1 Espacios normados, métricos y Topología

Durante todo este curso "espacio vectorial" querrá decir "espacio vectorial real", es decir que el cuerpo de escalares será siempre $\mathbb{R}$.
Informalmente la palabra "espacio" quiere decir "conjunto con alguna estructura". El vacío $\varnothing$ sí que es un conjunto, pero no se le pone ninguna estructura. Por lo tanto, a todos los tipos de espacio que se definan se les exigirá ser no vacíos.

1.1 Espacios normados

La idea de un espacio normado es que se trata de un espacio vectorial en el que sabemos medir longitudes de vectores. Empezamos con el ejemplo básico.

1.1.1 Productos escalares

Definición 1. *Un* **producto escalar** *en un espacio vectorial* $\mathbb{V}$ *es una función de dos variables*

$$\begin{array}{rccc} \langle\cdot,\cdot\rangle & : & \mathbb{V}\times\mathbb{V} & \longrightarrow & \mathbb{R} \\ & & (v,w) & \longmapsto & \langle v,w\rangle \end{array}$$

cumpliendo las siguientes condiciones:

1. *Bilineal: lineal en la variable v cuando se congela el valor de w y lineal en w cuando se congela el valor de v.*
2. *Simétrica: $\langle v,w\rangle = \langle w,v\rangle$.*
3. *Definida positiva: $\langle v,v\rangle > 0$ para todo $v \neq \mathbf{0}$.*

En el caso particular $\mathbb{V} = \mathbb{R}^n$, todos los productos escalares en $\mathbb{R}^n$ vienen dados de la manera siguiente:

$$\langle x,y\rangle \;=\; x^t\,A\,y \quad , \quad \text{donde} \left\{ \begin{array}{ll} x,y & \text{son vectores columna} \\ A & \text{es una matriz simétrica definida positiva} \end{array} \right.$$

Recordemos que una matriz simétrica A es **definida positiva** si cumple esto:

$$\text{para todo } \; x \in \mathbb{R}^n \setminus \{\mathbf{0}\} \quad x^t A x \;>\; 0 \,.$$

El **producto escalar estándar en $\mathbf{R^n}$** corresponde a la matriz identidad $A = I_n$, o sea

$$x\cdot y \;\stackrel{\text{def}}{=}\; x^t y \;=\; x_1y_1 + \cdots + x_ny_n \,.$$

1.1.2 Normas euclídeas

Definición 2. *La* **longitud euclídea** *o* **norma euclídea** *asociada al producto escalar* $\langle\cdot,\cdot\rangle$ *es la siguiente función:*

$$\|\cdot\| : \mathbb{V}\to\mathbb{R} \quad , \quad \|v\| \;=\; \sqrt{\langle v,v\rangle} \;\geq\; 0 \,.$$

La **norma euclídea estándar** en $\mathbb{R}^n$ es: $\|(x_1,\ldots,x_n)\| = \sqrt{x_1^2+\cdots+x_n^2}$.
Para todo producto escalar las siguientes propiedades son obvias:

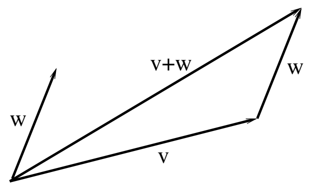
1. $\|v\| \geq 0$,
2. $\|v\| = 0 \Longleftrightarrow v = \mathbf{0}$,
3. $\|\lambda v\| = |\lambda|\,\|v\|$,

y se les añaden otras dos, no tan obvias:

Desigualdad de Cauchy-Schwarz: $|\langle v, w\rangle| \leq \|v\| \, \|w\|$.

Desigualdad triangular: $\|v+w\| \leq \|v\| + \|w\|$.

Para efectuar geométricamente la suma de vectores $v+w$, podemos trasladar paralelamente w hasta que su origen coincida con la punta de v, entonces $v+w$ une el origen de v con la punta (trasladada) de w. Colocados así, los tres vectores v, w y $v+w$ son los *lados de un triángulo.* La desigualdad triangular se llama así porque afirma que en cualquier triángulo la longitud de un lado no supera la suma de las longitudes de los otros dos lados.



Demostración algebraica de la desigualdad de Cauchy-Schwarz

La **matriz de Gram** de una sucesión de vectores $v_1, \ldots, v_k$ es la "tabla de multiplicar"

$$G \;=\; \big[\, \langle v_i, v_j\rangle \,\big]_{1\leq i,j\leq k} \,,$$

obviamente simétrica $k \times k$. Tiene la siguiente propiedad:

$$v_1, \ldots, v_k \;\; \text{linealmente independientes} \;\Longrightarrow\; G \text{ es definida positiva} \;\Longrightarrow\; \det G > 0.$$

Definida positiva. Si $x \neq \mathbf{0}$, entonces $v = x_1 v_1 + \cdots x_k v_k \neq \mathbf{0}$, luego $0 < v^t v = x^t G x$.
Determinante positivo. Como G es real simétrica, hay una base ortonormal $\{u_1, \ldots, u_k\}$ de $\mathbb{R}^k$ formada por autovectores de G. Si $\lambda_1, \ldots, \lambda_k$ son los correspondientes autovalores, entonces

$$\mathbb{R}^k \ni x \;=\; y_1 u_1 + \cdots + y_k u_k \;\;\Longrightarrow\;\; x^t G x \;=\; \lambda_1\, y_1^2 + \cdots + \lambda_k\, y_k^2 \,,$$

y está claro que $x^t G x > 0$ para todo $x \neq \mathbf{0}$ si y sólo si $\lambda_1, \ldots, \lambda_k > 0$, de donde

$$\det G \;=\; \lambda_1 \cdots \lambda_k \;>\; 0 \,.$$

Caso particular de dos vectores v, w:

$$G \;=\; \begin{bmatrix} \langle v,v\rangle & \langle v,w\rangle \\ \langle w,v\rangle & \langle w,w\rangle \end{bmatrix} \quad , \quad \det G \;=\; \langle v,v\rangle\langle w,w\rangle - \langle v,w\rangle^2 \;=\; \|v\|^2\|w\|^2 - |\langle v,w\rangle|^2 \,.$$

Si v, w son linealmente independientes, es $\|v\|^2\|w\|^2 - |\langle v,w\rangle|^2 > 0$ y $|\langle v,w\rangle| < \|v\| \, \|w\|$.
Si v, w son linealmente dependientes, es trivial ver que $|\langle v,w\rangle| = \|v\| \, \|w\|$.

Demostración geométrica de la desigualdad de Cauchy-Schwarz

Supuestos v, w linealmente independientes, les aplicamos el *proceso de Gram-Schmidt* y llegamos a una pareja ortonormal v_1, v_2 y números a_1, b, a_2 tales que

$$v \;=\; a_1 v_1 \quad , \quad w \;=\; b v_1 + a_2 v_2 \quad , \quad a_1, a_2 \;>\; 0 \quad , \quad b \;\; \text{cualquiera} \,.$$

Entonces $|\langle v,w\rangle| = a_1\,|b|$, y también:

$$\|v\| \, \|w\| \;=\; a_1 \sqrt{a_2^2 + b^2} \;>\; a_1 \sqrt{b^2} \;=\; a_1\,|b| \,.$$

En el plano vectorial generado por v, w hemos tomado unos *ejes ortogonales adaptados* al par v, w: el primer eje (el del vector v_1) contiene al vector v.

Demostración algebraica de la desigualdad triangular

Como corolario de la desigualdad de Cauchy-Schwarz :

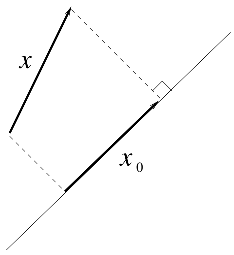
$$\langle v+w, v+w\rangle \;=\; \langle v,v\rangle + \langle w,w\rangle + 2\,\langle v,w\rangle \;\leq\; \|v\|^2 + \|w\|^2 + 2\,\|v\|\,\|w\| \;=\; \big(\|v\| + \|w\|\big)^2\ ,$$

es decir $\|v+w\|^2 \leq (\|v\|+\|w\|)^2$. Tomando raices cuadradas, sale $\|v+w\| \leq \|v\| + \|w\|$.

Demostración geométrica de la desigualdad triangular

Es obvia si $v+w=\mathbf{0}$. Supongamos $v+w\neq\mathbf{0}$.

Observación. Si x_0 es la *proyección ortogonal* de un vector x sobre cualquier subespacio vectorial, entonces $\|x_0\| \leq \|x\|$, es decir que *las proyecciones ortogonales acortan longitudes* (o las dejan igual).



Sean v_0, w_0 las proyecciones ortogonales respectivas de v, w sobre la recta vectorial $L = \langle v+w\rangle$, con lo cual $v+w = v_0 + w_0$.
Mirando sólo L, vemos que la longitud euclídea de v_0+w_0 es uno de los tres números siguientes:

$$\|v_0\| + \|w_0\| \quad , \quad \|v_0\| - \|w_0\| \quad , \quad \|w_0\| - \|v_0\|\ ,$$

en todo caso, un valor menor o igual que $\|v\| + \|w\|$.

1.1.3 Funciones homogéneas

Definiciones 3. *Sea $\mathbb{V}$ un espacio vectorial de dimensión no nula y sea $f : \mathbb{V}\setminus\{\mathbf{0}\} \to \mathbb{R}$ una función escalar (que puede estar definida o no en el vector nulo).*
Para un $k\in\mathbb{Z}$, decimos que f es **homogénea de grado k** *si*

$$v\in\mathbb{V}\setminus\{\mathbf{0}\}\,,\ \lambda\neq 0 \implies f(\lambda v) \;=\; \lambda^k\, f(v)\ .$$

Para un $a\in\mathbb{R}$ (positivo, nulo o negativo) decimos que f es **positivamente homogénea de grado a** *si*

$$v\in\mathbb{V}\setminus\{\mathbf{0}\}\,,\ \lambda > 0 \implies f(\lambda v) \;=\; \lambda^a\, f(v)\ .$$

Ejemplos homogéneos con $\mathbb{V}=\mathbb{R}^2$:

$f(x,y) \equiv 5x^3y^2 - 9x^4y + y^5 \ \ ,\ \ k=5.$

$f(x,y) \equiv \dfrac{x}{x^4+y^4} \ \ ,\ \ k=-3.$

$f(x,y) = \sqrt[3]{x^3+y^3} \ \ ,\ \ k=1.$

Ejemplos positivamente homogéneos (pero no homogéneos):

$f(x,y) = |x| + 3|y| \ \ ,\ a=1.$

$f(x,y) = \sqrt[6]{x^2+9y^2} \ \ ,\ a=1/3.$

$f(x,y) = |x| + \dfrac{y^3}{x^2+y^2} \ \ ,\ \ a=1.$

Cuando el grado es positivo (entero o fraccionario) extendemos estas funciones al origen poniendo $f(\mathbf{0}) = 0$. Entonces el **grafo** de f es el siguiente subconjunto de $\mathbb{V}\times\mathbb{R}$ (de $\mathbb{R}^{n+1}$ cuando $\mathbb{V} = \mathbb{R}^n$):

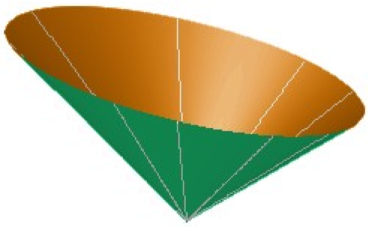
$$\text{grafo}\,(f) \;=\; \big\{ \big(x, f(x)\big) \;:\; x \in \mathbb{V} \big\} \,.$$

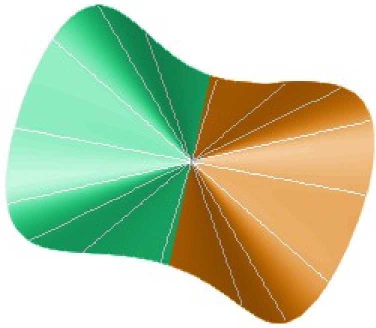
Que f sea positivamente homogénea de grado 1 significa que su grafo es un **cono:**

una unión de *semirrectas* en $\mathbb{V} \times \mathbb{R}$ que salen del origen $(\mathbf{0}, 0)$.

Si g es homogénea de grado 1, su grafo es un **cono simétrico respecto del origen:**
una unión de *rectas* pasando por $(\mathbf{0}, 0)$.

La siguiente figura muestra a la izquierda el grafo de $f(x,y) = 0'3\,x + \sqrt{x^2+y^2}$ y a la derecha el grafo de $g(x,y) = (1/3)\sqrt[3]{x^3 - 3xy^2}$





1.1.4 Normas en general

Definiciones 4. *Sea $\mathbb{V}$ un espacio vectorial de dimensión no nula. Una* **norma** *en $\mathbb{V}$ es cuaquier función*

$$\begin{array}{rcccc} \|\cdot\| & : & \mathbb{V} & \longrightarrow & \mathbb{R} \\ & & v & \longmapsto & \|v\| \end{array}$$

cumpliendo las siguientes condiciones:

[1]. $\|\mathbf{0}\| = 0$ *y* $\|v\| > 0$ *cuando* $v \neq \mathbf{0}$.

[h]. *Positivamente homogénea de grado 1:* $\lambda > 0 \Longrightarrow \|\lambda v\| = \lambda\,\|v\|$.

[2]. *Función par:* $\|-v\| = \|v\|$.

[3]. *Desigualdad triangular:* $\|v + w\| \leq \|v\| + \|w\|$.

Un **espacio normado** *es un par* $(\mathbb{V}, \|\cdot\|)$ *con* $\mathbb{V}$ *espacio vectorial y* $\|\cdot\|$ *norma en* $\mathbb{V}$.

Lo que esperamos de una norma es que es sirva para *medir longitudes* de vectores.

La propiedad [1] responde a dos ideas:

– No queremos longitudes negativas.

– Un buen criterio para saber si un vector es nulo es mirar si su longitud es nula.

Las propiedades [h] y [2] suelen juntarse en la fórmula $\|\lambda v\| = |\lambda|\,\|v\|$, con lo que los axiomas para normas suelen darse así:

1. $\|\mathbf{0}\| = 0$ y $\|v\| > 0$ cuando $v \neq \mathbf{0}$.
2. $\|\lambda v\| = |\lambda|\,\|v\|$.
3. $\|v + w\| \leq \|v\| + \|w\|$.

1.1.5 Bolas y conjuntos acotados

Sea $(\mathbb{V}, \|\cdot\|)$ un espacio normado.

Dado $r > 0$ (desigualdad estricta), la **bola abierta de centro el origen y radio r** es el conjunto:

$$B(\mathbf{0}, r) \;=\; \{\, v \in \mathbb{V} \,:\, \|v\| < r, \} \;.$$

Dado $r \geq 0$, la **bola cerrada de centro el origen y radio r** es el conjunto:

$$\overline{B}(\mathbf{0}, r) \;=\; \{\, v \in \mathbb{V} \,:\, \|v\| \leq r \,\} \;.$$

En particular, la bola cerrada de radio nulo $\overline{B}(\mathbf{0}, 0)$ es el conjunto $\{\mathbf{0}\}$.
La **bola unidad abierta** es $B(\mathbf{0}, 1)$.
La **bola unidad cerrada** es $\overline{B}(\mathbf{0}, 1)$. Su "cáscara" es el conjunto $\{v \in \mathbb{V} \,:\, \|v\| = 1\}$ de los **vectores unitarios** para la norma $\|\cdot\|$.

