}1, & \text{si }\ i=m,\,j=n,\\ 0, & \text{si } 1\le i<m,\ \ 1\le j<n.\end{cases}$$

Por ejemplo si $m=2$, $n=3$, la base canónica del espacio vectorial de matrices $M_{2\times3}[\mathbb{R}]$ está formada por las siguientes matrices:

$$A_1=\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix},\quad A_2=\begin{bmatrix}0&1&0\\0&0&0\end{bmatrix},\quad A_3=\begin{bmatrix}0&0&1\\0&0&0\end{bmatrix},$$

$$A_4=\begin{bmatrix}0&0&0\\1&0&0\end{bmatrix},\quad A_5=\begin{bmatrix}0&0&0\\0&1&0\end{bmatrix},\quad A_6=\begin{bmatrix}0&0&0\\0&0&1\end{bmatrix}.$$

Sea $A=(a_{ij})\in M_{2\times3}[\mathbb{R}]$, entonces $A=a_{11}A_1+...+a_{23}A_6$.

4. El espacio vectorial $C([a,b])$ de funciones continuas en $[a,b]$ es de dimensión infinita.

5. El espacio vectorial $\mathcal{I}([a,b])$ de funciones integrables en $[a,b]$ es de dimensión infinita.

6. El espacio $\mathbb{K}_n([a,b])$ es de dimensión finita $n+1$. La base canónica de $\mathbb{K}_n([a,b])$ está constituido por el conjunto de funciones $\{P_0,P_1,...,P_n\}$ con $P_0(x)=1$, $P_j(x)=x^j\quad x\in[a,b]$, $j=1,...,n$.

12.2. Definición de espacio normado.

Definición 18 *Sea V un espacio vectorial sobre $\mathbb{R}$. Una norma en V es una función N de V en $\mathbb{R}$ que satisface las siguientes propiedades:*

i. $N(x)\ge0\quad\forall x\in V$,

ii. $N(x)=0\Leftrightarrow x=0$,

iii. $N(\lambda x)=|\lambda|\,N(x)\qquad\forall\lambda\in\mathbb{R},\ \forall x\in V$,

iv. $N(x+y)\le N(x)+N(y)\qquad\forall x,y\in V$ *(desigualdad triangular).*

El número real no negativo $N(x)$ se llama norma de x. El par (V, N) se llama espacio normado.

Observación

Si la función N de V en $\mathbb{R}$ verifica las propiedades i), iii) y iv) de la definición de norma, pero no se verifica ii), la función N se dice seminorma en V.

Note en iv) que $x + y$ es la suma de los elementos x, y de V, mientras que $N(x) + N(y)$ es la suma de los números reales no negativos $N(x)$ y $N(y)$. En iii), λx es el producto del escalar λ (número real) por el elemento x de V y $|\lambda| N(x)$ es el producto de los números reales no negativos $|\lambda|$ y $N(x)$.

En ii), $x = 0$ denota el elemento neutro o nulo de V y $N(x) = 0$ es el elemento neutro o nulo de $\mathbb{R}$.

Notación

Si N es una norma en V, es usual escribir esta función con el símbolo $\|\cdot\|$ en vez de N y el espacio normado se escribirá $(V, \|\cdot\|)$ o se dirá V espacio normado provisto de la norma $\|\cdot\|$. Para $x \in V$, la norma de x se escribirá $\|x\|$.

Si en V se han definido varias normas, es preciso señalar que norma se está utilizando.

Proposición 5 *Sea V un espacio normado con $\|\cdot\|$ su norma. Se verifican las siguientes propiedades:*

i. $\|x - y\| = \|y - x\| \quad \forall x, y \in V.$

ii. $|\, \|x\| - \|y\| \,| \leq \|x - y\| \quad \forall x, y \in V.$

Demostración.

i. Sean x, $y \in V$. Entonces $\|x - y\| = \|(-1)(x - y)\| = |-1| \|y - x\| = \|y - x\|$.

ii. Sean x, $y \in V$. De la desigualdad triangular, se tiene $\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\|$, de donde $\|x\| - \|y\| \leq \|x - y\|$. Además,
$$\|y\| = \|(y - x) + x\| \leq \|y - x\| + \|x\| = \|x - y\| + \|x\|$$
y de esta desigualdad se obtiene la siguiente: $\|y\| - \|x\| \leq \|x - y\|$ que multiplicándola por -1 resulta $-\|x - y\| \leq \|x\| - \|y\|$.

Por lo tanto, $-\|x - y\| \leq \|x\| - \|y\| \leq \|x - y\|$, que es equivalente a la desigualdad $|\ \|x\| - \|y\|\ | \leq \|x - y\|$.

Nota: Recuerde que si $a \geq 0$, $|t| \leq a \Leftrightarrow -a \leq t \leq a$.

■

12.3. Ejemplos de espacios normados.

Comenzamos esta sección considerando el ejemplo más simple de espacio normado: el espacio vectorial $\mathbb{R}$ provisto de la función valor absoluto $|\cdot|$.

Sea $V = \mathbb{R}$. Se define la función $\|\cdot\|$ de $\mathbb{R}$ en $\mathbb{R}$ como sigue: $\|x\| = |x| \quad \forall x \in \mathbb{R}$. Entonces, la función $\|\cdot\|$ definida en $\mathbb{R}$ es una norma en $\mathbb{R}$. La verificación de las propiedades i) a iv) siguen inmediatamente de las propiedades del valor absoluto siguientes:

i. $|x| \geq 0 \quad \forall x \in \mathbb{R}$.

ii. $|x| = 0 \Leftrightarrow x = 0$.

iii. $|\lambda x| = |\lambda| |x| \quad \forall \lambda, x \in \mathbb{R}$.

iv. $|x + y| \leq |x| + |y| \quad \forall x, y \in \mathbb{R}$, (desigualdad triangular).

12.3.1. Normas en $\mathbb{R}^n$.

En el espacio vectorial real $\mathbb{R}^n$ se consideran dos normas importantes: la del máximo que se denota $\|\cdot\|_\infty$ y las hölderianas $\|\cdot\|_p$ con $p \in [1,\ \infty[$.

Norma $\|\cdot\|_\infty$.

Sea $V = \mathbb{R}^n$. Se define la función $\|\cdot\|_\infty$ de $\mathbb{R}^n$ en $\mathbb{R}$ como $\|\overrightarrow{x}\|_\infty = \underset{i=1,\dots,n}{\mathrm{Max}} |x_i| \quad \forall \overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. Se verifica que $\|\cdot\|_\infty$ es una norma en $\mathbb{R}^n$.

i) Es claro que para todo $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, $\|\overrightarrow{x}\|_\infty \geq 0$.

ii) Si $\overrightarrow{x} = \overrightarrow{0} = (0, ..., 0) \in \mathbb{R}^n$, se tiene $\|\overrightarrow{x}\|_\infty = 0$. Recíprocamente, si $\|\overrightarrow{x}\|_\infty = 0$ se sigue que $\underset{i=1,\dots,n}{\mathrm{Max}} |x_i| = 0$ y como $0 \leq |x_i| \leq \underset{i=1,\dots,n}{\mathrm{Max}} |x_i| = 0 \quad i = 1, ..., n$, resulta que $x_i = 0$, $i = 1, ..., n$, esto es, $\overrightarrow{x} = \overrightarrow{0}$. Así, $\|\overrightarrow{x}\|_\infty = 0 \Leftrightarrow \overrightarrow{x} = \overrightarrow{0}$.

iii) Sean $\lambda \in \mathbb{R}$, $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. Entonces $\lambda\overrightarrow{x} = (\lambda x_1, ..., \lambda x_n)$, y

$$\|\lambda\overrightarrow{x}\|_\infty = \underset{i=1,\dots,n}{\mathrm{Max}} |\lambda x_i| = \underset{i=1,\dots,n}{\mathrm{Max}} |\lambda|\,|x_i| = |\lambda| \underset{i=1,\dots,n}{\mathrm{Max}} |x_i| = |\lambda|\,\|\overrightarrow{x}\|_\infty.$$

Luego $\|\lambda\overrightarrow{x}\|_\infty = |\lambda|\,\|\overrightarrow{x}\|_\infty$.

iv) Sean $\overrightarrow{x} = (x_1, ..., x_n)$, $\overrightarrow{y} = (y_1, ..., y_n) \in \mathbb{R}^n$. Puesto que

$$\overrightarrow{x} + \overrightarrow{y} = (x_1 + y_1, ..., x_n + y_n), \qquad |x_i + y_i| \leq |x_i| + |y_i| \quad i = 1, ..., n,$$

y de la definición de la función $\|\cdot\|_\infty$, se sigue

$$\begin{aligned} \|\overrightarrow{x} + \overrightarrow{y}\|_\infty &= \underset{i=1,\dots,n}{\mathrm{Max}} |x_i + y_i| \leq \underset{i=1,\dots,n}{\mathrm{Max}} \{|x_i| + |y_i|\} \\ &= \underset{i=1,\dots,n}{\mathrm{Max}} |x_i| + \underset{i=1,\dots,n}{\mathrm{Max}} |y_i| = \|\overrightarrow{x}\|_\infty + \|\overrightarrow{y}\|_\infty . \end{aligned}$$

Luego, $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$.

Conclusión: $\|\cdot\|_\infty$ es una norma sobre $\mathbb{R}^n$.

Sean $n = 2$, $\overrightarrow{x} = (x, y) \in \mathbb{R}^2$, la norma $\|\cdot\|_\infty$ en $\mathbb{R}^2$ está definida como $\|\overrightarrow{x}\|_\infty = \mathrm{Max}\{|x|, |y|\}$. Note que $|x| \leq \|\overrightarrow{x}\|_\infty$ y $|y| \leq \|\overrightarrow{x}\|_\infty$.

Sean $n = 3$, $\overrightarrow{x} = (x, y, z) \in \mathbb{R}^3$, la norma $\|\cdot\|_\infty$ en $\mathbb{R}^3$ está definida como $\|\overrightarrow{x}\|_\infty = \mathrm{Max}\{|x|, |y|, |z|\}$. Además, se verifican las siguientes desigualdades: $|x| \leq \|\overrightarrow{x}\|_\infty$, $|y| \leq \|\overrightarrow{x}\|_\infty$, $|z| \leq \|\overrightarrow{x}\|_\infty$.

Si $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. Se tiene $|x_i| \leq \|\overrightarrow{x}\|_\infty$, $i = 1, ..., n$.

Normas hölderianas en el espacio $\mathbb{R}^n$.

Sea $p \in [1, \infty[$. Se define la función $\|\cdot\|_p$ de $\mathbb{R}^n$ en $\mathbb{R}$ como sigue:

$$\|\overrightarrow{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \qquad \forall \overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n.$$

Entonces $\|\cdot\|_p$ es una norma en $\mathbb{R}^n$, llamada norma de Hölder o norma hölderiana. Verifiquemos las propiedades i) a iv) de la definición de norma.

i)Sea $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. Puesto que $|x_i| \geq 0 \quad i = 1, ..., n$, de la definición de $\|\cdot\|_p$, se sigue $\|x\|_p \geq 0$.

ii) Si $\overrightarrow{x} = \overrightarrow{0} = (0, ..., 0)$ se tiene $\|\overrightarrow{x}\|_p = 0$. Supongamos que $\|\overrightarrow{x}\|_p = 0$ entonces $\sum\limits_{i=1}^{n} |x_i|^p = 0$. Se tiene la siguiente desigualdad: $0 \leq |x_i|^p \leq \sum\limits_{i=1}^{n} |x_i|^p = 0 \quad i = 1, ..., n$, consecuentemente $x_i = 0$, $i = 1, ..., n$, o sea $\overrightarrow{x} = (0, ..., 0)$. Luego, $\|\overrightarrow{x}\|_p = 0 \Leftrightarrow \overrightarrow{x} = \overrightarrow{0}$.

iii) Sean $\lambda \in \mathbb{R}$, $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. Entonces $\lambda \overrightarrow{x} = (\lambda x_1, ..., \lambda x_n)$ y por la definición de $\|\cdot\|_p$, se tiene

$$\begin{aligned} \|\lambda \overrightarrow{x}\|_p &= \left(\sum_{i=1}^n |\lambda x_i|^p \right)^{\frac{1}{p}} = \left(\sum_{i=1}^n |\lambda|^p |x_i|^p \right)^{\frac{1}{p}} = \left(|\lambda|^p \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \\ &= |\lambda| \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} = |\lambda| \, \|\overrightarrow{x}\|_p . \end{aligned}$$

Por lo tanto, $\|\lambda \overrightarrow{x}\|_p = |\lambda| \, \|\overrightarrow{x}\|_p$.

iv) Para probar la desigualdad triangular, se requieren de dos resultados preliminares: la desigualdad de Young y la desigualdad de Hölder.

