Apuntes de Complementos Matemáticos de la Ingeniería Industrial

Tema 1. Motivación y Fundamentos

Versión 3
26 de julio de 2020

Departamento de Matemática Aplicada I. UNED

Índice general

1. Motivación y Fundamentos

1.2. Espacios vectoriales con producto escalar y espacios normados

Producto escalar

Partimos de un espacio vectorial sobre $\mathbb{R}$ que llamamos V. Puede tener dimensión finita o ser de dimensión infinita. Sus elementos en general los vamos a designar como $\mathbf{v}$ o como $\bar{v}$.

Definición 1. *Un* ***producto escalar*** *definido sobre un espacio vectorial V es una aplicación $\langle\cdot,\cdot\rangle : V \times V \longrightarrow \mathbb{R}$ que cumple las siguientes propiedades:*

1. $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ *si* $\mathbf{v} \neq 0$*,* $\langle \mathbf{0}, \mathbf{0} \rangle = \mathbf{0}$.
2. $\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$ *para todo* $\mathbf{v}, \mathbf{u} \in V$.
3. $\langle \lambda\mathbf{v} + \mu\mathbf{u}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle + \mu \langle \mathbf{u}, \mathbf{w} \rangle$ *para todo* $\lambda, \mu \in \mathbb{R}$*,* $\mathbf{v}, \mathbf{u}, \mathbf{w} \in V$.

Aplicando lo que conocemos de las formas, podemos decir que un producto escalar es una forma bilineal no degenerada (definida positiva). El producto escalar también se puede llamar producto interior. En $\mathbb{R}^n$ es habitual denotar el producto escalar de dos vectores $\mathbf{v}, \mathbf{u}$ como $\mathbf{v} \cdot \mathbf{u}$.

Obsérvese que la definición anterior es válida para espacios tanto de dimensión finita como infinita (en general, espacios de funciones).

Vamos a especificar una notación que utilicemos a lo largo del texto. La base canónica de $\mathbb{R}^n$ es la formada por los vectores $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, donde cada $\mathbf{e}_i$ es el vector cuyas coordenadas son 0 excepto la que está en el lugar $i-$ésimo. En $\mathbb{R}^3$ la base canónica se suele denotar también como $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Vamos a usar ambas notaciones.

Ejemplo 1. Consideramos el espacio $\mathbb{R}^n$ y la base canónica del mismo, es decir, la formada por los vectores$\{(1,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,\ldots,0,1)\}$. Entonces la aplicación

$$\langle \mathbf{v}, \mathbf{u} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i,$$

para elementos $\mathbf{v} = (x_1, \ldots, x_n)$, $\mathbf{u} = (y_1, \ldots, y_n)$ de $\mathbb{R}^n$ es un producto escalar.

Comprobamos que se cumplen las siguientes propiedades:

1. $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ si $\mathbf{v} \neq (0, \ldots, 0)$ y $\langle (0, \ldots, 0), (0, \ldots, 0) \rangle = 0$:

$$\langle \mathbf{v}, \mathbf{v} \rangle = x_1 x_1 + \cdots + x_n x_n = x_1^2 + \cdots + x_n^2 > 0 \text{ si } x \neq (0, \ldots, 0),$$
$$\langle (0, \ldots, 0), (0, \ldots, 0) \rangle = 0 \cdot 0 + \cdots + 0 \cdot 0 = 0.$$

2. $\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$ para cualquier $\mathbf{v}, \mathbf{u} \in \mathbb{R}^n$:

$$\langle \mathbf{v}, \mathbf{u} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = y_1 x_1 + y_2 x_2 + \cdots + y_n x_n = \langle \mathbf{u}, \mathbf{v} \rangle.$$

3. $\langle \lambda \mathbf{v} + \mu \mathbf{u}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle + \mu \langle \mathbf{u}, \mathbf{w} \rangle$ para todo $\lambda, \mu \in \mathbb{R}$, $\mathbf{v}, \mathbf{u}, \mathbf{w} \in \mathbb{R}^n$, denotando $\mathbf{w} = (z_1, \ldots, z_n)$:

$$\begin{aligned} \langle \lambda \mathbf{v} + \mu \mathbf{u}, \mathbf{w} \rangle &= \sum_{i=1}^{n} (\lambda x_i + \mu y_i) z_i = \sum_{i=1}^{n} (\lambda x_i z_i + \mu y_i z_i) \\ &= \sum_{i=1}^{n} \lambda x_i z_i + \sum_{i=1}^{n} \mu y_i z_i = \lambda \sum_{i=1}^{n} x_i z_i + \mu \sum_{i=1}^{n} y_i z_i \\ &= \lambda \langle \mathbf{v}, \mathbf{w} \rangle + \mu \langle \mathbf{u}, \mathbf{w} \rangle. \end{aligned}$$

☻

Ejemplo 2. Consideramos el espacio $C[0,1]$ de las funciones de variable real con valores en $\mathbb{R}$ continuas en el intervalo $[0,1]$. Vamos a comprobar que

$$\langle f, g \rangle = \int_0^1 f(x) g(x) \, dx,$$

donde f y g son funciones reales continuas en $[0,1]$, es un producto escalar.

Hay que comprobar que se cumplen las siguientes propiedades:

1. $\langle f, f \rangle > 0$ si $f \neq 0$, $\langle 0, 0 \rangle = 0$, para cualquier $f \in C[0,1]$.
2. $\langle f, g \rangle = \langle g, f \rangle$, para cualquier $f, g \in C[0,1]$.
3. $\langle \lambda f + \mu g, h \rangle = \lambda \langle f, h \rangle + \mu \langle g, h \rangle$, para cualquier $\lambda, \mu \in \mathbb{R}$, y cualquier $f, g, h \in C[0,1]$.

Vamos a hacerlo.

1. Sea $f \neq 0$. Se tiene:

$$\langle f, f \rangle = \int_0^1 f(x) f(x) \, dx = \int_0^1 f^2(x) \, dx > 0$$

ya que la función $F(x) = f^2(x)$ verifica que $F \neq 0$, $F(x) \geq 0$ para $x \in [0,1]$ y $F \in C[0,1]$, por lo que la integral definida en $[0,1]$ de F representa entonces el área limitada por la gráfica de F y el eje x en el intervalo $[0,1]$.

Obsérvese que, en particular

$$\int_0^1 f(x) f(x)\, dx = 0$$

si y sólo si $f^2(x)$ es idénticamente nula y, por tanto, si y sólo si f es la función nula en $[0,1]$.

Además,

$$\langle 0,0\rangle = \int_0^1 0 \cdot 0 dx = \int_0^1 0 dx = 0.$$

2. Comprobamos la segunda propiedad:

$$\langle f,g\rangle = \int_0^1 f(x) g(x)\, dx = \int_0^1 g(x) f(x)\, dx = \langle g,f\rangle .$$

3. Nos queda la última propiedad:

$$\begin{aligned}\langle \lambda f + \mu g, h\rangle &= \int_0^1 (\lambda f + \mu g)(x) h(x) dx \\ &= \int_0^1 (\lambda f(x)h(x) + \mu g(x)h(x))\, dx \\ &= \int_0^1 \lambda f(x)h(x)dx + \int_0^1 \mu g(x)h(x)dx \\ &= \lambda \int_0^1 f(x)h(x)dx + \mu \int_0^1 g(x)h(x)dx \\ &= \lambda \langle f,h\rangle + \mu \langle g,h\rangle .\end{aligned}$$

Por tanto, esta aplicación sí define un producto escalar en las funciones continuas en el intervalo $[0,1]$.

Este espacio vectorial no tiene dimensión finita.

☻

Norma y distancia definidas a partir de un producto escalar

En $\mathbb{R}^n$, a partir del producto escalar, definimos la norma de un vector y, a partir de la norma, definimos la distancia entre dos puntos. También definimos ángulo que forman dos vectores y ortogonalidad (en espacios vectoriales de dimensión 2 y 3). Estos conceptos son más generales, y volveremos a ellos más adelante, pero antes vamos a recordarlos .

Recordamos que dado $\mathbf{u} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, definimos su norma o módulo como:

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = \sqrt{\sum_{i=1}^{n} (x_i)^2}.$$

Es decir, es la raíz cuadrada de su producto escalar consigo mismo. Intuitivamente, es la longitud del vector $\mathbf{u}$ en $\mathbb{R}^2$ o en $\mathbb{R}^3$ y el valor absoluto de un número si estamos en $\mathbb{R}$. Se llama **norma euclídea**.

Ejemplo 3. Consideramos el vector $\mathbf{u} = (3, 0, -4) \in \mathbb{R}^n$. Vamos a calcular su norma euclídea, que es

$$\begin{aligned} \|\mathbf{u}\| &= \|(3,0,-4)\| = ((3,0,-4) \cdot (3,0,-4))^{1/2} \\ &= \left(3^2 + 0^2 + (-4)^2\right)^{1/2} = \sqrt{9+16} = \sqrt{25} = 5. \end{aligned}$$

☻

Esta definición no sólo vale para vectores de $\mathbb{R}^n$, es válida para cualquier espacio vectorial con producto escalar, simplemente definiendo la norma como

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}.$$

Observamos que la norma así definida cumple las siguientes propiedades, para $\mathbf{u}, \mathbf{v} \in V$, $\lambda \in \mathbb{R}$:

- $\|\mathbf{u}\| \geq 0$ y $\|\mathbf{u}\| = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.
- $\|\lambda \mathbf{u}\| = |\lambda| \, \|\mathbf{u}\|$.
- $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ y $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$ si y sólo si $\mathbf{u}, \mathbf{v}$ son linealmente dependientes.

La primera propiedad indica que una norma (longitud) de un vector es siempre positiva o 0 y sólo es 0 si es la norma del vector nulo. La segunda implica que la norma de un vector proporcional a otro vector es la norma del vector multiplicada por la constante de proporcionalidad en valor absoluto. Y la tercera afirma que la norma de la suma de dos vectores va a ser siempre menor o igual que la suma de las normas de los vectores y que sólo va a ser igual si los vectores están alineados.

La demostración de estas propiedades es sencilla, y lo dejamos como ejercicio.

Ejemplo 4. Consideramos el espacio $C[0,1]$ y el producto escalar

$$\langle f,g\rangle = \int_0^1 f(x)\,g(x)\,dx.$$

Vamos a determinar la norma de $f(x) = \operatorname{sen}(\pi x)$. Es

$$\begin{aligned}
\|f\| &= \langle \operatorname{sen}(\pi x), \operatorname{sen}(\pi x)\rangle^{1/2} = \left(\int_0^1 \operatorname{sen}^2(\pi x)dx\right)^{1/2} \\
&= \left(\int_0^1 \frac{1-\cos 2\pi x}{2}dx\right)^{1/2} = \left(\int_0^1 \frac{1}{2}dx - \frac{1}{2}\int_0^1 \cos 2\pi x dx\right)^{1/2} \\
&= \left(\frac{1}{2}x\Big|_0^1 - \frac{1}{4\pi}\operatorname{sen}(2\pi x)\Big|_0^1\right)^{1/2} = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}.
\end{aligned}$$

porque $\operatorname{sen}^2 x = \dfrac{1-\cos(2x)}{2}$. ☻

Si consideramos que los elementos de $\mathbb{R}^n$ son puntos, la **distancia euclídea** entre dos puntos P y Q es la norma del vector PQ que une ambos puntos:

$$d(P,Q) = \left\|\overrightarrow{PQ}\right\|.$$

Ejemplo 5. La distancia entre los puntos $(1,0,1)$ y $(2,1,1)$ de $\mathbb{R}^3$ es la norma del vector

$$\mathbf{u} = (2,1,1) - (1,0,1) = (1,1,0).$$

Su norma es

$$\begin{aligned}
\|\mathbf{u}\| &= \|(1,1,0)\| = ((1,1,0)\cdot(1,1,0))^{1/2} \\
&= \sqrt{1^2+1^2} = \sqrt{2}.
\end{aligned}$$

Estos dos puntos, junto con el vector que los une, se representan en la siguiente figura

☻

Hemos repasado el producto escalar porque permite definir ángulo entre elementos de un espacio vectorial y ortogonalidad en $\mathbb{R}^2$ y $\mathbb{R}^3$. Estos conceptos coinciden con la idea intuitiva que tenemos de los mismos. Pero no es tan sencillo en espacios como $C[0,1]$. El **ángulo** que forman dos vectores no nulos $\mathbf{u},\mathbf{v} \in V$ es el número $\theta \in [0,\pi]$ cuyo coseno es

$$\cos\theta = \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\cdot\|\mathbf{v}\|}.$$

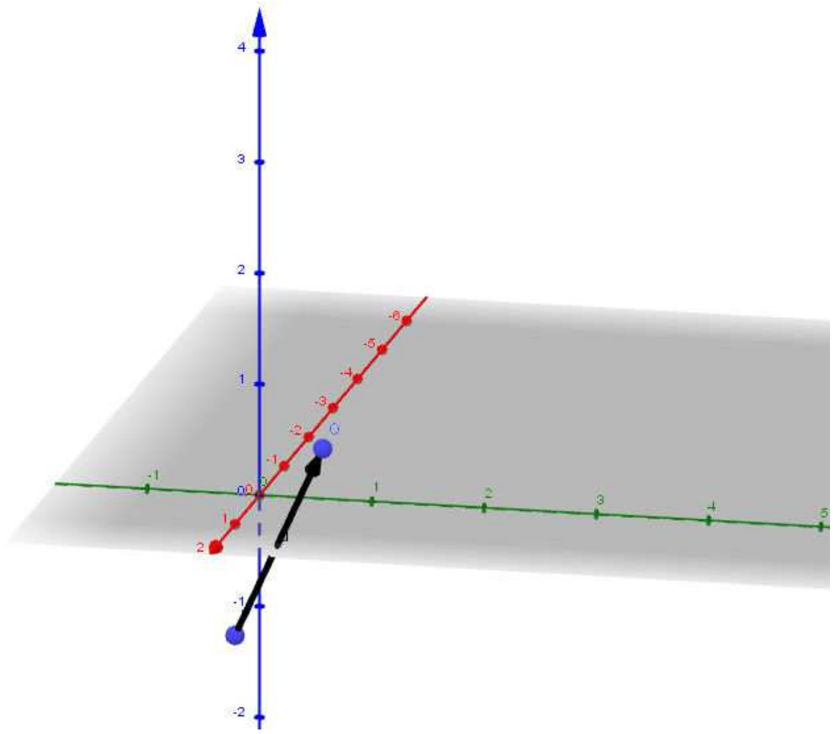


Figura 1: Distancia entre dos puntos

De esto, la expresión

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cos \theta$$

es una definición geométrica del producto escalar, y es independiente de la base y del sistema de coordenadas elegidos.

De esta interpretación geométrica del producto escalar se deduce otra, relacionada en este caso con la proyección de un vector $\mathbf{u}$ sobre otro vector $\mathbf{v}$, que es $proy\mathbf{u}_{\mathbf{v}} = \|\mathbf{u}\| \cos \theta$. Entonces el producto escalar es

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{v}\| \, proy\mathbf{u}_{\mathbf{v}}$$

Ya conocemos cuándo dos vectores son perpendiculares u ortogonales en $\mathbb{R}^2, \mathbb{R}^3$. Esta definición se extiende a cualquier espacio vectorial.

Veamos ahora dos definiciones de propiedades geométricas de vectores, en espacios vectoriales con producto escalar (también llamados prehilbertianos).

Definición 2. *Se dice que dos vectores no nulos* $\mathbf{u}, \mathbf{v} \in V$ *de un espacio vectorial son* ***vectores ortogonales*** *si* $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Se dice que dos vectores no nulos $\mathbf{u}, \mathbf{v} \in V$ *de un espacio vectorial son* ***vectores paralelos*** *si* $\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \cdot \|\mathbf{v}\|$.

Ejemplo 6. Consideramos en el espacio de las funciones continuas en el intervalo $[0,1]$, $C[0,1]$, el producto escalar

$$< f, g >= \int_0^1 f(x)\, g(x)\, dx.$$

Vamos a estudiar si las funciones $f(x) = \cos(\pi x)$ y $g(x) = \operatorname{sen}(\pi x)$ son ortogonales.