Definición 5. *En un espacio normado $\mathbb{V}$, un subconjunto $E \subset \mathbb{V}$ es* **acotado** *si está contenido en alguna bola, es decir si $E \subseteq \overline{B}(\mathbf{0}, r)$ para algún $r \geq 0$.*

1.1.6 Descomposición polar

Dado un espacio normado $(\mathbb{V}, \|\cdot\|)$, cada vector no nulo $v \in \mathbb{V} \setminus \{\mathbf{0}\}$ tiene una factorización

$$v \;=\; \lambda\, \omega \quad \text{con} \quad \begin{cases} \lambda & \text{escalar positivo} \\ \omega & \text{vector unitario} \end{cases}$$

Esta factorización resulta ser única, veámoslo. Se tiene que cumplir

$$\|v\| \;=\; \|\lambda \omega\| \;=\; \lambda\, \|\omega\| \;=\; \lambda \cdot 1 \;,$$

luego $\lambda = \|v\|$ es el único valor posible para λ. La única solución posible es, pues

$$\lambda \;=\; \|v\| \quad \text{y} \quad \omega \;=\; v/\|v\| \;,$$

que de hecho es solución, porque $\big\|(v/\|v\|)\big\| = (1/\|v\|)\|v\| = 1$.

Observaciones

(1) Esta factoriación es válida para cualquier función $\|\cdot\| : \mathbb{V} \to \mathbb{R}$ que cumpla los axiomas 1. y 2. de norma, no importa si cumple la desigualdad triangular o no.

(2) Una función que cumpla los axiomas 1. y 2. de norma está completamente determinada por su "bola unidad". Sea $x \in \mathbb{V}$ un vector cualquiera. Si $x = \mathbf{0}$, ya sabemos que $\|x\| = 0$. Si $x \neq \mathbf{0}$ y $x = \lambda \omega$ es su descomposición polar, entonces para $t > 0$ tenemos:

$$t < \lambda \;\Longrightarrow\; \frac{x}{t} \;=\; \frac{\lambda}{t}\omega \;\notin\; \overline{B}\big(\mathbf{0}, 1\big) \,,$$

$$t \geq \lambda \;\Longrightarrow\; \frac{x}{t} \;=\; \frac{\lambda}{t}\omega \;\in\; \overline{B}\big(\mathbf{0}, 1\big) \,,$$

es decir que el conjunto $\left\{ t > 0 \,:\, \dfrac{x}{t} \in \overline{B}\big(\mathbf{0}, 1\big) \right\}$ es el intervalo $[\lambda, +\infty)$, quedando el número $\|x\| = \lambda$ determinado como el extremo inferior de este intervalo:

$$\|x\| \;=\; \inf\big\{\, t > 0 \,:\, x/t \in \overline{B}(\mathbf{0}, 1) \,\big\} \;=\; \min\big\{\, t > 0 \,:\, x/t \in \overline{B}(\mathbf{0}, 1) \,\big\} \;.$$

1.1.7 Elipsoides sólidos

Definiciones 6. *Sea* $\mathbb{V}$ *un espacio vectorial de dimensión finita. Un subconjunto* $E \subset \mathbb{V}$ *es un* **elipsoide centrado en $\mathbf{x_0}$** *si existen una forma cuadrática* $Q : \mathbb{V} \to \mathbb{R}$ *definida positiva y un* $a > 0$ *tales que*

$$E \;=\; \big\{\, y \in \mathbb{V} \,:\, Q(y - x_0) = a \,\big\} \;=\; x_0 + \big\{ x \in \mathbb{V} \,:\, Q(x) = a \,\big\} \,.$$

Un subconjunto $B \subset \mathbb{V}$ *es un* **elipsoide sólido centrado en $\mathbf{x_0}$** *si existen una forma cuadrática* $Q : \mathbb{V} \to \mathbb{R}$ *definida positiva y un* $a > 0$ *tales que*

$$B \;=\; \big\{\, y \in \mathbb{V} \,:\, Q(y - x_0) \leq a \,\big\} \;=\; x_0 + \big\{ x \in \mathbb{V} \,:\, Q(x) \leq a \,\big\} \,.$$

Si $\|x\| \equiv \sqrt{\langle x, x\rangle}$ es una norma euclídea en un espacio vectorial $\mathbb{V}$, entonces las bolas cerradas de radio positivo

$$\overline{B}\big(\,\mathbf{0}\,,\,r\,\big) \;=\; \big\{\, x \in \mathbb{V}\,,\,\langle x, x\rangle \leq r^2\,\big\}\,,$$

son elipsoides sólidos centrados en el origen $\mathbf{0}$, porque $Q(x) \equiv \langle x, x\rangle$ es una forma cuadrática definida positiva.

Recíproco: si la bola unidad es un elipsoide sólido (centrado en $\mathbf{0}$), entonces la norma es euclídea. Supongamos que la bola unidad de una norma $\|\cdot\|$ es un elipsoide sólido

$$\big\{\, x \,:\, Q(x) \leq a \big\} = \left\{ x : \frac{1}{a}\, Q(x) \leq 1 \right\} \,.$$

Entonces la forma cuadrática $(1/a)Q$ procede de un producto escalar $\langle\cdot,\cdot\rangle$ en $\mathbb{V}$ y la correspondiente norma euclídea $\|\cdot\|'$ tiene al elipsoide sólido $\big\{\, x : Q(x) \leq a \big\}$ por bola unidad. Entonces $\|\cdot\| \equiv \|\cdot\|'$, porque la bola unidad determina la norma, y así $\|\cdot\|$ es euclídea.

Esta observación nos permite dar una infinidad de ejemplos de normas que no son euclídeas. En particular, es fácil comprobar que las siguientes fórmulas

$$\|(x,y)\|_1 \;\stackrel{\text{def}}{=}\; |x| + |y| \quad , \quad \|(x,y)\|_\infty \;\stackrel{\text{def}}{=}\; \max\big\{ |x|, |y| \big\} \,,$$

definen normas en $\mathbb{R}^2$ cuyas bolas son cuadrados, luego estas normas no son euclídeas.

Teorema 7. *Una norma* $\|\cdot\|$ *es euclídea si y sólo si cumple la* **regla del paralelogramo:**

$$2\big(\|x\|^2 + \|y\|^2\big) \;=\; \|x + y\|^2 + \|x - y\|^2 \quad , \quad \textit{para cualesquiera } \; x, y \in \mathbb{V} \,.$$

□

Corolario 8. *Una norma es euclídea si y sólo si el corte de la bola unidad con cada subespacio vectorial de dimensión 2 es una elipse sólida centrada en el origen.*

Demostración. Si cumple esa condición de corte, entonces en cada subespacio bidimensional cumple la regla del paralelogramo. Pero entonces la cumple siempre, porque dicha regla involucra sólo cuatro vectores: x , y , $x{+}y$, $x{-}y$, que están metidos en algún subespacio con $\dim \leq 2$. □

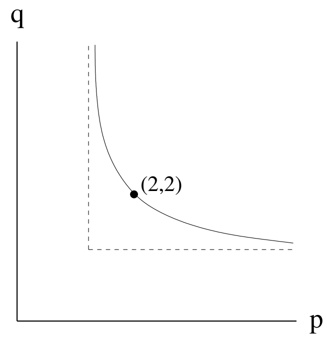
1.2 Las normas p en $\mathbf{R^n}$

Definiciones 9. *Sea $1 \le p < \infty$. La* **norma p** *en $\mathbb{R}^n$ es la función $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$ definida de la siguiente manera:*

$$\|(x_1, \ldots, x_n)\|_p \;=\; \left[\sum_{j=1}^{n} |x_j|^p\right]^{\frac{1}{p}} .$$

Dos números $1 < p, q < \infty$ son **exponentes conjugados** *si cumplen la ecuación* $\dfrac{1}{p} + \dfrac{1}{q} = 1.$

La ecuación $(1/p) + (1/q) = 1$ se transforma fácilmente en $(p-1)(q-1) = 1$ y, por lo tanto, define una curva en el plano pq que es el resultado de trasladar una unidad hacia la derecha y una unidad hacia arriba la rama de hipérbola $\{pq = 1, p > 0, q > 0\}$.



La aplicación que lleva un exponente a su conjugado es una biyección decreciente del intervalo $(1, \infty)$ consigo mismo:

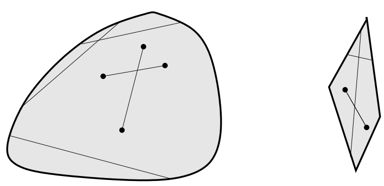
cuando p tiende a 1, su conjugado tiende a ∞ ; cuando p tiende a ∞, su conjugado tiende a 1.

El exponente $p = 2$ es el único que es conjugado de sí mismo.

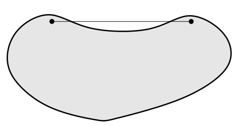
1.2.1 Convexidad

Definiciones 10. *Dados $x, y \in \mathbb{V}$, denotamos por $[x, y] = \{(1-\lambda)x + \lambda y : 0 \le \lambda \le 1\}$ el segmento rectilíneo cuyos extremos son x, y.*
Un subconjunto $E \subseteq \mathbb{V}$ es **convexo** *si siempre que $x, y \in E$ se tiene $[x, y] \subseteq E$.*

La siguiente figura muestra dos conjuntos convexos en el plano $\mathbb{R}^2$. También muestra los segmentos rectilíneos uniendo algunos pares de puntos.



Y ahora mostramos un subconjunto del plano que no es convexo, porque vemos que el segmento rectilíneo que une un par particular de puntos se sale del conjunto.



Intuitivamente, un conjunto convexo es el que no tiene ninguna "bahía": visto desde afuera, es "saliente por todas partes". Los subconjuntos convexos de la recta son los intervalos.

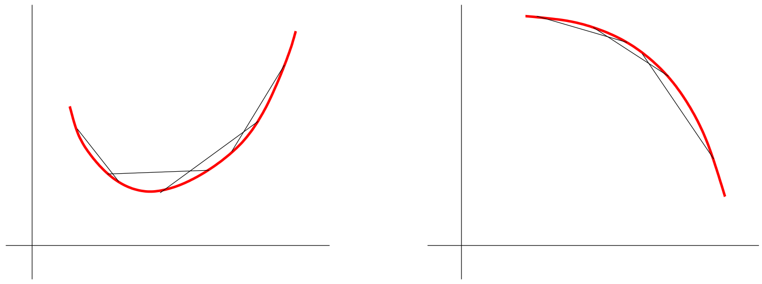
Definiciones 11. *Sean $E \subseteq \mathbb{V}$ conjunto convexo no vacío y $f : E \to \mathbb{R}$. Decimos que f es una* **función convexa** *si cumple:*

$$x, y \in E\,,\ \lambda \in [0,1] \implies f\big((1-\lambda)\,x + \lambda\,y\big) \ \le\ (1-\lambda)\,f(x) + \lambda\,f(y)\ . \tag{1}$$

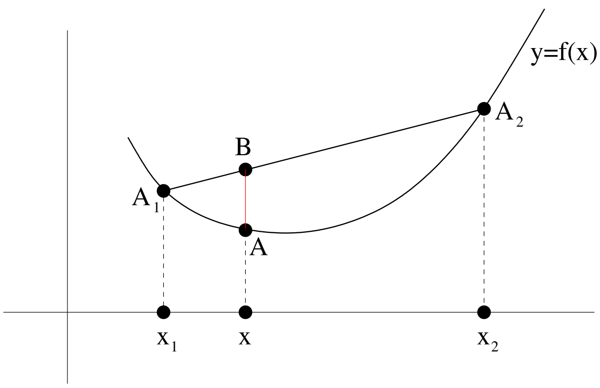
Análogamente decimos que f es una **función cóncava** *si cumple:*

$$x, y \in E\,,\ \lambda \in [0,1] \implies f\big((1-\lambda)\,x + \lambda\,y\big) \ \ge\ (1-\lambda)\,f(x) + \lambda\,f(y)\ .$$

Para una función convexa f de una variable, las **cuerdas** del grafo $\{(x, f(x)) : x \in I\}$ quedan por encima de él o forman parte de él. Para una función cóncava, las cuerdas del grafo quedan por debajo de él o forman parte de él.



Veamos, con una figura, que la condición de cuerdas por encima es equivalente a las desigualdades (1).



La función $f(x)$ está definida en un intevalo $I \subseteq \mathbb{R}$. Elegimos tres valores $x_1, x_2 \in I$ y $\lambda \in [0,1]$. El valor $x = (1-\lambda)x_1 + \lambda x_2$ es intermedio enre x_1 y x_2. Los puntos correspondientes sobre el grafo son

$$A_1 \ = \ \big(x_1, f(x_1)\big) \quad , \quad A \ = \ \big(x, f(x)\big) \quad , \quad A_2 \ = \ \big(x_2, f(x_2)\big)\ .$$

Con este mismo λ, definimos el punto $B = (1-\lambda)A_1 + \lambda A_2 \ \in\ [A_1, A_2]$. El siguiente cálculo

$$B \ = \ (1-\lambda)\,\big(x_1, f(x_1)\big) + \lambda\,\big(x_2, f(x_2)\big) \ = \ \Big((1-\lambda)x_1 + \lambda x_2\ ,\ (1-\lambda)f(x_1) + \lambda f(x_2)\Big)\ ,$$

muestra que B tiene la misma abscisa que A:

$$B \ = \ \big(x, y(x)\big) \quad , \quad \text{con}\ \ y(x) \ = \ (1-\lambda)f(x_1) + \lambda f(x_2)\ .$$

Que el segmento rectilíneo $[A_1, A_2]$ esté por encima del grafo equivale a que para cada x, intermedio entre x_1 y x_2, el punto $B = \big(x, y(x)\big)$ esté por encima del punto $A = \big(x, f(x)\big)$ (lo cual se indica en la figura por un segmento vertical rojo). Es decir, que para cualesquiera $x_1, x_2 \in I$ y $\lambda \in [0,1]$ se tenga

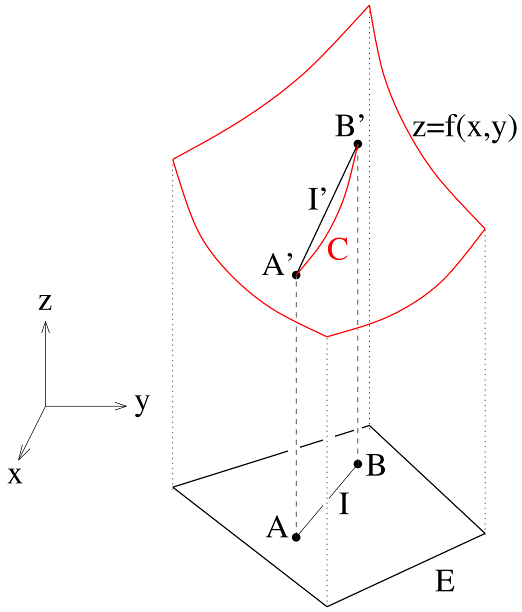
$$x \ = \ (1-\lambda)x_1 + \lambda x_2 \implies f(x) \ \le\ y(x) \ = \ (1-\lambda)f(x_1) + \lambda f(x_2)\ ,$$

que es lo mismo que la desigualdad (1).

Teorema 12. *Sean $I \subseteq \mathbb{R}$ un intervalo y $f : I \to \mathbb{R}$ una función derivable en todo punto de I. Si la derivada f' es monótona no decreciente, entonces f es convexa. Si f' es monótona no creciente, entonces f es cóncava.* $\square$

Ejemplos: x^2 es convexa en $I = \mathbb{R}$, mientras que $\log x$ es cóncava en $I = (0, +\infty)$.

La siguiente figura muestra el grafo (en $\mathbb{R}^3$) de una función convexa $f(x, y)$ definida en una región convexa E del plano xy.



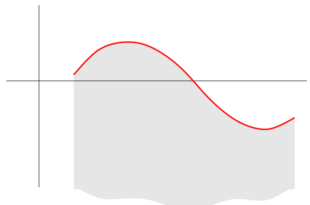
Sean $A, B \in E$, y sea $I \subseteq E$ el segmento rectilíneo $I = [A, B]$. Las verticales que pasan por I cortan a la superficie grafo $\{z = f(x, y)\}$ en una curva $C = \{(p, f(p)) : p \in I\}$. Si $A' = (A, f(A))$ y $B' = (B, f(B))$ son los extremos de C, entonces el segmento rectilíneo $I' = [A', B']$ queda por encima de la curva C, es decir que para todo $\lambda \in [0, 1]$ se tiene

$$(1 - \lambda) f(A) + \lambda f(B) \ \geq \ f\big((1 - \lambda) A + \lambda B \big) \ .$$

Por lo tanto, el segmento I' queda por encima de todo el grafo de f, ya que está en la vertical de la curva C. Podemos, pues, expresar la convexidad de $f(x, y)$ diciendo que, dado cuaquier par A', B' de puntos del grafo, el segmento $[A', B']$ queda por encima de dicho grafo.