Desigualdad de Young.

Esta se establece en los siguientes términos: sean $p, q \in \left]1, \infty\right[$ tales que $\frac{1}{p} + \frac{1}{q} = 1$, $\alpha \geq 0$, $\beta \geq 0$. Entonces

$$\alpha^{\frac{1}{p}} \beta^{\frac{1}{q}} \leq \frac{1}{p}\alpha + \frac{1}{q}\beta.$$

Para probar esta desigualdad, estudiemos la función siguiente:

$$f : \begin{cases} [0, \infty[& \to \mathbb{R} \\ x & \to \quad f(x) = \frac{1}{q} + \frac{1}{p}x - x^{\frac{1}{p}}. \end{cases}$$

Para $x > 0$ calculemos la derivada $f'(x)$ y determinemos los puntos críticos y los intervalos donde $f'(x) > 0$, $f'(x) < 0$, tenemos

$$\begin{aligned} f'(x) &= \frac{1}{p} - \frac{1}{p} x^{\frac{1}{p}-1} = \frac{1}{p} - \frac{1}{p} x^{-\frac{1}{q}} = \frac{1}{p}\left(1 - x^{-\frac{1}{q}}\right), \\ f'(x) &= 0 \Leftrightarrow x^{-\frac{1}{q}} = 1 \Leftrightarrow x = 1, \\ f'(x) &> 0 \Leftrightarrow 1 - x^{-\frac{1}{q}} > 0 \Leftrightarrow x > 1, \\ f'(x) &< 0 \Rightarrow 0 < x < 1. \end{aligned}$$

La función f es decreciente sobre $]0, 1[$ y creciente si $x \geq 1$. Puesto que $f(1) = 0$, la función f tiene un mínimo local en $x = 1$. Además, $f(0) = \frac{1}{q} > 0$ y $f(x) \underset{x \to \infty}{\to} +\infty$. Luego $f(1) = 0$ es un mínimo global. En consecuencia, para todo $x \geq 0$, $f(x) \geq 0$. En particular, para $x \geq 1$ y siendo f creciente, se tiene

$$0 = f(1) \leq f(x) = \frac{1}{q} + \frac{1}{p}x - x^{\frac{1}{p}}$$

y de esta desigualdad se obtiene $x^{\frac{1}{p}} \leq \frac{1}{q} + \frac{1}{p}x$.

Para $\alpha = 0$ o $\beta = 0$, la desigualdad de Young se verifica trivialmente. Supongamos que $\alpha > 0$, $\beta > 0$ y $x = \frac{\alpha}{\beta} \geq 1$ si $\alpha \geq \beta$ (en el caso contrario ponemos $x = \dfrac{\beta}{\alpha} \geq 1$). Remplazando $x = \frac{\alpha}{\beta}$ en la desigualdad precedente, resulta $\left(\frac{\alpha}{\beta}\right)^{\frac{1}{p}} \leq \frac{1}{q} + \frac{1}{p}\frac{\alpha}{\beta}$, luego $\alpha^{\frac{1}{p}} \beta^{-\frac{1}{p}+1} \leq \frac{1}{p}\alpha + \frac{1}{q}\beta$. Como $\frac{1}{p} + \frac{1}{q} = 1$, se tiene $\frac{1}{q} = 1 - \frac{1}{p}$, con lo que $\alpha^{\frac{1}{p}} \beta^{\frac{1}{q}} \leq \frac{1}{p}\alpha + \frac{1}{q}\beta$.

Desigualdad de Hölder.

Esta desigualdad se expresa como a continuación se indica:

$$\sum_{i=1}^n |x_i y_i| \leq \|\overrightarrow{x}\|_p \, \|\overrightarrow{y}\|_q \qquad \forall \overrightarrow{x} = (x_1, ..., x_n), \overrightarrow{y} = (y_1, ..., y_n) \in \mathbb{R}^n,$$

donde $p,\ q \in \left]1, \infty\right[$ tales que $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

Si $\overrightarrow{x} = 0$ o $\overrightarrow{y} = 0$, la desigualdad de Hölder se verifica trivialmente. Supongamos $\overrightarrow{x} \neq 0$, $\overrightarrow{y} \neq 0$.

Sean $\alpha = \dfrac{|x_i|^p}{\|\overrightarrow{x}\|_p^p}$, $\beta = \dfrac{|y_i|^q}{\|\overrightarrow{y}\|_q^q}$. Apliquemos la desigualdad de Young. Resulta

$$\left(\frac{|x_i|^p}{\|\overrightarrow{x}\|_p^p}\right)^{\frac{1}{p}} \left(\frac{|y_i|^q}{\|\overrightarrow{y}\|_q^q}\right)^{\frac{1}{q}} \leq \frac{1}{p}\frac{|x_i|^p}{\|\overrightarrow{x}\|_p^p} + \frac{1}{q}\frac{|y_i|^q}{\|\overrightarrow{y}\|_q^q} \qquad i = 1, ..., n,$$

con lo cual

$$\frac{|x_i|}{\|\overrightarrow{x}\|_p}\frac{|y_i|}{\|\overrightarrow{y}\|_q} \leq \frac{1}{p}\frac{|x_i|^p}{\|\overrightarrow{x}\|_p^p} + \frac{1}{q}\frac{|y_i|^q}{\|\overrightarrow{y}\|_q^q} \qquad i = 1, ..., n.$$

Sumando de 1 a n en cada miembro de la última desigualdad, obtenemos

$$\frac{1}{\|\overrightarrow{x}\|_p}\frac{1}{\|\overrightarrow{y}\|_q}\sum_{1=n}^{n}|x_i|\,|y_i| \leq \frac{1}{p\,\|\overrightarrow{x}\|_p^p}\sum_{i=1}^{n}|x_i|^p + \frac{1}{q\,\|\overrightarrow{y}\|_q^q}\sum_{i=1}^{n}|y_i|^q.$$

Puesto que $\|\overrightarrow{x}\|_p^p = \sum\limits_{i=1}^{n}|x_i|^p$, $\|\overrightarrow{y}\|_q^q = \sum\limits_{i=1}^{n}|y_i|^q$, entonces $\frac{1}{\|\overrightarrow{x}\|_p}\frac{1}{\|\overrightarrow{y}\|_q}\sum_{1=n}^{n}|x_i|\,|y_i| \leq \frac{1}{p} + \frac{1}{q} = 1$, y de esta desigualdad se deduce la lesigualdad de Hölder: $\sum_{1=n}^{n}|x_i y_i| \leq \|\overrightarrow{x}\|_p\,\|\overrightarrow{y}\|_q$.

Desigualdad triangular: $\|\overrightarrow{x} + \overrightarrow{y}\|_p \leq \|\overrightarrow{x}\|_p + \|\overrightarrow{y}\|_p \quad \forall \overrightarrow{x}, \overrightarrow{y} \in \mathbb{R}^n$.

De la definición de $\|\cdot\|_p$, obtenemos

$$\begin{aligned}
\|\overrightarrow{x} + \overrightarrow{y}\|_p^p &= \sum_{i=1}^{n}|x_i + y_i|^p = \sum_{i=1}^{n}|x_i + y_i|^{p-1}|x_i + y_i| \leq \sum_{i=1}^{n}|x_i + y_i|^{p-1}(|x_i| + |y_i|) \\
&= \sum_{i=1}^{n}|x_i + y_i|^{p-1}|x_i| + \sum_{i=1}^{n}|x_i + y_i|^{p-1}|y_i|.
\end{aligned}$$

Apliquemos la desigualdad de Hölder a cada sumando del lado derecho. Ya que $\frac{1}{p} + \frac{1}{q} = 1$, se obtiene $p + q = pq$ con lo que $p = q\,(p-1)$. Luego

$$\begin{aligned}
\sum_{i=q}^{n}|x_i + y_i|^{p-1}|x_i| &\leq \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}|x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}} = \|\overrightarrow{x}\|_p\left(\sum_{i=1}^{n}|x_i + y_i|^p\right)^{\frac{1}{q}} \\
&= \|\overrightarrow{x}\|_p\left(\|\overrightarrow{x} + \overrightarrow{y}\|_p^p\right)^{\frac{1}{q}},
\end{aligned}$$

Análogamente, $\sum_{i=1}^{n}|x_i + y_i|^{p-1}|y_i| \leq \|\overrightarrow{y}\|_p\left(\|\overrightarrow{x} + \overrightarrow{y}\|_p^p\right)^{\frac{1}{q}}$. Por lo tanto

$$\begin{aligned}
\|\overrightarrow{x} + \overrightarrow{y}\|_p^p &\leq \sum_{i=1}^{n}|x_i + y_i|^{p-1}|x_i| + \sum_{i=1}^{n}|x_i + y_i|^{p-1}|y_i| \\
&\leq \|\overrightarrow{x}\|_p\left(\|\overrightarrow{x} + \overrightarrow{y}\|_p^p\right)^{\frac{1}{q}} + \|\overrightarrow{y}\|_p\left(\|\overrightarrow{x} + \overrightarrow{y}\|_p^p\right)^{\frac{1}{q}} \\
&= \left(\|\overrightarrow{x}\|_p + \|\overrightarrow{y}\|_p\right)\|\overrightarrow{x} + \overrightarrow{y}\|_p^{\frac{p}{q}},
\end{aligned}$$

y de esta desigualdad, se deduce $\|\overrightarrow{x} + \overrightarrow{y}\|_p^{p-\frac{p}{q}} \leq \|\overrightarrow{x}\|_p + \|\overrightarrow{y}\|_p$. Nuevamente $\frac{1}{p} + \frac{1}{q} = 1$ entonces $1 = p - \frac{p}{q}$, con lo que se obtiene la desigualdad triangular: $\|\overrightarrow{x} + \overrightarrow{y}\|_p \leq \|\overrightarrow{x}\|_p + \|\overrightarrow{y}\|_p$.

Para $n = 2$, $\overrightarrow{x} = (x, y) \in \mathbb{R}^2$ la norma de Hölder está definida como sigue:

$$\|\overrightarrow{x}\|_p = \left(|x|^p + |y|^p\right)^{\frac{1}{p}} \quad \text{con} \quad p \in \left[1, \infty\right[.$$

Así, para $p = 1,$ la norma $\|\cdot\|_1$ está definida como $\|\overrightarrow{x}\|_1 = |x| + |y|$. Para $p = 2$, la norma $\|\cdot\|_2$ se escribe como $\|\overrightarrow{x}\|_2 = \left(x^2 + y^2\right)^{\frac{1}{2}}$. Esta se conoce como norma euclídea. Para $p = 3,$ la norma $\|\cdot\|_3$ se escribe como $\|\overrightarrow{x}\|_3 = \left(|x|^3 + |y|^3\right)^{\frac{1}{3}}$.

Para $n = 3$, $\quad \overrightarrow{x} = (x, y, z) \in \mathbb{R}^3$ la norma de Hölder $\|\cdot\|_p$ está definida como

$$\|\overrightarrow{x}\|_p = (|x|^p + |y|^p + |z|^p)^{\frac{1}{p}} \text{ con } p \in [1, \infty[.$$

Particularmente, para $p = 1$, $\quad \|\overrightarrow{x}\|_1 = |x| + |y| + |z|$. Para $p = 2$, $\quad \|\overrightarrow{x}\|_2 = \left(x^2 + y^2 + z^2\right)^{\frac{1}{2}}$. Esta es la norma euclídea. Para $p = 3{,}5 = \frac{7}{2}$, $\|\overrightarrow{x}\|_{\frac{7}{2}} = \left(|x|^{\frac{7}{2}} + |y|^{\frac{7}{2}} + |z|^{\frac{7}{2}}\right)^{\frac{2}{7}}$.