El producto escalar de $f(x) = \cos(\pi x)$ y $g(x) = \operatorname{sen}(\pi x)$ es

$$\begin{aligned} < f, g >&= \int_0^1 \cos(\pi x)\, \operatorname{sen}(\pi x) dx = \int_0^1 \frac{1}{2} \operatorname{sen}(2\pi x) dx \\ &= \frac{1}{4\pi} \left(-\cos 2\pi x\right)\Big|_0^1 = \frac{1}{4\pi} \left(-\cos 2\pi + \cos 0\right) = 0. \end{aligned}$$

Luego sí son ortogonales. ☻

Proposición 3. *Dada una familia* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *de vectores ortogonales dos a dos se cumple el* ***Teorema de Pitágoras***

$$\|\mathbf{v}_1 + \cdots + \mathbf{v}_n\|^2 = \|\mathbf{v}_1\|^2 + \cdots + \|\mathbf{v}_n\|^2.$$

Esta definición también se extiende a conjuntos de forma natural. Dos subconjuntos A, B de un espacio vectorial con producto escalar son **subespacios ortogonales** si $\langle x, y \rangle = 0$ para todo $x \in A, y \in B$.

A lo largo de esta asignatura vamos a trabajar frecuentemente con subespacios ortogonales (por ejemplo: plano tangente y recta normal a una superficie en un punto).

Señalamos que el conjunto de vectores ortogonales a un subespacio vectorial de un espacio vectorial V es un subespacio vectorial de V.

Ejemplo 7. Consideramos el espacio $\mathbb{R}^4$. Sean $M = \{(x, y, z, t) \in \mathbb{R}^4 : 2x - y + z = 0\}$ y N la recta que pasa por el origen y tiene vector director $(2, -1, 1, 0)$. Vamos a estudiar si estos espacios son ortogonales.

Tenemos que comprobar que el producto escalar de cualquier elemento de M con cualquier elemento de N es 0. Para ello, determinamos la expresión de cualquier elemento de M y N. Si $(x, y, z, t) \in M$, entonces como $y = 2x + z$ para los puntos de M, podemos escribirlo como

$$(\lambda, 2\lambda + \mu, \mu, \gamma), \quad \text{para } \lambda, \mu, \gamma \in \mathbb{R}.$$

Este espacio representa un hiperplano de $\mathbb{R}^4$. Por otro lado, si $(x, y, z, t) \in N$, existe $\alpha \in \mathbb{R}$ tal que:

$$(x, y, z, t) = (2\alpha, -\alpha, \alpha, 0),$$

que es la ecuación de una recta. Calculamos el producto escalar de un elemento de cada espacio:

$$\begin{aligned}(2\alpha, -\alpha, \alpha, 0) \cdot (\lambda, 2\lambda + \mu, \mu, \gamma) &= 2\alpha\lambda - \alpha\,(2\lambda + \mu) + \alpha\mu + 0 \cdot \gamma \\ &= 2\alpha\lambda - 2\alpha\lambda - \alpha\mu + \alpha\mu = 0.\end{aligned}$$

Por eso, sí podemos concluir que son espacios ortogonales. ☻

Vamos a enunciar la siguiente propiedad que se cumple en espacios vectoriales cuando se ha definido un producto escalar.

Proposición 4. *En un espacio vectorial V se cumple, para $\mathbf{u}, \mathbf{v} \in V$, la* ***desigualdad de Cauchy-Schwarz****:*

$$|\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle .$$

Esta desigualdad sólo es igualdad si $\mathbf{u}, \mathbf{v}$ son linealmente dependientes, es decir, si existe $\lambda \in \mathbb{R}$, tal que $\mathbf{u} = \lambda \mathbf{v}$.

Demostración. La demostración de este resultado no es complicada, se basa en calcular el producto escalar de $\lambda \mathbf{u} - \mu \mathbf{v}$ consigo mismo, para $\lambda, \mu \in \mathbb{R}$:

$$\begin{aligned}0 &\leq \langle \lambda\mathbf{u} - \mu\mathbf{v}, \lambda\mathbf{u} - \mu\mathbf{v} \rangle = \langle \lambda\mathbf{u}, \lambda\mathbf{u} - \mu\mathbf{v} \rangle - \langle \mu\mathbf{v}, \lambda\mathbf{u} - \mu\mathbf{v} \rangle \\ &= \lambda \langle \mathbf{u}, \lambda\mathbf{u} - \mu\mathbf{v} \rangle - \mu \langle \mathbf{v}, \lambda\mathbf{u} - \mu\mathbf{v} \rangle \\ &= \lambda^2 \langle \mathbf{u}, \mathbf{u} \rangle - \lambda\mu \langle \mathbf{u}, \mathbf{v} \rangle - \lambda\mu \langle \mathbf{v}, \mathbf{u} \rangle + \mu^2 \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \lambda^2 \langle \mathbf{u}, \mathbf{u} \rangle - 2\lambda\mu \langle \mathbf{u}, \mathbf{v} \rangle + \mu^2 \langle \mathbf{v}, \mathbf{v} \rangle .\end{aligned}$$

Si elegimos $\lambda = \langle \mathbf{u}, \mathbf{v} \rangle$, $\mu = \langle \mathbf{u}, \mathbf{u} \rangle$, tenemos:

$$\begin{aligned}&\lambda^2 \langle \mathbf{u}, \mathbf{u} \rangle - 2\lambda\mu \langle \mathbf{u}, \mathbf{v} \rangle + \mu^2 \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{v} \rangle^2 \langle \mathbf{u}, \mathbf{u} \rangle - 2 \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{v} \rangle^2 + \langle \mathbf{u}, \mathbf{u} \rangle^2 \langle \mathbf{v}, \mathbf{v} \rangle \\ &= - \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{v} \rangle^2 + \langle \mathbf{u}, \mathbf{u} \rangle^2 \langle \mathbf{v}, \mathbf{v} \rangle \geq 0.\end{aligned}$$

Si $\mathbf{u} = 0$, entonces la desigualdad se cumple de forma obvia. Si $\mathbf{u} \neq 0$, entonces $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ y

$$\langle \mathbf{u}, \mathbf{v} \rangle^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle .$$

□

Espacios normados. Norma

Ahora vamos a generalizar la idea de módulo de un vector que hemos visto en el espacio euclídeo $\mathbb{R}^n$, a cualquier espacio vectorial. Hemos definido la norma a partir de un producto escalar, y en el caso de $\mathbb{R}^n$ coincide con la idea intuitiva que tenemos de longitud de un vector. Para mantener esta idea intuitiva, vamos a pedirle algunas cosas, para que coincida con la idea intuitiva de longitud de vector que ya conocemos. Primero pediremos que sea siempre positiva y que sea 0 si y sólo si es la norma del vector nulo. Además, vamos a pedir que la norma de un vector proporcional a otro sea proporcional en la misma proporción. Y también vamos a pedir que la norma de la suma de dos vectores va a ser siempre menor o igual que la suma de las normas de los vectores y sea sólo igual cuando los vectores estén alineados. Estas son las propiedades que cumplía la norma que viene del producto escalar.

Con estas condiciones, podemos definir ya norma.

Definición 5. *Sea V un espacio vectorial sobre $\mathbb{R}$ y sean $\mathbf{u},\mathbf{v} \in V$, $\lambda \in \mathbb{R}$. Una **norma** es una aplicación $\|\cdot\| : V \longrightarrow \mathbb{R}$ que cumple las propiedades siguientes:*

- No degenerada: $\|\mathbf{u}\| \geq 0$ y $\|\mathbf{u}\| = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.
- Homogénea por homotecias: $\|\lambda\mathbf{u}\| = |\lambda| \|\mathbf{u}\|$.
- Desigualdad triangular: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Un espacio vectorial donde se ha definido una norma se llama **espacio normado**. A veces se indica $(V, \|\cdot\|)$.

Queremos destacar que la norma es una aplicación, no es un número.

Vamos a ver algunas propiedades de la norma.

- Para todo $\mathbf{u} \in V$, se tiene
$$\|\mathbf{u}\| \geq 0.$$
- Para $\mathbf{u},\mathbf{v} \in V$, se verifica
$$|\|\mathbf{u}\| - \|\mathbf{v}\|| \leq \|\mathbf{u} \pm \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$
- Para $\mathbf{u}_i \in V$, $\lambda_i \in \mathbb{R}$, $i = 1, 2, \cdots, n$, se cumple
$$\left\| \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \right\| < \sum_{i=1}^{n} |\lambda_i| \|\mathbf{u}_i\| .$$

Conocemos la norma euclídea en $\mathbb{R}^n$, definida para $\mathbf{x} = (x_1, \ldots, x_n)$ como

$$\|\mathbf{x}\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}.$$

En los ejercicios hemos comprobado que cumple las propiedades de una norma. En $\mathbb{R}$ esta norma determina, en cada número, su valor absoluto. En $\mathbb{R}^2$ o en $\mathbb{R}^3$ representa, para cada vector, su longitud.

Podemos definir más normas en $\mathbb{R}^n$, como las normas del máximo $\|\cdot\|_\infty$ (también se denota, a veces, como $\|\cdot\|_0$) y de la suma $\|\cdot\|_1$. Estas normas se definen, para $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, como

$$\begin{aligned} \|(x_1, \ldots, x_n)\|_\infty &= \max_{i=1\ldots n}\{|x_i|\}, \\ \|(x_1, \ldots, x_n)\|_1 &= \sum_{i=1}^{n} |x_i|. \end{aligned}$$

Vamos a comprobar que la norma del máximo es una norma y dejamos como ejercicio comprobar que lo es la norma de la suma. Comenzamos con la propiedad de no degeneración:

$$\|(x_1, \ldots, x_n)\|_\infty = \max_{i=1\ldots n}\{|x_i|\} \geq 0$$

obviamente, porque cada $|x_i| \geq 0$. Además,

$$0 = \|(x_1, \ldots, x_n)\|_\infty = \max_{i=1,\ldots n}\{|x_i|\} \Longleftrightarrow |x_i| = 0 \text{ para todo } i.$$

Es decir, $\|\mathbf{x}\| = 0$ si y sólo si $\mathbf{x} = \mathbf{0}$.

Vamos a comprobar ahora que es homogénea por homotecias:

$$\begin{aligned} \|\lambda\mathbf{x}\|_\infty &= \|\lambda(x_1, \ldots, x_n)\|_\infty = \max_{i=1,\ldots n}\{|\lambda x_i|\} = \max_{i=1,\ldots n}\{|\lambda||x_i|\} \\ &= |\lambda| \max_{i=1,\ldots n}\{|x_i|\} = \lambda \|\mathbf{x}\|_\infty. \end{aligned}$$

Nos queda por comprobar la desigualdad triangular:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_\infty &= \|(x_1 + y_1, \ldots, x_n + y_n)\|_\infty = \max_{i=1,\ldots n}\{|x_i + y_i|\} \\ &\leq \max_{i=1,\ldots n}\{|x_i| + |y_i|\} \leq \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty \end{aligned}$$

porque $|x_i + y_i| \leq |x_i| + |y_i|$ para cualquiera números $x_i, y_i \in \mathbb{R}$.

Sin embargo, esta norma no proviene de un producto escalar. Se puede comprobar.

De la misma forma, en espacios de funciones, como por ejemplo el espacio de las funciones continuas definidas sobre el intervalo $[0,1]$ y con valores en $\mathbb{R}$, $C[0,1]$, se puede definir la norma del supremo $\|\cdot\|_\infty$ mediante

$$\|f\|_\infty = \sup_{x \in [0,1]} \{|f(x)|\}.$$

Sabemos que este supremo existe, porque es una función continua, con valores en $\mathbb{R}$, definida sobre un conjunto cerrado.

Otra norma usual es la siguiente norma:

$$\|f\|_1 = \int_0^1 |f(x)|\, dx.$$

Ejemplo 8. Vamos a comprobar que la norma del supremo

$$\|f\|_\infty = \sup_{x\in[0,1]} \{|f(x)|\}$$

definida sobre el espacio vectorial $C[0,1] = \{f : [0,1] \longrightarrow \mathbb{R} : f \text{ es continua}\}$ es una norma. Tenemos que comprobar:

1. No degenerada: $\|f\|_\infty \geq 0$ y $\|f\|_\infty = 0$ si y sólo si $f(x) = 0$ para todo x. En efecto:

$$\|f\|_\infty = \sup_{x\in[0,1]} \{|f(x)|\} = 0 \Longleftrightarrow f(x) = 0 \text{ si } x \in [0,1].$$

2. Homogénea por homotecias: $\|\lambda f\|_\infty = |\lambda| \|f\|_\infty$:

$$\begin{aligned}\|\lambda f\|_\infty &= \sup_{x\in[0,1]} \{|\lambda f(x)|\} = \sup_{x\in[0,1]} \{|\lambda| |f(x)|\} = |\lambda| \sup_{x\in[0,1]} \{|f(x)|\} \\ &= |\lambda| \|f\|_\infty.\end{aligned}$$

3. Desigualdad triangular: $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. Tenemos:

$$\begin{aligned}\|f+g\|_\infty &= \sup_{x\in[0,1]} \{|f(x)+g(x)|\} \leq \sup_{x\in[0,1]} \{|f(x)| + |g(x)|\} \\ &\leq \sup_{x\in[0,1]} \{|f(x)|\} + \sup_{x\in[0,1]} \{|g(x)|\} = \|f\|_\infty + \|g\|_\infty.\end{aligned}$$

☻

Si tenemos un subespacio vectorial W de un espacio normado $(V, \|\cdot\|)$, entonces $\|\cdot\|$ restringida a W también es una norma, que se llama **norma inducida**. A veces se indica como $\|\cdot\|_W$. De este modo, $(W, \|\cdot\|_W)$ es un espacio normado. Se dice que W es un **subespacio normado** de V.

Tenemos que normas distintas pueden dar lugar a conjuntos distintos, o a vectores unitarios distintos. Vamos a ver la primera de estas situaciones con un ejemplo. Vamos a considerar la norma euclídea y la norma del supremo de $\mathbb{R}^2$ y vamos a determinar, para estas normas, los conjuntos

$$\begin{aligned}A &= \left\{(x,y) \in \mathbb{R}^2 : \|(x,y)\| \leq 1\right\}, \\ B &= \left\{(x,y) \in \mathbb{R}^2 : \|(x,y)\|_\infty \leq 1\right\}.\end{aligned}$$

El conjunto es:

$$\begin{aligned} A &= \{(x,y) \in \mathbb{R}^2 : \|(x,y)\| \leq 1\} = \{(x,y) \in \mathbb{R}^2 : \sqrt{x^2+y^2} \leq 1\} \\ &= \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq 1\}. \end{aligned}$$

Dicho conjunto se corresponde con el círculo centrado en $(0,0)$ y de radio 1, y se representa en la Figura 2.

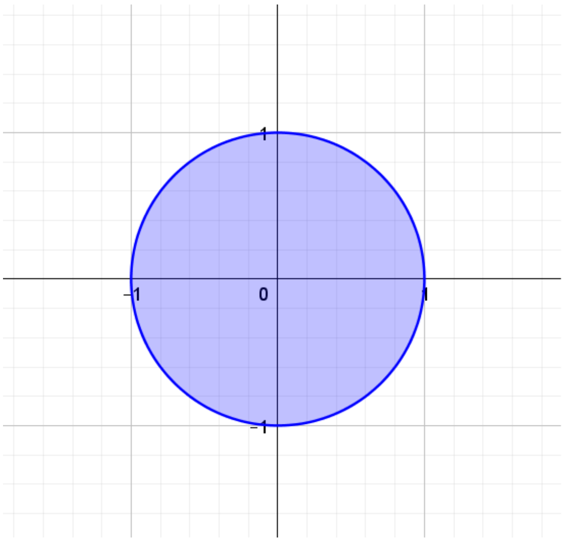


Figura 2: Conjunto A

Para el conjunto B, tenemos:

$$\begin{aligned} B &= \{(x,y) \in \mathbb{R}^2 : \|(x,y)\|_\infty \leq 1\} = \{(x,y) \in \mathbb{R}^2 : \max\{|x|,|y|\} \leq 1\} \\ &= \{(x,y) \in \mathbb{R}^2 : |x| \leq 1, |y| \leq 1\} \\ &= \{(x,y) \in \mathbb{R}^2 : -1 \leq x \leq 1, -1 \leq y \leq 1\}. \end{aligned}$$

Por tanto, B se corresponde con la Figura 3, en el plano.

A partir de una norma, tal como hacíamos con el producto escalar, se puede definir una distancia como

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|.$$

Se llama distancia asociada a la norma.