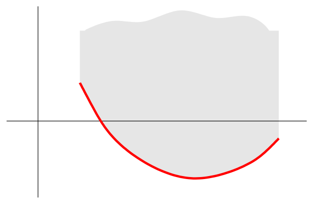
Definiciones 13. *Sean un conjunto no vacío E y una función $f : E \to \mathbb{R}$. El* **subgrafo** *de f es el siguiente subonjunto de $E \times \mathbb{R}$:*

$$\big\{ (p, y) \in E \times \mathbb{R} \ : \ y \ \leq \ f(p) \big\}$$



El **supergrafo** *de f es este otro conjunto:*

$$\big\{ (p, y) \in E \times \mathbb{R} \ : \ y \ \geq \ f(p) \big\}$$



Ejercicio 14. *Sean $E \subseteq \mathbb{R}^n$ convexo y $f : E \to \mathbb{R}$. Demuestra que f es convexa si y sólo si el supergrafo de f es un conjunto convexo en $\mathbb{R}^{n+1}$. Demuestra que f es cóncava si y sólo si el subgrafo de f es un conjunto convexo en $\mathbb{R}^{n+1}$.*

Ejercicio 15. *Sean $E \subseteq \mathbb{R}^n$ convexo y $f : E \to \mathbb{R}$ convexa. Demuestra que para todo $c \in \mathbb{R}$ el* **subnivel** $\big\{ p \in E \ : \ f(p) \ \leq \ c \big\}$ *es o vacío o un conjunto convexo en $\mathbb{R}^n$.*

1.2.2 Desigualdades de Young y Hölder

Teorema 16. *Dados exponentes conjugados p, q, se cumplen las desigualdades siguientes:*

Desigualdad de Young: $a, b \geq 0 \Longrightarrow ab \leq \dfrac{a^p}{p} + \dfrac{b^q}{q}.$

Desigualdad de Hölder: $x, y \in \mathbb{R}^n \Longrightarrow |x \cdot y| \leq \|x\|_p \, \|y\|_q.$

La desigualdad de Young con $p = 2$ nos da $\sqrt{a^2b^2} \leq \frac{a^2}{2} + \frac{b^2}{2}$, es decir la *desigualdad aritmético-geométrica* $\sqrt{xy} \leq \frac{x+y}{2}$ con $(x, y) = (a^2, b^2)$ números positivos cualesquiera.
La norma $\|\cdot\|_2$ es la norma euclídea estándar en $\mathbb{R}^n$ y la desigualdad de Hölder con $p = 2$ es la desigualdad de Cauchy-Schwarz.

Demostración de la desigualdad de Young. Hacemos $\lambda = \dfrac{1}{q} \in [0, 1]$, de donde $1 - \lambda = \dfrac{1}{p}$.
Como $\log x$ es una función cóncava, si $a, b > 0$ tenemos:

$$\log\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \;\geq\; \frac{1}{p}\,\log(a^p) + \frac{1}{q}\,\log(b^q) \;=\; \log a + \log b \;=\; \log(ab)\;,$$

y, tomando exponenciales, sale la desigualdad de Young. Si $a = 0$ o $b = 0$, es evidente.

Demostración de la desigualdad de Hölder. Es obvia cuando $x = \mathbf{0}$ o $y = \mathbf{0}$. Para $x \neq \mathbf{0} \neq y$, tomamos las descomposiciones polares $x = \|x\|_p\, \alpha$ e $y = \|y\|_q\, \beta$, con $\|\alpha\|_p = 1 = \|\beta\|_q$, lo que convierte la desigualdad de Hölder en $\|x\|_p\|y\|_q\, |\alpha \cdot \beta| \leq \|x\|_p\|y\|_q$. Basta, pues, con demostrar lo siguiente:

$$\|\alpha\|_p = 1 = \|\beta\|_q \;\Longrightarrow\; |\alpha \cdot \beta| \;\leq\; 1\;. \tag{2}$$

Si $\alpha = (\alpha_1, \cdots, \alpha_n)$ y $\beta = (\beta_1, \ldots, \beta_n)$, aplicamos la desigualdad de Young a cada producto $|\alpha_j\beta_j|$ y obtenemos:

$$|\alpha \cdot \beta| \;=\; \left|\sum_{j=1}^{n} \alpha_j\beta_j\right| \;\leq\; \sum_{j=1}^{n} |\alpha_j\beta_j| \;\leq\; \frac{1}{p}\sum_{j=1}^{n} |\alpha_j|^p + \frac{1}{q}\sum_{j=1}^{n} |\beta_j|^q \;=\; \frac{1}{p}\cdot 1^p + \frac{1}{q}\cdot 1^q \;=\; 1\;,$$

lo que demuestra (2) y, por lo explicado, la desigualdad de Hölder en general. □

1.2.3 Desigualdad de Minkowski

Dicha desigualdad dice que para $1 < p < \infty$ y $x, y \in \mathbb{R}^n$ se verifica:

$$\|x + y\|_p \;\leq\; \|x\|_p + \|y\|_p\;.$$

Como es fácil ver que esto también se cumple para $p = 1$, resulta que las normas p son todas ellas verdaderas normas en $\mathbb{R}^n$.

Teorema 17. (Hermann Minkowski). *Sea $\|\cdot\|$ una función en $\mathbb{V}$ que cumple los axiomas 1. y 2. de las normas. Entonces $\|\cdot\|$ cumple la desigualdad triangular si y sólo si el conjunto $\{x \in \mathbb{V} : \|x\| \leq 1\}$ es convexo.* □

Demostración. Si $\|\cdot\|$ es una norma, entonces es fácil comprobar que es una función convexa y, por el ejercicio 15, la bola unidad $\{v : \|v\| \leq 1\}$ es convexa.
Suponiendo ahora que $\|\cdot\|$ cumple los axiomas 1. y 2. de las normas y que el conjunto $\{v : \|v\| \leq 1\}$ es convexo, vamos a ver que $\|x + y\| \leq \|x\| + \|y\|$. Esto último es obvio si $x = \mathbf{0}$ o si $y = \mathbf{0}$. Si $x \neq \mathbf{0} \neq y$, tomamos las descomposiciones polares

$$x \;=\; a\alpha \begin{cases} a = \|x\| \\ \|\alpha\| = 1 \end{cases} \qquad y \;=\; b\beta \begin{cases} b = \|y\| \\ \|\beta\| = 1 \end{cases}$$

Por otra parte, dado el número $\lambda = \dfrac{b}{a+b} \in [0,1]$ se tiene $1-\lambda = \dfrac{a}{a+b}$. Como α, β pertencen al conjunto convexo $\{v : \|v\| \leq 1\}$, se tiene $(1-\lambda)\alpha + \lambda\beta \in \{v : \|v\| \leq 1\}$, es decir,

$$1 \;\geq\; \big\|\,(1-\lambda)\alpha + \lambda\beta\,\big\| \;=\; \left\|\frac{a}{a+b}\alpha + \frac{b}{a+b}\beta\right\| \;=\; \left\|\frac{a\alpha + b\beta}{a+b}\right\| \;=\; \left\|\frac{x+y}{a+b}\right\| \;=\; \frac{\|x+y\|}{a+b} \;,$$

en definitiva $\|x+y\| \leq a+b = \|x\| + \|y\|$. □

Vamos a demostrar la desigualdad de Minkowski a partir de este teorema. Como $p > 1$, la función $f(t) \equiv |t|^p$ es convexa. Se sigue que la siguiente función también es convexa:

$$g(x_1, \ldots, x_n) \;\equiv\; |x_1|^p + \cdots + |x_n|^p \;\equiv\; \|x\|_p^p \,.$$

En efecto, para cualesquiera $x, y \in \mathbb{R}^n$ y $\lambda \in [0,1]$ se cumple:

$$\begin{aligned}
g\big(\,(1-\lambda)x + \lambda y\,\big) \;&=\; g\Big(\,(1-\lambda)x_1 + \lambda y_1 \;,\; \ldots \;,\; (1-\lambda)x_n + \lambda y_n\,\Big) \;\stackrel{\text{def}}{=} \\
&=\; f\big(\,(1-\lambda)x_1 + \lambda y_1\,\big) + \cdots + f\big(\,(1-\lambda)x_n + \lambda y_n\,\big) \;\leq \\
&\leq\; (1-\lambda)f(x_1) + \lambda f(y_1) + \cdots + (1-\lambda)f(x_n) + \lambda f(y_n) \;= \\
&=\; (1-\lambda)\,\big[\,f(x_1) + \cdots + f(x_n)\,\big] + \lambda\,\big[\,f(y_1) + \cdots + f(y_n)\,\big] \;= \\
&=\; (1-\lambda)g(x) + \lambda g(y)\;.
\end{aligned}$$

Entonces el conjunto $\{x \in \mathbb{R}^n : g(x) \leq 1\} = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$ es convexo por el ejercicio 15 y, por el teorema 17, la función $\|\cdot\|_p$ cumple la desigualdad triangular.

Otra demostración de la desigualdad de Minkowski. Dados $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, definimos los siguientes vectores:

$$x' \;=\; (|x_1|, \ldots, |x_n|) \;\;,\;\; y' = (|y_1|, \ldots, |y_n|) \;\;,\;\; z \;=\; (|x_1+y_1|^{p-1}, \ldots, |x_n+y_n|^{p-1})\;.$$

Empezamos con la siguiente estimación:

$$\|x+y\|_p^p \;=\; \sum_{j=1}^n |x_j + y_j|^p \;\leq\; \sum_{j=1}^n (|x_j| + |y_j|)\,|x_j + y_j|^{p-1} \;=\; x' \cdot z + y' \cdot z\;,$$

aplicamos la desigualdad de Hölder:

$$x' \cdot z + y' \cdot z \;\leq\; \|x'\|_p\,\|z\|_q + \|y'\|_p\,\|z\|_q\;,$$

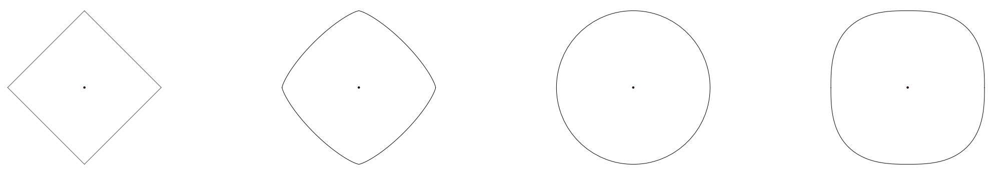
y comprobamos que $\|z\|_q = \|x+y\|_p^{p-1}$ (aquí se usa la fórmula $q = p/(p-1)$). Como es obvio que $\|x'\|_p = \|x\|_p$ y que $\|y'\|_p = \|y\|_p$, llegamos a:

$$\|x+y\|_p^p \leq \big(\|x\|_p + \|y\|_p\big)\,\|x+y\|_p^{p-1}\;.$$

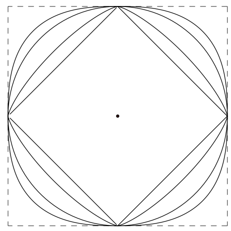
Si $x + y \neq \mathbf{0}$, dividimos por $\|x+y\|_p^{p-1}$ y ya está. Si $x + y = \mathbf{0}$ la desigualdad de Minkowski es obvia.

1.2.4 Más sobre las normas p

Mostramos a continuación dibujos de la bola unidad $\{v \in \mathbb{R}^2 \,:\, \|v\|_p \leq 1\}$ para los valores (de izquierda a derecha) $p=1$, $p=1'3$, $p=2$ y $p=2'7$:



Por supuesto, todas son subconjuntos convexos del plano. Ahora mostramos esas mismas bolas superpuestas, y, en línea de trazos, la "forma límite" a la que tienden cuando $p \to \infty$:



La bola aumenta en todas las direcciones a medida que p aumenta, excepto en las direcciones de los ejes coordenados. Esto se debe a una disminución monótona del valor de la norma:

Proposición 18. *Fijado un vector $v_0 \in \mathbb{R}^n$ que no esté en los ejes coordenados, el número $\|v_0\|_p$ es función estrictamente decreciente de p. Si por el contrario v_0 está en uno de los ejes coordenados, entonces $\|v_0\|_p$ no depende de p.*

En ambos casos $p \longmapsto \|v_0\|_p$ es una función monótona y tiene un límite finito (y no negativo) cuando $p \to \infty$. Es fácil demostrar que:

$$\lim_{p\to\infty} \|(x_1, \ldots, x_n)\|_p \;=\; \max(|x_1|, \ldots, |x_n|)\ .$$

Este resultado define una norma en $\mathbb{R}^n$, que denotamos $\|\cdot\|_\infty$ y de manera natural llamamos "norma infinito" en $\mathbb{R}^n$:

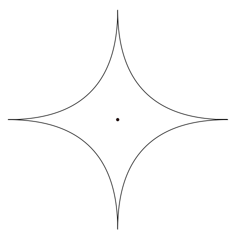
$$\|(x_1, \ldots, x_n)\|_\infty \overset{\text{def}}{=} \max_{1\le j\le n} |x_j|\ .$$

La bola unidad de $\|\cdot\|_\infty$ es $[-1, 1]^n \subset \mathbb{R}^n$. Cuando $n = 2$, es el cuadrado límite mostrado arriba.

Los dibujos que hemos mostrado en el plano son las intersecciones con el plano x_1x_2 de las bolas unidad de la norma p en $\mathbb{R}^n$. Es fácil demostrar que no son elipses sólidas, excepto para $n = 2$, luego sólo la bola unidad de $\|\cdot\|_2$ es un elipsoide sólido en $\mathbb{R}^n$:

De todas las normas $\|\cdot\|_p$ en $\mathbb{R}^n$, solamente la $\|\cdot\|_2$ es euclídea.

¿Por qué $\mathbf{p \ge 1}$? Cuando $0 < p < 1$, la fórmula $\varphi_p(x_1, \ldots, x_n) = \left[\sum_{j=1}^n |x_j|^p\right]^{1/p}$ todavía tiene sentido y la función $\varphi_p(x)$ así definida cumple los axiomas 1. y 2. de las normas; pero no cumple la desigualdad triangular. Por ejemplo, para $n = 2$ y $p = 1/2$ la "bola" $\{x : \varphi_{1/2}(x) \le 1\}$ tiene la siguiente forma no convexa:



1.3 Normas de operador

Cuando vemos $\mathbb{R}^{mn}$ como el espacio de las aplicaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^m$, se distinguen en él unas normas que van a sernos especialmente útiles. Son las que definimos aquí.

Definición 19. *Sean* $(\mathbb{V}, \|\cdot\|_{\mathbb{V}}), (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ *espacios normados. Para todo* $r > 0$ *definimos las bolas*

$$\overline{B}_{\mathbb{V}}(\mathbf{0}, r) \;=\; \big\{\, x \in \mathbb{V} \;:\; \|x\|_{\mathbb{V}} \;\leq\; r \,\big\} \quad , \quad \overline{B}_{\mathbb{W}}(\mathbf{0}, r) \;=\; \big\{\, x \in \mathbb{W} \;:\; \|x\|_{\mathbb{W}} \;\leq\; r \,\big\} \, .$$

Llamamos **aplicación lineal acotada de** $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ **a** $(\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ *a cualquier aplicación lineal* $L : \mathbb{V} \to \mathbb{W}$ *tal que el conjunto imagen* $L\big(\,\overline{B}_{\mathbb{V}}(\mathbf{0}, 1)\,\big)$ *es acotado en* $(\mathbb{W}, \|\cdot\|_{\mathbb{W}})$*. Es decir, que existe una constante* $M \geq 0$ *tal que*

$$L\big(\,\overline{B}_{\mathbb{V}}(\mathbf{0}, 1)\,\big) \;\subseteq\; \overline{B}_{\mathbb{W}}(\mathbf{0}, M) \, . \tag{3}$$

El conjunto de las aplicaciones lineales acotadas se denota por $\mathcal{L}\big(\,(\mathbb{V}, \|\cdot\|_{\mathbb{V}})\,,\,(\mathbb{W}, \|\cdot\|_{\mathbb{W}})\,\big)$ o, cuando no hay duda de cuales son las normas, por $\mathcal{L}(\mathbb{V}, \mathbb{W})$.