Consecuencias

1. Para $p = 1$ la norma $\|\cdot\|_1$ está definida como $\|\overrightarrow{x}\|_1 = \sum_{i=1}^n |x_i| \quad \overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$. Además, para $\overrightarrow{y} = (y_1, ..., y_n) \in \mathbb{R}^n$, se verifica la desigualdad de Hölder para $p = 1$ y $q = \infty$:

$$\sum_{i=1}^{n} |x_i y_i| \leq \|\overrightarrow{x}\|_1 \|\overrightarrow{y}\|_\infty .$$

En efecto, de la desigualdad $|y_i| \leq \underset{i=1,n}{\text{Max}} |y_i| = \|\overrightarrow{y}\|_\infty \quad i = 1, \ldots, n$, se sigue que

$$\sum_{i=1}^{n} |x_i y_i| = \sum_{i=1}^{n} |x_i| \, |y_i| \leq \sum_{i=1}^{n} (|x_i| \, \|\overrightarrow{y}\|_\infty) \leq \|\overrightarrow{y}\|_\infty \sum_{i=1}^{n} |x_i| = \|\overrightarrow{y}\|_\infty \|\overrightarrow{x}\|_1 .$$

Así, la desigualdad de Hölder es válida para $p = 1$ y $q = \infty$.

2. Sean $p \in]1, \infty[$ y $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ con $\overrightarrow{x} \neq 0$. Mostremos que $\lim\limits_{p \to \infty} \|\overrightarrow{x}\|_p = \|\overrightarrow{x}\|_\infty$. Primeramente $\|\overrightarrow{x}\|_\infty \leq \|\overrightarrow{x}\|_p \quad \forall \overrightarrow{x} \in \mathbb{R}^n$. En efecto,

$$\|\overrightarrow{x}\|_\infty = \underset{i=1,\ldots,n}{\text{Max}} |x_i| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = \|\overrightarrow{x}\|_p .$$

Por otro lado,

$$\begin{aligned} \|\overrightarrow{x}\|_p &= \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} \left(\underset{i=1,\ldots,n}{\text{Max}} |x_i|\right)^p\right)^{\frac{1}{p}} = \left(\sum_{i=1}^{n} \|\overrightarrow{x}\|_\infty^p\right)^{\frac{1}{p}} \\ &= \|\overrightarrow{x}\|_\infty \left(\sum_{i=1}^{n} 1\right)^{\frac{1}{p}} = n^{\frac{1}{p}} \|\overrightarrow{x}\|_\infty . \end{aligned}$$

De los dos resultados previos, se deduce la siguiente desigualdad: $\|\overrightarrow{x}\|_\infty \leq \|\overrightarrow{x}\|_p \leq n^{\frac{1}{p}} \|\overrightarrow{x}\|_\infty$. Tomando límite cuando $p \to \infty$, se obtiene

$$\lim_{p \to \infty} \|\overrightarrow{x}\|_\infty \leq \lim_{p \to \infty} \|\overrightarrow{x}\|_p \leq \lim_{p \to \infty} n^{\frac{1}{p}} \|\overrightarrow{x}\|_\infty ,$$

y como $\lim\limits_{p \to \infty} n^{\frac{1}{p}} = 1$, se deduce $\|\overrightarrow{x}\|_\infty \leq \lim_{p \to \infty} \|\overrightarrow{x}\|_p \leq \|\overrightarrow{x}\|_\infty$. Así, $\quad \lim_{p \to \infty} \|\overrightarrow{x}\|_p = \|\overrightarrow{x}\|_\infty$.

3. Relación entre $\|\cdot\|_p$ y $\|\cdot\|_q$ con $1 \leq p < q$. Mostremos que $\|\overrightarrow{x}\|_p \leq n^{\frac{q-p}{pq}} \|\overrightarrow{x}\|_q \quad \forall \overrightarrow{x} \in \mathbb{R}^n$ con $q > p \geq 1$.

   Sean $p,\, q \in [1,\infty[$ con $q > p$. Sea $r = \dfrac{q}{p} > 1$ y $s \in]1,\infty[$ tal que $\dfrac{1}{r} + \dfrac{1}{s} = 1$. Por la desigualdad de Hölder, para cada $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ se tiene

$$\begin{aligned}\|\overrightarrow{x}\|_p^p &= \sum_{i=1}^{n} |x_i|^p \leq \left(\sum_{i=1}^{n} 1^s\right)^{\frac{1}{s}} \left(\sum_{i=1}^{n} (|x_i|^p)^r\right)^{\frac{1}{r}} = n^{\frac{1}{s}} \left(\sum_{i=1}^{n} |x_i|^{pr}\right)^{\frac{1}{r}} = n^{\frac{1}{s}} \left(\sum_{i=1}^{n} |x_i|^q\right)^{\frac{1}{r}} \\ &= n^{\frac{1}{s}} \left(\|\overrightarrow{x}\|_q^q\right)^{\frac{1}{r}} = n^{\frac{1}{s}} \|\overrightarrow{x}\|_q^{\frac{q}{r}} = n^{\frac{1}{s}} \|\overrightarrow{x}\|_q^p .\end{aligned}$$

   Note que $r = \frac{q}{p}$ entonces $pr = q$ y $\frac{q}{r} = p$. Se tiene $\|\overrightarrow{x}\|_p^p \leq n^{\frac{1}{s}} \|\overrightarrow{x}\|_q^p$ y tomando la raíz p-ésima se deduce $\|\overrightarrow{x}\|_p \leq n^{\frac{1}{sp}} \|\overrightarrow{x}\|_q$. Como $\dfrac{1}{r} + \dfrac{1}{s} = 1$ y $r = \dfrac{q}{p}$ se sigue que $s = \frac{q}{q-p}$ y en consecuencia $\|\overrightarrow{x}\|_p \leq n^{\frac{q-p}{pq}} \|\overrightarrow{x}\|_q$.

   Así, si $p,\, q \in [1,\infty[$ tales que $q > p \geq 1$, $\quad \|\overrightarrow{x}\|_p \leq n^{\frac{q-p}{pq}} \|\overrightarrow{x}\|_q \quad \forall \overrightarrow{x} \in \mathbb{R}^n$.

4. Si $p = 2$ y $\overrightarrow{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, la norma $\|\cdot\|_2$ viene dada por $\|\overrightarrow{x}\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}}$ que se conoce con el nombre de norma euclídea. Además, para $p = q = 2$, $\overrightarrow{x} = (x_1, ..., x_n)$, $\overrightarrow{y} = (y_1, ..., y_n) \in \mathbb{R}^2$, la desigualdad de Hölder se escribe $\sum_{i=1}^n |x_i y_i| \leq \|\overrightarrow{x}\|_2 \|\overrightarrow{y}\|_2$, que coincide con la conocida desigualdad de Cauchy-Schwarz que se verá más adelante.

5. Sean $n \in \mathbb{Z}^+$, $\alpha_j \in \mathbb{R}^+$, $j = 1, ..., n$, las siguientes son normas en $\mathbb{R}^n$:

   a) $\|\overrightarrow{v}\| = \underset{j=1,\ldots,n}{\mathrm{Max}} \{\alpha_j |v_j|\} \qquad \overrightarrow{v} = (v_1, ..., v_n) \in \mathbb{R}^n$.

   b) $\|\overrightarrow{v}\| = \sum\limits_{j=1}^{n} \alpha_j |v_j| \qquad \overrightarrow{v} = (v_1, ..., v_n) \in \mathbb{R}^n$.

   c) $\|\overrightarrow{v}\| = \left(\sum\limits_{j=1}^{n} \alpha_j |v_j|^2\right)^{\frac{1}{2}} \qquad \overrightarrow{v} = (v_1, ..., v_n) \in \mathbb{R}^n$.

   d) $\|\overrightarrow{v}\| = \underset{j=1,\ldots,n}{\mathrm{Max}} \left\{\sum\limits_{i=1}^{j} \alpha_j |v_j|\right\} \qquad \overrightarrow{v} = (v_1, ..., v_n) \in \mathbb{R}^n$.

12.3.2. Normas geométricas de matrices.

Sean $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, se designa con $\mathcal{B}_V = \{\overrightarrow{e}_1, ..., \overrightarrow{e}_n\}$ y $\mathcal{B}_W = \left\{\overrightarrow{f}_1, ..., \overrightarrow{f}_m\right\}$ las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$, $p,\, q \in [1,\infty]$ y $\|\cdot\|_p$, $\|\cdot\|_q$ normas en $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente. Se denota con $M_{m\times n}[\mathbb{R}]$ el espacio vectorial de las matrices reales de $m \times n$ y $A \in M_{m\times n}[\mathbb{R}]$. Se define la aplicación lineal T de $\mathbb{R}^n$ en $\mathbb{R}^m$ como

$$T(\overrightarrow{x}) = A\overrightarrow{x} \quad \forall \overrightarrow{x} \in \mathbb{R}^n.$$

Entonces T es continua en todo punto $\overrightarrow{x_o} \in \mathbb{R}^n$. Más aún, debido a la linealidad de T, se tiene $T(\overrightarrow{x} - \overrightarrow{x}_o) = T(\overrightarrow{x}) - T(\overrightarrow{x}_o) \quad \forall \overrightarrow{x}, \overrightarrow{x}_o \in \mathbb{R}^n$, por lo que la continuidad de T en $\overrightarrow{x_o}$ es equivalente a la continuidad de T en el origen. Por lo tanto, dado $\epsilon > 0$, $\exists\ \delta > 0$ tal que $\forall \overrightarrow{x} \in \mathbb{R}^n$ con $\|\overrightarrow{x}\|_p < \delta \Longrightarrow \|T(\overrightarrow{x})\|_q < \epsilon$.

Sea $\overrightarrow{v} \in \mathbb{R}^n$ tal que $\|\overrightarrow{v}\|_p = 1$ y sea $\overrightarrow{x} = \frac{\delta}{2}\overrightarrow{v}$. Entonces,

$$\|\overrightarrow{x}\|_p = \left\|\frac{\delta}{2}\overrightarrow{v}\right\|_p = \frac{\delta}{2}\|\overrightarrow{v}\|_p = \frac{\delta}{2} < \delta,$$

y por la linealidad de T, se tiene $T(\overrightarrow{x}) = T(\frac{\delta}{2}\overrightarrow{v}) = \frac{\delta}{2}T(\overrightarrow{v})$, en consecuencia

$$\|T(\overrightarrow{x})\|_q = \frac{\delta}{2}\|T(\overrightarrow{v})\|_q < \epsilon,$$

de donde $\|T(\overrightarrow{v})\|_q < \frac{2\epsilon}{\delta} = M$.

Así, $\|T(\overrightarrow{v})\|_q \leq M \quad \forall \overrightarrow{v} \in \mathbb{R}^n$ con $\|\overrightarrow{v}\|_p = 1$, y de la definición de T se sigue que el conjunto $\left\{\|A\overrightarrow{x}\|_q \mid \overrightarrow{x} \in \mathbb{R}^n \text{ con } \|\overrightarrow{x}\|_p = 1\right\}$ es acotado superiormente. Este resultado nos permite definir las normas geométricas de matrices que a continuación se propone.

Definición 19 *Sea $A = (a_{ij}) \in M_{m\times n}[\mathbb{R}]$. La norma geométrica de la matriz A se denota con $|||A|||$ y se define como* $|||A||| = \sup_{\|\overrightarrow{x}\|_p\leq 1}\|A\overrightarrow{x}\|_q$.