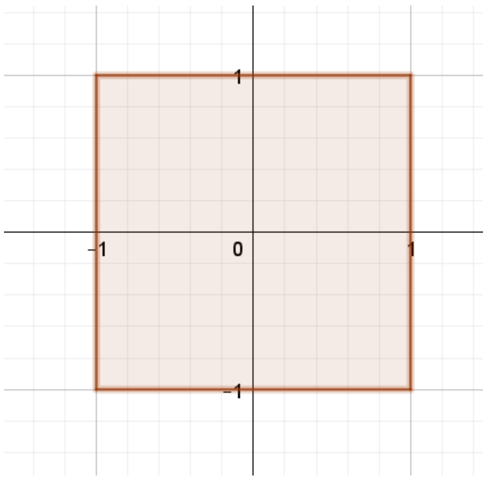


Figura 3: Conjunto B

Ejemplo 9. Consideramos en el espacio de los polinomios definidos en el intervalo $[0,1]$, la norma definida, para $p(x)$, mediante:

$$\|p\| = \int_0^1 (p(x))^2 \, dx.$$

Vamos a determinar la distancia, definida a través de la norma, entre los polinomios $f(x) = x^2 + x + 1$ y $g(x) = x - 2$.

La distancia entre dos polinomios es la norma de su diferencia, es decir, se determina mediante:

$$\begin{aligned}
\|f - g\| &= \int_0^1 (f(x) - g(x))^2 \, dx = \int_0^1 \left(x^2 + x + 1 - (x-2)\right)^2 dx \\
&= \int_0^1 \left(x^2 + 3\right)^2 dx = \int_0^1 \left(x^4 + 6x^2 + 9\right) dx \\
&= \frac{1}{5}x^5 + 2x^3 + 9x\Big|_0^1 = \frac{1}{5} + 2 + 9 = \frac{56}{5}.
\end{aligned}$$

☻

Normas equivalentes

Aunque los conjuntos que definan dos normas en un espacio vectorial V no sean iguales, este hecho no es relevante para trabajar con convergencia de sucesiones o continuidad de funciones en V. Lo que sí es importante es que si una sucesión tiene límite con una norma lo tenga también con la otra.

Esto, de forma más formal, es decir que las estructuras topológicas de las dos normas son las mismas (los mismos conjuntos abiertos y cerrados). O, lo que es lo mismo, que las dos normas son equivalentes.

Dos normas $\|\cdot\|$, $\|\cdot\|_1$ definidas en un espacio vectorial V, son **normas equivalentes** si existen $\lambda, \mu \in \mathbb{R}, \lambda, \mu > 0$ tales que para todo $\mathbf{u} \in V$ se cumple

$$\lambda \|\mathbf{u}\| \leq \|\mathbf{u}\|_1 \leq \mu \|\mathbf{u}\| .$$

Un resultado importante es que en $\mathbb{R}^n$ todas las normas son equivalentes. Esto significa que podemos trabajar con cualquiera de ellas. En general, a no ser que digamos lo contrario, trabajaremos con la norma euclídea.

Observación 6. *Cualquier espacio vectorial normado V sobre $\mathbb{R}$ de dimensión finita n es isomorfo a $\mathbb{R}^n$. Por tanto, el estudio de cualquier espacio normado de dimensión n se reduce al estudio de $\mathbb{R}^n$.*

Vamos a demostrarlo, pero en una primera lectura se puede obviar esta demostración. Partimos de V, un espacio vectorial de dimensión n y de un base suya de vectores unitarios $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ con la norma $\|\cdot\|^*$. Sea $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base canónica de $\mathbb{R}^n$.

Entonces podemos establecer un isomorfismo algebraico[1] entre $\mathbb{R}^n$ y V:

$$f(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

Puesto que el espacio V es de dimensión finita, se cumple además que F es un isomorfismo topológico u homeomorfismo (biyectiva, continua y con inversa continua) entre $(\mathbb{R}^n, \|\cdot\|)$ y $(\mathbb{V}, \|\cdot\|^*)$, en donde $\|\cdot\|$ es una norma cualquiera[2] de $\mathbb{R}^n$. Además, cualquier propiedad que posea un subconjunto $M \subset \mathbb{R}^n$ la tiene también $f(M)$ y cualquier propiedad que tenga un subconjunto $F \subset V$ la tiene también $f^{-1}(F)$.

Al hablar de isomorfismo queremos decir aplicación biyectiva entre dos espacios. Un isomorfismo algebraico es una aplicación lineal biyectiva entre dos espacios vectoriales

Se puede demostrar fácilmente sin más que tomar una base $\{a_1, \ldots, a_n\}$ de vectores unitarios con la norma $\|\cdot\|^*$. Si $\{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$ es la base canónica de $\mathbb{R}^n$, entonces podemos establecer un isomorfismo algebraico entre $\mathbb{R}^n$ y V:

$$f(x_1\mathbf{e_1} + \cdots + x_n\mathbf{e_n}) = x_1 f(a_1) + \cdots + x_n f(a_n) .$$

[1]Se recuerda que un isomorfismo algebraico es una aplicación lineal biyectiva tal que su inversa también es lineal.

[2]Recuerde que todo espacio normado es topológico, pues la norma induce una topología.

Entonces esta aplicación es un homeomorfismo (biyectiva, continua y con inversa continua) entre $(\mathbb{R}^n, \|\cdot\|)$ y $\left(\mathbb{V}, \|\cdot\|^*\right)$, en donde $\|\cdot\|$ es una base cualquiera de $\mathbb{R}^n$. Entonces, cualquier propiedad que posea un subconjunto $M \subset \mathbb{R}^n$ la tiene también $f(M)$ y cualquier propiedad que tenga un subconjunto $F \subset V$ la tiene también $f^{-1}(F)$. Por tanto, el estudio de cualquier espacio de dimensión n se reduce al estudio de $\mathbb{R}^n$.

Sabemos que podemos definir una norma a partir de un producto escalar, pero no siempre es posible definir un producto escalar a partir de una norma. La norma asociada a un producto escalar cumple la **identidad del paralelogramo**

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\right)$$

para $\mathbf{u}, \mathbf{v} \in V$. Esta identidad tiene una clara interpretación geométrica: si tomamos el paralelogramo determinado por los vectores $\mathbf{u}, \mathbf{v}$ entonces la suma de los cuadrados de la longitud de sus diagonales es igual a la suma de los cuadrados de sus cuadro lados (que son iguales dos a dos). Pero además, si una norma cumple esta igualdad, entonces la norma viene de un producto escalar y el espacio es vectorial y tiene producto escalar. Es decir, la identidad del paralelogramo es condición necesaria y suficiente para que una norma venga de un producto escalar. Este resultado se atribuye a Fréchet, von Neumann y Jordan y no vamos a demostrarlo.

Utilizando este resultado podemos comprobar con facilidad cuándo una norma proviene de un producto escalar. Vamos a comprobar que la norma del supremo en $\mathbb{R}^n$ no viene de un producto escalar. Recordamos que esta norma es:

$$\|(x_1, \ldots, x_n)\|_\infty = \max_{i=1,\ldots n}\{|x_i|\}.$$

Si $\|\cdot\|_\infty$ cumpliese la identidad del paralelogramo, entonces para todo $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ se tendría que

$$\|\mathbf{x}+\mathbf{y}\|_\infty^2 + \|\mathbf{x}-\mathbf{y}\|_\infty^2 = 2\left(\|\mathbf{x}\|_\infty^2 + \|\mathbf{y}\|_\infty^2\right).$$

Es decir,

$$\begin{aligned}\left(\max_{i=1,\ldots,n}\{|x_i+y_i|\}^2\right) + \left(\max_{i=1,\ldots,n}\{|x_i-y_i|\}\right)^2 \\ = 2\left(\left(\max_{i=1,\ldots n}\{|x_i|\}\right)^2 + \left(\max_{i=1,\ldots n}\{|y_i|\}\right)^2\right).\end{aligned}$$

Sin embargo, si elegimos $\mathbf{x} = \mathbf{e}_1, \mathbf{y} = \mathbf{e}_1 + \mathbf{e}_2$, entonces

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|_\infty^2 &= \left(\max_{i=1,\ldots n}\{|x_i+y_i|\}\right)^2 = 2^2 = 4,\\
\|\mathbf{x}-\mathbf{y}\|_\infty^2 &= \left(\max_{i=1,\ldots n}\{|x_i-y_i|\}\right)^2 = 1^2 = 1,\\
\|\mathbf{x}\|_\infty^2 &= \max_{i=1,\ldots n}\{|x_i|\} = 1,\\
\|\mathbf{y}\|_\infty^2 &= \max_{i=1,\ldots n}\{|y_i|\} = 1,\\
\|\mathbf{x}+\mathbf{y}\|_\infty^2 + \|\mathbf{x}-\mathbf{y}\|_\infty^2 &= 5,\\
2\left(\|\mathbf{x}\|_\infty^2 + \|\mathbf{y}\|_\infty^2\right) &= 4,
\end{aligned}$$

luego no se cumple la identidad del paralelogramo. Por tanto, $\|\cdot\|_\infty$ no proviene de un producto escalar.

Con este ejemplo vemos que los espacios vectoriales con producto escalar son espacios normados, pero que en los espacios normados no siempre se puede definir un producto escalar. Es decir, la idea de espacio normado generaliza los espacios vectoriales con producto escalar.

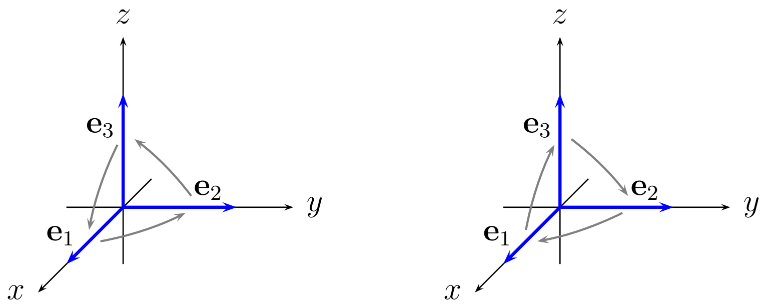
1.3. Orientación, producto vectorial y producto mixto

Como sabemos, para obtener las coordenadas de un vector de un espacio vectorial de dimensión finita respecto a una base distinta a la que está expresado, se multiplica el vector de sus coordenadas respecto a la base primera por una matriz, que es la matriz de cambio de base. Esta matriz nos permite clasificar de alguna forma, las bases de espacios vectoriales de dimensión finita que son, como veremos más adelante, isomorfos a $\mathbb{R}^n$.

En $\mathbb{R}^n$ se dice que **dos bases tienen la misma orientación** si el determinante de la matriz de cambio de base es positivo. Así, se pueden clasificar las bases en $\mathbb{R}^n$ según su orientación con respecto a una base fijada. Como el determinante de la matriz de cambio de base puede ser positivo o negativo, se obtienen dos clases de bases según el valor del determinante: la clase de orientación positiva (misma orientación) y la clase de orientación negativa (distinta orientación).

Ejemplo 10. Normalmente se considera que la base canónica de $\mathbb{R}^3$ con su orden natural, que es $B = \{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\} = \{(1,0,0),(0,1,0),(0,0,1)\}$ tiene orientación positiva.

Según esto, la base $B' = \{\mathbf{e_1}, \mathbf{e_3}, \mathbf{e_2}\} = \{(1,0,0),(0,0,1),(0,1,0)\}$ tiene orientación negativa. Y $B'' = \{\mathbf{e_2}, \mathbf{e_3}, \mathbf{e_1}\} = \{(0,1,0),(0,0,1),(1,0,0)\}$ tiene orientación positiva.



Orientación positiva Orientación negativa

Figura 4: Orientaciones positiva y negativa

☻

La orientación de una base es un tema importante y que vamos a utilizar al trabajar con curvas y superficies. Allí hablaremos de orientación positiva o negativa y entre otras cosas, en nuestra vida cotidiana es la formulación matemática de izquierda y derecha y da significado matemático a la expresión “recorrer una curva en un sentido”.

Ahora trabajamos en $\mathbb{R}^3$. Una terna de vectores linealmente independientes $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ se dice que es derecha o dextrógira si tiene el mismo tipo de orientación que los dedos pulgar, índice y corazón de la mano derecha, o si al ir de $\mathbf{v}_1$ a $\mathbf{v}_2$ el vector $\mathbf{v}_3$ tiene el sentido de avance de un tornillo. En otro caso, se dice que la terna es levógira o izquierda.

Ejemplo 11. La base canónica $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ de $\mathbb{R}^3$ es dextrógira.

El conjunto de vectores linealmente independientes B formado por los vectores $\{(1,-1,0),(1,1,-1),(0,0,1)\}$ es dextrógiro (los vectores están dados respecto a la base canónica) y el conjunto $\{(1,-1,0),(0,1,-1),(0,0,-1)\} = B'$ es levógiro.

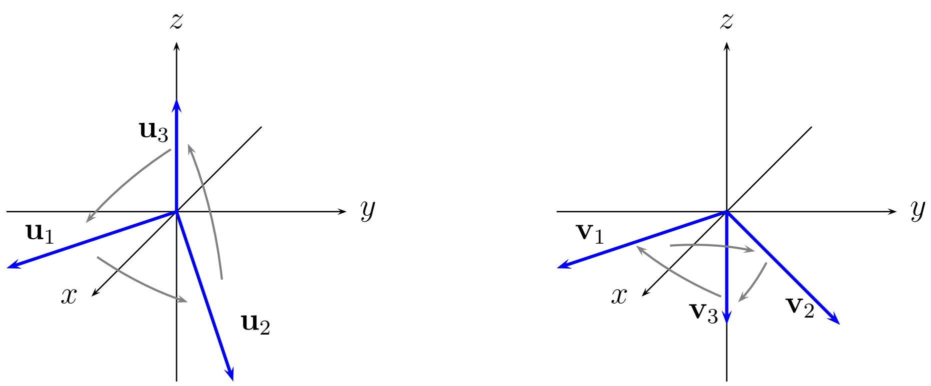
En efecto, la matriz que expresa B respecto a la base canónica es

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

Su determinante es 2, se comprueba de forma inmediata. Así pues, B tiene la misma orientación que la base canónica y, por tanto, B es dextrógiro.

De la misma forma, el determinante de la matriz asociada a B' respecto a la base canónica es -1, por eso, la base es levógira. Ambas bases se han representado en la Figura 5. Los vectores de B se han llamado $\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}$ y los de B' se han indicado con $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$.

☻



Elementos de B Elementos de B'

Figura 5: Orientaciones de las bases

Definición 7. *Si* $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ *son dos vectores de* $\mathbb{R}^3$*, el* ***producto vectorial*** *de* $\mathbf{u}$ *y* $\mathbf{v}$*, denotado por* $\mathbf{u} \wedge \mathbf{v}$ *o* $\mathbf{u} \times \mathbf{v}$*, es el único vector de* $\mathbb{R}^3$ *que cumple*

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \det(\mathbf{u}, \mathbf{v}, \mathbf{w})$$

para todo $\mathbf{w} \in \mathbb{R}^3$.

El producto vectorial también se llama producto exterior.

Determinar qué vector es el producto vectorial de dos vectores puede parecer complicado, pero no lo es tanto. Baste observar que si $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$, entonces se cumple:

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{e}_1 = \begin{vmatrix} 1 & 0 & 0 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}.$$

Teniendo en cuenta que el producto escalar de un vector por $\mathbf{e}_1$ es la primera coordenada del vector, entonces la primera coordenada de $\mathbf{u} \times \mathbf{v}$ es el determinate anterior. De forma similar, se obtienen las coordenadas segunda y tercera de $\mathbf{u} \wedge \mathbf{v}$ respecto a la base canónica de $\mathbb{R}^3$. Y así, se cumple

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_1 \end{vmatrix} \mathbf{e}_3.$$

Otra forma de expresarlo es abusando de la notación y escribiendo

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix},$$

donde tenemos que calcular un determinante cuya primera fila no son números, sino son los vectores de la base canónica.

Ejemplo 12. Vamos a determinar el producto vectorial de los vectores $(0,-1,3)$ y $(2,-1,1)$.

Hacemos

$$(0,-1,3)\times(2,-1,1)=\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 0 & -1 & 3 \\ 2 & -1 & 1 \end{vmatrix}=2\mathbf{e}_1+6\mathbf{e}_2+2\mathbf{e}_3.$$

Luego $(0,-1,3)\times(2,-1,1)=(2,6,2)$ en la base canónica. ☻

Proposición 8. *Propiedades del producto vectorial.*

1. $\mathbf{u}\times\mathbf{v}=-\mathbf{v}\times\mathbf{u}$, *no conmutativo.*
2. $(\lambda\mathbf{u}+\mu\mathbf{w})\times\mathbf{v}=(\lambda\mathbf{u})\times\mathbf{v}+(\mu\mathbf{w})\times\mathbf{v}=\lambda(\mathbf{u}\times\mathbf{v})+\mu(\mathbf{w}\times\mathbf{v})$, *linealidad.*
3. $\mathbf{u}\times\mathbf{v}=0$ *si y sólo si* $\mathbf{u}$ *y* $\mathbf{v}$ *son linealmente dependientes o cero.*
4. $(\mathbf{u}\times\mathbf{v})\times\mathbf{w}\neq\mathbf{u}\times(\mathbf{v}\times\mathbf{w})$, *no asociativo.*
5. $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{u}=(\mathbf{u}\times\mathbf{v})\cdot\mathbf{v}=\mathbf{0}$.