La fórmula (3) equivale a $\|x\|_{\mathbb{V}} \;\leq\; 1 \implies \|L(x)\|_{\mathbb{W}} \;\leq\; M$.

Lema. Dados una aplicación Lineal $L : \mathbb{V} \to \mathbb{W}$ y un número $M \geq 0$, son equivalentes

1. $\|x\|_{\mathbb{V}} \;\leq\; 1 \implies \|L(x)\|_{\mathbb{W}} \;\leq\; M$.
2. $\|\omega\|_{\mathbb{V}} \;=\; 1 \implies \|L(\omega)\|_{\mathbb{W}} \;\leq\; M$.
3. Se tiene $\|L(x)\|_{\mathbb{W}} \;\leq\; M \, \|x\|_{\mathbb{V}}$ para todo $x \in \mathbb{V}$.

Demostración. Es evidente que 1. implica 2., porque $\|\omega\|_{\mathbb{V}} = 1 \implies \|\omega\|_{\mathbb{V}} \leq 1$.
Supongamos que se cumple 2. y veamos que entonces se cumple 3. Es trivial para $x = \mathbf{0}$, porque

$$\|L(\mathbf{0})\|_{\mathbb{W}} \;=\; \|\mathbf{0}\|_{\mathbb{W}} \;=\; 0 \;=\; M\,0 \;=\; M\,\|\mathbf{0}\|_{\mathbb{V}} \, .$$

Para cada $x \in \mathbb{V} \setminus \{\mathbf{0}\}$ tenemos una descomposición polar $x = \|x\|_{\mathbb{V}}\,\omega$, con $\|\omega\|_{\mathbb{V}} = 1$. Entonces

$$\|L(x)\|_{\mathbb{W}} \;=\; \big\|L\big(\|x\|_{\mathbb{V}}\,\omega\big)\big\|_{\mathbb{W}} \;=\; \big\|\|x\|_{\mathbb{V}}\,L(\omega)\big\|_{\mathbb{W}} \;=\; \|x\|_{\mathbb{V}}\,\big\|L(\omega)\big\|_{\mathbb{W}} \;\leq\; \|x\|_{\mathbb{V}}\,M \, .$$

Finalmente, es trivial ver que 3. implica 1. □

El lema nos dice que si L es lineal acotada entonces los tres números siguientes son iguales:

– la mínima cota superior para $\{\, \|L(x)\|_{\mathbb{W}} \;:\; \|x\|_{\mathbb{V}} \leq 1 \,\}$,

– la mínima cota superior para $\{\, \|L(\omega)\|_{\mathbb{W}} \;:\; \|\omega\|_{\mathbb{V}} = 1 \,\}$,

– la mínima constante $M \geq 0$ tal que $\|L(v)\|_{\mathbb{W}} \leq M\,\|v\|_{\mathbb{V}}$ para todo $v \in \mathbb{V}$.

Definición 20. *Dada* $L : (\mathbb{V}, \|\cdot\|_{\mathbb{V}}) \to (\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ *lineal acotada, su* **norma de operador** *(respecto de las normas de vectores* $\|\cdot\|_{\mathbb{V}}$ *y* $\|\cdot\|_{\mathbb{W}}$*) es el número:*

$$\|L\| \;=\; \sup\{\, \|L(v)\|_{\mathbb{W}} \;:\; \|v\|_{\mathbb{V}} \leq 1 \,\} \;=\; \sup\{\, \|L(\omega)\|_{\mathbb{W}} \;:\; \|\omega\|_{\mathbb{V}} = 1 \,\} \, .$$

El número $\|L\|$ verifica (omitidos subíndices):

$$\|L(v)\| \;\leq\; \|L\|\,\|v\| \quad \text{para todo} \;\; v \in \mathbb{V} \, , \tag{4}$$

y no hay ninguna constante M menor que $\|L\|$ que nos dé $\|L(v)\| \leq M\|v\|$ para todo $v \in \mathbb{V}$.

Proposición 21. *El conjunto $\mathcal{L}(\mathbb{V},\mathbb{W})$ es cerrado para la suma y producto por constante, por lo tanto es un espacio vectorial. La norma de operador es, efectivamente, una norma en este espacio vectorial.*
Dados tres espacios normados $\mathbb{V}_1,\mathbb{V}_2,\mathbb{V}_3$ y aplicaciones $\mathbb{V}_1 \xrightarrow{L_1} \mathbb{V}_2 \xrightarrow{L_2} \mathbb{V}_3$ lineales acotadas, la compuesta $L_2\circ L_1 : \mathbb{V}_1 \to \mathbb{V}_3$ es lineal acotada y se verifica:

$$\|L_2\circ L_1\| \;\leq\; \|L_2\|\,\|L_1\| \;. \tag{5}$$

Ejercicio 22. *Demuestra esta proposición.*

Teorema 23. *Si $\mathbb{V}$ y $\mathbb{W}$ tienen dimensiones finitas, entonces toda aplicación lineal $\mathbb{V}\to\mathbb{W}$ es acotada respecto de cualquier par de normas que demos a $\mathbb{V}$ y $\mathbb{W}$.*

Demostraremos este teorema más adelante. De momento, lo vamos a utilizar.

Supongamos elegidas normas en tres espacios numéricos: $(\mathbb{R}^n,\|\cdot\|'),(\mathbb{R}^m,\|\cdot\|''),(\mathbb{R}^k,\|\cdot\|''')$ y sean dadas matrices $A\in \mathrm{M}_{k\times m}(\mathbb{R}), B\in\mathrm{M}_{m\times n}(\mathbb{R})$. Podemos verlas como operadores:

$$(\mathbb{R}^n,\|\cdot\|') \xrightarrow{B} (\mathbb{R}^m,\|\cdot\|'') \xrightarrow{A} (\mathbb{R}^k,\|\cdot\|''')\;,$$

con lo cual cada una tiene su norma de operador (respecto de las correspondientes normas de vectores). Además existe el producto AB y podemos verlo como un operador

$$AB : (\mathbb{R}^n,\|\cdot\|') \to (\mathbb{R}^k,\|\cdot\|''')\;.$$

Entonces:

$$\|AB\| \;\leq\; \|A\|\,\|B\|\;. \tag{6}$$

Esta desigualdad es lo que hace útiles las normas de operador: la mayoría de las normas que podemos definir en los espacios de matrices no cumplen (6).

Si fijamos una norma $\|\cdot\|$ en $\mathbb{R}^n$ y consideramos las matrices $n\times n$ como operadores

$$(\mathbb{R}^n,\|\cdot\|) \to (\mathbb{R}^n,\|\cdot\|)\;,$$

utilizando *la misma norma de vectores* en salida y en llegada, queda definida una norma de operador para las matrices $n\times n$ que cumple lo siguiente:

$$\|A_1A_2\cdots A_s\| \;\leq\; \|A_1\|\,\|A_2\|\cdots\|A_s\|\;,$$

y también cumple $\|I_n\|=1$, lo que refuerza la idea de que es una norma especial.

1.3.1 Elipsoides y norma de operador

Para A matriz *invertible* $n\times n$, y suponiendo que tanto en salida como en llegada tenemos la norma euclídea estándar

$$A : (\mathbb{R}^n,\|\cdot\|_2) \to (\mathbb{R}^n,\|\cdot\|_2)\;,$$

vamos a calcular la norma de operador $\|A\|$ explícitamente.

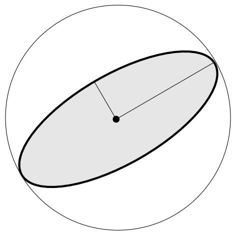
Lema 24. *Cuando A es invertible, la imagen $A\cdot\overline{B}(\mathbf{0},1)$ de la bola unidad estándar es un elipsoide sólido (centrado en el origen).*

Demostración. Como A es invertible existe la inversa $B=A^{-1}$. Entonces:

$$v\in A\cdot\overline{B}(\mathbf{0},1) \iff Bv\in\overline{B}(\mathbf{0},1) \iff \|Bv\|\leq 1 \iff 1\geq \|Bv\|^2=(Bv)^t(Bv)=v^t(B^tB)v\;.$$

Al ser B invertible, la matriz simétrica B^tB es definida positiva. Por lo tanto, el conjunto $\{v : v^t(B^tB)v\leq 1\}$ es un elipsoide sólido, centrado en el origen, que coincide con $A\cdot\overline{B}(\mathbf{0},1)$. □

La norma de operador $\|A\|$ es el mínimo radio M tal que la bola estándar $\overline{B}(\mathbf{0}, M)$ contiene al elipsoide sólido $E = A \cdot \overline{B}(\mathbf{0}, 1)$. Este radio coincide con el valor del **semieje principal máximo** de E.



Vamos a hallar los semiejes principales de E y a demostrar que, efectivamente, el mayor de ellos es igual a $\|A\|$.
La matriz B^tB es simétrica y definida positiva, por lo tanto existe una base ortonormal $\{u_1, \ldots, u_n\}$ de $\mathbb{R}^n$ que diagonaliza B^tB con autovalores reales positivos $0 < \mu_1 \leq \mu_2 \leq \cdots \leq \mu_n$.
Escribiendo el vector general v en esa base ortonormal, obtenemos:

$$v \;=\; y_1u_1 + \cdots + y_nu_n \implies v^t(B^tB)v \;=\; \mu_1\, y_1^2 + \cdots + \mu_n\, y_n^2 \geq \mu_1 \|v\|_2^2 \,.$$

Deducimos que $\|v\|_2^2 \leq (1/\mu_1)\, v^t(B^tB)v$, luego si $v^t(B^tB)v \leq 1$ entonces $\|v\|_2 \leq \sqrt{1/\mu_1}$. Esto demuestra que $E \subseteq \overline{B}\big(\,\mathbf{0}\,,\, \sqrt{1/\mu_1}\,\big)$.
Por otra parte, el vector $v_0 = \sqrt{1/\mu_1}\, u_1$ está en E y tiene $\|v_0\|_2 = \sqrt{1/\mu_1}$. Concluimos que el mínimo radio M tal que $E \subseteq \overline{B}(\mathbf{0}, M)$ es $\sqrt{1/\mu_1}$, es decir que $\|A\| = \sqrt{1/\mu_1}$.
De la sigiente equivalencia

$$y_1u_1 + \cdots + y_nu_n \in E \iff 1 \geq \mu_1\, y_1^2 + \cdots + \mu_n\, y_n^2 \;=\; \left(\frac{y_1}{1/\sqrt{\mu_1}}\right)^2 + \cdots + \left(\frac{y_n}{1/\sqrt{\mu_n}}\right)^2 \,, \quad (7)$$

deducimos que E es un elipsoide sólido cuyos semiejes principales miden

$$\sqrt{1/\mu_n} \;\leq\; \sqrt{1/\mu_{n-1}} \;\leq\; \cdots \;\leq\; \sqrt{1/\mu_1}\,,$$

con direcciones respectivas las de las rectas generadas por $u_1, \ldots, u_n$. Queda visto que $\|A\| = \sqrt{1/\mu_1}$ es el mayor de estos semiejes principales, tal como habíamos afirmado.

La matriz B^tB es la inversa de la matriz AA^t, luego los autovalores de AA^t son:

$$\lambda_1 = 1/\mu_1\,,\ \lambda_2 = 1/\mu_2\,,\ \ldots\,,\ \lambda_n = 1/\mu_n\,,$$

y la fórmula (7) podemos escribirla así:

$$y_1u_1 + \cdots + y_nu_n \in E \iff \left(\frac{y_1}{\sqrt{\lambda_1}}\right)^2 + \cdots + \left(\frac{y_n}{\sqrt{\lambda_n}}\right)^2 \;\leq\; 1\,.$$

Visto así, los semiejes principales $\sqrt{\lambda_n} \leq \sqrt{\lambda_{n-1}} \leq \cdots \leq \sqrt{\lambda_1}$ son las raíces cuadradas de los autovalores de AA^t. La norma $\|A\| = \sqrt{1/\mu_1} = \sqrt{\lambda_1}$ es la la raíz cuadrada del mayor de los autovalores de AA^t.

Observación. De la igualdad $A^{-1}(AA^t)A = A^tA$ se deduce que las matrices AA^t y A^tA, aunque suelen ser distintas, tienen los mismos autovalores.

Proposición 25. *Si A es una matriz cuadrada invertible, entonces su norma como operador de $(\mathbb{R}^n, \|\cdot\|_2)$ a $(\mathbb{R}^n, \|\cdot\|_2)$ viene dada por $\|A\| = \sqrt{\lambda}$, donde λ es, indistintamente, el máximo autovalor de AA^t o de A^tA.*

1.4 Espacios métricos y Topología

Las estructuras que vamos a definir aquí no sólo tienen sentido en espacios vectoriales, sino también en subconjuntos cualesquiera de un $\mathbb{R}^n$, que pueden no parecerse en nada a un espacio vectorial.

1.4.1 Funciones distancia

Definiciones 26. *Sea X un conjunto no vacío. Una* **distancia** *o* **métrica** *en X es cualquier función*

$$\begin{array}{rccc} d & : & X \times X & \longrightarrow & \mathbb{R} \\ & & (x,y) & \longmapsto & d(x,y) \end{array}$$

que cumpla las siguientes condiciones para cualesquiera $x, y, z \in X$:

dist1. $d(x,y) \geq 0$ *y* $d(x,y) = 0 \Longleftrightarrow x = y$.

dist2. $d(x,y) = d(y,x)$.

dist3. $d(x,z) \leq d(x,y) + d(y,z)$.

Un **espacio métrico** *es un par (X,d) formado por un conjunto X y una distancia d en X.*

La propiedad **dist1** significa que no queremos distancias negativas y que un buen *criterio de igualdad* para dos puntos es ver si están a distancia nula. La propiedad **dist2** se llama *simetría.* La propiedad **dist3** se llama *desigualdad triangular (para distancias).*

Decir que el espacio métrico es el par (X,d) significa, en particular, que si mantenemos el conjunto X pero cambiamos de una distancia a otra entonces tenemos *otro* espacio métrico.

Distancia procedente de una norma. Si $(\mathbb{V}, \|\cdot\|)$ es un espacio normado, podemos hallar la distancia entre dos puntos midiendo la longitud del vector que los une:

$$d(v,w) \;=\; \|v-w\| \quad , \quad \text{para cualesquiera } \; v,w \in \mathbb{V}\,. \tag{8}$$

Esto, efectivamente, define una distancia:

- d cumple **dist1** porque $\|v-w\| \geq 0$ y $\|v-w\| = 0 \Longleftrightarrow v - w = \mathbf{0}$.
- d cumple **dist2** porque $\|v-w\| = \|(-1)(v-w)\| = \|w-v\|$.
- d cumple **dist3** porque $\|u-w\| = \|u-v+v-w\| \leq \|u-v\| + \|v-w\|$.

Advertencias. (1) Las condiciones **dist1**, **dist2** y **dist3** tienen sentido en *cualquier* conjunto $X \neq \varnothing$, sin necesidad de que tenga estructura de espacio vectorial o de cualquier otro tipo. Por ejemplo, si $\|\cdot\|$ es una norma en $\mathbb{R}^n$ y $E \subset \mathbb{R}^n$ es cualquier subconjunto no vacío, entonces tenemos una distancia d_E definida en E de la manera siguiente:

$$d_E(p,q) \;=\; \|p-q\| \quad , \quad \text{para } \; p,q \in E\,. \tag{9}$$

(2) Incluso en el caso en que X es un espacio vectorial, pueden darse en él muchas distancias que no vengan de ninguna norma. Damos ejemplos de esto en el apartado 1.4.2.