Se verifica inmediatamente que $|||\cdot|||$ es una norma en $M_{m\times n}[\mathbb{R}]$. La prueba se deja como ejercicio. Además, de la definición de norma geométrica de una matriz se sigue inmediatamente que para toda matriz $A \in M_{m\times n}[\mathbb{R}]$ se verifica la desigualdad siguiente:

$$\|A\overrightarrow{x}\|_q \leq |||A|||\,\|\overrightarrow{x}\|_p \quad \forall \overrightarrow{x} \in \mathbb{R}^n.$$

Teorema 6 *Sea $A = (a_{ij}) \in M_{m\times n}[\mathbb{R}]$. Entonces,*

i. $|||A|||_1 = \sup\limits_{\|\overrightarrow{x}\|_1\leq 1}\|A\overrightarrow{x}\|_1 = \max\limits_{j=1,\ldots,n}\sum\limits_{i=1}^{m}|a_{ij}|$ *(máximo por columnas de A).*

ii. $|||A|||_\infty = \sup\limits_{\|\overrightarrow{x}\|_\infty\leq 1}\|A\overrightarrow{x}\|_\infty = \max\limits_{i=1,\ldots,m}\sum\limits_{j=1}^{n}|a_{ij}|$ *(máximo por filas de A).*

iii. $|||A|||_2 = \sup\limits_{\|\overrightarrow{x}\|_2\leq 1}\|A\overrightarrow{x}\|_2 = \left(\max\limits_{i=1,\ldots,n}|\lambda_i|\right)^{\frac{1}{2}}$, *donde $\lambda_1,\ldots\lambda_n$ son los valores propios de A^TA y A^T denota la matriz transpuesta de A.*

Demostración.

i. a) Probemos que $|||A|||_1 \leq \max\limits_{j=1,\ldots,n}\sum\limits_{i=1}^{m}|a_{ij}|$. Sea $\overrightarrow{x}^T = (x_1,\ldots,x_n) \in \mathbb{R}^n$ tal que $\|\overrightarrow{x}\|_1 = \sum\limits_{j=1}^{n}|x_j| = 1$. Entonces,

$$A\overrightarrow{x} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n\end{bmatrix} = \begin{bmatrix} \sum\limits_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{mj}x_j\end{bmatrix},$$

y por la definición de la norma $\|\cdot\|_1$ en $\mathbb{R}^m$, se sigue que

$$\begin{aligned}\|A\overrightarrow{x}\|_1 &= \sum_{i=1}^{m}\left(\left|\sum_{j=1}^{n}a_{ij}x_j\right|\right) \leq \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|\,|x_j| = \sum_{j=1}^{n}\left(|x_j|\sum_{i=1}^{m}|a_{ij}|\right) \\ &\leq \sum_{j=1}^{n}\left(|x_j|\max_{j=1,\ldots,n}\sum_{i=1}^{m}|a_{ij}|\right) = \left(\max_{j=1,\ldots,n}\sum_{i=1}^{m}|a_{ij}|\right)\sum_{j=1}^{n}|x_j| = \left(\max_{j=1,\ldots,n}\sum_{i=1}^{m}|a_{ij}|\right).\end{aligned}$$

Por lo tanto, de la definición de la norma geométrica $|||\cdot|||_1$, se obtiene la desigualdad siguiente:

$$|||A|||_1 = \sup_{\|\overrightarrow{x}\|_1\leq 1}\|A\overrightarrow{x}\|_1 \leq \max_{j=1,\ldots,n}\sum_{i=1}^{m}|a_{ij}|.$$

b) Probemos que $\max\limits_{j=1,\ldots,n}\sum\limits_{i=1}^{m}|a_{ij}| \leq |||A|||_1$. Para el efecto, sea k la columna para la cual se verifica la igualdad $\max\limits_{j=1,\ldots,n}\sum\limits_{i=1}^{m}|a_{ij}| = \sum\limits_{i=1}^{m}|a_{ik}|$. Para $\overrightarrow{x}=\overrightarrow{e_k}$, el k-ésimo vector de la base canónica de $\mathbb{R}^n$, se tiene $A\overrightarrow{e_k}=\begin{bmatrix} a_{1k}\\ \vdots \\ a_{mk}\end{bmatrix}$, y en consecuencia $\|A\overrightarrow{e}_k\|_1=\sum_{i=1}^{m}|a_{ik}|=\max_{j=1,\ldots,n}\sum_{i=1}^{m}|a_{ik}|$.

Nuevamente, de la definición de la norma geométrica $|||\cdot|||_1$, se obtiene la desigualdad siguiente:

$$\max_{j=1,\ldots,n}\sum_{i=1}^{n}|a_{ij}| = \|A\overrightarrow{e}_k\|_1 \leq \sup_{\|\overrightarrow{x}\|_1\leq 1}\|A\overrightarrow{x}\|_1 = |||A|||_1 .$$

De las desigualdades obtenidas en las partes a) y b) se deduce finalmente el resultado buscado:

$$|||A|||_1 = \sup_{\|\overrightarrow{x}\|\leq 1}\|A\overrightarrow{x}\|_1 = \max_{j=1,\ldots,n}\sum_{i=1}^{n}|a_{ij}| .$$

ii. Primeramente, obtenemos la desigualdad $|||A|||_\infty \leq \max\limits_{i=1,\ldots,m}\sum\limits_{j=1}^{n}|a_{ij}|$. Sea $\overrightarrow{x}^T=(x_1,\ldots,x_n)\in\mathbb{R}^n$ tal que $\|\overrightarrow{x}\|_\infty=\max\limits_{j=1,\ldots,n}|x_j|=1$. Entonces, $|x_j|\leq\|\overrightarrow{x}\|_\infty,\ \ j=1,\ldots,n$, y

$$\|A\overrightarrow{x}\|_\infty = \max_{i=1,\ldots,m}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \leq \max_{i=1,\ldots,m}\sum_{j=1}^{n}|a_{ij}|\,|x_j| \leq \max_{i=1,\ldots,m}\sum_{j=1}^{n}|a_{ij}| ,$$

consecuentemente

$$\|A\overrightarrow{x}\|_\infty \leq \sup_{\|\overrightarrow{x}\|_\infty\leq 1}\|A\overrightarrow{x}\|_\infty = |||A|||_\infty \leq \max_{i=1,\ldots,m}\sum_{j=1}^{n}|a_{ij}| .$$

Mostremos a continuación la desigualdad $\max\limits_{i=1,\ldots,m}\sum\limits_{j=1}^{n}|a_{ij}| \leq |||A|||_\infty$. Para el efecto, sea k el índice para el cual la fila k-ésima de A es tal que $\max_{i=1,\ldots,m}\sum_{j=1}^{n}|a_{ij}|=\sum_{j=1}^{n}|a_{kj}|$. Sea $\overrightarrow{x}^T=(x_1,\ldots,x_n)\in\mathbb{R}^n$ definido como sigue: $x_j=\begin{cases}\dfrac{|\,a_{kj}\,|}{a_{kj}}, & \text{si }\ a_{kj}\neq 0,\\ 0, & \text{si }\ a_{kj}=0.\end{cases}$ Se tiene $\|\overrightarrow{x}\|_\infty=1$, y

$$\|A\overrightarrow{x}\|_\infty = \max_{i=1,\ldots,m}\left|\sum_{j=1}^{n}a_{ij}x_j\right| = \sum_{j=1}^{n}|a_{kj}| ,$$

de donde

$$\sum_{j=1}^{n}|a_{kj}| = \|A\overrightarrow{x}\|_\infty \leq \sup_{\|\overrightarrow{x}\|_\infty\leq 1}\|A\overrightarrow{x}\|_\infty = |||A|||_\infty .$$

Por lo tanto, $|||A|||_\infty=\max_{i=1,\ldots,m}\sum_{j=1}^{n}|a_{ij}|$.

iii. Las normas euclídeas en $\mathbb{R}^n$ y $\mathbb{R}^m$ están definidas como

$$\begin{aligned}\|\overrightarrow{x}\|_2 &= \left(\overrightarrow{x}^T\ \overrightarrow{x}\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}} \quad \forall\overrightarrow{x}\in\mathbb{R}^n,\\ \|A\overrightarrow{x}\|_2^2 &= \left(A\overrightarrow{x}\right)^T\ A\overrightarrow{x} = \overrightarrow{x}^TA^TA\overrightarrow{x} \quad \forall\overrightarrow{x}\in\mathbb{R}^n.\end{aligned}$$

Definimos la función $\overrightarrow{g}$ de $\mathbb{R}^n$ en $\mathbb{R}$ como sigue $g(x)=\overrightarrow{x}^TA^TA\overrightarrow{x}\quad \overrightarrow{x}\in\mathbb{R}^n$. Esta función es diferenciable y el gradiente de g está definido como $\nabla g(\overrightarrow{x})=2A^TA\overrightarrow{x}\quad\forall\overrightarrow{x}\in\mathbb{R}^n$.

Sea $S = \left\{ \overrightarrow{x} \in \mathbb{R}^n \mid \overrightarrow{x}^T \overrightarrow{x} = 1 \right\}$, y consideramos el problema siguiente: $\max_{\overrightarrow{x} \in S} g(x)$. Note que $g(\overrightarrow{x}) = \|A\overrightarrow{x}\|_2^2$, de modo que

$$|||A|||_2 = \sup_{\|\overrightarrow{x}\|_2 \leq 1} \|A\overrightarrow{x}\|_2 = \sup_{\|\overrightarrow{x}\|_2 \leq 1} [g(\overrightarrow{x})]^{\frac{1}{2}}.$$

Apliquemos el método de los multiplicadores de Lagrange. Definimos

$$\Phi(\overrightarrow{x}, \lambda) = g(\overrightarrow{x}) + \lambda\left(1 - \overrightarrow{x}^T \overrightarrow{x}\right) \quad \overrightarrow{x} \in \mathbb{R}^n,$$

y λ es el multiplicador de Lagrange. Por las condiciones necesarias de extremo, tenemos el par de ecuaciones siguiente:

$$\begin{aligned} \nabla_{\overrightarrow{x}} \Phi(\overrightarrow{x}, \lambda) &= \nabla g(\overrightarrow{x}) - 2\lambda\overrightarrow{x} = 2A^T A\overrightarrow{x} - 2\lambda\overrightarrow{x} = 0, \\ \nabla_\lambda \Phi(\overrightarrow{x}, \lambda) &= 1 - \overrightarrow{x}^T \overrightarrow{x} = 0. \end{aligned}$$

Se obtiene el siguiente sistema de ecuaciones: $\overrightarrow{x} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, $\begin{cases} \left(A^T A - \lambda I\right)\overrightarrow{x} = 0, \\ \overrightarrow{x}^T \overrightarrow{x} = 1. \end{cases}$ Así, la determinación de los puntos críticos de $\Phi(\overrightarrow{x}, \lambda)$ se transforma en el clásico problema de valores propios: $A^T A\overrightarrow{x} = \lambda\overrightarrow{x}$.

Puesto que $A^T A$ es simétrica, se sabe que los valores propios son reales. Sean $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales valores propios, y $\overrightarrow{x}_1, \ldots, \overrightarrow{x}_n \in \mathbb{R}^n$ los respectivos vectores propios tales que$\|\overrightarrow{x}_i\|_2 = 1$ $i = 1, \ldots, n$; esto es,

$$\begin{cases} A^T A\overrightarrow{x}_i = \lambda_i x_i \quad i = 1, \ldots, n \\ \|\overrightarrow{x}_i\|_2^2 = 1. \end{cases}$$

De la definición de la función g se deduce

$$0 \leq g(\overrightarrow{x}_i) = \overrightarrow{x}_i^T A^T A\overrightarrow{x}_i = \overrightarrow{x}_i^T \lambda_i \overrightarrow{x}_i = \lambda_i \quad i = 1, \ldots, n.$$

Por lo tanto $|||A|||_2 = \sup_{\|\overrightarrow{x}\|_2 \leq 1} \|A\overrightarrow{x}\|_2 = \left(\max_{i=1,\ldots,n} |\lambda_i|\right)^{\frac{1}{2}}$.

■

Observación

1. Sea $\overrightarrow{A} = (a_1, ..., a_n) \in \mathbb{R}^n$. Se tiene $A = (a_1, ..., a_n) \in M_{1\times n}[\mathbb{R}]$. Entonces $|||A|||_1 = \left\|\overrightarrow{A}\right\|_\infty$,y $|||A|||_\infty = \left\|\overrightarrow{A}\right\|_1$.