A partir de estar propiedades, obtenemos mucha información interesante. Primero observamos que $\mathbf{u}\times\mathbf{v}$ es un vector perpendicular al plano determinado por $\mathbf{u}$ y $\mathbf{v}$ (es decir, al subespacio generado por estos vectores). Se demuestra teniendo en cuenta la última propiedad y la linealidad del producto vectorial.

Por otro lado, como

$$\det(\mathbf{u},\mathbf{v},(\mathbf{u}\times\mathbf{v}))=(\mathbf{u}\times\mathbf{v})\cdot(\mathbf{u}\times\mathbf{v})=|(\mathbf{u}\times\mathbf{v})|^2>0,$$

entonces la base formada por $\mathbf{u},\mathbf{v},(\mathbf{u}\times\mathbf{v})$ tiene orientación positiva.

Además, el módulo del producto vectorial $\mathbf{u}\times\mathbf{v}$ es el área del paralelogramo de lados $\mathbf{u}$ y $\mathbf{v}$, y viene dado por:

$$|\mathbf{u}\times\mathbf{v}|=|\mathbf{u}|\,|\mathbf{v}|\operatorname{sen}\theta,$$

donde θ es el ángulo que forman $\mathbf{u}$ y $\mathbf{v}$. Este resultado se demuestra teniendo en cuenta la relación entre seno y coseno, la expresión del módulo del producto escalar de $\mathbf{u}$ y $\mathbf{v}$ en función del ángulo que forman, y el módulo de $|(\mathbf{u}\times\mathbf{v})|^2$. No lo vamos a demostrar aquí.

Definición 9. *El* ***producto mixto*** *de tres vectores* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *es* $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$.

A partir del cálculo del producto vectorial con determinantes, es fácil ver que si $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$, $\mathbf{w} = (w_1, w_2, w_3)$, entonces

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

Como interpretación del producto mixto, diremos que el valor absoluto de $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$ es el volumen del paralelepípedo dado por los tres vectores, si son linealmente independientes. Es sencillo de demostrar.

Por otro lado, supongamos que la base del paralelepípedo es el paralelogramo cuyos lados son $\mathbf{u}, \mathbf{v}$, y que $\mathbf{w}$ determina la arista que no está en ese plano.

A partir de esto podemos determinar fácilmente el volumen del tetraedro que forman los vectores. El volumen de un tetraedro es

$$\frac{1}{3} bh,$$

si b es el área de la base y h es la altura. Si consideramos el prisma triangular cuya base viene determinada por $\mathbf{u}$ y $\mathbf{v}$ el volumen de dicho prisma es la mitad del volumen del paralelepípedo determinado por $\mathbf{u}, \mathbf{v}$ y $\mathbf{w}$, luego dicho volumen es

$$\frac{1}{2} (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}.$$

Como el prisma triangular se descompone en tres tetraedros de igual volumen (ver Figura 6), se deduce que el volumen del tetraedro definido por $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ es

$$\frac{1}{6} (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}.$$

Ejemplo 13. Determine el volumen del tetraedro de vértices

$$A = (0, 0, 0), \quad B = (1, 1, 1), \quad C = (2, 0, -1), \quad D = (-1, 0, 1).$$

Si consideramos los vectores

$$\mathbf{u} = AB = (1, 1, 1), \quad \mathbf{v} = AC = (2, 0, -1), \quad \mathbf{w} = AD = (-1, 0, 1),$$

entonces el volumen del paralelepípedo formado por los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w}$ es $|(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}|$. Y el volumen del tetraedro es la sexta parte del área de este paralelepípedo:

$$V = \frac{1}{6} |(AB \times AC) \cdot AD| = \frac{1}{6} \left\| \begin{matrix} 1 & 1 & 1 \\ 2 & 0 & -1 \\ -1 & 0 & 1 \end{matrix} \right\| = \left| -\frac{1}{6} \right| = \frac{1}{6}.$$

☻

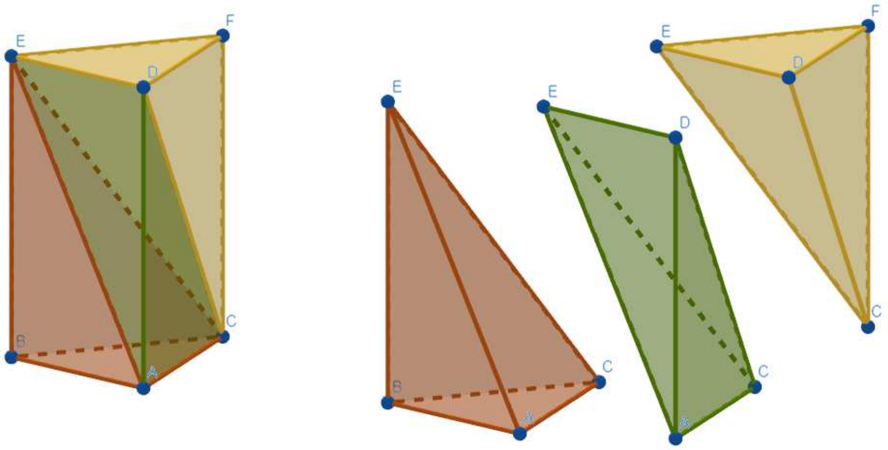


Figura 6: Descomposición del prisma triangular en tres tetraedros de igual volumen

Terminamos adelantando una última propiedad. Si **u** y **v** son funciones de $\mathbb{R}$ en $\mathbb{R}^3$ (son expresiones de la forma $(u_1(t), u_2(t), u_3(t))$, que repasamos en el siguiente apartado) que dependen de un parámetro t, y que son derivables, entonces su producto vectorial también es derivable y

$$\frac{d}{dt}(\mathbf{u}(t) \times \mathbf{v}(t)) = \frac{d\mathbf{u}(t)}{dt} \times \mathbf{v}(t) + \mathbf{u}(t) \times \frac{d\mathbf{v}(t)}{dt}.$$

1.4. Funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

Partimos de funciones $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Son las aplicaciones que a cada $\mathbf{x} = (x_1, \ldots, x_n) \in A$ le asignan un elemento

$$\begin{aligned} f(\mathbf{x}) = f(x_1, \ldots, x_n) &= (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \\ &= (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})), \end{aligned}$$

donde cada una de las funciones $f_i : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función de varias variables, que se llama **componente** de f. En esta definición hemos indicado las posibles notaciones que nos vamos a encontrar para hablar de funciones de varias variables definidas sobre un espacio de varias variables (considerando siempre dimensiones finitas).

Son conocidas las propiedades de las componentes, al ser cada componente una función de varias variables que toma valores en $\mathbb{R}$. También se ha trabajado con derivadas direccionales, derivadas parciales y diferencial de estas funciones.

Vamos a distinguir algunos casos particulares de funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ en los siguientes ejemplos.

Ejemplo 14. Si $n = m = 1$, tenemos las funciones $f : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ de variable real sobre $\mathbb{R}$. La función $f : [0, 60] \longrightarrow \mathbb{R}$ que a cada $t \in [0, 60]$ le asigna los kilómetros recorridos por un coche Fórmula 1 durante los primeros t minutos de una carrera es un ejemplo de estas funciones.

☻

Ejemplo 15. Si $n > 1, m = 1$, tenemos las funciones $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ de varias variables con las que hemos trabajado frecuentemente. Un ejemplo es la función $T : [0, 1] \times [0, 1] \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ que asigna a cada punto de coordenadas (x, y) de una placa cuadrada de lado 1, la temperatura $T(x, y)$ en ese punto.

☻

Ejemplo 16. Si $n = 1, m > 1$, la función se suele llamar función vectorial. Casos particulares son:

- Curva plana, considerando el movimiento de una partícula que depende sólo del instante (tiene un grado de libertad, $n = 1, m = 2$). La función $\sigma : I \subset \mathbb{R} \longrightarrow \mathbb{R}^2$ da la posición de la partícula dependiente del instante de tiempo.
- Curva en el espacio cuando $n = 1, m = 3$, que describe la posición de una partícula que se mueve con un grado de libertad en el espacio tridimensional.

☻

Ejemplo 17. Si $n > 1, m > 1$, las funciones se suelen llamar campo vectorial. Citaremos varios ejemplos de estas funciones:

- Simetría en un espacio de dimensión $n > 1$ respecto al origen, que es la función $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ definida por $f(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n)$.
- Superficie, con $n = 2, m = 3$, donde la posición de una partícula en el espacio está determinada a partir de su movimiento con dos grados de libertad.
- Cambio de coordenadas polares a cartesianas y viceversa, en el plano, con $n = 2, m = 2$.
- Cambio de coordenadas cilíndricas a cartesianas y viceversa, para $n = 3, m = 3$, en el espacio.

- En el espacio, cambio de coordenadas esféricas a cartesianas y viceversa, con $n = 3, m = 3$.

☻

Ejemplo 18. Si $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ tiene todas las derivadas parciales en todos los puntos de A, entonces el gradiente de f es la aplicación que a cada $\mathbf{x} \in A$ le asigna

$$\nabla f(\mathbf{x}) = (D_1 f(\mathbf{x}), \ldots, D_i f(\mathbf{x}), \ldots, D_n f(\mathbf{x})),$$

y donde $\nabla f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$.

☻

Las definiciones, resultados y procedimientos relativos a funciones de varias variables estudiados anteriormente se pueden extrapolar fácilmente a las funciones objeto de nuestro estudio, teniendo en cuenta que cada una de las componentes es una función de varias variables que toma valores en $\mathbb{R}$.

De esta forma, para funciones $f, g : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ se pueden definir su función suma y función resta, es decir, considerando estas operaciones componente a componente.

Ejemplo 19. Sean $f, g : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dadas por

$$f(x,y,z) = \left(xe^y, z^2 + 1\right), \quad g(x,y,z) = \left(\frac{x-z}{x^2+1}, z\right).$$

Entonces

$$\begin{aligned}(f+g)(x,y,z) &= f(x,y,z) + g(x,y,z) = \left(xe^y + \frac{x-z}{x^2+1}, z^2 + 1 + z\right),\\ (f-g)(x,y,z) &= f(x,y,z) - g(x,y,z) = \left(xe^y - \frac{x-z}{x^2+1}, z^2 + 1 - z\right),\end{aligned}$$

☻

La composición de funciones de varias variables con valores en espacios multidimensionales se hace considerando las dimensiones de los espacios imagen de la primera función y dominio de la segunda. Consideramos $A \subset \mathbb{R}^n, B \subset \mathbb{R}^m$ y $C \subset \mathbb{R}^l$, con $f(A) \subset B$ y $g(B) \subset C$. Si para $f : A \longrightarrow B$ y $g : B \longrightarrow C$ se cumple que para todo $\mathbf{x} \in A, f(\mathbf{x}) \in B$, entonces la función $h = g \circ f$ está definida como

$$h(\mathbf{x}) = g \circ f(\mathbf{x}) = g(f(\mathbf{x})).$$

Ejemplo 20. Sean $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dadas por:

$$f(x,y) = \left(y^3, x\right), \quad g(x,y) = \left(\cos(x-y), x^2 + y - 1, x\right).$$

Entonces

$$\begin{aligned} h(x,y) &= g \circ f(x,y) = g(f(x,y)) = g\left(y^3, x\right) \\ &= \left(\cos\left(y^3 - x\right), \left(y^3\right)^2 + x - 1, y^3\right) = \left(\cos\left(y^3 - x\right), y^6 + x - 1, y^3\right). \end{aligned}$$

☻

Continuidad

Una función $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es continua en un punto $\mathbf{a} \in A$ si existe $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x})$ y además $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. Esto es equivalente a decir que cada una de las componentes f_i de f es continua en el punto $\mathbf{a}$. Además, sabemos cómo estudiar la continuidad de cada f_i, al ser funciones de varias variables con valores en $\mathbb{R}$.

Ejemplo 21. La función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ con $f(x,y) = (x^2, y - x + 1, \operatorname{sen} \pi y)$ es continua en $(0,0,0)$, porque

$$\begin{aligned} f(0,0) &= \left(0^2, 0 - 0 + 1, \operatorname{sen} \pi \cdot 0\right) = (0,1,0), \\ \lim_{(x,y)\to(0,0)} f(x,y) &= \lim_{(x,y)\to(0,0)} \left(x^2, y - x + 1, \operatorname{sen} \pi y\right) \\ &= \left(\lim_{(x,y)\to(0,0)} x^2, \lim_{(x,y)\to(0,0)} (y - x + 1), \lim_{(x,y)\to(0,0)} \operatorname{sen} \pi y\right) \\ &= \left(0^2, 0 - 0 + 1, \operatorname{sen} \pi \cdot 0\right) = (0,1,0), \\ \lim_{(x,y)\to(0,0)} f(x,y) &= f(0,0). \end{aligned}$$

☻

Una función $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es continua en el conjunto A si es continua en todos los puntos $\mathbf{a} \in A$. Esto es equivalente a decir que cada una de las componentes f_i de f es continua en el conjunto A.

Si $f, g : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ son funciones continuas en A, entonces también lo son su suma, resta.

Si $f : A \subset \mathbb{R}^n \longrightarrow B \subset \mathbb{R}^m$, $g : B \subset \mathbb{R}^m \longrightarrow \mathbb{R}^l$ son funciones continuas en A y B respectivamente, entonces la función $h = g \circ f$ definida como

$$h(\mathbf{x}) = g \circ f(\mathbf{x}) = g(f(\mathbf{x}))$$

también es continua en cualquier $\mathbf{x} \in A$.

Diferenciabilidad de funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

Sea $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$, dada por $f = (f_1, \dots, f_i, \dots, f_m)$, donde cada $f_i : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. Entonces se puede extender el concepto de derivada según un vector $\mathbf{v}$ a esta función mediante:

$$D_{\mathbf{v}} f(\mathbf{a}) = (D_{\mathbf{v}} f_1(\mathbf{a}), \dots, D_{\mathbf{v}} f_i(\mathbf{a}), \dots, D_{\mathbf{v}} f_m(\mathbf{a})).$$

Obsérvese que la notación que utilizamos es la habitual para indicar derivada según un vector.

Como $f_i : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, podemos utilizar las definiciones y resultados ya conocidos para las funciones de varias variables sobre $\mathbb{R}$. En concreto, podemos considerar sus derivadas parciales según $\mathbf{e}_j$, denotadas como $D_j f_i$ o $\frac{\partial f_i}{\partial x_j}$. A partir de ellas, podemos definir un equivalente a la aplicación gradiente, para funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$. Es la apliación dada por la siguiente matriz de orden $m \times n$:

$$\begin{pmatrix} D_1 f_1(\mathbf{a}) & D_2 f_1(\mathbf{a}) & \dots & D_n f_1(\mathbf{a}) \\ D_1 f_2(\mathbf{a}) & D_2 f_2(\mathbf{a}) & \dots & D_n f_2(\mathbf{a}) \\ \vdots & \vdots & \dots & \vdots \\ D_1 f_m(\mathbf{a}) & D_2 f_m(\mathbf{a}) & \dots & D_n f_m(\mathbf{a}) \end{pmatrix},$$

que es la **matriz jacobiana** de f en $\mathbf{a}$ y se denota $f'(\mathbf{a})$.