1.4.2 Bolas métricas

Definiciones 27. *Sean (X,d) un espacio métrico y $x_0 \in X$ un punto suyo. La* **bola abierta** *con centro $x_0 \in X$ y radio $r > 0$ es el conjunto*

$$B_d(x_0, r) \;=\; \big\{ x \in X \,:\, d(x, x_0) < r \big\}\,.$$

La **bola cerrada** *con centro x_0 y radio $r \geq 0$ es el conjunto*

$$\overline{B}_d(x_0, r) \;=\; \big\{ x \in X \,:\, d(x, x_0) \leq r \big\}\,.$$

Cuando no hay duda de cuál es la función distancia, se omite el subíndice d.

Nótese que en la definición de bola cerrada permitimos el valor $r = 0$, siendo $\overline{B}(x_0, 0) = \{x_0\}$.

Para que una distancia d en un espacio vectorial $\mathbb{V}$ venga de una norma según la fórmula (8), debe cumplir una serie de condiciones (no las vamos a enumerar todas). Una condición es que bolas de un mismo radio r sean **trasladadas** unas de otras:

$$\text{si } \ p = p_0 + v\,, \ \text{ entonces } \ \overline{B}(p,r) \ = \ v + \overline{B}(p_0,r) \ \stackrel{\text{def}}{=} \ \big\{ v + q \ : \ q \in \overline{B}(p_0,r) \,\big\}\ .$$

Veamos un ejemplo de una distancia en $\mathbb{V} = \mathbb{R}^2$ que no cumple esto. Dada la biyección

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \quad , \quad \varphi(x,y) \ = \ (x^3, 2y)\ ,$$

y la norma $\|\cdot\| = \|\cdot\|_\infty$, definimos:

$$d(p,q) \ = \ \|\varphi(p) - \varphi(q)\|\ . \tag{10}$$

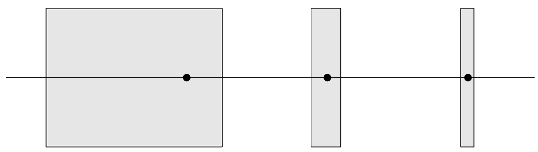
Es inmediato comprobar que d es una distancia, y que sus bolas de radio 1 vienen dadas por

$$\overline{B}\big(\,(x,y)\,,\,1\,\big) \ = \ \left[\,\sqrt[3]{x^3-1} \ , \ \sqrt[3]{x^3+1}\,\right] \times \left[\,y - \frac{1}{2}\,,\, y + \frac{1}{2}\,\right]\ .$$

En particular, las bolas de radio 1 centradas respectivamente en los puntos $(1,0), (2,0), (3,0)$ son, aproximadamente,

$$[\,0\,,\,1'26\,] \times \left[-\frac{1}{2},\frac{1}{2}\right] \quad , \quad [\,1'91\,,\,2'08\,] \times \left[-\frac{1}{2},\frac{1}{2}\right] \quad , \quad [\,2'96\,,\,3'04\,] \times \left[-\frac{1}{2},\frac{1}{2}\right]\ ,$$

y se ve en la siguiente figura que no son trasladadas unas de otras

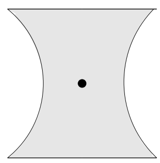


luego esta distancia no procede de ninguna norma.

Otra condición, necesaria para que la distancia se construya por la fórmula (8), es que las bolas sean conjuntos convexos. Si en la fórmula (10) dejamos $\|\cdot\| = \|\cdot\|_\infty$ pero cambiamos φ a esta otra biyección

$$\varphi(x,y) \ = \ \big(\,x/(1+y^2)\,,\,y\,\big)\ ,$$

entonces queda definida en $\mathbb{R}^2$ una distancia cuyas bolas no son convexas. Por ejemplo, la bola $\overline{B}\big((0,0),1\big)$ tiene la siguiente forma:



Si (X,d) es cualquier espacio métrico, podemos darle otra distancia d' definida así:

$$d'(x,y) \ = \ \min\big\{\,d(x,y)\,,\,1\,\big\}\ . \tag{11}$$

Con esta distancia es $\overline{B}_{d'}(x,1) = X$ para todo $x \in X$. Si $X = \mathbb{V}$ es un espacio vectorial, entonces para distancias definidas por la fórmula (8) o por la fórmula (10) se tiene $\overline{B}(x,1) \neq \mathbb{V}$. Luego d' no procede de ninguna norma, ni por la fórmula (8) ni por la fórmula (10).

1.4.3 Conjuntos acotados

Definición 28. *Un subconjunto* $E \subseteq X$ *es* **acotado** *si está contenido en alguna bola, es decir si* $E \subseteq \overline{B}(x_0, r)$ *para algún punto* $x_0 \in X$ *y algún radio* $r \geq 0$.

A partir de la desigualdad triangular se deduce fácilmente que para todo $x \in X$ se tiene

$$\overline{B}(x_0, r) \subseteq \overline{B}\big(x\,,\, r + d(x, x_0)\big)\,,$$

luego un subconjunto acotado está contenido en bolas *centradas en cualquier punto* $x \in X$.

1.4.4 Abiertos y cerrados

Definición 29. *Sea* (X, d) *un espacio métrico. Un subconjunto* $U \subseteq X$ *es* **abierto** *(para la distancia d) si se cumple:*

$$x \in U \implies \text{existe un } r > 0 \text{ tal que } B(x, r) \subseteq U\,.$$

Dado un punto $x_0 \in X$*, un* **entorno de** $\mathbf{x_0}$ *es cualquier abierto* U *tal que* $x_0 \in U$*. Dado un subconjunto* $E \subseteq X$*, un* **entorno de E** *es cualquier abierto* U *tal que* $E \subseteq U$*.*
Un subconjunto $E \subseteq X$ *es* **cerrado** *si su complemento* $X \setminus E$ *es abierto. De manera equivalente, los cerrados son los complementos de los abiertos.*

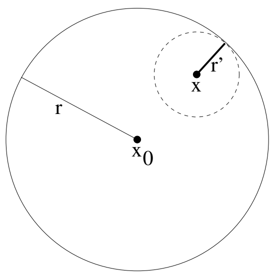
Recordemos que en Matemáticas una implicación $\mathcal{A} \implies \mathcal{B}$ se considera verdadera cuando la premisa $\mathcal{A}$ es falsa. Por lo tanto al vacío $\varnothing \subset X$ lo consideramos abierto. El total X es claramente un abierto en (X, d). Tanto $\varnothing$ como X son también cerrados, porque cada uno es el complementario del otro.

Es con frecuencia útil considerar a un abierto $U \subseteq \mathbb{R}^n$ como un *conjunto en el que es cómodo estar,* porque dado cualquier punto $x \in U$ podemos desplazarnos al menos una pequeña distancia r *en cualquier dirección* sin salirnos de U.

En la recta real, con la distancia $d(x, y) = |x - y|$, el intervalo $(0, 1)$ es abierto, el intervalo $[0, 1]$ es cerrado y $[0, 1)$ no es ni abierto ni cerrado. Los subconjuntos de un espacio métrico no son "puertas": pueden no ser ni abiertos ni cerrados. De hecho, *tanto los abiertos como los cerrados son subconjuntos muy especiales:* la inmensa mayoría de los subconjuntos no son de ninguno de esos dos tipos.

Proposición 30. *Las bolas abiertas* $B(x_0, r)$ *son todas conjuntos abiertos.*
En consecuencia, una bola abierta es entorno de todos sus puntos.

Demostración. Dado $x \in B(x_0, r)$, se define un número positivo $r' > 0$ por la igualdad $r' = r - d(x, x_0)$. El siguiente dibujo, que representa esa situación en el plano euclídeo, sirve para guiar nuestra intuición pero no es una demostración, ya que hemos visto que las bolas métricas pueden tener formas extrañas.



Para probar que, sea cual sea el espacio métrico, se cumple $B(x, r') \subseteq B(x_0, r)$, razonamos así:

$$y \in B(x, r') \iff d(x, y) < r' \implies d(x_0, y) \overset{*}{\leq} d(x_0, x) + d(x, y) < d(x_0, x) + r' = r\,,$$

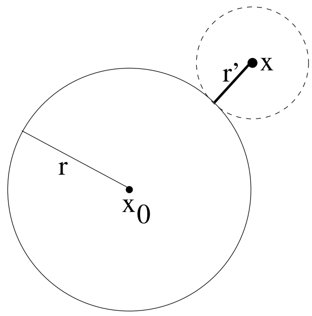
donde se ha marcado con $*$ el lugar donde se usa la desigualdad triangular. Esto prueba que $y \in B(x, r') \implies y \in B(x_0, r)$, es decir $B(x, r') \subseteq B(x_0, r)$.
Observa: cómo hemos usado la desigualdad triangular para acotar distancias *por arriba.* $\square$

Proposición 31. *Las bolas cerradas $\overline{B}(x_0, r)$, con $r \geq 0$, son subconjuntos cerrados.*

Demostración. Hay que demostrar que el complemento

$$X \setminus \overline{B}(x_0, r) \;=\; \{x \in X \;:\; d(x_0, x) > r\} \;,$$

es un abierto: para cada x con $d(x, x_0) > r$ hay que hallar un $r' > 0$ tal que la bola abierta $B(x, r')$ esté toda ella contenida en $X \setminus \overline{B}(x_0, r)$. Dado un tal x, se define un número positivo $r' > 0$ por la fórmula $r' = d(x_0, x) - r$. El siguiente dibujo muestra esa situación en el plano euclídeo:



y sugiere que, con la definición que hemos dado de r', se tiene $B(x, r') \subseteq X \setminus B(x_0, r)$. Insistimos en que el dibujo guía nuestra intuición pero no es una demostración.
Sea $y \in B(x, r')$. Razonamos de la manera siguiente, usando la simetría y la desigualdad triangular:

$$d(x,y) < r' \implies d(x_0, x) \;\leq\; d(x_0, y) + d(y, x) \;<\; d(x_0, y) + r' \implies r \;=\; d(x_0, x) - r' \;<\; d(x_0, y)\;.$$

en definitiva:

$$d(x,y) < r' \implies d(x_0, y) \;>\; r\;,$$

es decir $B(x, r') \subseteq X \setminus \overline{B}(x_0, r)$, como se quería demostrar.
Observa: cómo ahora hemos usado la desigualdad triangular para acotar distancias *por abajo.* □

Proposición 32. *Fijamos un espacio métrico (X, d).*

1. *La unión (finita o infinita) de abiertos produce un abierto.*
2. *La intersección finita de abiertos produce un abierto.*

Como consecuencia, son cerrados:

1. *La intersección finita o infinita de cerrados,*
2. *La unión finita de cerrados.*

Demostración. Sea $(U_i)_{i \in I}$ una familia de abiertos de (X, d), con el conjunto de índices I finito o infinito, y $U = \bigcup_{i \in I} U_i$. Por definición de unión, dado $x \in U$ existe un $i_0 \in I$ con $x \in U_{i_0}$. Al ser U_{i_0} un abierto, existe un $r > 0$ tal que

$$B(x, r) \;\subseteq\; U_{i_0} \;\subseteq\; U\;,$$

y queda visto que U es un abierto.
Para demostrar que la intersección finita de abiertos es abierta, es suficiente demostrarlo para la intersección de dos. Sean, pues U_1, U_2 abiertos en (X, d). Si $x \in U_1 \cap U_2$ entonces $x \in U_1$ y $x \in U_2$, luego existen radios $r_1, r_2 > 0$ con $B(x, r_1) \subseteq U_1$ y $B(x, r_2) \subseteq U_2$. Si $r = \min\{r_1, r_2\}$, entonces:

$$r > 0 \quad,\quad B(x,r) \subseteq U_1 \quad,\quad B(x,r) \subseteq U_2\;,$$

luego $B(x, r) \subseteq U_1 \cap U_2$. Esto prueba que $U_1 \cap U_2$ es un abierto de (X, d) □

Proposición 33. *Dado un espacio métrico (X, d), un subconjunto $U \subseteq X$ es un abierto si y sólo si es una unión de bolas abiertas.*

Demostración. Ya hemos visto que las bolas abiertas son abiertos y que la unión de abiertos es abierta. Sea ahora U un abierto cualquiera y para cada $x \in U$ elijamos un $r_x > 0$ tal que $B(x, r_x) \subseteq U$. Esto da lugar a una familia de bolas abiertas $\big(B(x, r_x)\big)_{x \in U}$ con U como conjunto de índices. Entonces:

$$U = \bigcup_{x \in U} \{x\} \subseteq \bigcup_{x \in U} B(x, r_x) \subseteq U ,$$

donde la última inclusión se debe a que cada $B(x, r_x)$ está contenida en U. Así $U = \bigcup_{x \in U} B(x, r_x)$.

$\square$

1.4.5 Abiertos y cerrados relativos

Sean (X, d) un espacio métrico e $Y \subseteq X$ un subconjunto no vacío. La restricción $d_Y = d|_{Y \times Y}$ es una distancia en Y que produce el nuevo espacio métrico (Y, d_Y).

Teorema-definición 34. *Un subconjunto $V \subseteq Y$ es un* **abierto relativo de Y** *si es abierto en (Y, d_Y). Esto equivale a $V = Y \cap U$ para algún U abierto en (X, d).*
Un subconjunto $F \subseteq Y$ es **cerrado relativo de Y** *si es un cerrado de (Y, d_Y). Esto equivale a $F = Y \cap E$ para algún E cerrado en (X, d).*

La demostración de las equivalencias se deja como ejercicio.

Ejemplo. Sea $\mathbb{R}$ con $d(x, y) = |x - y|$. El conjunto $A = (1, 2]$ no es abierto ni cerrado en $\mathbb{R}$, pero sí es cerrado en $Y_1 = (1, \infty)$ y sí es abierto en $Y_2 = (-\infty, 2]$.

Si conocemos la clase de los abiertos de X (y por lo tanto también los cerrados en X) entonces ya conocemos, mediante la operación de intersección, los abiertos y cerrados relativos de Y.

1.4.6 Interior, exterior, cierre y frontera

Definiciones 35. *Sean (X, d) un espacio métrico y $E \subseteq X$.*
Un punto $x \in X$ es **interior a** *E si existe un $r > 0$ tal que $B(x, r) \subseteq E$. El conjunto de estos puntos se llama* **interior de E** *y se denota $\operatorname{int} E$.*
Un punto $x \in X$ es **exterior a E** *si es interior a $X \setminus E$: existe un $r > 0$ con $B(x, r) \cap E = \varnothing$. El conjunto de los puntos exteriores a E es el* **exterior de E**.
Decimos que $x \in X$ es un **punto adherente a E** *si no es exterior a E, es decir si* toda bola abierta centrada en x corta a E: *para todo $r > 0$ tenemos $B(x, r) \cap E \neq \varnothing$. El conjunto de estos puntos se denota $\overline{E}$ y se llama* **adherencia de E** *o* **cierre de E**.
Un punto $x \in X$ es un **punto frontera de** *E si no es ni interior ni exterior a E. Es decir, toda bola abierta centrada en x corta tanto a E como a $X \setminus E$. El conjunto de estos puntos se llama* **frontera topológica de E** *y se denota $\operatorname{Fr} E$.*

Si meditas un momento sobre estas definiciones, verás que se tienen las siguientes relaciones:

$$\operatorname{int} E \subseteq E \subseteq \overline{E} \quad , \quad \text{exterior } E = \operatorname{int}(X \setminus E) = X \setminus \overline{E} \quad , \quad \operatorname{Fr} E = \overline{E} \setminus \operatorname{int} E .$$

Intuitivamente, que x sea exterior a E significa que no sólo no está en E sino que además está "un poco alejado" de E.

Lema 36. *El interior de cualquier conjunto $E \subseteq X$ es un abierto de X.*
Por lo tanto el cierre $\overline{E} = X \setminus \operatorname{int}(X \setminus E)$ y la frontera $\operatorname{Fr} E = \overline{E} \setminus \operatorname{int} E$ son ambos cerrados.