   Las normas geométricas de matrices son submultiplicativas, como se muestra en el siguiente teorema.

Teorema 7 *Sean A, $B \in M_{n\times n}[\mathbb{R}]$. Entonces*

i. $|||AB|||_1 \leq |||A|||_1 \, |||B|||_1$,

ii. $|||AB|||_\infty \leq |||A|||_\infty \, |||B|||_\infty$.

iii. $|||AB|||_2 \leq |||A|||_2 \, |||B|||_2$.

Demostración.

i. Sea $\overrightarrow{x} \in \mathbb{R}^n$ con $\overrightarrow{x} \neq 0$ tal que $\overrightarrow{y} = B\overrightarrow{x} \neq 0$. Entonces $\frac{\|AB\overrightarrow{x}\|_1}{\|\overrightarrow{x}\|_1} = \frac{\|A\overrightarrow{y}\|_1}{\|\overrightarrow{y}\|_1} \cdot \frac{\|\overrightarrow{y}\|_1}{\|\overrightarrow{x}\|_1}$. Luego

$$\begin{aligned} |||AB|||_1 &= \sup_{\|\overrightarrow{x}\|_1 \leq 1} \frac{\|AB\overrightarrow{x}\|_1}{\|\overrightarrow{x}\|_1} = \sup_{\|\overrightarrow{x}\|_1 \leq 1} \frac{\|A\overrightarrow{y}\|_1}{\|\overrightarrow{y}\|_1} \cdot \frac{\|\overrightarrow{y}\|_1}{\|\overrightarrow{x}\|_1} \\ &\leq \sup_{\|\overrightarrow{y}\|_1 \leq 1} \frac{\|A\overrightarrow{y}\|_1}{\|\overrightarrow{y}\|_1} \cdot \sup_{\|\overrightarrow{x}\|_1 \leq 1} \frac{\|B\overrightarrow{y}\|_1}{\|\overrightarrow{x}\|_1} = |||A|||_1 \, |||B|||_1. \end{aligned}$$

Otra forma de obtener este resultado se muestra a continuación:

$$\|AB\overrightarrow{x}\|_1 = \|AB\overrightarrow{x}\|_1 \leq |||A|||_1 \, |||B\overrightarrow{x}|||_1 \leq |||A|||_1 \, |||B|||_1 \, \|\overrightarrow{x}\|_1 \, ,$$

de donde

$$|||A|||_1 = \max_{\|\overrightarrow{x}\|_1 \leq 1} \|AB\overrightarrow{x}\|_1 \leq |||A|||_1 \, |||B|||_1 \, .$$

ii. Sea $\overrightarrow{x} \in \mathbb{R}^n$. Entonces,

$$\|AB\overrightarrow{x}\|_\infty = \|A\,(B\overrightarrow{x})\|_\infty \leq |||A|||_\infty \, |||B\overrightarrow{x}|||_\infty \leq |||A|||_\infty \, |||B|||_\infty \, \|\overrightarrow{x}\|_\infty \, ,$$

y para $\|\overrightarrow{x}\|_\infty \leq 1$, se obtiene $|||AB|||_\infty \leq |||A|||_\infty \, |||B|||_\infty \, .$

iii. Es inmediata.

■

Otra clase de normas en $M_{m\times n}[\mathbb{R}]$ se definen a continuación, donde $A = (a_{ij})_{m\times n} \in M_{m\times n}[\mathbb{R}]$.

a) $N(A) = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} |a_{ij}| \, .$

b) $N(A) = \left(\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}} .$

c) $N(A) = \operatorname*{Max}\limits_{i=1,\ldots,m} \operatorname*{Max}\limits_{j=1,\ldots,n} |a_{ij}| \, .$

d) $N(A) = \operatorname*{Max}\limits_{i=1,\ldots,m} \sum\limits_{j=1}^{n} |a_{ij}| \, .$

e) $N(A) = \operatorname*{Max}\limits_{j=1,\ldots,n} \sum\limits_{i=1}^{m} |a_{ij}| \, .$

f) Sean $p \in [1, \infty[$ y $A = (a_{ij})_{m\times n} \in M_{m\times n} \in [\mathbb{R}]$. Se define $\|A\|_p = \left(\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} |a_{ij}|^p \right)^{\frac{1}{p}}$, entonces $\|\cdot\|_p$ es una norma de Hölder sobre $M_{m\times n}[\mathbb{R}]$.

12.3.3. Normas en el espacio de funciones continuas $C([a,b])$.

Sean $a, b \in \mathbb{R}$ tales que $a < b$. Se denota con $C([a,b])$ al espacio de todas las funciones continuas en $[a,b]$. En este espacio se van a definir dos normas: la norma de Chebyshev notada $\|\cdot\|_\infty$ y la norma de Hölder que se denota con $\|\cdot\|_p$, donde $p \in [1, \infty[$.

Norma de Chebyshev.

Se define la función $\|\cdot\|_\infty$ de $C([a,b])$ en $\mathbb{R}$ como se indica a continuación:

$$\|f\|_\infty = \operatorname*{Max}_{x\in[a,b]} |f(x)| \quad \forall f \in C([a,b]) \, .$$

Esta se conoce como norma de Chebyshev. Probemos que $\|\cdot\|_\infty$ es una norma sobre $C([a,b])$. En efecto,

i) De la definición de $\|\cdot\|_\infty$, se tiene $\|f\|_\infty \geq 0$.

ii) Si $f = 0$, esto es, $f(x) = 0 \quad \forall x \in [a,b]$, $\|f\|_\infty = 0$. Recíprocamente, si $\|f\|_\infty = 0 = \operatorname*{Max}\limits_{x\in[a,b]} |f(x)|$, entonces $f(x) = 0$ $\forall x \in [a,b]$, es decir que $f = 0$.

iii) Sean $\lambda \in \mathbb{R}$, $f \in C([a,b])$. Entonces $\|\lambda f\|_\infty = \operatorname*{Max}\limits_{x\in[a,b]} |\lambda f(x)| = |\lambda| \, \|f\|_\infty \, .$

iv) Sean $f,\, g \in C\left([a,b]\right)$. Entonces $\left|f\left(x\right)+g\left(x\right)\right| \leq \left|f\left(x\right)\right| + \left|g\left(x\right)\right| \quad \forall x \in [a,b]$. Resulta,

$$\|f+g\|_\infty = \underset{x\in[a,b]}{\text{Max}} \left|f\left(x\right)+g\left(x\right)\right| \leq \underset{x\in[a,b]}{\text{Max}} \left|f\left(x\right)\right| + \underset{x\in[a,b]}{\text{Max}} \left|g\left(x\right)\right| \leq \|f\|_\infty + \|g\|_\infty .$$

Conclusión**:** $\|\cdot\|_\infty$ es una norma en $C\left([a,b]\right)$.

Notación**:** El espacio $C\left([a,b]\right)$ provisto de la norma $\|\cdot\|_\infty$ se le nota $\mathcal{L}^\infty\left([a,b]\right)$.

Normas hölderianas en el espacio de funciones continuas $C([a,b])$.

Sea $p \in [1,\infty[$. Se define la función $\|\cdot\|_p$ de $C\left([a,b]\right)$ en $\mathbb{R}$ como sigue:

$$\|f\|_p = \left(\int_a^b \left|f\left(x\right)\right|^p dx\right)^{\frac{1}{p}} \qquad \forall f \in C\left([a,b]\right).$$

Probemos que $\|\cdot\|_p$ es una norma sobre $C\left([a,b]\right)$ denominada norma hölderiana. Para el efecto mostremos que $\|\cdot\|_p$ satisface las cuatro propiedades de la definición de norma.

i) Es claro que $\|f\|_p \geq 0 \quad \forall f \in C\left([a,b]\right)$.

ii) Si $f=0$ se tiene $\|f\|_p = 0$. Supongamos que $\|f\|_p = 0$ y probemos que $f=0$, o lo que es equivalente a probar que $f \neq 0 \Rightarrow \|f\|_p \neq 0$. Recuerde que si u, v son proposiciones, se tiene la siguiente tautología: $(u \Longrightarrow v) \Longleftrightarrow [(\sim v) \Longrightarrow (\sim u)]$. Efectivamente, si $f \neq 0$, existe $x_0 \in [a,b]$ tal que $f\left(x_0\right) \neq 0$ y como f es continua, existe $[\alpha,\beta] \subset [a,b]$ tal que $f\left(x\right) \neq 0 \quad \forall x \in [\alpha,\beta]$. Luego

$$0 < \int_\alpha^\beta \left|f\left(x\right)\right|^p dx \leq \int_a^b \left|f\left(x\right)\right|^p dx = \|f\|_p^p ,$$

es decir que $\|f\|_p > 0$. Así, $f \neq 0 \Rightarrow \|f\|_p > 0$, o lo que es lo mismo $\|f\|_p = 0 \Rightarrow f = 0$.

iii) Sean $\lambda \in \mathbb{R},\ f \in C\left([a,b]\right)$. Se verifica inmediatamente que $\|\lambda f\|_p = |\lambda|\, \|f\|_p$.

iv) Mediante un procedimiento análogo al de la demostración de la desigualdad triangular para la norma hölderiana en $\mathbb{R}^n$, se obtiene de la desigualdad de Young, la desigualdad de Hölder que para funciones continuas se establece del modo siguiente: para todo $p,\, q \in [1,\infty[$ tales que $\dfrac{1}{p}+\dfrac{1}{q}=1$, $f, g \in \mathcal{C}\left([a,b]\right)$, entonces $fg \in \mathcal{C}(\,[a,b])$ y

$$\|fg\|_1 = \int_a^b \left|f\left(x\right)g\left(x\right)\right| dx \leq \|f\|_p \|g\|_q ,$$

o lo que es lo mismo

$$\int_a^b \left|f\left(x\right)g\left(x\right)\right| dx \leq \left(\int_a^b \left|f\left(x\right)\right|^p dx\right)^{\frac{1}{p}} \left(\int_a^b \left|g\left(x\right)\right|^q dx\right)^{\frac{1}{q}} .$$

Para probar esta desigualdad se pone $\alpha = \frac{|f(x)|^p}{\|f\|_p^p}$, $\beta = \frac{|g(x)|^q}{\|g\|_q^q}$ con $f \neq 0,\ g \neq 0$ en la desigualdad de Young que ha sido establecida anteriormente. Se obtiene

$$\frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|g(x)|^q}{\|g\|_q^q} \qquad x \in [a,b],$$

y luego se integra sobre el intervalo $[a,b]$, esto es

$$\frac{1}{\|f\|_p\|g\|_q}\int_a^b \left|f\left(x\right)g\left(x\right)\right| dx \leq \frac{1}{p}\frac{1}{\|f\|_p^p}\int_a^b \left|f\left(x\right)\right|^p dx + \frac{1}{q}\frac{1}{\|g\|_q^q}\int_a^b \left|g\left(x\right)\right|^q dx = 1,$$

y de esta desigualad se obtiene la desigualdad de Hölder.

Sean $f,\, g \in C\left([a,b]\right)$. Utilizando la desigualdad de Hölder se obtiene la desigualdad triangular siguiente:

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p .$$

Conclusión: $\|\cdot\|_p$ es una norma sobre $C([a,b])$, $p \in [1,\infty[$.

Notación: El espacio $C([a,b])$ provisto de la norma $\|\cdot\|_p$ se le nota $\mathcal{L}^p([a,b])$.

Sea $f \in C([a,b])$. Para $p=1$, la norma $\|\cdot\|_1$ está definida como $\|f\|_1 = \int_a^b |f(x)|\,dx$. Note que $f \in \mathcal{L}^1([a,b])$. Para $p=2$, $\|f\|_2 = \left(\int_a^b |f(x)|^2\,dx\right)^{\frac{1}{2}}$ es la norma euclídea. Se tiene $f \in \mathcal{L}^2([a,b])$. Si $p=4$, la función $\|\cdot\|_4$ está definida como $\|f\|_4 = \left(\int_a^b |f(x)|^4\,dx\right)^{\frac{1}{4}}$. Note que $f \in \mathcal{L}^4([a,b])$.