Ejemplo 22. Dada $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$f(x, y, z) = \left(x + \cos\left(x^2 yz\right), (x + y + z) e^z\right),$$

vamos a determinar su matriz jacobiana. Hacemos

$$\begin{aligned} f'(\mathbf{x}) &= \begin{pmatrix} D_1 f_1(\mathbf{x}) & D_2 f_1(\mathbf{x}) & D_3 f_1(\mathbf{x}) \\ D_1 f_2(\mathbf{x}) & D_2 f_2(\mathbf{x}) & D_3 f_2(\mathbf{x}) \end{pmatrix} \\ &= \begin{pmatrix} 1 - 2xyz \operatorname{sen}(x^2 yz) & -x^2 z \operatorname{sen}(x^2 yz) & -x^2 y \operatorname{sen}(x^2 yz) \\ e^z & e^z & e^z (1 + x + y + z) \end{pmatrix}. \end{aligned}$$

☻

Definimos formalmente la diferencial de una función f de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Definición 10. *Una función $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es diferenciable en un punto $\mathbf{a} \in A$ si existe una aplicación lineal $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ tal que*

$$\lim_{h \to 0} \frac{\| f(\mathbf{a} + \mathbf{h}t) - f(a) - \lambda(\mathbf{h}) \|}{\| h \|} = 0.$$

*La función λ se denomina **diferencial** de f en $\mathbf{a}$. Si f es diferenciable en todos los puntos de A se dice que es diferenciable en A.*

La diferencial tiene las siguientes propiedades:

1. Si existe, la diferencial de f en $\mathbf{a}$ es única. Se denota $Df(\mathbf{a})$, es decir $\lambda(\mathbf{x}) = Df(\mathbf{a})(\mathbf{x})$ y se tiene que la matriz asociada a la aplicación lineal $Df(\mathbf{a})$ es la matriz jacobiana $f'(\mathbf{a})$.

2. Condición necesaria de diferenciabilidad: Si la función f diferenciable en $\mathbf{a}$, entonces f es continua en $\mathbf{a}$.

Además, al igual que pasaba con la derivada y con las derivadas parciales, hay varias reglas de cálculo que nos permiten determinar la diferencial de una función. Algunas son conocidas del estudio de las funciones de varias variables que toman valores en $\mathbb{R}$.

Proposición 11. *Propiedades de funciones diferenciables.*

1. *Toda función constante en A es diferenciable en A y su diferencial en cualquier punto $\mathbf{a}$ de A es la aplicación lineal nula.*

2. *Toda aplicación lineal es diferenciable y su diferencial en cualquier punto es ella misma, es decir, si $T : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es lineal, entonces $DT(\mathbf{a})(\mathbf{x}) = T(\mathbf{x})$ para todo $\mathbf{a} \in A$.*

3. *Si $f, g : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ son dos funciones diferenciables en $\mathbf{a} \in A$ sobre $\mathbb{R}$ y $\alpha, \beta \in \mathbb{R}$ son dos números reales, entonces las funciones $\alpha f + \beta g$ y $f \cdot g$ son diferenciables en $\mathbf{a}$ y:*

$$
\begin{aligned}
D(\alpha f + \beta g)(\mathbf{a}) &= \alpha Df(\mathbf{a}) + \beta Dg(\mathbf{a}), \\
D(f \cdot g)(\mathbf{a}) &= f(a)\, Dg(\mathbf{a}) + g(a)\, Df(\mathbf{a}).
\end{aligned}
$$

4. *Si $f, g : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ son dos funciones diferenciables en $\mathbf{a} \in A$ sobre $\mathbb{R}$ y $g(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in A$, entonces la función cociente es diferenciable en $\mathbf{a}$ y se cumple:*

$$
D\left(\frac{f}{g}\right)(\mathbf{a}) = \frac{g(\mathbf{a})\, Df(\mathbf{a}) - f(\mathbf{a})\, Dg(\mathbf{a})}{(g(\mathbf{a}))^2}.
$$

5. ***Regla de la cadena.*** *Sean A un abierto de $\mathbb{R}^n$ y B un abierto de $\mathbb{R}^m$, f una función de A en B y g una función de B, en $\mathbb{R}^p$. Si f es diferenciable en $\mathbf{a} \in A$ y g es diferenciable en $f(\mathbf{a}) \in B$, entonces la aplicación compuesta $h = g \circ f$ es diferenciable en a y:*

$$
Dh(\mathbf{a}) = Dg(f(\mathbf{a})) \circ Df(\mathbf{a}).
$$

Ejemplo 23. Sean $f:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ y $g:\mathbb{R}^2 \longrightarrow \mathbb{R}$ las funciones dadas por

$$f(x,y) = \left(x^2y^4, \cos xy^3\right),\ g(x,y) = ye^x.$$

Vamos a estudiar si $g \circ f$ es diferenciable y en caso de que lo sea, vamos a determinar su diferencial, tanto para un punto genérico (x,y) como en el punto $(0,1)$.

En efecto, es diferenciable, por ser composición de funciones diferenciables. Además:

$$\begin{aligned}
Df(x,y) &= \begin{pmatrix} D_1f_1(x,y) & D_2f_1(x,y) \\ D_1f_2(x,y) & D_2f(x,y) \end{pmatrix} \\
&= \begin{pmatrix} 2xy^4 & 4x^2y^3 \\ -y^3 \operatorname{sen} xy^3 & -3xy^2 \operatorname{sen} xy^3 \end{pmatrix}, \\
Dg(x,y) &= \begin{pmatrix} D_1g(x,y) & D_2g(x,y) \end{pmatrix} = (ye^x, e^x),
\end{aligned}$$

$$\begin{aligned}
Dg(f(x,y)) &= Dg\left(\left(x^2y^4, \cos xy^3\right)\right) = \left(e^{x^2y^4}\cos xy^3, e^{x^2y^4}\right), \\
D(g\circ f)(x,y) &= Dg(f(x,y)) \circ Df(x,y) = Dg(f(x,y)) \cdot Df(x,y) \\
&= \left(e^{x^2y^4}\cos xy^3, e^{x^2y^4}\right)\begin{pmatrix} 2xy^4 & 4x^2y^3 \\ -y^3 \operatorname{sen} xy^3 & -3xy^2 \operatorname{sen} xy^3 \end{pmatrix} \\
&= (2xy^4e^{x^2y^4}\cos xy^3 - y^3e^{x^2y^4}\operatorname{sen} xy^3, \\
&\quad 4x^2y^3e^{x^2y^4}\cos xy^3 - 3xy^2e^{x^2y^4}\operatorname{sen} xy^3).
\end{aligned}$$

Para determinar la diferencial en $(0,1)$ hacemos:

$$\begin{aligned}
D(g\circ f)(0,1) &= (2\cdot 0\cdot 1^4e^{0^21^4}\cos(0\cdot 1^3) - 1^3e^{0^21^4}\operatorname{sen}(0\cdot 1^3), \\
&\quad 4\cdot 0^21^3e^{0^21^4}\cos(0\cdot 1^3) - 3\cdot 0\cdot 1^2e^{0^21^4}\operatorname{sen})(0\cdot 1^3)) \\
&= (0,0).
\end{aligned}$$

Si lo hubiéramos hecho directamente, habríamos hecho:

$$g(f(x,y)) = g\left(x^2y^4, \cos xy^3\right) = \left(\cos xy^3\right)e^{x^2y^4}.$$

Y asó podemos obtener las derivadas parciales:

$$\begin{aligned}
D_1g(f(x,y)) &= 2\left(\cos xy^3\right)xy^4e^{x^2y^4} - y^3e^{x^2y^4}\operatorname{sen} xy^3, \\
D_2g(f(x,y)) &= 4x^2y^3e^{x^2y^4}\cos xy^3 - 3xy^2e^{x^2y^4}\operatorname{sen} xy^3.
\end{aligned}$$

☻

Finalmente, el siguiente teorema nos dice cuál es la diferencial de una función.

Teorema 12. *Una aplicación* $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$*,* $f = (f_1, \ldots, f_i, \ldots, f_m)$ *es diferenciable en* $\mathbf{a} \in A$ *si y sólo si cada* f_i *es diferenciable en* $\mathbf{a}$*. Además, la diferencial de* f *en* $\mathbf{a}$ *es la aplicación lineal* $Df(\mathbf{a})$ *cuyas componentes son las diferenciales* $Df_i(\mathbf{a})$ *de las componentes de* f*, es decir:*

$$Df(\mathbf{a}) = (Df_1(\mathbf{a}), \ldots, Df_i(\mathbf{a}), \ldots, Df_m(\mathbf{a})).$$

Este resultado indica que la diferencial está dada por la matriz jacobiana.

Ejemplo 24. Según vimos en un ejemplo anterior, la matriz jacobiana de $f : \mathbb{R}^3 \to \mathbb{R}^2$, con

$$f(\mathbf{x}) = f(x, y, z) = \left(x + \cos\left(x^2 yz\right), (x + y + z) e^z\right)$$

es

$$\begin{aligned} f'(\mathbf{x}) &= \begin{pmatrix} D_1 f_1(\mathbf{x}) & D_2 f_1(\mathbf{x}) & D_3 f_1(\mathbf{x}) \\ D_1 f_2(\mathbf{x}) & D_2 f_2(\mathbf{x}) & D_3 f_2(\mathbf{x}) \end{pmatrix} \\ &= \begin{pmatrix} 1 - 2xyz \operatorname{sen}(x^2yz) & -x^2 z \operatorname{sen}(x^2yz) & -x^2 y \operatorname{sen}(x^2yz) \\ e^z & e^z & e^z(1 + x + y + z) \end{pmatrix}. \end{aligned}$$

Entonces la diferencial de la función f en el punto $(0, 0, 0)$ es la aplicación lineal $Df(\mathbf{0}) : \mathbb{R}^3 \to \mathbb{R}^2$ dada por

$$Df(\mathbf{0})(x, y, z) = f'(\mathbf{0})(x, y, z) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x, x + y + z).$$

Es decir, es la aplicación lineal dada por la matriz jacobiana de f en $(0, 0, 0)$.

☻

Función implícita. Derivación implícita

Consideramos una función $f : A \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$. Se dice que la ecuación $f(x, y) = 0$ define a y como función implícita de x en un entorno $U \subset A$ si existe un entorno $V \subset \mathbb{R}$ y una aplicación $g : V \longrightarrow \mathbb{R}$ tales que los puntos $(x, y) \in U$ con $x \in V, y = g(x)$ satisfacen $f(x, y) = f(x, g(x)) = 0$. Expresado matemáticamente:

$$\{(x, y) \in U : f(x, y) = 0\} \supset \{(x, y) \in U : x \in V, y = g(y)\}.$$

Esto es lo mismo que decir si podemos escribir y en función de x (para $x \in V$), o si podemos "despejar" y (es decir, cuando $y = g(x)$ o, abusando de la notación $y = y(x)$), entonces en estos puntos se cumple la ecuación $f(x,y) = 0$, para $(x,y) \in U$.

Obsérvese que esta definición se puede reescribir para "despejar" x.

Ejemplo 25. Se tiene la función $f : [0,\infty) \times [0,\infty) \longrightarrow \mathbb{R}$ definida por $f(x,y) = x^2 + y^2 - 5$. Entonces si consideramos la ecuación

$$f(x,y) = x^2 + y^2 - 5 = 0,$$

vemos que la variable x se puede expresar en función de la variable y como

$$x = \sqrt{5 - y^2},$$

en el conjunto $[0,\sqrt{5}] \times [0,\sqrt{5}]$. Esto significa que la función f define a la variable x como función implícita de y mediante $x = g(y) = \sqrt{5 - y^2}$ y que

$$f\left(\sqrt{5 - y^2}, y\right) = 0.$$

Es decir, hemos despejado una de las variables en función de la otra.

Podíamos, de la misma forma, haber expresado y en función de x.

☻

Sea $f : A \subset \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$. Entonces el sistema

$$\begin{gathered} f_1(\mathbf{x},\mathbf{y}) = 0, \\ \cdots \\ f_m(\mathbf{x},\mathbf{y}) = 0, \end{gathered}$$

define a $\mathbf{y} = (y_1,\ldots,y_m)$ como **función implícita** de $\mathbf{x} = (x_1,\ldots,x_n)$ en un entorno $U \subset A$ si existen un abierto $V \subset \mathbb{R}^n$ y una aplicación $g = (g_1,\ldots,g_m) : V \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ tales que $(\mathbf{x}, g(\mathbf{x}) \in U$ y $f(\mathbf{x}, g(\mathbf{x})) = 0$ para todo $\mathbf{x} \in V$.

Ejemplo 26. Se tiene la función $f : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}$ definida por $f(\mathbf{x},t) = f(x,y,t) = e^x - 3y + t^3$. De forma evidente, a partir de la ecuación

$$f(\mathbf{x},t) = f(x,y,t) = e^x - 3y + t^3 = 0$$

podemos despejar la variable t a partir de (x,y):

$$t = \sqrt[3]{3y - e^x}.$$

Por eso, se dice que f define a la variable t como función implícita de (x, y) mediante $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $g(x, y) = \sqrt[3]{3y - e^x}$ y que

$$f\left(x, y, \sqrt[3]{3y - e^x}\right) = 0.$$

En este ejemplo, hemos despejado una de las variables en función de las otras dos.

☻

La definición anterior nos dice que cuando tenemos un sistema de ecuaciones, podemos despejar unas de las incógnitas en función de las otras. Lo difícil es saber cuándo podemos hacerlo (cuándo existe la función g) y aún en el caso de poder hacerlo, no siempre es fácil encontrar la expresión de g.

Si la función f es lineal, encontrar una función implícita se reduce a estudiar si un sistema tiene solución única.

Teorema 13. ***Teorema de la función implícita.*** *Sea $A \subset \mathbb{R}^n \times \mathbb{R}^m$ un abierto, $(\mathbf{a}, \mathbf{b})$ un punto de A y $f : A \subset \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$ una función de clase q en A (es decir, tiene derivadas parciales hasta orden q y son continuas en A). Si se cumple:*

1. $f(\mathbf{a}, \mathbf{b}) = 0$,
2. $\det(D_{n+i} f_j(\mathbf{a}, \mathbf{b})) \neq 0$, *para* $i, j \in \{1, 2, \ldots, m\}$,

Entonces existe un subconjunto abierto $U \subset A$ que contiene a $(\mathbf{a}, \mathbf{b})$, un entorno abierto $V \subset \mathbb{R}^n$ de $\mathbf{a}$ y una única función $g : V \longrightarrow \mathbb{R}^m$ de clase q en V tales que:

1. $\det(D_{n+i} f_j(\mathbf{x}, \mathbf{y})) \neq 0$, *para todo* $(\mathbf{x}, \mathbf{y}) \in U$,
2. $\{(\mathbf{x}, \mathbf{y}) \in U : f(\mathbf{x}, \mathbf{y}) = 0\} = \{(\mathbf{x}, \mathbf{y}) \in U : \mathbf{x} \in V, \mathbf{y} = g(\mathbf{x})\}$.