Demostración. Empecemos por observar lo siguiente:

$$\big(U \text{ abierto y } U \subseteq E\big) \implies U \subseteq \operatorname{int} E . \tag{12}$$

En efecto, para cada $y \in U$ existe un $r > 0$ tal que $B(y,r) \subseteq U$, luego $B(y,r) \subseteq E$ y por lo tanto $y \in \operatorname{int} E$. Como y era cualquier punto de U, queda visto que $U \subseteq \operatorname{int} E$.
Sea ahora x un punto cualquiera de $\operatorname{int} E$. Se verifica $B(x,r) \subseteq E$ para algún $r > 0$. Pero hemos visto en la proposición 30 que una bola abierta es un abierto, luego por (12) sabemos que $B(x,r) \subseteq \operatorname{int} E$. Como x era cualquier punto de $\operatorname{int} E$, hemos probado que $\operatorname{int} E$ es abierto. □

Proposición 37. *El interior de E es el abierto más grande contenido en E.*
El cierre de E es el cerrado más pequeño que contiene a E.
Por lo tanto, **un subconjunto $E \subseteq X$ es cerrado si y sólo si $\overline{E} = E$.**

Demostración. Como (12) afirma que $\operatorname{int} E$ contiene a cualquier abierto contenido en E, y él mismo es abierto por el lema 36, es el mayor abierto que cabe dentro de E.
Ya hemos comentado que el cerrado $\overline{E}$ contiene a E. Sea ahora C cualquier cerrado de (X,d) que contenga a E. El abierto $X \setminus C$ está contenido en $X \setminus E$, luego contenido en $\operatorname{int}(X \setminus E)$. Razonamos así:

$$X \setminus C \;\subseteq\; \operatorname{int}(X \setminus E) \;=\; X \setminus \overline{E} \;\Longrightarrow\; C \;\supseteq\; \overline{E}\,,$$

y así $\overline{E} \subseteq C$: el cierre $\overline{E}$ es el cerrado más pequeño en el que cabe E. □

Corolario 38. *Una vez que conocemos la clase de los abiertos en (X,d), conocemos el interior, el cierre y la frontera de cada subconjunto $E \subseteq X$.*
Dicho de otra manera, si dos funciones de distancia en X definen los mismos abiertos, entonces para cada subconjunto $E \subseteq X$ definen el mismo interior, el mismo cierre y la misma frontera.

Primer ejemplo: la recta $\mathbb{R}$ con $d(x,y) = |x-y|$. Estudiemos el caso de $E = (a,b]$. Es fácil convencerse de que (a,b) es el abierto más grande de $\mathbb{R}$ contenido en E, luego $\operatorname{int}(a,b] = (a,b)$. También es fácil ver que $[a,b]$ es el menor cerrado de $\mathbb{R}$ que contiene a E, luego $\overline{(a,b]} = [a,b]$. La frontera topológica $\operatorname{Fr}(a,b] = [a,b] \setminus (a,b) = \{a\} \cup \{b\}$ está formada por los dos puntos que separan a E del resto de la recta real.

Segundo ejemplo: un espacio normado $(\mathbb{V}, \|\cdot\|)$ y los subconjuntos $E = B(x,r)$ y $F = \overline{B}(x,r)$. Se comprueba fácilmente que:

$$\operatorname{int} B(x,r) \;=\; \operatorname{int} \overline{B}(x,r) \;=\; B(x,r) \quad , \quad \overline{B(x,r)} \;=\; \overline{\overline{B}(x,r)} \;=\; \overline{B}(x,r)\,,$$

y por lo tanto la "esfera generalizada" $\{\, y \,:\, \|x-y\| = r \,\}$ es la "cáscara" que separa la bola del resto del espacio:

$$\{\, y \,:\, \|x-y\| = r \,\} \;=\; \operatorname{Fr} B(x,r) \;=\; \operatorname{Fr} \overline{B}(x,r)\,, \tag{13}$$

Advertencias. (1) La frontera topológica es una "cáscara separadora" para *conjuntos muy sencillos en espacios sencillos,* como es una bola en un espacio normado. Pero para conjuntos o espacios métricos más complicados eso ya no es así. Un ejemplo es el conjunto $\mathbb{Q}$ de los racionales en la recta real (dotada de la distancia usual); enseguida se ve que:

$$\operatorname{int} \mathbb{Q} \;=\; \varnothing \quad , \quad \overline{\mathbb{Q}} \;=\; \mathbb{R}\,,$$

luego $\operatorname{Fr} \mathbb{Q} = \mathbb{R}$ sí que es un cerrado (tal como afirma el lema 36) pero no tiene nada que ver con la idea de una cáscara separadora.

(2) En la recta real $\mathbb{R}$ consideramos dos distancias: $d(x,y) = |x-y|$ y $d'(x,y) = \min\{1, |x-y|\}$.
Para ti: comprueba que ambas distancias definen los mismos conjuntos abiertos, y por lo tanto los mismos conceptso de interior, cierre y frontera para subconjuntos de $\mathbb{R}$.
La bola $B_d(0,1) = B_{d'}(0,1) = (-1,1)$ tiene frontera $\{-1,1\}$, mientras que la "esfera generalizada" $\{x : d'(0,x) = 1\} = (-\infty,-1] \cup [1,+\infty)$ *es mucho más grande,* en violento contraste con las igualdades (13) que se cumplen en espacios normados.
Así, el espacio métrico $(\mathbb{R}, d')$ nos presenta un fenómeno que no ocurre ni en el plano euclídeo ni en nigún espacio normado. Una prueba más de que los dibujos en el plano euclídeo, si bien pueden ser de gran ayuda (como en las demostraciones de las proposiciones 30 y 31), no son una

"garantía": si queremos mostrar que algo es verdad en *todos* los espacios métricos, es obligatorio elaborar una demostración basada sólo en los tres axiomas de las distancias.

Dado un espacio métrico (X, d) y un punto $x \in X$, el conjunto $\{x\}$ es cerrado porque coincide con la bola cerrada $\overline{B}(x, 0)$. De hecho, es un cerrado relativo de todo conjunto en el que esté x.

Definición 39. *Sea $E \subseteq X$ un subconjunto. Decimos que $x \in E$ es un* **punto aislado** *de E si existe $r_x > 0$ (dependiente de x) tal que $B(x, r_x) \cap E = \{x\}$, es decir que el conjunto $\{x\}$ es, además de cerrado, un abierto relativo de E.*
Decimos que E es **discreto** *si todos sus puntos son aislados.*

El conjunto $E = \{1/n : n = 1, 2, 3, \dots\}$ es discreto en la recta real. En cambio $\overline{E} = E \cup \{0\}$ no es discreto porque 0 es un punto suyo que no es aislado en E. Vemos, de paso, que un conjunto discreto puede no ser cerrado.

Ejercicio. Demuestra que un espacio métrico (X, d) es discreto si y sólo si todos sus subconjuntos son abiertos.

1.5 Convergencia de sucesiones

Definiciones 40. *Dado un conjunto X, una* **sucesión de elementos de X** *es una aplicación*

$$\{1, 2, 3, \dots\} \longrightarrow X \quad , \quad n \longmapsto x_n \ ,$$

del conjunto de los enteros positivos a X. Se la indica como una lista infinita $x_1, x_2, x_3, \dots$, también en la forma abreviada $\{x_n\}_{n=1}^{\infty}$ o simplemente $\{x_n\}$.
Dado un subconjunto $E \subseteq X$, escribimos $\{x_n\} \subset E$ para indicar que $x_n \in E$ para todo n.
Una **cola de la sucesión** *es el resultado $x_k, x_{k+1}, x_{k+2}, \dots$ de surpimir los $k - 1$ primeros términos para algún k.*

La aplicación $\{1, 2, 3, \dots\} \to X$ no tiene por qué ser inyectiva, es decir que se permiten *repeticiones* en las sucesiones. Por ejemplo $1, 2, 1, 1, 3, 1, 1, 1, 4, 1, 1, 1, 1, 5, 1, 1, \dots$ es una sucesión de números perfectamente válida; algunos de sus valores son $x_3 = x_4 = 1, x_5 = 3, x_{12} = 1$, etc.

Definición 41. *Sea (X, d) un espacio métrico. Una sucesión $\{x_n\}_{n=1}^{\infty} = \{x_1, x_2, x_3, \dots\}$ de puntos de X* **converge al punto $\mathbf{x_0} \in \mathbf{X}$**, *lo cual indicamos escribiendo $\{x_n\} \to x_0$, si cada bola $B(x_0, r)$ contiene una cola $x_k, x_{k+1}, x_{k+2}, \cdots$ de dicha sucesión.*
Dicho con más precisión, $\{x_n\}$ converge a x_0 si y sólo si para cada $r > 0$ existe un k (dependiente de r y de la sucesión) tal que $m \geq k \Longrightarrow x_m \in B(x, r)$.

Ejercicio: si la intersección $B(x, r) \cap B(x', r')$ no es vacía, entonces $d(x, x') < r + r'$.

Veamos, a partir de este resultado, que **una sucesión no puede converger a dos puntos diferentes** $x_0 \neq y_0$. Hacemos $r = d(x_0, y_0)/3 > 0$ y, como $d(x_0, y_0) > r + r$, se tiene $B(x_0, r) \cap B(y_0, r) = \varnothing$. Si cada bola contuviese una cola:

$$x_k, x_{k+1}, x_{k+2}, \dots \in B(x_0, r) \quad , \quad x_h, x_{h+1}, x_{h+2}, \dots \in B(y_0, r) \ ,$$

entonces para $n = \max\{k, h\}$ el punto x_n tendría que estar en las dos bolas a la vez, imposible.

Cada sucesión o bien no converge o lo hace a un único punto.

La unicidad es, precisamente, lo que da importancia al concepto de límite. Muchos objetos importantes en Matemáticas (números, funciones, conjuntos) se construyen como límites y, gracias a la unicidad, quedan perfectamente definidos por tal construcción.

Definiciones 42. *Una sucesión $\{x_n\}$, de puntos de (X, d), es* **convergente en X** *si existe un punto $x_0 \in X$ al cual converge. Entonces x_0 es único y se llama* **límite** *de la sucesión. Además de $\{x_n\} \to x_0$, también podemos escribir $\lim x_n = x_0$.*

Ejercicio: para que una sucesión $\{x_n\}$ sea convergente, es necesario que sea una **sucesión de Cauchy,** es decir que para cada $\varepsilon > 0$ debe haber un k (dependiendo de ε) tal que

$$n, m \geq k \implies d\big(x_n\,,\, x_m\big) \leq \varepsilon\,.$$

Es fáci ver que $\{x_n\} \to x_0$ si y sólo si todo entorno de x_0 contiene una cola de la sucesión $\{x_n\}$. Por lo tanto, la clase de los abiertos en (X, d) determina qué sucesiones convergen así como el límite de cada una de ellas.

Proposición 43. *Dado un subconjunto $E \subseteq X$, el cierre $\overline{E}$ es el conjunto de los límites de sucesiones $\{x_n\} \subset E$ convergentes en X. Por lo tanto E es cerrado si y sólo si contiene todos esos límites.*

Demostración. Sea F el conjunto de esos límites.
Primera parte: $F \subseteq \overline{E}$. Dado $x_0 \in F$, hay una sucesión $\{x_n\} \subset E$ convergente a x_0. Para cada $r > 0$ la bola $B(x_0, r)$ contiene una cola de la sucesión, luego contiene puntos de E y así $B(x_0, r) \cap E \neq \varnothing$. Esto prueba que $x_0 \in \overline{E}$ y que $F \subseteq \overline{E}$.
Segunda parte: $\overline{E} \subseteq F$. Fijamos un $y_0 \in \overline{E}$. Para cada entero positivo n tenemos $B(y_0, 1/n) \cap E \neq \varnothing$ y elegimos un punto x_n en esta intersección, formando así una sucesión $\{x_n\}_{n=1}^{\infty} \subset E$. Dado ahora cualquier $r > 0$, hay un n tal que $1/n < r$, con lo cual:

$$m \geq n \implies d(x_m, y_0) < \frac{1}{m} \leq \frac{1}{n} < r \implies x_m \in B(y_0, r)\,,$$

luego $B(y_0, r)$ contiene la cola $x_n, x_{n+1}, x_{n+2}, \ldots$ Así $\{x_n\} \to y_0 \in F$. Como y_0 era un punto arbitrario de $\overline{E}$, tenemos $\overline{E} \subseteq F$. □

Corolario 44. *Dados $E \subseteq Y \subseteq X$, el conjunto E es cerrado relativo de Y si y sólo si contiene los límites de todas las sucesiones $\{x_n\} \subset E$ que son convergentes en Y.*

Por ejemplo $E = (0, 1]$ es cerrado relativo en $Y = (0, 2)$. La sucesión $\big\{1/n\big\}_{n=1}^{\infty} \subset E$ *converge en* $\mathbb{R}$ *pero no en* Y, por eso no impide a E ser cerrado relativo en Y.

¿Por qué los conjuntos cerrados se llaman así? Primero, cuando un conjunto tiene una operación (suma, producto, etc) se dice que un subconjunto E es *cerrado para esa operación* si al operar con elementos de E siempre resulta un elemento de E. El conjunto $\mathbb{N}$ es cerrado para la suma pero no para la resta. Segundo, hay operaciones de más de dos argumentos, por ejemplo dados números $a, b, c \in \mathbb{R}$ el máximo $\max\{a, b, c\}$ es una operación que produce un número a partir de esos tres. Tercero, hay operaciones que no se pueden efectuar con cualesquiera elementos, por ejemplo la división $(x, y) \mapsto x/y$ no está definida en los pares $(x, 0)$. El conjunto $\mathbb{Z}$ no es cerrado para la divisón pero $\mathbb{Q}$ sí lo es. El *paso al límite* puede verse como una operación de infinitos argumentos

$$\{x_1, x_2, x_3, \ldots\} \longmapsto \lim x_n\,,$$

que no está definida para todas las sucesiones (sólo para las convergentes) y decir que *"E es cerrado para esta operación"* equivale a $E = \overline{E}$. Por eso llamamos *subconjuntos cerrados* a los que cumplen $E = \overline{E}$. Es un pequeño "milagro" que estos conjuntos coincidan con los complementarios de los abiertos.

1.6 Continuidad

Sean $(X, d), (X', d')$ dos espacios métricos cualesquiera y $f : X' \to X''$ una aplicación.

Definición 45. *Dado un punto $x_0 \in X$, decimos que f es* **continua en $\mathbf{x_0}$** *si para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que*

$$f\big(\,B(x_0\,,\,\delta)\big) \subseteq B\big(\,f(x_0)\,,\,\varepsilon\,\big)\,.$$

Es decir, para toda bola B' centrada en $f(x_0)$ hay una bola B, centrada en x_0, que es enviada por f dentro de B'.

Ejercicio. Cada una de las condiciones siguientes equivale a que f sea continua en x_0:

1. Para cada entorno V de $f(x_0)$ existe un entorno U de x_0 tal que $f(U) \subseteq V$.
2. Para cada entorno V de $f(x_0)$ el conjunto $f^{-1}(V)$ contiene un entorno de x_0.
3. Para todo $\varepsilon' > 0$ existe un $\delta' > 0$ tal que $f\big(\,\overline{B}(x_0, \delta')\,\big) \subseteq \overline{B}\big(\,f(x_0), \varepsilon'\,\big)$.

Proposición 46. *Sean espacios métricos* $(X, d), (X', d'), (X'', d'')$ *y aplicaciones* $X \xrightarrow{f} X' \xrightarrow{g} X''$. *Si* f *es continua en* $x_0 \in X$ *y* g *es continua en* $x_0' = f(x_0) \in X'$, *entonces la compuesta* $g \circ f$ *es continua en* x_0.