Para $p=q=2$, la desigualdad de Hölder se expresa como sigue:

$$\|fg\|_1 = \int_a^b |f(x)\,g(x)|\,dx \leq \|f\|_2\,\|g\|_2 \qquad \forall f,g \in C([a,b]),$$

que coincide con la conocida desigualdad de Cauchy-Schwarz.

Es importante observar el significado de los espacios $\mathcal{L}^p([a,b])$ que aquí hemos dado: simplemente es el espacio $C([a,b])$ provisto de la norma $\|\cdot\|_p$. Estos espacios $\mathcal{L}^p([a,b])$ son diferentes de los espacios $L^p(a,b)$ que designan a los espacios de Lebesgue que se tratan en los cursos de Análisis Funcional, Teoría de Integración, etc. Para información $\mathcal{L}^p([a,b]) \subset L^p(a,b)$. Similarmente $\mathcal{L}^\infty([a,b]) \subset L^\infty(a,b)$, donde $L^\infty(a,b)$ pertenece a la clase de los espacios de Lebesgue.

Sean $\omega \in C([a,b])$ tal que $\omega(x) > 0 \quad \forall x \in [a,b]$, las siguientes son normas en $C([a,b])$:

a) $\|f\| = \int_a^b \omega(x)\,|f(x)|\,dx \quad f \in C([a,b])$.

b) $\|f\| = \left(\int_a^b \omega(x)\,f^2(x)\,dx\right)^{\frac{1}{2}} \quad f \in C([a,b])$.

Otras normas en $C^1([0,10])$ se definen a continuación:

a) $N(u) = \underset{x\in[0,10]}{\mathrm{Max}} \{|u(x)|, |u'(x)|\}$.

b) $M(u) = \int_0^{10} (|u(x)| + |u'(x)|)\,dx$.

c) $R(u) = \left(\int_0^{10} (|u(x)|^p + |u'(x)|^p)\,dx\right)^{\frac{1}{p}}$ para $p \in]1,\infty[$.

12.4. Espacios con producto interno.

En esta sección revisamos brevemente una clase de espacios vectoriales reales V en los que se define un producto escalar (dicho también producto interno o producto punto) que los denominaremos espacios con producto escalar o espacios con producto punto, o espacios euclídeos. En esta clase de espacios se introducirán las nociones geométricas de ángulo, de perpendicularidad u ortogonalidad, la conocida ley del paralelogramo y el teorema de Pitágoras.

Enfatizaremos en dos clases de espacios $\mathbb{R}^n$ y $C([a,b])$.

Definición 20 *Sea V un espacio vectorial real. Un producto interno o producto escalar en V es una función denotada $\langle\cdot,\cdot\rangle$ de $V \times V$ en $\mathbb{R}$ que satisface las siguientes propiedades:*

i. $\langle x,y\rangle = \langle y,x\rangle \quad \forall x,\, y \in V$,

ii. $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle \quad \forall x,\, y,\, z \in V$,

iii. $\langle \lambda x,y\rangle = \lambda\langle x,y\rangle \quad \forall \lambda \in \mathbb{R}, \quad \forall x,\, y \in V$,

iv. $\langle x,x\rangle = 0 \Leftrightarrow x = 0$,
$\langle x,x\rangle > 0 \Leftrightarrow x \neq 0 \quad x \in V$.
Para $x,\, y \in V$, el número real $\langle x,y\rangle$ se llama producto escalar o producto interno de x con y.

Definición 21 *Un espacio vectorial V en el que se ha definido un producto escalar $\langle \cdot, \cdot \rangle$ se denomina espacio con producto interior, espacio con producto escalar o espacio prehilbertiano.*

Observación

Si V es un espacio vectorial complejo y $\langle \cdot, \cdot \rangle$ denota un producto escalar en V, la propiedad i) se escribe como $\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \forall x,\, y \in V,$ donde el lado derecho de la igualdad designa el número complejo conjugado de $\langle y, x \rangle$. Las propiedades ii), iii) y iv) de la definición de producto escalar permanecen invariables.

Ejemplos

1. Sea $V = \mathbb{R}^n$. Un producto escalar $\langle \cdot, \cdot \rangle$ en $\mathbb{R}^n$ se define como $\langle \overrightarrow{x}, \overrightarrow{y} \rangle = \sum_{i=1}^{n} x_i y_i$, donde $\overrightarrow{x} = (x_1, ..., x_n)$, $\overrightarrow{y} = (y_1, ..., y_n) \in \mathbb{R}^n$. En notación matricial, esto es, los elementos de $\mathbb{R}^n$ se escriben como vectores columna, el producto escalar $\langle \cdot, \cdot \rangle$ en $\mathbb{R}^n$ se escribe como

$$\langle \overrightarrow{x}, \overrightarrow{y} \rangle = \overrightarrow{x}^T \overrightarrow{y} = \sum_{i=1}^{n} x_i y_i,$$

donde $\overrightarrow{x}^T = (x_1, ..., x_n)\,,\ \overrightarrow{y}^T = (y_1, ..., y_n) \in \mathbb{R}^n$ denotan los vectores transpuestos de los vectores columna $\overrightarrow{x}$ e $\overrightarrow{y}$. Las propiedades i) a iv) de la definición de producto escalar se verifican fácilmente y se dejan como ejercicio.

Para $n = 2$, el producto $\langle \cdot, \cdot \rangle$ en $\mathbb{R}^2$ está dado como sigue: si $\overrightarrow{x} = (a_1, b_1)$, $\overrightarrow{y} = (a_2, b_2) \in \mathbb{R}^2$, entonces $\langle \overrightarrow{x}, \overrightarrow{y} \rangle = a_1 a_2 + b_1 b_2$.

Para $n = 3$, el producto $\langle \cdot, \cdot \rangle$ en $\mathbb{R}^3$ está dado como $\langle \overrightarrow{x}, \overrightarrow{y} \rangle = a_1 a_2 + b_1 b_2 + c_1 c_2$ con $\overrightarrow{x} = (a_1, b_1, c_1)$, $\overrightarrow{y} = (a_2, b_2, c_2) \in \mathbb{R}^3$.

2. Sea $V = C\left([a, b]\right)$. Un producto escalar $\langle \cdot, \cdot \rangle$ en $C\left([a, b]\right)$ se define como

$$\langle f, g \rangle = \int_a^b f\left(x\right) g\left(x\right) dx \quad \forall f,\, g \in C\left([a, b]\right).$$

Por ejemplo, si $f,\, g \in C\left([-1, 1]\right)$ están dadas como $f\left(x\right) = x^3$, $g\left(x\right) = x^2 + 1 \quad x \in [-1, 1]$. Entonces

$$\langle f, g \rangle = \int_{-1}^{1} f\left(x\right) g\left(x\right) dx = \int_{-1}^{1} x^3 \left(x^2 + 1\right) dx = 0.$$

Probemos que la función $\langle \cdot, \cdot \rangle$ de $C\left([a, b]\right) \times C\left([a, b]\right)$ en $\mathbb{R}$ satisface las cuatro propiedades de la definición de producto escalar. Para ello utilicemos algunas propiedades de las funciones integrables y de las funciones continuas. Sean $f,\, g,\, h \in C\left([a, b]\right)$ y $\lambda \in \mathbb{R}$. Entonces

i. $\langle f, g \rangle = \int_a^b f\left(x\right) g\left(x\right) dx = \int_a^b g\left(x\right) f\left(x\right) dx = \langle g, f \rangle$.

ii. $\langle f + g, h \rangle = \int_a^b \left[f\left(x\right) + g\left(x\right)\right] h\left(x\right) dx = \int_a^b f\left(x\right) h\left(x\right) dx + \int_a^b g\left(x\right) h\left(x\right) dx = \langle f, h \rangle + \langle g, h \rangle$.

iii. $\langle \lambda f, g \rangle = \int_a^b \lambda f\left(x\right) g\left(x\right) dx = \lambda \int_a^b f\left(x\right) g\left(x\right) dx = \lambda \langle f, g \rangle$.

iv. Si $f = 0$ es claro que $\langle f, f \rangle = 0$. Mostremos que si $\langle f, f \rangle = 0$ entonces $f = 0$. Para ello, haciendo uso de la tautología $(p \Rightarrow q) \Leftrightarrow [(\sim q) \Rightarrow (\sim p)]$ en la que p, q son las proposiciones siguientes $p:\ f = 0,\ \ q: \langle f, f \rangle = 0$; tenemos la proposición siguiente: $f \neq 0 \Rightarrow \langle f, f \rangle > 0$.

Si $f \neq 0$, existe $x_0 \in [a, b]$ tal que $f\left(x_0\right) \neq 0$. Por hipótesis f es continua, por lo tanto es continua en x_0 y siendo $f\left(x_0\right) \neq 0$, existe un intervalo $[\alpha, \beta] \subset [a, b]$ tal que $f\left(x\right) \neq 0$ $\forall x \in [\alpha, \beta]$. Luego

$$0 < \int_\alpha^\beta f^2\left(x\right) dx \leq \int_a^b f^2\left(x\right) dx = \langle f, f \rangle.$$

Así, $f \neq 0 \Rightarrow \langle f, f \rangle > 0$ y por la tautología antes citada se deduce $\langle f, f \rangle = 0 \Rightarrow f = 0$. Consecuentemente, $\langle f, f \rangle = 0 \Leftrightarrow f = 0$. Además, del resultado precedente, es claro que $\langle f, f \rangle > 0 \quad \forall f \in C\left([a, b]\right)$ con $f \neq 0$.

3. Sean $V = M_{n\times n}[\mathbb{R}]$, $A = (a_{ij}) \in M_{n\times n}[\mathbb{R}]$. La traza de la matriz A se nota $tr(A)$ y se define como $tr(A) = \sum_{i=1}^{n} a_{ii}$. Una función $\langle \cdot,\cdot\rangle$ de $M_{n\times n}[\mathbb{R}] \times M_{n\times n}[\mathbb{R}]$ en $\mathbb{R}$ definida como
$$\langle A, B\rangle = tr\left(B^T A\right) \quad \forall A,\, B \in M_{n\times n}[\mathbb{R}]$$
es un producto escalar en $M_{n\times n}[\mathbb{R}]$.

Propiedades adicionales del producto escalar.

En un espacio prehilbertiano real V se verifican las propiedades siguientes:

i. $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle \quad \forall x,\, y,\, z \in V,$

ii. $\langle x, \lambda y\rangle = \lambda \langle x, y\rangle \quad \forall \lambda \in \mathbb{R},\ \forall x,\, y \in V,$

iii. $\langle x - y, z\rangle = \langle x, z\rangle - \langle y, z\rangle \quad \forall x,\, y,\, z \in V,$
$\langle x, y - z\rangle = \langle x, y\rangle - \langle x, z\rangle \quad \forall x,\, y,\, z \in V,$

iv. $\langle 0, x\rangle = \langle x, 0\rangle = 0 \quad \forall x \in V.$

v. Sean $x,\, y \in V$, si para todo $z \in V$, $\langle x, z\rangle = \langle y, z\rangle$, entonces $x = y$.

vi. Sean $x_1, ..., x_n,\ y \in V$, $\alpha_1, ..., \alpha_n \in \mathbb{R}$. Entonces
$$\left\langle \sum_{i=1}^{n} \alpha_i x_i,\ y \right\rangle = \sum_{i=1}^{n} \alpha_i \langle x_i,\ y\rangle, \quad \text{y,} \quad \left\langle y, \sum_{i=1}^{n} \alpha_i x_i \right\rangle = \sum_{i=1}^{n} \alpha_i \langle x_i,\ y\rangle.$$

En un espacio vectorial real V se pueden definir una infinidad de productos escalares. En los ejercicios se exhiben algunos productos escalares definidos en $\mathbb{R}^2$ y en $C([0,1])$.