Observación 14. *Señalemos lo siguiente respecto al teorema de la función implícita:*

- *Este teorema nos da condiciones suficientes que nos permiten asegurar que podemos despejar unas de las incógnitas en función de las otras (existe la función g), pero no nos dice cómo calcular g.*
- *Escribiéndolo así, podemos escribir las últimas m incógnitas en función de las n primeras, pero valdría en cualquier orden. En el caso descrito en el teorema, podemos despejar las variables $y_1, \ldots, y_m$ en función de $x_1, \ldots, x_n$.*

- *Si no se cumple la segunda condición, puede haber una, varias o ninguna función implícita, no lo sabemos (las condiciones no son necesarias).*
- *El teorema tiene carácter local, no nos asegura que se pueda hacer globalmente con expresiones únicas.*
- *Como veremos a continuación, el teorema de la función implícita nos permite determinar las derivadas parciales (de cualquier orden) de g.*

A partir del teorema de la función implícita, podemos determinar las derivadas parciales de g. Abusando de la notación, podemos escribir, suponiendo que se cumplen las condiciones del teorema de la función implícita, que:

$$\mathbf{y} = g(\mathbf{x}) = \mathbf{y}(\mathbf{x}) = (y_1(\mathbf{x}), \cdots y_m(\mathbf{x})).$$

En ese caso, tenemos que para los puntos $(\mathbf{x}, \mathbf{y}(\mathbf{x})) \in U$, se cumple $f(\mathbf{x}, \mathbf{y}(\mathbf{x})) = 0$ o, para cada componente

$$f_i(\mathbf{x}, \mathbf{y}(\mathbf{x})) = 0.$$

Así, queremos determinar $\dfrac{\partial y_i}{\partial x_k}$, para $i = 1, \ldots, m,\ k = 1, \ldots, n$. Con este fin, derivamos la última expresión respecto a la variable x_k, resulta:

$$\frac{\partial f_i}{\partial x_k} + \frac{\partial f_i}{\partial y_1}\frac{\partial y_1}{\partial x_k} + \frac{\partial f_i}{\partial y_2}\frac{\partial y_2}{x_k} + \cdots + \frac{\partial f_i}{\partial y_m}\frac{\partial y_m}{\partial x_k} = 0.$$

Para cada k fijo, como $i = 1, \ldots, m$, tenemos un sistema lineal de m ecuaciones y m incógnitas. La matriz de los coeficientes es $\left(\dfrac{\partial f_i}{\partial y_j}\right)_{i,j=1,\ldots m}$ y su fila i es el vector formado por las derivadas parciales de f_i respecto a las variables $(y_1, \ldots, \ldots, y_m)$. Si lo escribimos utilizando la notación matricial, tenemos

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial y_1} & \cdots & \dfrac{\partial f_m}{\partial y_m} \end{pmatrix} \begin{pmatrix} \dfrac{\partial y_1}{\partial x_k} \\ \vdots \\ \dfrac{\partial y_m}{\partial x_k} \end{pmatrix} = - \begin{pmatrix} \dfrac{\partial f_1}{\partial x_k} \\ \vdots \\ \dfrac{\partial f_m}{\partial x_k} \end{pmatrix}.$$

Si llamamos A a la matriz de las derivadas parciales de f_i con respecto a y_j, entonces el determinante de A es distinto de 0 por el teorema de la función implícita, y por eso, tiene inversa, A^{-1}. La matriz columna de los términos independientes es $\left(-\dfrac{\partial f_i}{\partial x_k}\right)_{i=1,\dots m}$. Pero además, considerando $k = 1,\dots,n$, tenemos la siguiente ecuación matricial:

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial y_1} & \cdots & \dfrac{\partial f_m}{\partial y_m} \end{pmatrix} \begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial y_m}{\partial x_1} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{pmatrix} = -\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix},$$

donde las incógnitas son $\left(\dfrac{\partial y_i}{\partial x_k}\right)_{\substack{i=1,\dots,m \\ k=1,\dots,n}}$. Como A tiene inversa, podemos obtener su solución:

$$A\begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial y_m}{\partial x_1} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{pmatrix} = -\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix} \Longrightarrow$$

$$\Longrightarrow \begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial y_m}{\partial x_1} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{pmatrix} = -A^{-1}\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Otra forma de resolverla es utilizando la regla de Cramer para k fijo. En ese caso:

$$\frac{\partial y_i}{\partial x_k} = \frac{\dfrac{\partial(f_1,\dots,f_m)}{\partial(y_1,\dots,y_{i-1},-x_i,y_{i+1},\dots,y_m)}(\mathbf{x},y(\mathbf{x}))}{\dfrac{\partial(f_1,\dots,f_m)}{\partial(y_1,\dots,\dots,y_m)}(\mathbf{x},y(\mathbf{x}))},$$

donde $\dfrac{\partial(f_1,\dots,f_m)}{\partial(y_1,\dots,\dots,y_m)}(\mathbf{x},y(\mathbf{x}))$ es el determinante de la matriz de los coeficientes y $\dfrac{\partial(f_1,\dots,f_m)}{\partial(y_1,\dots,y_{i-1},-x_i,y_{i+1},\dots,y_m)}$ es el determinante de la matriz cuya fila j es el vector formado por las derivadas parciales de f_j respecto a $(y_1,\dots,y_{i-1},-x_i,y_{i+1},\dots,y_m)$. Esta es una expresión general, pero el sistema resultante se puede resolver por cualquier método.

Para derivadas parciales de orden 2 o superior, seguimos este mismo procedimiento.

Ejemplo 27. Sea $f:\mathbb{R}^5 \to \mathbb{R}^2$ definida por

$$f(x,y,z,u,v) = \left(u+v+x^2-y^2+z^2,\, u^2+v^2+u-2xyz\right)$$

1. Estudiar si f define a u y v como funciones implícitas diferenciables $(u,v) = (g_1(x,y,z), g_2(x,y,z))$ en un entorno de $\left(0,0,0,-\frac{1}{2},\frac{1}{2}\right)$.

2. Obtener la matriz jacobiana de dicha función implícita en caso de que exista.

Solución:

1. Tenemos $f(x,y,z,u,v) = (f_1(x,y,z,u,v), f_2(x,y,z,u,v))$, donde

$$f_1(x,y,z,u,v) = u+v+x^2-y^2+z^2, \quad f_2(x,y,z,u,v) = u^2+v^2+u-2xyz,$$

que son funciones continuas en $\mathbb{R}^5$. Además,

$$f_1\left(0,0,0,-\frac{1}{2},\frac{1}{2}\right) = 0, \qquad f_2\left(0,0,0,-\frac{1}{2},\frac{1}{2}\right) = 0,$$

$$\left|\begin{array}{cc} \dfrac{\partial f_1}{\partial u} & \dfrac{\partial f_1}{\partial v} \\ \dfrac{\partial f_2}{\partial u} & \dfrac{\partial f_2}{\partial v} \end{array}\right|_{\left(0,0,0,-\frac{1}{2},\frac{1}{2}\right)} = \left|\begin{array}{cc} 1 & 1 \\ 2u+1 & 2v \end{array}\right|_{\left(0,0,0,-\frac{1}{2},\frac{1}{2}\right)} = \left|\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right| \neq 0.$$

Luego se cumplen las condiciones para la existencia de función implícita y existe un subconjunto U de $\mathbb{R}^5$ que contiene a $\left(0,0,0,-\frac{1}{2},\frac{1}{2}\right)$, un entorno V de $(0,0,0)$ y una función $g(x,y,z) = (g_1(x,y,z), g_2(x,y,z))$ tales que:

$$\begin{aligned} \{(x,y,z,u,v) \in U : f(x,y,z,u,v) = 0\} &= \\ &= \{(x,y,z,u,v) \in U : (x,y,z) \in V, (u,v) = g(x,y,z,u,v)\} \\ &= \{(x,y,z,g(x,y,z)) : (x,y,z) \in V\}. \end{aligned}$$

2. Llamando a las componentes de la función implícita g_1 y g_2, la matriz jacobiana en $P = \left(0,0,0,-\frac{1}{2},\frac{1}{2}\right)$ será

$$\begin{aligned} \begin{pmatrix} \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} & \dfrac{\partial g_1}{\partial z} \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} & \dfrac{\partial g_2}{\partial z} \end{pmatrix} &= \begin{pmatrix} \dfrac{\partial f_1}{\partial u} & \dfrac{\partial f_1}{\partial v} \\ \dfrac{\partial f_2}{\partial u} & \dfrac{\partial f_2}{\partial v} \end{pmatrix}_P^{-1} \begin{pmatrix} -\dfrac{\partial f_1}{\partial x} & -\dfrac{\partial f_1}{\partial y} & -\dfrac{\partial f_1}{\partial z} \\ -\dfrac{\partial f_2}{\partial x} & -\dfrac{\partial f_2}{\partial y} & -\dfrac{\partial f_2}{\partial z} \end{pmatrix}_P \\ &= \begin{pmatrix} 1 & 1 \\ 2u+1 & 2v \end{pmatrix}_P^{-1} \begin{pmatrix} -2x & 2y & -2z \\ 2yz & 2xz & 2xy \end{pmatrix}_P \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Podíamos haber derivada implícitamente y luego haber resuelto el sistema. Partimos de

$$
\begin{aligned}
&0 = f\left(x, y, z, g_1\left(x, y, z\right), g_2\left(x, y, z\right)\right) \Longrightarrow \\
&\Longrightarrow \left\{ \begin{array}{l} 0 = g_1\left(x, y, z\right) + g_2\left(x, y, z\right) + x^2 - y^2 + z^2, \\ 0 = g_1^2\left(x, y, z\right) + g_2^2\left(x, y, z\right) + g_1\left(x, y, z\right) - 2xyz. \end{array} \right.
\end{aligned}
$$

Derivamos estas expresiones respecto a x, y, z:

$$
\begin{aligned}
(0,0) &= \left(\frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial x} + 2x, 2g_1\frac{\partial g_1}{\partial x} + 2g_2\frac{\partial g_2}{\partial x} + \frac{\partial g_1}{\partial x} - 2yz\right), \\
(0,0) &= \left(\frac{\partial g_1}{\partial y} + \frac{\partial g_2}{\partial y} - 2y, 2g_1\frac{\partial g_1}{\partial y} + 2g_2\frac{\partial g_2}{\partial y} + \frac{\partial g_1}{\partial y} - 2xz\right), \\
(0,0) &= \left(\frac{\partial g_1}{\partial z} + \frac{\partial g_2}{\partial z} + 2z, 2g_1\frac{\partial g_1}{\partial z} + 2g_2\frac{\partial g_2}{\partial z} + \frac{\partial g_1}{\partial z} - 2xy\right).
\end{aligned}
$$

Tenemos un sistema de 6 ecuaciones y 6 incógnitas, que particularizando en $x = y = z = 0$, son:

$$
\begin{aligned}
0 &= \frac{\partial g_1}{\partial x}(0,0,0) + \frac{\partial g_2}{\partial x}(0,0,0), \\
0 &= -\frac{\partial g_1}{\partial x}(0,0,0) + \frac{\partial g_2}{\partial x}(0,0,0) + \frac{\partial g_1}{\partial x}(0,0,0) = \frac{\partial g_2}{\partial x}(0,0,0), \\
0 &= \frac{\partial g_1}{\partial y}(0,0,0) + \frac{\partial g_2}{\partial y}(0,0,0), \\
0 &= -\frac{\partial g_1}{\partial y}(0,0,0) + \frac{\partial g_2}{\partial y}(0,0,0) + \frac{\partial g_1}{\partial y}(0,0,0) = \frac{\partial g_2}{\partial y}(0,0,0), \\
0 &= \frac{\partial g_1}{\partial z}(0,0,0) + \frac{\partial g_2}{\partial z}(0,0,0), \\
0 &= -\frac{\partial g_1}{\partial z}(0,0,0) + \frac{\partial g_2}{\partial z}(0,0,0) + \frac{\partial g_1(0,0,0)}{\partial z} = \frac{\partial g_2}{\partial z}(0,0,0).
\end{aligned}
$$

Observe que aquí se ha tenido en cuenta que $g_1(0,0,0) = -\frac{1}{2}$, $g_2(0,0,0) = \frac{1}{2}$. Obviamente, la solución de estas ecuaciones son

$$
\begin{aligned}
\frac{\partial g_1}{\partial x}(0,0,0) &= \frac{\partial g_2}{\partial x}(0,0,0) = 0, \\
\frac{\partial g_1}{\partial y}(0,0,0) &= \frac{\partial g_2}{\partial y}(0,0,0) = 0, \\
\frac{\partial g_1}{\partial z}(0,0,0) &= \frac{\partial g_2}{\partial z}(0,0,0) = 0.
\end{aligned}
$$

☻

Ejemplo 28. Se considera la función $f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ definida por

$$f(x, y, z) = 1 - e^{(x^7+y^2-xz-1)}.$$

Vamos a:

1. Encontrar el valor de a tal que en un entorno del punto $(1, 0, a)$ la ecuación $f(x, y, z)$ defina implícitamente una función $z = g(x, y)$.
2. Calcular $\dfrac{\partial g}{\partial x}(1, 0)$ y $\dfrac{\partial g}{\partial y}(1, 0)$.
3. Calcular $\dfrac{\partial^2 g}{\partial x^2}(1, 0)$, $\dfrac{\partial^2 g}{\partial y^2}(1, 0)$ $\dfrac{\partial^2 g}{\partial x \partial y}(1, 0)$ y $\dfrac{\partial^2 g}{\partial y \partial x}(1, 0)$.

Comencemos:

1. Para que se verifiquen las hipótesis, en primer lugar f debe ser una función diferenciable con continuidad en un entorno de $(1, 0, a)$, lo que se verifica ya que

$$f'(x, y, z) = -\left(7x^6 - z \quad 2y \quad - x\right) e^{(x^7+y^2-xz-1)}.$$

Además, debe ser $f(1, 0, a) = 0$, que es lo mismo que $1 - e^{-a} = 0$, lo que implica $a = 0$. Finalmente, en el punto $(1, 0, 0)$ debe ser $D_3 f(1, 0, 0) \neq 0$. Como $D_3 f(x, y, z) = xe^{(x^7+y^2-xz-1)}$, entonces $D_3 f(1, 0, 0) = 1$, por lo que la función f verifica las hipótesis del teorema de la función implícita en este punto.

2. Al verificar las hipótesis de la función implícita, en un entorno U del punto $(1, 0, 0)$ la función f define a z como función implícita de las otras variables. Si $z = g(x, y)$, entonces en U

$$f(x, y, g(x, y)) = 1 - e^{(x^7+y^2-xg(x,y)-1)} = 0.$$

Si derivamos esta expresión respecto a x y a y, tenemos

$$-\left(7x^6 - g(x, y) - x\frac{\partial g}{\partial x}(x, y)\right) e^{(x^7+y^2-xg(x,y)-1)} = 0,$$
$$-\left(2y - x\frac{\partial g}{\partial y}(x, y)\right) e^{(x^7+y^2-xg(x,y)-1)} = 0.$$

Particularizando para $x = 1$, $y = 0$ y $g(1, 0) = 0$, si sustituimos estas condiciones en la ecuación anterior obtenemos los valores de las derivadas parciales de g respecto a x y a y

$$\frac{\partial g}{\partial x}(1, 0) = 7 \qquad \frac{\partial g}{\partial y}(1, 0) = 0.$$

3. Derivando las ecuaciones del punto anterior respecto a las dos variables obtenemos:

$$
\begin{aligned}
0 &= \Bigg(42x^5 - 2\frac{\partial g}{\partial x}(x,y) - x\frac{\partial^2 g}{\partial x^2}(x,y) \\
&\quad + \left(7x^6 - g(x,y) - x\frac{\partial g}{\partial x}(x,y)\right)^2\Bigg) e^{(x^7+y^2-xg(x,y)-1)}, \\
0 &= \Bigg(\frac{\partial g}{\partial y}(x,y) - x\frac{\partial^2 g}{\partial x \partial y}(x,y) \\
&\quad + \left(7x^6 - g(x,y) - x\frac{\partial g}{\partial x}(x,y)\right)\left(2y - x\frac{\partial g}{\partial y}(x,y)\right)\Bigg) e^{(x^7+y^2-xg(x,y)-1)}, \\
0 &= \left(2 - x\frac{\partial^2 g}{\partial y^2}(x,y) + \left(2y - x\frac{\partial g}{\partial y}(x,y)\right)^2\right) e^{(x^7+y^2-xg(x,y)-1)}, \\
0 &= \Bigg(\frac{\partial g}{\partial y}(x,y) - x\frac{\partial^2 g}{\partial x \partial y}(x,y) \\
&\quad + \left(2y - x\frac{\partial g}{\partial y}(x,y)\right)\left(7x^6 - g(x,y) - x\frac{\partial g}{\partial x}(x,y)\right)\Bigg) e^{(x^7+y^2-xg(x,y)-1)}.
\end{aligned}
$$

Sustituyendo $(x,y,g(x,y)) = (1,0,0)$, considerando los valores obtenidos en el segundo punto, se obtiene

$$
\frac{\partial^2 g}{\partial x^2}(x,y) = 28 \qquad \frac{\partial^2 g}{\partial x \partial y}(x,y) = 0 = \frac{\partial^2 g}{\partial y \partial x}(x,y) \qquad \frac{\partial^2 g}{\partial y^2}(x,y) = 2.
$$

☻

Función inversa. Cambio de variable.

Consideramos dos espacios vectoriales, que llamamos E, F donde hay definida una distancia, y sea $f : E \longrightarrow F$ una aplicación entre ellos. En general, está aplicación no es biyectiva y por eso, no existe f^{-1}. Pero, dado un punto $a \in E$, sí que podemos dar condiciones para que existan un entorno V de $\mathbf{a}$ y otro entorno W de $f(\mathbf{a})$ para que haya una biyección entre V y W, y que así, exista f^{-1} en W.

Vamos a centrarnos en espacios normados de dimensión n, es decir, en $\mathbb{R}^n$.

Sea $f = (f_1, \ldots, f_n) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, donde

$$
f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)) = (y_1, \ldots, y_n).
$$

Si f es inyectiva, entonces su función inversa definida sobre $f(\mathbb{R}^n)$ se puede obtener si sabemos resolver el siguiente sistema (expresando las x_i en función de las y_i):

$$\begin{aligned} y_1 &= f_1(x_1, \ldots, x_n), \\ &\cdots \\ y_n &= f_n(x_1, \ldots, x_n). \end{aligned}$$

Como esto es difícil en general, buscamos condiciones que nos permitan asegurar que esto se puede hacer (o afirmar que f tiene función inversa). Si f es una función lineal, en principio, podemos hacerlo (es un sistema de n ecuaciones lineales y n incógnitas). Si f es diferenciable, entonces podemos utilizando el teorema de Taylor sabemos que $F(\mathbf{x}) = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x} - \mathbf{a})$ es una aproximación lineal de f en un entorno V de $\mathbf{a}$, porque

$$\lim_{x \to a} \frac{\|F(\mathbf{x}) - f(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{a}\|} = 0.$$

Como la matriz de la aplicación lineal $Df(\mathbf{a})$ es su matriz jacobiana, entonces parece lógico pensar que f tiene inversa en un entorno de $\mathbf{a}$ si $\det f'(\mathbf{a}) \neq 0$.