Demostración. Sea $\varepsilon_3 > 0$. Como g es continua en x_0', existe $\varepsilon_2 > 0$ tal que

$$g\big(\,B\big(x_0', \varepsilon_2\big)\,\big) \;\subseteq\; B\big(\,g(x_0')\,,\,\varepsilon_3\,\big) \;=\; B\big(\,(g \circ f)(x_0)\,,\,\varepsilon_3\,\big)\;.$$

Como f es continua en x_0, dado ε_2 existe $\varepsilon_1 > 0$ tal que $f\big(B(x_0, \varepsilon_1)\big) \subseteq B(x_0', \varepsilon_2)$ y así:

$$(g \circ f)\big(B(x_0, \varepsilon_1)\big) \;=\; g\big(\,f\big(B(x_0, \varepsilon_1)\big)\,\big) \;\subseteq\; g\big(\,B\big(x_0', \varepsilon_2\big)\,\big) \;\subseteq\; B\big(\,(g \circ f)(x_0)\,,\,\varepsilon_3\,\big)\;.$$

Como ε_3 era arbitrario, queda probada la continuidad de $g \circ f$ en x_0. □

Definición 47. *Una aplicación entre espacios métricos* $f : X \to X'$ *es* **continua** *si es continua en cada punto* $x_0 \in X$.

$$\boxed{\text{La compuesta de aplicaciones continuas es continua.}}$$

Teorema 48. *Para una aplicación* $f : (X, d) \to (X', d')$, *son equivalentes:*

1. f *es continua.*
2. *Para todo abierto* $V \subseteq X'$ *la preimagen* $f^{-1}(V)$ *es un abierto de* X.
3. *Para todo cerrado* $C \subseteq X'$ *la preimagen* $f^{-1}(C)$ *es un cerrado de* X.
4. *Para todo* $x_0 \in X$ *y toda sucesión* $\{x_n\} \subset X$ *con* $x_n \to x_0$ *se tiene* $\big\{f(x_n)\big\} \to f(x_0)$.

Demostración. Supongamos f continua y $V \subseteq X'$ abierto. Dado $x \in f^{-1}(V)$ es $f(x) \in V$ y existe un $\varepsilon > 0$ tal que $B\big(\,f(x)\,,\,\varepsilon\,\big) \subseteq V$. Entonces existe un $\delta > 0$ tal que

$$f\big(B(x, \delta)\big) \;\subseteq\; B\big(\,f(x)\,,\,\varepsilon\,\big) \;\subseteq\; V\,,$$

luego $B(x, \delta) \subseteq f^{-1}(V)$. Esto prueba que $f^{-1}(V)$ es abierto y que 1. $\Longrightarrow$ 2.
Supongamos ahora que f satisface 2. Dados $x \in X$ y $\varepsilon > 0$, el conjunto $f^{-1}\big(\,B\big(f(x), \varepsilon\big)\,\big)$ es un abierto de X al que pertenece x. Por lo tanto existe un $\delta > 0$ tal que

$$B(x, \delta) \;\subseteq\; f^{-1}\big(\,B\big(f(x), \varepsilon\big)\,\big)\;.$$

Aplicando f en ambos lados sale $f\big(B(x, \delta)\big) \subseteq B\big(f(x), \varepsilon\big)$. Luego f es continua y 2. $\Longrightarrow$ 1.

Para todo $A \subseteq X'$ se tiene $f^{-1}(X' \setminus A) = X \setminus f^{-1}(A)$. Si f satisface 2. y C es un cerrado de X', entonces $X' \setminus C$ es abierto de X' y su preimagen $X \setminus f^{-1}(C)$ es abierto de X, luego $f^{-1}(C)$ es un cerrado de X. Esto prueba que 2. $\Longrightarrow$ 3. Análogamente 3. $\Longrightarrow$ 2.

Sea ahora f continua y $\{x_n\} \to x_0$ en X. Dado $r > 0$ existe un $\varepsilon > 0$ tal que $f\big(B(x_0, \varepsilon)\big) \subseteq B\big(f(x_0), r\big)$. A su vez la bola $B(x_0, \varepsilon)$ contiene una cola $x_n, x_{n+1}, x_{n+2}, \ldots$ de la sucesión $\{x_n\}$ y, aplicando f:

$$f(x_n), f(x_{n+1}), f(x_{n+2}), \ldots \;\in\; f\big(B(x_0, \varepsilon)\big) \;\subseteq\; B\big(f(x_0), r\big)\,,$$

es decir que para todo $r > 0$ la bola $B\big(f(x_0), r\big)$ contiene una cola de la sucesión $\big\{f(x_n)\big\}$, lo cual equivale a $\big\{f(x_n)\big\} \to f(x_0)$. Esto prueba que 1. $\Longrightarrow$ 4.

Supongamos, por último, que f satisface 4. pero hay un punto $x_0 \in X$ en el que es discontinua. Entonces habrá un $\varepsilon_0 > 0$ "malo", en el sentido de que para ningún $\delta > 0$ estará la imagen $f\big(B(x_0,\delta)\big)$ contenida en $B\big(f(x_0),\varepsilon_0\big)$. En particular

$$f\big(B(x_0,1/n)\big) \not\subseteq B\big(f(x_0),\varepsilon_0\big) \quad , \quad n = 1,2,3,\ldots$$

Para cada entero positivo n podremos elegir un $x_n \in B(x_0,1/n)$ tal que $f(x_n) \notin B\big(f(x_0),\varepsilon_0\big)$. Estas elecciones formarán una sucesión $\{x_n\} \subset X$ con $\{x_n\} \to x_0$, mientras que la bola $B\big(f(x_0),\varepsilon_0\big)$ no contendrá ningún elemento de la sucesión $\big\{f(x_n)\big\}$, con lo cual $\big\{f(x_n)\big\} \not\to f(x_0)$, en contradicción con 4. Por reducción al absurdo, tal punto x_0 no existe y f es continua. Esto prueba que 4. $\Longrightarrow$ 1. □

Corolario 49. *Si conocemos la clase de los abiertos de X y la clase de los abiertos de X', entonces ya sabemos qué aplicaciones $X \to X'$ son continuas y cuáles no.*

Ejercicio. Demuestra las siguientes equivalencias.

1. Dadas funciones escalares $f_1,\ldots,f_k : (X,d) \to \mathbb{R}$, la correspondiente función vectorial

   $$f : (X,d) \longrightarrow (\mathbb{R}^k, \|\cdot\|_\infty) \quad , \quad x \longmapsto \big(f_1(x),\ldots,f_k(x)\big) \,,$$

   es continua si y sólo si las $f_1,\ldots,f_k$ son todas continuas.

2. Una sucesión de vectores $\{x_n\} \subset \mathbb{R}^k$ converge a x en $(\mathbb{R}^k, \|\cdot\|_\infty)$ si y sólo si para $i = 1,\ldots,k$ la sucesión $\{x_n^i\}_{n=1}^\infty$ de las i-ésimas coordenadas de los x_n converge a la i-ésima coordenada de x.

Si en cada subconjunto de $\mathbb{R}$ ponemos la distancia $|x-y|$ y en cada subconjunto de $\mathbb{R}^2$ ponemos la distancia $\|x-y\|_\infty$, entonces las siguientes funciones son todas continuas:

Suma: $\mathbb{R}^2 \to \mathbb{R} \quad , \quad (x,y) \longmapsto x+y$.
Multiplicación: $\mathbb{R}^2 \to \mathbb{R} \quad , \quad (x,y) \longmapsto xy$.
División: $\mathbb{R} \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R} \quad , \quad (x,y) \longmapsto \dfrac{x}{y}$.
Directas: $\mathbb{R} \to \mathbb{R} \quad , \quad x \longmapsto e^x = \exp(x)\,,\ \operatorname{sen} x\,,\ \cos x\,,$
Logaritmo: $(0,+\infty) \to \mathbb{R} \quad , \quad x \longmapsto \log x$.
Seno inversa: $[-1,1] \to [-\pi/2,\pi/2] \quad , \quad x \longmapsto \operatorname{arcsen} x$.
Coseno inversa: $[-1,1] \to [0,\pi] \quad , \quad x \longmapsto \arccos x$.
Raíz impar: $\mathbb{R} \to \mathbb{R} \quad , \quad x \longmapsto \sqrt[n]{x}\ ,\ n = 3,5,7,\ldots$
Raíz positiva: $[0,+\infty) \to [0,+\infty) \quad , \quad x \longmapsto \sqrt[n]{x},\ n = 2,3,4,5,\ldots$
Valor absoluto: $\mathbb{R} \to \mathbb{R} \quad , \quad x \longmapsto |x|$.

A esa lista se podrían añadir muchas más. Combinando la lista anterior con el hecho de que la compuesta de continuas es continua, obtenemos:

Si $f,g : (X,d) \to \mathbb{R}$ son continuas entonces $f+g : (X,d) \to \mathbb{R}$ es continua.
Si $f,g : (X,d) \to \mathbb{R}$ son continuas entonces $fg : (X,d) \to \mathbb{R}$ es continua.
Si $f,g : (X,d) \to \mathbb{R}$ son continuas y g nunca se anula entonces $\dfrac{f}{g} : (X,d) \to \mathbb{R}$ es continua.
Si $f : (X,d) \to \mathbb{R}$ es continua entonces $e^f, \operatorname{sen} f, \cos f : (X,d) \to \mathbb{R}$ son continuas.
Si $f : (X,d) \to \mathbb{R}$ es continua y siempre positiva entonces $\log f : (X,d) \to \mathbb{R}$ es continua.
Si $f,g : (X,d) \to \mathbb{R}$ son continuas y f siempre positiva entonces $f^g = \exp(g \log f)$ es continua.
Si $f : (X,d) \to [-1,1]$ es continua entonces $\operatorname{arcsen} f : (X,d) \to [-\pi/2,\pi/2]$ es continua.
Si $f : (X,d) \to [-1,1]$ es continua entonces $\arccos f : (X,d) \to [0,\pi]$ es continua.
Si $f : (X,d) \to \mathbb{R}$ es continua y n es impar entonces $\sqrt[n]{f} : (X,d) \to \mathbb{R}$ es continua.
Si $f : (X,d) \to [0,+\infty)$ es continua entonces $\sqrt[n]{f} : (X,d) \to [0,+\infty)$ es continua.
Si $f : (X,d) \to \mathbb{R}$ es continua entonces $|f| : (X,d) \to \mathbb{R}$ es continua.

Haciendo combinaciones sucesivas de esos casos se consigue lo siguiente:

> Sea $\mathcal{A}(x)$ una *fórmula elemental* en las variables $x = (x_1, \ldots, x_n)$, que en un conjunto $E \subseteq \mathbb{R}^n$ *no plantea ningún problema:*
> cuando $x \in E$ entonces cada vez que en $\mathcal{A}(x)$ hay un cociente el denominador es no nulo, cada vez que hay un logaritmo el logaritmando es estrictamente positivo, cada vez que hay una raíz de índice par el radicando es no negativo, etc.
> Entonces $x \mapsto \mathcal{A}(x)$ define una función $E \to \mathbb{R}$ continua respecto de $\|\cdot\|_\infty$.
> Lo mismo vale para una función vectorial $E \to \mathbb{R}^k$ cada una de cuyas k componentes se define por una fórmula elemental sin problemas cuando $x \in E$.

Aviso. Aparte de las elementales, existen muchas más funciones continuas.

Importante. La condición 2. del teorema 48 dice que **una función continua es una máquina de generar abiertos por medio de desigualdades estrictas.** Concretamente, si $f : X \to \mathbb{R}$ es continua entonces $\{x : f(x) \neq 0\}$, $\{x : f(x) > a\}$, $\{x : f(x) < b\}$ y $\{x : a < f(x) < b\}$ son abiertos de X. Las desigualdades $f(x) < g(x)$ se convierten en $(g - f)(x) > 0$ y se aplica lo dicho.
Análogamente, la condición 3. del teorema 48 dice que **una función continua es una máquina de generar cerrados por medio de ecuaciones o desigualdades no estrictas.** De este modo, si $f : X \to \mathbb{R}$ es continua entonces $\{x : f(x) = 0\}$, $\{x : f(x) \geq a\}$, $\{x : f(x) \leq b\}$ y $\{x : a \leq f(x) \leq b\}$ son cerrados de X. Las desigualdades $f(x) \leq g(x)$ se convierten en $(g - f)(x) \geq 0$ y se aplica lo dicho.

Combinando eso con lo que hemos dicho sobre la continuidad de fórmulas elementales, ahora podemos decir de manera inmediata que un conjunto como el siguiente

$$\left\{ (x,y,z) \in \mathbb{R}^3 \, : \, \operatorname{sen} x + e^{3z} > 7 \; , \; 5y < \frac{x^2 + 8xyz^4 - 10y^3}{x^2 + y^4 + e^z} < 3 \right\} ,$$

es intersección de dos abiertos y por lo tanto es abierto, sin tener que pensarlo mucho.
De los siguientes subconjuntos de $\mathbb{R}^2$

$$C \;=\; \{\, (x,y) : |x{+}y| \leq 4 \,\} \;\; , \;\; E \;=\; \{\, (x,y) : |x{+}y| \leq 4 \,,\, xy < 45 \,\} \;\; , \;\; U \;=\; \{\, (x,y) : xy < 45 \,\} \,,$$

podemos decir de inmediato que C es cerrado y U es abierto; mientras que $E = C \cap U$ es, a la vez, abierto relativo en C y cerrado relativo en U.

Definición 50. *Un subconjunto $E \subseteq X$ es* **denso en** X *si* $\overline{E} = X$.

Por ejemplo, los conjuntos $\mathbb{Q}$ y $\mathbb{R} \setminus \mathbb{Q}$ son ambos densos en la recta real (con la distancia usual).

Proposición 51. (Principio de prolongación de las identidades). *Sean $f, g : (X,d) \to (X', d')$ aplicaciones continuas y $E \subseteq X$ denso en X. Si $f|_E \equiv g|_E$ entonces $f \equiv g$.*

La prueba se deja como ejercicio. Este resultado se utiliza, por ejemplo, en el documento donde estudiamos la regla del paralelogramo.

1.7 Compacidad

Definiciones 52. *Las* **subsucesiones** *de una sucesión $\{x_n\}_{n=1}^\infty$ son el resultado de tomar una sucesión estrictamente creciente de enteros positivos $n_1 < n_2 < n_3 < \cdots$ y quedarse sólo con los correspondientes términos $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$. Obsérvese que $n_k \geq k$ para todo k.*
Dado un subconjunto $K \subseteq X$ los **recubrimientos** *de K son las familias $(A_i)_{i \in I}$ de subconjuntos de X tales que $K \subseteq \bigcup_{i \in I} A_i$.*

Dado un espacio métrico (X,d), es fácil ver que si una sucesión $S=\{x_n\}\subset X$ converge a $x_0\in X$ entonces todas las subsucesiones de S convergen también a x_0.

Teorema-definición 53. *Sea (X,d) un espacio métrico. Dado un subconjunto $K\subset X$, las dos propiedades siguientes son equivalentes:*

> **Propiedad de sucesiones, o de Bolzano-Weierstrass:** *Toda sucesión $\{x_n\}\subset K$ tiene una subsucesión $\{x_{n_k}\}$ convergente a algún punto de K.*
>
> **Propiedad de recubrimiento, o de Heine-Borel:** *Todo recubrimiento $(U_i)_{i\in I}$ de K por abiertos de X tiene una subfamilia finita $U_{i_1},\dots,U_{i_N}$ que también recubre K, es decir $K\subseteq U_{i_1}\cup\dots\cup U_{i_N}$.*

Un **subconjunto compacto de X** *es cualquier $K\subseteq X$ que tiene estas dos propiedades equivalentes.*
Un **espacio métrico compacto** *es cualquier espacio métrico (X,d) que es subconjunto compacto de sí mismo.*

La equivalencia enunciada en este teorema se demuestra en otro documento.

Recordemos que $d_K=d|_{K\times K}$ es la distancia inducida por d en K.