Longitud o norma de un vector

Definición 22 *Sea V un espacio vectorial real provisto de un producto escalar $\langle \cdot,\cdot\rangle$. La longitud o norma de $x \in V$ se nota $\|x\|$ y se define como* $\|x\| = (\langle x, x\rangle)^{\frac{1}{2}}$.

Esta norma $\|\cdot\|$ se dice asociada al producto escalar $\langle \cdot,\cdot\rangle$ y se le denomina norma euclídea.

Ejemplos

1. En el caso en que $V = \mathbb{R}^n$, la norma del vector $\overrightarrow{x}^T = (x_1, ..., x_n) \in \mathbb{R}^n$ asociada al producto escalar definido como $\overrightarrow{x}^T \overrightarrow{y} = \sum\limits_{i=1}^{n} x_i y_i$, se escribe $\|\overrightarrow{x}\|_2 = \left(\overrightarrow{x}^T \overrightarrow{x}\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$. Esta norma coincide con la norma hölderiana en $\mathbb{R}^n$ con $p = 2$.

2. Sea $V = C([-1,1])$. La norma de $f \in C([-1,1])$ asociada al producto escalar $\langle \cdot,\cdot\rangle$ antes definido, está definida como $\|f\|_2 = (\langle f, f\rangle)^{\frac{1}{2}} = \left(\int_{-1}^{1} f^2(x)\, dx\right)^{\frac{1}{2}}$. Esta norma coincide con la norma hölderiana en $C([a,b])$ con $p = 2$.

3. Se denota $\mathbb{K}_n([a,b])$ al espacio vectorial de los polinomios reales de grado $\leq n$ restringidos al intervalo $[a,b] \subset \mathbb{R}$ El espacio $\mathbb{K}_n([a,b])$ es un subespacio de $C([a,b])$ de dimensión $n+1$. Definido un producto escalar $\langle \cdot,\cdot\rangle$ en $C([a,b])$, este es un producto escalar en $\mathbb{K}_n([a,b])$ y la norma asociada se escribe
$$\|p\|_2 = (\langle p, p\rangle)^{\frac{1}{2}} = \left(\int_a^b p^2(t)\, dt\right)^{\frac{1}{2}} \qquad \forall p \in \mathbb{K}_n([a,b]).$$
En tal caso diremos que $\mathbb{K}_n([a,b])$ es un espacio con producto intermo inducido por el de $C([a,b])$ y que la norma $\|\cdot\|_2$ en $\mathbb{K}_n([a,b])$ es la inducida por la norma $\|\cdot\|_2$ en $C([a,b])$.

4. Sea $V = C^1([-1,1])$. Un producto escalar $\langle \cdot, \cdot \rangle$ en $C^1([-1,1])$ se define como

$$\langle f, g \rangle = \int_{-1}^{1} \left[f(x) g(x) + \frac{df}{dx}(x) \frac{dg}{dx}(x) \right] dx$$

y la norma de $f \in C^1([-1,1])$ asociada a este producto escalar se define como

$$\|f\|_{1,2} = \left(\int_{-1}^{1} \left(|f(x)|^2 + \left| \frac{df}{dx}(x) \right|^2 \right) dx \right)^{\frac{1}{2}}.$$

5. Sean $\Omega = [-1,1] \times [-1,1] \subset \mathbb{R}^2$. Se denota con $C(\Omega)$ al espacio vectorial de funciones continuas en Ω. Un producto escalar $\langle \cdot, \cdot \rangle$ en $C(\Omega)$ se define como

$$\langle f, g \rangle = \int_{-1}^{1} \int_{-1}^{1} f(x,y) g(x,y) \, dxdy \qquad \forall f, g \in C(\Omega).$$

Se propone como ejercicio probar que efectivamente la función $\langle \cdot, \cdot \rangle$ definida en $C(\Omega) \times C(\Omega)$ es un producto escalar. La norma de $f \in C(\Omega)$ asociada a este producto escalar se define como

$$\|f\| = \left(\int_{\Omega} |f(x,y)|^2 \, dxdy \right)^{\frac{1}{2}}.$$

6. Considerar el espacio de funciones $C^1([-1,1])$ que poseen derivada continua en $[-1,1]$. Se define la función $\langle \cdot, \cdot \rangle_1$ de $C^1([-1,1]) \times C^1([-1,1])$ en $\mathbb{R}$ como sigue:

$$\langle u, v \rangle_1 = \int_{-1}^{1} \left[u(x) v(x) + u'(x) v'(x) \right] dx \qquad \forall u, \, v \in C^1([-1,1]).$$

entonces $\langle \cdot, \cdot \rangle$ es un producto escalar en $C^1([-1,1])$.

7. Sean $\Omega = [-1,1] \times [-1,1] \subset \mathbb{R}^2$. Se denota con $C(\Omega)$ al espacio vectorial de funciones continuas en Ω. Un producto escalar $\langle \cdot, \cdot \rangle$ en $C(\Omega)$ se define como

$$\langle f, g \rangle = \int_{-1}^{1} \int_{-1}^{1} f(x,y) g(x,y) \, dxdy \qquad \forall f, g \in C(\Omega).$$

La norma de $f \in C(\Omega)$ asociada a este producto escalar se define como $\|f\| = \left(\int_{\Omega} |f(x,y)|^2 \, dxdy \right)^{\frac{1}{2}}$.

8. Sea $\Omega = [-1,1] \times [-1,1] \subset \mathbb{R}^2$. Se designa con $C^1(\Omega)$ al espacio de funciones reales que poseen derivadas parciales primeras continuas en Ω. En $C^1(\Omega)$ se define la función real $\langle \cdot, \cdot \rangle_{1,2}$ como a continuación se indica:

$$\langle f, g \rangle_{1,2} = \int_{\Omega} \left(f(x,y) g(x,y) + \frac{\partial f}{\partial x}(x,y) \frac{\partial g}{\partial x}(x,y) + \frac{\partial f}{\partial y}(x,y) \frac{\partial g}{\partial y}(x,y) \right) dxdy \qquad \forall f, g \in C^1(\Omega).$$

entonces $\langle \cdot, \cdot \rangle_{1,2}$ es un producto escalar en $C^1(\Omega)$. Este producto escalar se escribe en forma abreviada como $\langle f, g \rangle_{1,2} = \int_{\Omega} (fg + \nabla f \cdot \nabla g) \quad \forall f, g \in C^1(\Omega)$ y la norma asociada a este producto escalar se escribe como $\|f\|_{1,2} = \left(\int_{\Omega} \left(f^2 + |\nabla f|^2 \right) \right)^{\frac{1}{2}} \quad \forall f \in C^1(\Omega)$.

Teorema 8 *Sea V un espacio vectorial real provisto de un producto escalar $\langle \cdot, \cdot \rangle$. La longitud o norma $\|\cdot\|$ en V satisface las siguientes propiedades:*

i. $\|x\| \geq 0 \quad \forall x \in V.$

ii. $\|x\| = 0 \Leftrightarrow x = 0.$

iii. $\|\lambda x\| = |\lambda| \, \|x\| \quad \forall \lambda \in \mathbb{R}, \ x \in V.$

iv. $|\langle x, y \rangle| \leq \|x\| \, \|y\| \quad \forall x, \, y \in V$ *(desigualdad de Cauchy-Schwarz).*

v. $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, \, y \in V$ *(desigualdad triangular).*

vi. $|\|x\| - \|y\|| \leq \|x - y\| \quad \forall x, \, y \in V.$

vii. $\|x + y\|^2 + \|x - y\|^2 = 2 \|x\|^2 + 2 \|y\|^2 \quad \forall x, \, y \in V$ *(ley del paralelogramo).*

viii. $\langle x, y \rangle = \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right) \quad \forall x, \, y \in V$ *(identidad de polarización).*

Demostración.

i. Puesto que la función $\langle \cdot, \cdot \rangle$ de $V \times V$ en $\mathbb{R}$ es un producto escalar, esta tiene la propiedad siguiente: para $x \in V$, $\langle x, x \rangle = 0 \Leftrightarrow x = 0, \quad \langle x, x \rangle > 0 \Leftrightarrow x \neq 0$, y de la definición de norma, se tiene $\|x\| \geq 0 \quad \forall x \in V$.

ii. Como $\langle x, x \rangle = 0 \Leftrightarrow x = 0$, resulta $\|x\| = (\langle x, x \rangle)^{\frac{1}{2}} = 0 \Leftrightarrow x = 0$.

iii. Sean $\lambda \in \mathbb{R}$, $x \in V$. Entonces

$$\|\lambda x\| = (\langle \lambda x, \lambda x \rangle)^{\frac{1}{2}} = \left(\lambda^2 \langle x, x \rangle\right)^{\frac{1}{2}} = |\lambda| \, (\langle x, x \rangle)^{\frac{1}{2}} = |\lambda| \, \|x\| \, .$$

Así, $\|\lambda x\| = |\lambda| \, \|x\| \quad \forall \lambda \in \mathbb{R}, \quad \forall x \in V.$

iv. Sean $x, \, y \in V$. Por la parte i) de este teorema se tiene $\|x + \lambda y\|^2 \geq 0 \quad \forall \lambda \in \mathbb{R}$. Luego, por la definición de norma y las propiedades del producto escalar, se tiene

$$\begin{aligned} 0 \quad &\leq \quad \|x + \lambda y\|^2 = \langle x + \lambda y, x + \lambda y \rangle = \langle x, x \rangle + \langle x, \lambda y \rangle + \langle \lambda y, x \rangle + \langle \lambda y, \lambda y \rangle \\ &= \quad \langle x, x \rangle + \lambda \langle x, y \rangle + \lambda \langle y, x \rangle + \lambda^2 \langle y, y \rangle = \|x\|^2 + 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2 \, . \end{aligned}$$

Sea P el polinomio de grado ≤ 2 definido por $\quad P(\lambda) = \|x\|^2 + 2 \langle x, y \rangle \lambda + \|y\|^2 \lambda^2, \quad \lambda \in \mathbb{R}$. Por la parte precedente se verifica $P(\lambda) \geq 0 \quad \forall \lambda \in \mathbb{R}$, con lo que el discriminante $d = (2 \langle x, y \rangle)^2 - 4 \|x\|^2 \|y\|^2 \leq 0$ de donde $\quad 4 (\langle x, y \rangle)^2 \leq 4 \|x\|^2 \|y\|^2$. Tomando la raíz cuadrada y considerando que la norma $\|\cdot\|$ es no negativa, se deduce la desigualdad de Cauchy-Schwarz: $|\langle x, y \rangle| \leq \|x\| \, \|y\| \, .$

v. Sean $x, y \in V$. Entonces

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2 \, .$$

Se tiene $\langle x, y \rangle \leq |\langle x, y \rangle|$ y por la desigualdad de Cauchy-Schwarz resulta

$$\langle x, y \rangle \leq |\langle x, y \rangle| \leq \|x\| \, \|y\| \quad \forall x, \, y \in V.$$

Luego,

$$\|x + y\|^2 = \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2 \|x\| \, \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \, .$$

Tomando la raíz cuadrada y considerando que la norma es no negativa, se obtiene la desigualdad triangular

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, \, y \in V.$$

vi. La desigualdad $|\|x\| - \|y\|| \leq \|x - y\| \quad \forall x, y \in V$ ya fue probada en la sección de los espacios normados.

vii. De la definición de norma y de las propiedades del producto escalar, se tiene

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ &= \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle + \langle x,x\rangle - 2\langle x,y\rangle + \langle y,y\rangle \\ &= 2\left(\|x\|^2 + \|y\|^2\right).\end{aligned}$$

viii. Se propone como ejercicio.

■

Definición 23 *Sean V un espacio prehilbertiano, $x, y \in V$. La distancia de x a y se denota y se define como* $d(x,y) = \|x - y\|$.