Teorema 15. *Teorema de la función inversa. Sean A un abierto de $\mathbb{R}^n$, $f = (f_1, \ldots, f_n) : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una función de clase q en A, $\mathbf{a} \in A$ y $\det f'(\mathbf{a}) \neq 0$. Entonces existen un entorno abierto V de $\mathbf{a}$ y un entorno abierto W de $f(\mathbf{a})$ tales que*

1. *$f(V) = W$,*
2. *La restricción de f a V tiene inversa $f^{-1} : W \longrightarrow V$ de clase q en W.*
3. *Además, para cada $y \in W$ se cumple $(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1}$.*

Ejemplo 29. Sea f la función $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida como

$$f(x, y) = (\cos x + \cos y, \operatorname{sen} x + \operatorname{sen} y).$$

Vamos a buscar la o las opciones correctas:

a. Es localmente inversible en cada punto de $\mathbb{R}^2$.

b. Posee inversa global.

c. No posee inversa global.

d. Ninguna de las anteriores.

Observamos que

$$f(x,y) = f(x+2\pi, y+2\pi).$$

Por tanto, no es inyectiva y no posee inversa global en $\mathbb{R}^2$.

Además, no en todos los puntos ha inversa local. Por ejemplo, en los puntos de la forma $(\varepsilon, -\varepsilon)$ y $(-\varepsilon, \varepsilon)$, para ε suficientemente pequeño y positivo, tenemos:

$$f(\varepsilon, -\varepsilon) = f(-\varepsilon, \varepsilon),$$

por lo que la función no es inyectiva. Como no es inyectiva en un entorno de $(0,0)$, entonces no va tener inversa local cerca de este punto. Por eso, no son ciertas ni la opción a) ni la opción b). Sólo es correcta la opción c).

☻

Ejemplo 30. Considérense las aplicaciones $f, g : \mathbb{R}^3 \to \mathbb{R}^3$ definidas por

$$f(x,y,z) = (e^{2x}, -ze^{y^2}, e^{-yz}),$$
$$g(x,y,z) = (x+y, x^2+z, xz).$$

Tenemos las opciones siguientes:

a. La función f admite inversa diferenciable en el punto $(0,0,0)$.

b. La función f admite inversa diferenciable en el punto $(1,0,1)$.

c. La función $g \circ f$ admite inversa diferenciable en el punto $(1,0,1)$.

d. Ninguna de las anteriores.

Veamos que las opciones correctas son la b) y c).

Comprobamos dónde es diferenciable la función f. Es de clase C^∞ en $\mathbb{R}^3$, por ser cada una de las tres componentes de f de clase C^∞ en $\mathbb{R}$ (por tratarse de productos y composiciones de la función exponencial con funciones polinómicas). Según el teorema de la función inversa, f admite inversa diferenciable en un entorno de cada punto donde el determinante jacobiano de f sea distinto de 0. Puesto que

$$\det f'(x,y,z) = \begin{vmatrix} 2e^{2x} & 0 & 0 \\ 0 & -2yze^{y^2} & -e^{y^2} \\ 0 & -ze^{-yz} & -ye^{-yz} \end{vmatrix} = 4y^2ze^{2x+y^2-yz} - 2ze^{2x+y^2-yz}$$
$$= 2e^{2x+y^2-yz}z(2y^2-1),$$

f admite inversa diferenciable en los puntos $(x,y,z) \in \mathbb{R}^3$ tales que $z(2y^2-1) \neq 0$, es decir, en el conjunto

$$A = \{(x,y,z) \in \mathbb{R}^3 : z \neq 0 \text{ o } y \neq \pm\frac{\sqrt{2}}{2}\}.$$

Por otro lado, las funciones f y g son de clase C^∞ en $\mathbb{R}^3$, por tanto, su composición $f \circ g$ también lo es. Calculemos el jacobiano de g:

$$\det g'(x,y,z) = \begin{vmatrix} 1 & 1 & 0 \\ 2x & 0 & 1 \\ z & 0 & x \end{vmatrix} = z - 2x^2.$$

En virtud de la regla de la cadena, se tiene que

$$\det(g \circ f)'(x,y,z) = \det g'(f(x,y,z)) \det f'(x,y,z).$$

Por consiguiente, $g \circ f$ admite inversa diferenciable en los puntos $(x,y,z) \in \mathbb{R}^3$ tales que $\det g'(f(x,y,z)) \neq 0$ y $\det f'(x,y,z) \neq 0$. Obsérvese que

$$\det g'(f(x,y,z)) = \det g'(e^{2x}, -ze^{y^2}, e^{-yz}) = e^{-yz} - 2e^{4x} = 0 \Longleftrightarrow e^{4x+yz} = \frac{1}{2}$$
$$\Longleftrightarrow 4x + yz = -\ln 2,$$

luego $g \circ f$ admite inversa diferenciable en el conjunto

$$A = \{(x,y,z) \in \mathbb{R}^3 : 4x + yz \neq -\ln 2, z \neq 0, y \neq \pm\frac{\sqrt{2}}{2}\}.$$

☻

El teorema de la función inversa se puede aplicar en casos muy concretos de funciones con una finalidad muy concreta. Son los **cambios de variable.**

Proposición 16. *Si $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una función de clase 1 en A y $\det f'(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in A$ entonces la imagen de cualquier subconjunto abierto de A es un subconjunto abierto.*

Definición 17. *Cambio de variable. Sea A un abierto de $\mathbb{R}^n$. Una aplicación $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ se dice que es un cambio de variable, aplicación regular o difeomorfismo de clase q si:*

1. $f \in C^q(A)$, $q \geq 1$.
2. $\det f'(\mathbf{x}) \neq 0$ *para todo* $x \in A$.

3. f es inyectiva en A.

Proposición 18. *Si f es una aplicación regular de clase q en A (es decir, continua y con derivadas parciales hasta orden q continuas), entonces $f(A)$ es abierto y $f^{-1} : f(A) \longrightarrow A$ es de clase q en $f(A)$.*

Ejemplo 31. Cambio de coordenadas polares a cartesianas y viceversa. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por

$$f(\rho, \theta) = (x, y) = (\rho \cos\theta, \rho \operatorname{sen}\theta).$$

Como $\cos\theta = \cos(\theta + 2\pi)$ y $\operatorname{sen}\theta = \operatorname{sen}(\theta + 2\pi)$ no es inyectiva global. Es inyectiva considerando que está definida sobre

$$A = \left\{(\rho, \theta) \in \mathbb{R}^2 : \rho > 0, -\pi < \theta < \pi\right\}.$$

Y en ese caso, f es un cambio de variable entre A y $f(A)$, donde $f(A)$ es $\mathbb{R}^2$ menos el semieje x negativo.

Las ecuaciones inversas son:

$$\rho = \sqrt{x^2 + y^2}, \ \theta = \operatorname{arctg}\frac{y}{x} \ \text{para} \ (x, y) \in f(A).$$

☻

Ejemplo 32. Cambio de coordenadas cilíndricas a cartesianas y viceversa. Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$f(\rho, \theta, z) = (x, y, z) = (\rho \cos\theta, \rho \operatorname{sen}\theta, z).$$

La función f es de clase infinito y si jacobiano

$$\det f'(\rho, \theta, z) = \begin{vmatrix} \cos\theta & -\rho \operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho$$

es distinto de 0 si $\rho \neq 0$.

Como $f(\rho, \theta, z) = f(\rho, \theta + 2\pi, z)$, no es inyectiva. Sin embargo, con el mismo A de las coordenadas polares, es inyectiva en $A \times \mathbb{R}$ y f es un cambio de variable de $A \times \mathbb{R}$ sobre $f(A \times \mathbb{R})$.

Las ecuaciones inversas son:

$$\rho = \sqrt{x^2 + y^2}, \ \theta = \operatorname{arctg}\frac{y}{x}, \ z = z \ \text{para} \ (x, y, z) \in f(A \times \mathbb{R}).$$

☻

Ejemplo 33. Cambio de coordenadas esféricas a cartesianas y viceversa. Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$f(\rho,\theta,\varphi) = (x,y,z) = (\rho\cos\theta\operatorname{sen}\varphi, \rho\operatorname{sen}\theta\operatorname{sen}\varphi, \rho\cos\varphi).$$

La función f es de clase infinito y su jacobiano cumple:

$$\det f'(\rho,\theta,z) = \begin{vmatrix} cos\theta\operatorname{sen}\varphi & -\rho\operatorname{sen}\theta\operatorname{sen}\varphi & \rho\cos\theta\cos\varphi \\ \operatorname{sen}\theta\operatorname{sen}\varphi & \rho\cos\theta\operatorname{sen}\varphi & \rho\operatorname{sen}\theta\cos\varphi \\ \cos\varphi & 0 & -\rho\operatorname{sen}\varphi \end{vmatrix} = -\rho^2\operatorname{sen}\varphi$$

es distinto de 0 si $\rho \neq 0$ o $\varphi \neq k\pi$.

Como $f(\rho,\theta,\varphi) = f(\rho,\theta+2\pi,\varphi)$ (por ejemplo) no es inyectiva. Pero si consideramos

$$A = \left\{(\rho,\theta,\varphi) \in \mathbb{R}^3 : \rho > 0, -\pi < \theta < \pi, 0 < \varphi < \pi\right\},$$

entonces f es inyectiva en A y f es un cambio de variable de A sobre $f(A)$.

Las ecuaciones inversas son:

$$\rho = \sqrt{x^2+y^2+z^2},\ \theta = \operatorname{arctg}\frac{y}{x},\ \varphi = \arccos\frac{z}{\sqrt{x^2+y^2+z^2}}$$

para $(x,y,z) \in f(A)$.

☻

Ejemplo 34. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ una función definida en el conjunto $A = [-1,1] \times [-1,1] \subset \mathbb{R}^2$ por $f(x,y) = (x^2,y)$. Estudiamos si f es un cambio de variable de orden 1 en A.

Para ello, partimos del determinante de la matriz jacobiana:

$$\det f'(x,y) = \begin{vmatrix} 2x & 0 \\ 0 & 1 \end{vmatrix} = 2x.$$

Como $\det f'(0,y) = 0$ para todo $y \in [-1,1]$, podemos asegurar que no es un cambio de variable.

☻

Ejemplo 35. Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$ un cambio de variable. ¿Pueden existir $a_1, a_2 \in A$, con $a_1 \neq a_2$ tal que $f(a_1) = f(a_2)$?

La respuesta es no, porque en ese caso no sería inyectiva.

☻

1.4. Combinaciones baricéntricas

Nota importante
El estudio de los apartados Coordenadas baricéntricas y Transformaciones afines no deben hacerse de forma exclusiva con este documento. Deben estudiarse también los apartados correspondientes de "Notas de Geometría Diferencial con aplicaciones".

Cuando queremos describir un objeto matemáticamente, lo que es un primer paso si queremos trabajar con él a través de un ordenador, necesitamos un sistema de coordenadas. Un ejemplo conocido son las coordenadas de un vector de un espacio vectorial a partir de los elementos de una base suya, o los cambios de coordenadas. Además nos interesa el objeto en sí y no su relación con un sistema de coordenadas determinado. Esto significa que vamos a trabajar con herramientas que son independientes del sistema de coordenadas elegido.

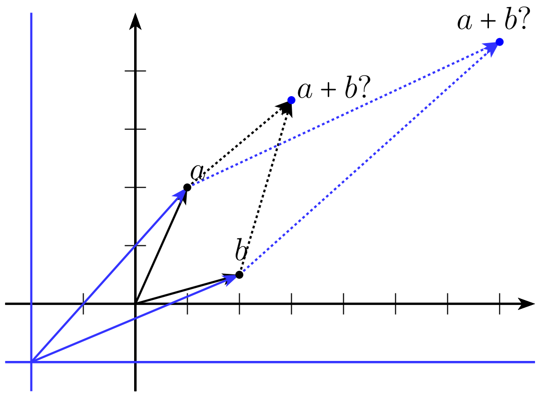
Vamos a trabajar en un espacio bidimensional (o tridimensional), que vamos a llamar espacio afín. Los denotamos, respectivamente, $\mathbb{E}^2$ y $\mathbb{E}^3$. Sus elementos son puntos, que vamos a denotar con las letras (minúsculas en negrita) $\mathbf{a}$, $\mathbf{b}$, $\mathbf{x}$, $\mathbf{y}$ y sucesivas. Un punto identifica una posición, como por ejemplo, el punto medio de un segmento, el baricentro de un triángulo o el centro de una esfera.

Pero además, consideramos $\mathbb{R}^2$ y $\mathbb{R}^3$, que son espacios vectoriales donde hay un producto escalar, y están definidos distancia y ángulos. Sus elementos son vectores, que denotamos con las letras, también minúsculas en negrita, $\mathbf{u}$,$\mathbf{v}$, $\mathbf{w}$, y sucesivas.

Podemos escribir los puntos y los vectores como matrices fila o como matrices columna.

Aunque escribimos igual los puntos y los vectores, no son lo mismo. Los puntos indican posición y los vectores indican movimiento o desplazamiento. Por eso, podemos sumar y restar vectores, y podemos multiplicar vectores por escalares (por números reales, en nuestro caso).

No podemos sumar puntos ni multiplicarlos por escalares, como se ve en la siguiente figura.



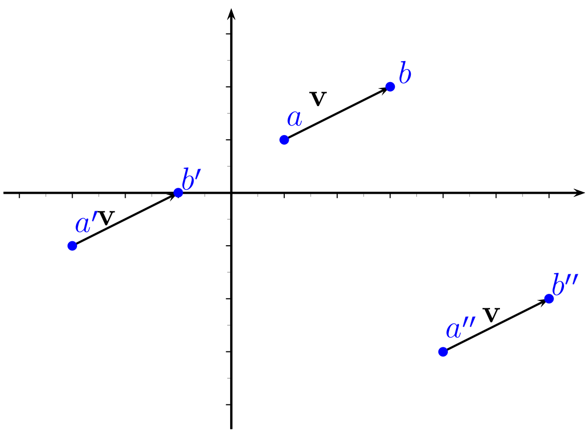
Pero podemos "restarlos" y el resultado es un vector. Por un lado, dos puntos cualesquiera, **a** y **b** determinan un único vector que va de **a** a **b**. Lo podemos determinar restando las coordenadas una a una:

$$\mathbf{v} = \mathbf{b} - \mathbf{a}.$$

Si **a** y **b** están en $\mathbb{E}^2$, entonces $\mathbf{v} \in \mathbb{R}^2$ y si **a** y **b** están en $\mathbb{E}^3$, $\mathbf{v} \in \mathbb{R}^3$.

Pero, por otro lado, dado un vector $\mathbf{v} \in \mathbb{R}^n$, con $n = 2, 3$, hay infinitos pares de puntos **a** y **b** tales que $\mathbf{v} = \mathbf{b} - \mathbf{a}$. De hecho, si tenemos un par **a** y **b** con esta característica, también lo van a cumplir $\mathbf{a}+\mathbf{w}$ y $\mathbf{b}+\mathbf{w}$ para cualquier vector $\mathbf{w} \in \mathbb{R}^3$.

Un ejemplo de esto lo tenemos en la siguiente figura.



Sin embargo, se pueden definir operaciones similares con los puntos. Por ejemplo, si tenemos un segmento definido por su extremos $\mathbf{a}$ y $\mathbf{b}$ en $\mathbb{E}^2$ o en $\mathbb{E}^3$, entonces su punto medio está dado por la posición

$$\mathbf{p}_m = \frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b} = \mathbf{a} + \frac{1}{2}\left(\mathbf{b} - \mathbf{a}\right).$$

Parece una suma de puntos, pero en realidad es una posición más un desplazamiento (un punto más un vector), al haberlo reescrito como $\mathbf{a} + \frac{1}{2}\left(\mathbf{b} - \mathbf{a}\right)$. Observamos que se hace una "combinación lineal" con los puntos, donde la suma de los coeficientes es 1. De la misma forma, el segmento que une $\mathbf{a}$ y $\mathbf{b}$ está descrito por la ecuación

$$t\mathbf{a} + (1 - t)\,\mathbf{b}, \text{ para } t \in [0, 1].$$

Para cualquier t, el punto resultante está en el segmento y además, para cualquier punto del segmento existe un t de tal forma que se para ese t tenemos el punto. Observamos que en las dos ecuaciones anteriores, la clave está en que la suma de los coeficientes es 1, porque de nuevo podemos escribir

$$t\mathbf{a} + (1 - t)\,\mathbf{b} = \mathbf{b} + \mathbf{t}\left(\mathbf{a} - \mathbf{b}\right),$$

que es de la suma de un punto y un vector.