Proposición 54. *K es subconjunto compacto de X si y sólo si (K,d_k) es un espacio compacto.*

Demostración. Supongamos que K es subconjunto compacto de X. Sea $(V_i)_{i\in I}$ un recubrimiento de K por abiertos relativos. Para cada $i\in I$ hay un abierto U_i de X tal que $V_i=K\cap U_i$. Supuesto elegido un U_i para cada i (aquí estamos utilizando el axioma de elección), se forma una familia $(U_i)_{i\in I}$ de abiertos de X tal que

$$K\ =\ \bigcup_{i\in I}V_i\ =\ \bigcup_{i\in I}(K\cap U_i)\ =\ K\cap\bigcup_{i\in I}U_i\ ,$$

lo cual implica $K\subseteq\bigcup_{i\in I}U_i$, luego $(U_i)_{\imath\in I}$ es un recubrimiento de K por abiertos de X. Como estamos suponiendo que K es un subconjunto compacto de X, existe una subfamilia finita $U_{i_1},\dots,U_{i_N}$ tal que $K\subseteq U_{i_1}\cup\dots\cup U_{i_N}$. Entonces

$$K\ =\ K\cap\left(U_{i_1}\cup\dots\cup U_{i_N}\right)\ =\ \bigcup_{j=1}^{N}(k\cap U_{i_j})\ =\ \bigcup_{j=1}^{N}V_{i_j}\ .$$

de esta manera, de cualquier recubrimiento $(V_i)_{i\in I}$ de K por abiertos relativos hemos extraido una subfamilia finita que también recubre K. Esto significa que (K,d_K) es un espacio compacto. Queda demostrado que si K es subconjunto compacto entonces es un espacio compacto.
La demostración del recíproco (que si K es un espacio compacto entonces es un subconjunto compacto de X) es casi igual y queda como ejercicio. □

Una consecuencia del teorema 53 es que la clase de los abiertos en X, por sí sola, ya determina qué subconjuntos son compactos y cuáles no.

Teorema 55. *Para que un subconjunto de un espacio métrico $K\subseteq X$ sea compacto es necesario que sea cerrado y acotado.*
Si K es compacto, entonces todo cerrado contenido en K es también compacto.

Demostración. Sean $K\subseteq X$ compacto y $\{x_n\}\subset K$ sucesión con $\{x_n\}\to x\in X$. Existe una subsucesión $\{x_{n_k}\}$ convergente a un punto $y\in K$. Como esta subsucesión también converge a x, tenemos $x=y\in K$. Es decir que K es cerrado para el paso al límite, luego es cerrado.
La familia de bolas $\big(B(x,1)\big)_{x\in K}$ es un recubrimiento de K por abiertos de X. Tomamos un subrecubrimiento finito $K\subseteq B(x_1,1)\cup\dots\cup B(x_N,1)$ y haciendo

$$r\ =\ 1+\max\left\{\,d(x_1,x_2)\ \dots,d(x_1,x_N)\,\right\}\ ,$$

se llega a $K\subseteq B(x_1,r)$, luego K es acotado.
Se deja como ejercicio la demostración de que todo cerrado contenido en K es compacto. □

Teorema 56. *En* $(\mathbb{R}^n, \|\cdot\|_\infty)$ *todo subconjunto cerrado y acotado es compacto.*
Si $(\mathbb{V}, \|\cdot\|)$ *es un espacio normado de dimensión infinita, entones la bola* $\overline{B}(\mathbf{0},1)$ *no es compacta.*

Demostración. Sea $K \subset \mathbb{R}^n$, cerrado y acotado, y $\{x_j\}_{j=1}^\infty \subset K$. Para $i = 1, \ldots, n$ la sucesión $\{x_j^i\}_{j=1}^\infty$ de las i-ésimas coordenadas de los x_j es acotada en $\mathbb{R}$. La primera sucesión $\{x_j^1\}$ tiene una subsucesión $\{x_{j_k}^1\}$ convergente en $\mathbb{R}$ a un número x^1. Dada ahora la sucesión $\{x_j^2\}$ de las segundas coordenadas, consideramos la subsucesión $\{x_{j_k}^2\}$ correspondiente a la misma sucesión de índices $j_1 < j_2 < j_3 < \cdots$ utilizada para definir $\{x_{j_k}^1\}$. Existe una sub-subsucesión $\left\{ x_{j_{k_h}}^2 \right\}_{h=1}^\infty$ convergente en $\mathbb{R}$ a un número x^2 y la correspondiente sub-subsucesión $\left\{ x_{j_{k_h}}^1 \right\}$ de las primeras coordenadas todavía converge a x^1. Siguiendo así, después de n pasos tenemos una sucesión creciente de índices $m_1 < m_2 < m_3 < \cdots$ y números $x^1, \ldots, x^n \in \mathbb{R}$ tales que para cada $i = 1, \ldots, n$ la sucesión de números $\left\{ x_{m_s}^i \right\}_{s=1}^\infty$ converge a x^i. Entonces la sucesión de vectores $\left\{ x_{m_s} \right\}_{s=1}^\infty$ converge al punto $x = (x^1, \ldots, x^n)$ en $(\mathbb{R}^n, \|\cdot\|_\infty)$ y además $x \in K$ por ser K cerrado.
Esto prueba que K tiene la propiedad de las sucesiones, luego es compacto.
El segundo enunciado se demuestra en otro documento. □

Teorema 57. *Si* $f : X \to Y$ *es continua y* $K \subseteq X$ *es compacto, entonces* $f(K) \subseteq Y$ *es compacto.*
Si $K \subseteq X$ *es compacto y* $f : K \to \mathbb{R}$ *es continua (para la distancia usual en* $\mathbb{R}$*), entonces* f *alcanza su máximo y su mínimo. Es decir que existen* $x_1, x_2 \in K$ *tales que* $f(x_1) \leq f(x) \leq f(x_2)$ *para todo* $x \in K$*.*

Demostración. La primera parte es por la caracterización de los compactos por sucesiones y la propiedad $f\big(\lim x_n\big) = \lim f(x_n)$ de las aplicaciones continuas.
Para la segunda parte definimos el conjunto de números $A = f(K) \subset \mathbb{R}$, del que sabemos que es cerrado y acotado. Al ser acotado existen $\inf A, \sup A$ y son valores finitos. Además son límites de sucesiones en A y, como A también es cerrado, resulta $\inf A, \sup A \in A$. Es decir que existen $x_1, x_2 \in K$ tales que $f(x_1) = \inf A$ y $f(x_2) = \sup A$, luego $f(x_1) \leq f(x) \leq f(x_2)$ para todo $x \in K$. □

1.8 Normas y distancias equivalentes

Teorema-definición 58. *Dos normas* $\|\cdot\|, \|\cdot\|'$ *en un espacio vectorial* $\mathbb{V}$ *son* **equivalentes** *si dan lugar a los mismos conjuntos abiertos en* $\mathbb{V}$*.*
Esto resulta ser equivalente a la existencia de constantes $c, C > 0$ *tales que*

$$c\,\|v\| \leq \|v\|' \leq C\,\|v\| \quad \text{para todo } \; v \in \mathbb{V}\,. \tag{14}$$

Por lo tanto, dos normas equivalentes también definen los mismos conjuntos acotados.

Demostración. Veamos que las dos condiciones siguientes son equivalentes:

1. Todo abierto para $\|\cdot\|'$ es un abierto para $\|\cdot\|$.

2. Existe una $C > 0$ tal que $\|v\|' \leq C\,\|v\|$ para todo $v \in \mathbb{V}$.

Sea $\mathrm{id} : \mathbb{V} \to \mathbb{V}$ la identidad $v \longmapsto v$. Expresando la condición 1. así:
Si U es un abierto en $(\mathbb{V}, \|\cdot\|')$ entonces $\mathrm{id}^{-1}(U)$ es abierto en $(\mathbb{V}, \|\cdot\|)$,
vemos que equivale a que id sea continua vista como $(\mathbb{V}, \|\cdot\|) \to (\mathbb{V}, \|\cdot\|')$ y, por lo tanto, lineal acotada que es lo que expresa la condición 2.

Intercambiando $\|\cdot\|$ con $\|\cdot\|'$, sabemos también que la existencia de una $C' > 0$ tal que $\|v\| \leq C'\,\|v\|'$ para todo $v \in \mathbb{V}$ equivale a que todo abierto para $\|\cdot\|$ sea un abierto para $\|\cdot\|'$.
Poniendo $c = 1/C'$, esto último equivale a su vez a tener $c\,\|v\| \leq \|v\|'$ para todo $v \in \mathbb{V}$. □

Teorema-definición 59. *En $\mathbb{R}^n$ todas las normas son equivalentes entre sí.*
Llamamos **abiertos estándar** *de $\mathbb{R}^n$ a los definidos por cualquier norma.*

Demostración. Basta con probar que toda norma es equivalente a la $\|\cdot\|_\infty$. Denotaremos por $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base estándar de $\mathbb{R}^n$. Para cualquier norma $\|\cdot\|$ se tiene:

$$\|x\| \;=\; \|x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n\| \;\leq\; |x_1|\,\|\mathbf{e}_1\| + \cdots + |x_n|\,\|\mathbf{e}_n\| \;\leq\; C\,\|x\|_\infty \;,$$

donde hemos tomado $C = \|\mathbf{e}_1\| + \cdots + \|\mathbf{e}_n\| > 0$. Esto implica que para cualesquiera $x, y \in \mathbb{R}^n$ es:

$$\big|\, \|x\| - \|y\| \,\big| \;\leq\; \|x - y\| \;\leq\; C\,\|x - y\|_\infty \;,$$

luego si escribimos $f(x) = \|x\|$ para todo $x \in \mathbb{R}^n$ entonces:

$$\varepsilon \;>\; 0 \;\Longrightarrow\; f\big(B(x, \varepsilon/C)\big) \;\subseteq\; \big(\, f(x) - \varepsilon\,,\, f(x) + \varepsilon\,\big) \;,$$

donde B denota bolas abiertas respecto de $\|\cdot\|_\infty$. Esto nos dice que $f \equiv \|\cdot\|$ es una función escalar continua respecto de $\|\cdot\|_\infty$.
La cáscara $K = \{\omega \,:\, \|\omega\|_\infty = 1\}$ es un conjunto cerrado y acotado en $(\mathbb{R}^n, \|\cdot\|_\infty)$, luego compacto por el teorema 55. Por el teorema 57 existe un $\omega_0 \in K$ tal que $c = f(\omega_0)$ es el mínimo de f en K. Entonces c cumple las dos condiciones siguientes:

$$c \;>\; 0 \quad , \quad \|\omega\|_\infty \;=\; 1 \;\Longrightarrow\; \|\omega\| \;\geq\; c\,.$$

Por el truco habitual de la descomposición polar, obtenemos $\|x\| \geq c\,\|x\|_\infty$ para todo $x \in \mathbb{R}^n$. En definitiva $c\,\|x\|_\infty \leq \|x\| \leq C\,\|x\|_\infty$ para todo $x \in \mathbb{R}^n$, como se quería demostrar. □

Una consecuencia del teorema 59 es que las funciones y aplicaciones $\mathbb{R}^n \supseteq E \xrightarrow{f} \mathbb{R}^m$ que hemos llamado *elementales* en el apartado 1.6 son continuas respecto de cualesquiera normas que pongamos en $\mathbb{R}^n$ y en $\mathbb{R}^m$.
En particular, toda aplicación lineal $\mathbb{R}^n \to \mathbb{R}^m$ es acotada respecto de cualesquiera normas que pongamos en $\mathbb{R}^n$ y en $\mathbb{R}^m$. Dicho de otra manera, una vez elegidas normas en $\mathbb{R}^n$ y en $\mathbb{R}^m$ cada matriz real $m \times n$ tiene una norma finita y bien definida respecto de ellas.

Definición 60. *Dos funciones distancia en un mismo conjunto X son* **distancias equivalentes** *si dan lugar a los mismos abiertos.*

Ahora sabemos que dos distancias equivalentes producen:

Los mismos cerrados.
Las mismas nociones de interior, exterior, cierre y frontera de cada subconjunto.
La misma clase de sucesiones convergentes y, para cada una de éstas, el mismo punto límite.
La misma clase de subconjuntos compactos.

De igual modo, si en X tenemos dos distancias equivalentes y en X' otras dos distancias equivalentes, entonces la clase de las aplicaciones continuas $X \to X'$ es la misma para los cuatro pares posibles de distancias (una en X y otra en X').

Hemos visto que dos normas equivalentes definen la misma clase de conjuntos acotados. Esto ya no es verdad para distancias equivalentes, como muestra el siguiente ejempo en $X = \mathbb{R}$:

$$d(x,y) \;=\; |x - y| \quad , \quad d'(x,y) \;=\; \min\big\{\, 1\,,\, |x - y|\,\big\}\,.$$

Corolario importante:

La clase de los abiertos no determina, ella sola, la clase de los subconjuntos acotados.

1.9 Conexión

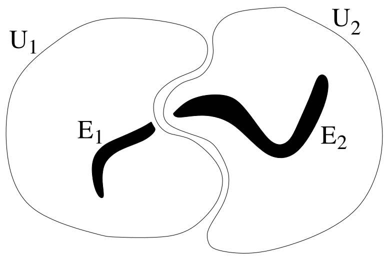
Definiciones 61. *Un* **camino** *en un espacio métrico* (X, d) *es cualquier aplicación continua* $\alpha : I \to X$ *cuyo dominio* I *es un intervalo de la recta real.*
Un subconjunto no vacío $E \subseteq X$ *es* **conexo por caminos** *o* **conexo por arcos** *si cada par de puntos* $p, q \in E$ *se puede* **unir por un camino en E:** *existe un camino* $\alpha(t) : [0, 1] \to E$ *contenido en* E*, que empieza en* p *y termina en* q*:* $\alpha(0) = p$, $\alpha(1) = q$.

Proposición 62. *1. Los subconjuntos conexos por caminos de la recta real son los intervalos y los conjuntos de un elemento* $\{a\}$*.*
2. La imagen de un conexo por caminos por una aplicación continua es conexa por caminos.
3. La unión no disjunta de dos conexos por caminos es conexa por caminos.

El siguiente teorema viene a decir que un conjunto conexo por caminos "es de una sola pieza".

Teorema 63. *Un conjunto conexo por caminos nunca puede ser unión disjunta de dos abiertos relativos no vacíos.*

De hecho, si E es una unión disjunta $E_1 \cup E_2$, con $E_1 = E \cap U_1$ y $E_2 = E \cap U_2$ abiertos relativos disjuntos, entonces ningún punto de E_1 puede unirse a ningún punto de E_2 por caminos contenidos en E.



Por supuesto, un conexo por caminos sí puede ser unión disjunta de un abierto relativo con un cerrado relativo: $\mathbb{R}$ es conexo por caminos y es la unión disjunta de $(-\infty, 0)$ con $[0, +\infty)$.

Corolario 64. *Si* E *es conexo por caminos, entonces los únicos subconjunto de* E *que son a la vez abiertos relativos y cerrados relativos son* $\varnothing$ *y* E*.*

Dado un punto $p \in E$, el más grande subconjunto de E conexo por caminos y conteniendo a p es el conjunto:

$$\big\{\, q \in E \; : \; p \text{ se une con } q \text{ por un camino contenido en } E \,\big\} \; .$$

Estos conjuntos se llaman **componentes conexas por caminos** de E. La propiedad 3. de la proposición 62 hace que estas componentes formen una **partición:** su unión es todo E y dos cualesquiera que sean distintas son disjuntas.

Proposición 65. *Si* U *es un abierto de* $\mathbb{R}^n$ *entonces sus componentes conexas por caminos son también abiertos, y a lo más hay una cantidad numerable de ellas. De este modo, todo abierto de* $\mathbb{R}^n$ *tiene una única partición en abiertos conexos por caminos, en cantidad finita o numerable.*

Aviso. Es importante no confundir las palabras "conexo" y "convexo". Todo subconjunto convexo $C \subseteq \mathbb{R}^n$ es conexo por caminos, pues si $p, q \in C$ entonces

$$\alpha : [0, 1] \longrightarrow C \quad , \quad \alpha(t) = (1 - t)p + tq \; ,$$

es un camino (rectilíneo) en C que une p con q. Pero la mayoría de los conexos por caminos no son convexos. Un ejemplo muy sencillo es un abierto del plano con la forma de la letra C.