Si $\langle \cdot, \cdot \rangle$ es un producto escalar definido en V y $\|\cdot\|$ la norma asociada, se tiene

$$d(x,y) = \|x - y\| = (\langle x-y, x-y\rangle)^{\frac{1}{2}} \quad \forall x, y \in V.$$

Teorema 9 *Sea V un espacio prehilbertiano.*
La función d de $V \times V$ en $\mathbb{R}$ definida como $d(x,y) = \|x - y\| \quad \forall x, y \in V$ es una métrica en V, es decir que satisface las propiedades siguientes:

i. $d(x,y) \geq 0 \quad \forall x, y \in V.$

ii. $d(x,y) = 0 \Leftrightarrow x = y, \; x, y \in V.$

iii. $d(x,y) = d(y,x) \quad \forall x, y \in V.$

iv. $d(x,z) \leq d(x,y) + d(y,z) \quad \forall x, y, z \in V$ *(desigualdad triangular).*

Demostración. La prueba es inmediata y se propone como ejercicio. ■

Si V es un espacio vectorial real provisto de un producto interior o producto escalar $\langle \cdot, \cdot \rangle$, la norma asociada a este producto escalar está dado por $\|x\| = (\langle x, x\rangle)^{\frac{1}{2}} \quad x \in V$, y la métrica asociada a la norma $\|\cdot\|$ está definida como $d(x,y) = \|x - y\| \quad x, y \in V$, con lo cual V es un espacio métrico que escribimos (V, d). En la siguiente sección trataremos más en detalle las métricas sobre un conjunto no vacío E.

12.4.1. Ortogonalidad o perpendicularidad.

Definición 24 *Sea V un espacio vectorial provisto del porducto escalar* $\langle \cdot, \cdot \rangle$.

i) *Sean $x, y \in V$. Se dice que x es ortogonal o perpendicular a y, que se nota $x \perp y$, si y solo si* $\langle x, y\rangle = 0$.

ii) *Sean $x \in V$, $M \subset V$ con $M \neq \phi$. Se dice que x es ortogonal a M, que se escribe $x \perp M$, si y solo si* $\langle x, y\rangle = 0 \quad \forall y \in M$.

iii) *Sean M, N dos subconjuntos no vacíos de V. Se dice que M es ortogonal a N, que se nota $M \perp N$, si y solo si* $\langle x, y\rangle = 0 \quad \forall x \in M, \forall y \in N$.

iv) *Sea $M \subset V$ con $M \neq 0$. Se dice que M es ortogonal si y solo si* $\langle x, y\rangle = 0 \quad \forall x, y \in M, \; x \neq y$.

v) *Se dice que M es ortonomal si y solo si M es ortogonal, y* $\forall x \in M, \; \|x\| = 1$.

Ejemplos

1. Sea $V = \mathbb{R}^n$. El conjunto $M = \{\overrightarrow{e}_1, \ldots, \overrightarrow{e}_n\}$, donde $\overrightarrow{e}_1^T = (1, 0, \ldots, 0), \ldots, \overrightarrow{e}_n^T = (0, \ldots, 0, 1)$ son los vectores de la base canónica de $\mathbb{R}^n$, es un conjuto ortogonal, pués
$$\overrightarrow{e}_j^T \overrightarrow{e}_k = 0 \text{ si } j \neq k, \quad \text{y,} \quad \|\overrightarrow{e}_j\| = 1 \quad j = 1, \ldots, n.$$

2. Sean $L > 0$. Se denota con $C([-L, L])$ al espacio vectorial de las funciones continuas en $[-L, L]$. Proveemos a $C([-L, L])$ del producto escalar $\langle \cdot, \cdot \rangle$ definido por
$$\langle u, v \rangle = \int_{-L}^{L} u(x) v(x) \, dx \quad \forall u, v \in C([-L, L]).$$
Los siguientes conjuntos de funciones son muy importantes en el desarrollo en series de Fourier de funciones reales periódicas de período $2L$ y continuas a trozos en el intervalo $[-L, L]$. Sean M, N los subconjuntos de $C([-L, L])$ definidos como $M = \{\varphi_k \mid k \in \mathbb{N}\}$, $N = \{\psi_k \mid k \in \mathbb{Z}^+\}$, donde
$$\begin{aligned} \varphi_0(x) &= 1 \quad x \in [-L, L], \quad \varphi_k(x) = \cos\left(\frac{k\pi x}{L}\right) \quad x \in [-L, L], \quad k = 1, 2, \ldots, \\ \psi_k(x) &= \operatorname{sen}\left(\frac{k\pi x}{L}\right) \quad x \in [-L, L], \ k = 1, 2, \ldots \end{aligned}$$
Se tiene

i) M es un conjunto ortogonal.

ii) N es un conjunto ortogonal.

iii) $M \perp N$.

i) Probemos que M es ortogonal, esto es,
$$\begin{aligned} \langle \varphi_0, \varphi_k \rangle &= 0 \quad \forall k = 1, 2, \ldots \\ \langle \varphi_j, \varphi_k \rangle &= 0 \quad \forall j, k \in \mathbb{Z} \text{ con } j \neq k. \end{aligned}$$
En efecto, de la definición de φ_0 y φ_k, se tiene
$$\langle \varphi_0, \varphi_k \rangle = \int_{-L}^{L} \varphi_0(x) \varphi_k(x) \, dx = \int_{-L}^{L} \cos\left(\frac{k\pi x}{L}\right) dx = \frac{L}{k\pi} \operatorname{sen}\left(\frac{k\pi x}{L}\right)\Big|_{-L}^{L} = 0 \quad k = 1, 2, \ldots$$
Por otro lado,
$$\langle \varphi_j, \varphi_k \rangle = \int_{-L}^{L} \varphi_j(x) \varphi_k(x) \, dx = \int_{-l}^{L} \cos\left(\frac{j\pi x}{L}\right) \operatorname{sen}\left(\frac{k\pi x}{L}\right) dx \quad \forall x \in [-L, L].$$
Como $\cos\left(-\frac{j\pi x}{L}\right) = \cos\left(\frac{j\pi x}{L}\right)$, $\operatorname{sen}\left(-\frac{k\pi x}{L}\right) = -\operatorname{sen}\left(\frac{k\pi x}{L}\right)$ entonces la función $\varphi_j \varphi_k$ es impar para $j \neq k$. Luego $\langle \varphi_j, \varphi_k \rangle = 0$.

ii) Pasemos a probar que N es ortogonal, es decir,
$$\langle \psi_j, \psi_k \rangle = \int_{-l}^{L} \operatorname{sen}\left(\frac{j\pi x}{L}\right) \operatorname{sen}\left(\frac{k\pi x}{L}\right) dx = 0 \quad \forall j, k \in \mathbb{Z} \text{ con } j \neq k.$$
Sean $a, b \in \mathbb{R}$, de las identidades trigonométricas
$$\cos(a + b) = \cos a \cos b - \operatorname{sen} a \operatorname{sen} b, \qquad \cos(a - b) = \cos a \cos b + \operatorname{sen} a \operatorname{sen} b,$$
se obtiene
$$\operatorname{sen} a \operatorname{sen} b = \frac{1}{2} [\cos(a - b) - \cos(a + b)],$$
y poniendo $a = \frac{j\pi x}{L}$, $b = \frac{k\pi x}{L}$, resulta
$$\begin{aligned} \langle \psi_j, \psi_k \rangle &= \int_{-L}^{L} \frac{1}{2} \left[\cos\left(\frac{\pi(j-k)x}{L}\right) - \cos\left(\frac{\pi(j+k)x}{L}\right)\right] dx \\ &= \frac{1}{2} \left(\frac{L}{\pi(j-k)} \operatorname{sen}\left(\frac{\pi(j-k)x}{L}\right) - \frac{L}{\pi(j+k)} \operatorname{sen}\left(\frac{\pi(j+k)x}{L}\right)\right)\Big|_{-L}^{L} \\ &= 0 \quad \text{si } j \neq k. \end{aligned}$$

iii) Para mostrar que $M \perp N$, se debe probar que $\langle \varphi_0, \psi_k \rangle = 0 \quad k = 1, 2, \ldots$ y $\langle \varphi_j, \psi_k \rangle = 0 \quad j, k \in \mathbb{Z}^+$. La verificación se propone como ejercicio.

Definición 25 *Sea V un espacio con producto interior, $x, y \in V$. El ángulo $\theta \in [0, \pi]$ que forman los vectores x e y se define como* $\cos\theta = \frac{\langle x,y \rangle}{\|x\|\|y\|} \quad x \neq 0,\ y \neq 0.$

Teorema 10 (de Pitágoras) *Sea V un espacio con producto interior, $x, y \in V$. Si $x \perp y$ se tiene*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 .$$

De manera mas general, sea $\{x_1, \cdots, x_n\}$ un conjunto ortogonal de V, entonces

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \| x_i \|^2 .$$

Demostración. De la definición de la norma $\|.\|$, se tiene

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2 \langle x, y \rangle + \langle y, y \rangle .$$

Por hipótesis $x \perp y$, luego $\langle x, y \rangle = 0$, y como $\langle x, x \rangle = \|x\|^2$, $\langle y, y \rangle = \|y\|^2$, se concluye

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 .$$

La prueba de la generalización del teorema de Pitágoras a una familia ortogonal finita se realiza por inducción y se propone como ejercicio. ■

En el caso de espacios vectoriales reales, se tiene que: si $x, y \in V$ tales que $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ entonces $x \perp y$. Efectivamente, si $x, y \in V$ entonces $\|x + y\|^2 = \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2$. Por hipótesis, $\|x + y\|^2 = \|x\|^2 + \|y\|^2$, y de la igualdad $\|x\|^2 + \|y\|^2 = \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2$ de donde $\langle x, y \rangle = 0$, o sea $x \perp y$.

12.5. Lecturas complementarias y bibliografía

1. Owe Axelsson, Iterative Solution Methods, Editorial Cambridge University Press, Cambridge, 1996.
2. E. K. Blum, Numerical Analysis and Computation. Theory and Practice, Editorial Addison-Wesley Publishing Company, Reading, Massachusetts, 1972.
3. Richard L. Burden, J. Douglas Faires, Análisis Numérico, Séptima Edición, International Thomson Editores, S. A., México,2002.
4. P. G. Ciarlet, Introduction á l'Analyse Numérique Matricielle et á l'Optimisation, Editorial Masson, París, 1990.
5. James W. Demmel, Applied Numerical Linear Algebra, Editorial Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997.
6. V. N. Faddeva, Métodos de Cálculo de Algebra Lineal, Editorial Paraninfo, Madrid, 1967.
7. Francis G. Florey, Fundamentos de Algebra Lineal y Aplicaciones, Editorial Prentice-Hall Hispanoamericana, S. A., México, 1980.
8. Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence, Algebra Lineal, Editorial Publicaciones Cultural, S. A., México, 1982.
9. Noel Gastinel, Análisis Numérico Lineal, Editorial Reverté, S. A., Barcelona, 1975.

10. Gene H. Golub, Charles F. Van Loan, Matrix Computations, Second Edition, The Johns Hopkins University Press, Baltimore, 1989.

11. Kenneth Hoffman, Ray Kunze, Algebra Lineal, Editorial Prentice-Hall Hispanoamericana, S. A., México, 1987.

12. Franz E. Hohn, Algebra de Matrices, Editorial Trillas, México, 1979.

13. Roger A. Horn, Charles R. Johnson, Matrix Analysis, Editorial Cambridge University Press, Cambridge, 1999.

14. A. N. Kolmogórov, S. V. Fomín, Elementos de la Teoría de Funciones y del Análisis Funcional. Editorial Mir, Moscú, 1972.

15. Peter Linz, Theoretical Numerical Analysis, Editorial Dover Publications, Inc., New York, 2001.

16. Anthony N. Michel, Charles J. Herget, Applied Algebra and Functional Analysis, Editorial Dover Publications, Inc., New York, 1981.

17. Ben Noble, James W. Daniel, Algebra Lineal Aplicada, Editorial Prentice-Hall Hispanoamericana, S. A., México, 1989.

18. Fazlollah Reza, Los Espacios Lineales en la Ingeniería, Editorial Reverté, S. A., Barcelona, 1977.

19. Gilbert Strang, Algebra Lineal y sus Aplicaciones, Editorial Fondo Educativo Interamericano, México, 1982.

20. Arthur Wouk, A Course of Applied Functional Analysis, Editorial John Wiley&Sons, New York, 1979.