Podemos extender esta idea a un número finito de puntos y a las **coordenadas baricéntricas**. Para los puntos $\mathbf{a}_i \in \mathbb{E}^2, \mathbb{E}^3$ y para $\lambda_i \in \mathbb{R}$, con $i = 0, \dots, m$ y con $\sum_{i=0}^{m} \lambda_i = 1$, decimos que

$$\mathbf{a} = \sum_{i=0}^{m} \lambda_i \mathbf{a}_i$$

es la **combinación baricéntrica** de los puntos $\mathbf{a}_i$ con pesos (o coordenadas baricéntricas) λ_i. Señalamos que aunque parece una suma de puntos, en realidad se puede reescribir como la suma de un punto más una combinación lineal de vectores, porque

$$\mathbf{a} = \sum_{i=0}^{m} \lambda_i \mathbf{a}_i = \mathbf{a}_0 + \sum_{i=1}^{m} \lambda_i \left(\mathbf{a}_i - \mathbf{a}_0\right).$$

Ejemplo 36. El centro de gravedad del cuadrado de vértices $(0,0), (1,0), (1,1)$ y $(0,1)$ es la combinación baricéntrica de estos puntos de coordenadas $\lambda_i = \frac{1}{4}$:

$$\mathbf{g} = \frac{1}{4}(0,0) + \frac{1}{4}(1,0) + \frac{1}{4}(1,1) + \frac{1}{4}(0,1) = \left(\frac{1}{2}, \frac{1}{2}\right).$$

De hecho, el baricentro es el centro de gravedad, y señalamos que el nombre de coordenadas baricéntricas se deriva del término baricentro. ☻

Para entender bien las coordenadas baricéntrica, vamos a definir referencia afín. Se dice los puntos $\mathbf{p}_0, ..., \mathbf{p}_n$ forman una **referencia afín** de $\mathbb{R}^n$ si los vectores $\mathbf{p}_0\mathbf{p}_1, \ldots, \mathbf{p}_0\mathbf{p}_n$ forman una base de $\mathbb{R}^n$. Si tenemos una referencia afín $\mathbf{p}_0, ..., \mathbf{p}_n$, entonces cualquier punto $\mathbf{p}$ se escribe de forma única como una combinación baricéntrica

$$\mathbf{p} = \sum_{i=0}^{n} \lambda_i \mathbf{p}_i.$$

Ejemplo 37. Vamos a comprobar que los puntos $\mathbf{p}_0 = (1,1)$, $\mathbf{p}_1 = (2,1)$ y $\mathbf{p}_2 = (2,-1)$ son una referencia afín?

Veamos que los siguientes vectores son linealmente independientes:

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{p}_1 - \mathbf{p}_0 = (2,1) - (1,1) = (1,0)\,, \\ \mathbf{v}_2 &= \mathbf{p}_2 - \mathbf{p}_0 = (2,-1) - (1,1) = (1,-2)\,. \end{aligned}$$

El determinante de la matriz que forman es

$$\begin{vmatrix} 1 & 0 \\ 1 & -2 \end{vmatrix} = -2 \neq 0$$

y, por tanto, son base de $\mathbb{R}^2$, ya que son dos vectores linealmente independientes y las bases de $\mathbb{R}^2$ están formadas por dos vectores.

☻

Ejemplo 38. Vamos a encontrar las coordenadas baricéntricas del punto $\mathbf{p} = (0,0)$ en la referencia afín $\mathbf{p}_0 = (1,1)$, $\mathbf{p}_1 = (2,1)$ y $\mathbf{p}_2 = (2,-1)$.

Tenemos que encontrar λ_0, λ_1 y λ_2 tales que

$$(0,0) = \lambda_0(1,1) + \lambda_1(2,1) + \lambda_2(2,-1),$$

con $1 = \lambda_0 + \lambda_1 + \lambda_2$. Entonces

$$\begin{aligned} (0,0) &= \lambda_0(1,1) + \lambda_1(2,1) + (1 - \lambda_0 - \lambda_1)(2,-1) \\ &= (-\lambda_0 + 2, 2\lambda_0 + 2\lambda_1 - 1)\,. \end{aligned}$$

Entonces obtenemos las ecuaciones

$$\left. \begin{aligned} 0 &= -\lambda_0 + 2, \\ 0 &= 2\lambda_0 + 2\lambda_1 - 1. \end{aligned} \right\}$$

Obviamente

$$\begin{aligned} \lambda_0 &= 2, \\ 2\lambda_1 &= -2\lambda_0 + 1 = -4 + 1 = -3 \Longrightarrow \lambda_1 = -\frac{3}{2}. \end{aligned}$$

Utilizar la relación entre las coordenadas baricéntricas obtenemos λ_2:

$$\lambda_2 = 1 - \lambda_0 - \lambda_1 = 1 - 2 + \frac{3}{2} = \frac{1}{2}.$$

Entonces

$$(0,0) = 2(1,1) - \frac{3}{2}(2,1) + \frac{1}{2}(2,-1).$$

☻

Un caso importante de coordenadas baricéntricas son la clausura convexa de un conjunto de puntos. Dado un conjunto de puntos $\mathbf{a}_i, i = 1 \cdots n$, la clausura convexa de estos puntos es el conjunto de combinaciones baricéntricas de los puntos $\{\mathbf{a}_i, i = 1 \cdots n\}$ de modo que las coordenadas baricéntricas (también llamadas pesos), además de sumar 1, son no negativas. Es decir, son los puntos $\mathbf{a}$ tales que

$$\mathbf{a} = \sum_{i=0}^{m} \lambda_i \mathbf{a}_i : \sum_{i=0}^{m} \lambda_i = 1,\ \lambda_i \geq 0 \text{ para todo } i = 0, \ldots, m.$$

En $\mathbb{E}^2$ nos podemos imaginar la clausura convexa de un conjunto de puntos como el conjunto delimitado por una goma elástica que bordea a los puntos, que están señalados con clavos en el plano.

1.5. Transformaciones afines

Una transformación afín o afinidad en el plano está dada por la ecuación

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b},$$

donde $\mathbf{A}$ es una matriz no singular de la forma

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix},$$

y $\mathbf{b} = (b_0, b_1)^t$ representa una traslación.

De la misma forma, podemos considerar las transformaciones afines o afinidades en el espacio como las transformaciones dadas por la ecuación

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b},$$

donde $\mathbf{A}$ es una matriz no singular de la forma

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{pmatrix},$$

y $\mathbf{b} = (b_0, b_1, b_2)^t$ es una traslación.

Si $\mathbf{A} = \mathbf{I}, \mathbf{b} = \mathbf{0}$ entonces la transformación es la **identidad**:

$$f(\mathbf{x}) = \mathbf{I}\mathbf{x} + \mathbf{0} = \mathbf{x}.$$

Si $\mathbf{A} = \mathbf{I}$, entonces la transformación es una **traslación** de vector $\mathbf{b}$:

$$f(\mathbf{x}) = \mathbf{I}\mathbf{x} + \mathbf{b} = \mathbf{x} + \mathbf{b}.$$

Ejemplo 39. La transformación

$$f(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x+1 \\ y \end{pmatrix}$$

es una traslación de vector $(1, 0)$. ☻

Ejemplo 40. La transformación

$$f(\mathbf{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} x+1 \\ y+2 \\ z-3 \end{pmatrix}$$

es una traslación de vector $(1, 2, -3)$.

☻

Ejemplo 41. La transformación en $\mathbb{E}^3$ dada por

$$f(\mathbf{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

es la proyección sobre el plano xy. No es una transformación afín, porque la matriz es singular.

☻

Ejemplo 42. Una homotecia es una transformación afín que a partir de un punto fijo, multiplica las distancias por el mismo número. Si el punto fijo es el origen, está representada por una matriz diagonal $\mathbf{A}$ y un vector $\mathbf{b} = \mathbf{0}$. Por ejemplo, la transformación

$$f(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2\mathbf{x}.$$

☻

Ejemplo 43. Si la matriz $\mathbf{A}$ es una matriz diagonal, y $\mathbf{b} = \mathbf{0}$, entonces la transformación f reescala cada uno de los ejes por los elementos de la diagonal de $\mathbf{A}$ correspondientes. Por ejemplo, la transformación

$$f(\mathbf{x}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x \\ -y \\ z \end{pmatrix}$$

multiplica la primera componente por 2, la segunda por -1 y la tercera permanece igual. ☻

Las transformaciones afines no respetan, en general, cualquier combinación lineal de puntos, porque

$$\begin{aligned} f(\lambda \mathbf{x} + \mu \mathbf{y}) &= \mathbf{A}(\lambda \mathbf{x} + \mu \mathbf{y}) + \mathbf{b} = \lambda \mathbf{A}\mathbf{x} + \mu \mathbf{A}\mathbf{y} + \mathbf{b}, \\ \lambda f(\mathbf{x}) + \mu f(\mathbf{y}) &= \lambda(\mathbf{A}\mathbf{x} + \mathbf{b}) + \mu(\mathbf{A}\mathbf{y} + \mathbf{b}) = \lambda \mathbf{A}\mathbf{x} + \mu \mathbf{A}\mathbf{y} + \lambda \mathbf{b} + \mu \mathbf{b} \\ &\neq f(\lambda \mathbf{x} + \mu \mathbf{y}) \end{aligned}$$

cuando $\lambda + \mu \neq 1$. Es decir, cuando hay una traslación, la afinidad no respeta las combinaciones lineales de dos puntos, excepto si tenemos una combinación baricéntrica.

Este resultado es válido en general, para una combinación baricéntrica. Dados $m+1$ puntos $\mathbf{x}_i$, $\lambda_i \in \mathbb{R}$, con $i = 0, \ldots, m$ y $\sum_{i=0}^{m} \lambda_i = 1$, para cualquier afinidad, se tiene

$$f\left(\sum_{i=0}^{m} \lambda_i \mathbf{x}_i\right) = \sum_{i=0}^{m} \lambda_i f(\mathbf{x}_i).$$

Por este motivo, las afinidades transforman segmentos en segmentos y rectas en rectas. Y además, respetan las proporciones entre segmentos de una recta.

Ejemplo 44. Sea f la transformación dada por

$$f(\mathbf{x}) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Entonces

$$2f(1,0) - f(0,1) = f(2(1,0) - (0,1)).$$

En efecto, como $2 - 1 = 1$, $2(1,0) - (0,1)$ es una combinación baricéntrica de los puntos $(1,0)$ y $(0,1)$ y la transformación es afín, debe cumplirse. Vamos a comprobar la afirmación:

$$\begin{aligned}
f(2(1,0)-(0,1)) &= f(2,-1) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \\
2f(1,0)-f(0,1) &= 2\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
&- \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
&= \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ -2 \end{pmatrix}.
\end{aligned}$$

☻

Si $\mathbf{A}$ es una matriz ortogonal (es decir $\mathbf{A}\mathbf{A}^t = \mathbf{I}$), entonces la transformación es una **isometría**. Conserva las distancias, es decir

$$d(\mathbf{x},\mathbf{y}) = d(f(\mathbf{x}), f(\mathbf{y})).$$

Vamos a demostrarlo viendo que el cuadrado de ambas distancias es el mismo, porque

$$\begin{aligned}
d^2(f(\mathbf{x}), f(\mathbf{y})) &= \|f(\mathbf{x}) - f(\mathbf{y})\|^2 = \langle f(\mathbf{x}) - f(\mathbf{y}), f(\mathbf{x}) - f(\mathbf{y})\rangle \\
&= \langle \mathbf{A}\mathbf{x} + \mathbf{b} - (\mathbf{A}\mathbf{y} + \mathbf{b}), \mathbf{A}\mathbf{x} + \mathbf{b} - (\mathbf{A}\mathbf{y} + \mathbf{b})\rangle \\
&= \langle \mathbf{A}(\mathbf{x}-\mathbf{y}), \mathbf{A}(\mathbf{x}-\mathbf{y})\rangle = (\mathbf{x}-\mathbf{y})^t \mathbf{A}^t\mathbf{A}(\mathbf{x}-\mathbf{y}) \\
&= (\mathbf{x}-\mathbf{y})^t(\mathbf{x}-\mathbf{y}) = \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle \\
&= \|\mathbf{x}-\mathbf{y}\|^2 = d(\mathbf{x},\mathbf{y}).
\end{aligned}$$

Además, si la transformación es una isometría, la matriz $\mathbf{A}$ debe ser ortogonal.

Señalamos que no toda matriz con determinante ± 1 es ortogonal.

Ejemplo 45. La transformación plana dada por

$$\mathbf{A} = \begin{pmatrix} \cos\pi & -\operatorname{sen}\pi \\ \operatorname{sen}\pi & \cos\pi \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

es un giro de π radianes, ya que

$$f(x,y) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}.$$

La transformación plana dada por

$$\mathbf{A} = \begin{pmatrix} \cos\frac{\pi}{2} & -\operatorname{sen}\frac{\pi}{2} \\ \operatorname{sen}\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

es un giro de $\frac{\pi}{2}$ radianes más una traslación con vector $(1,-1)$:

$$f(x,y) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -y+1 \\ x-1 \end{pmatrix}.$$

En general, la transformación plana dada por

$$\mathbf{A} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix},$$

es un giro de θ radianes. ☻

Ejemplo 46. Un giro de θ radianes respecto al eje x está dado por

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

y $\mathbf{b} = \mathbf{0}$. ☻

Un **movimiento** en $\mathbb{E}^2$ es una isometría donde el determinante de la matriz $\mathbf{A}$ es 1. La matriz $\mathbf{A}$ de cualquier movimiento plano es un giro de θ radianes y, por tanto, es de la forma

$$\mathbf{A} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

A menudo, estas transformaciones han sido llamadas movimientos directos y su característica es que conserva la orientación en el plano.

Las isometrías que no son movimientos tienen una matriz $\mathbf{A}$ de la forma

$$\mathbf{A} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}.$$

En estas isometrías no se conserva la orientación.

Ejemplo 47. Determinemos los movimientos que transforman el segmento $[(0,0),(0,1)]$ en el segmento $[(1,0),(2,0)]$. Como es un movimiento, sabemos que f debe estar dada por

$$f(x,y)^t = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\operatorname{sen}\theta + b_0 \\ x\operatorname{sen}\theta + y\cos\theta + b_1 \end{pmatrix}.$$

Además, sabemos que hay dos posibilidades: $(0,0)$ se transforma en $(1,0)$ y $(0,1)$ en $(2,0)$ o $(0,0)$ se transforma en $(2,0)$ y $(0,1)$ en $(1,0)$. El primer caso se reduce a

$$\begin{aligned} f(0,0) &= \begin{pmatrix} 0\cos\theta - 0\,\mathrm{sen}\,\theta + b_0 \\ 0\,\mathrm{sen}\,\theta + 0\cos\theta + b_1 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ f(0,1) &= \begin{pmatrix} 0\cos\theta - 1\,\mathrm{sen}\,\theta + b_0 \\ 0\,\mathrm{sen}\,\theta + 1\cos\theta + b_1 \end{pmatrix} = \begin{pmatrix} -1\,\mathrm{sen}\,\theta + b_0 \\ 1\cos\theta + b_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}. \end{aligned}$$

Tenemos las ecuaciones

$$\left.\begin{array}{r} b_0 = 1, \\ b_1 = 0, \\ -\,\mathrm{sen}\,\theta + 1 = 2, \\ \cos\theta + 0 = 0. \end{array}\right\} \Longrightarrow \left.\begin{array}{r} b_0 = 1, \\ b_1 = 0, \\ \mathrm{sen}\,\theta = -1, \\ \cos\theta = 0, \end{array}\right\}$$

$$\Longrightarrow b_0 = 1, b_1 = 0, \theta = \frac{3\pi}{2}.$$

En este caso, la transformación sería un giro de $\frac{3\pi}{2}$ radiantes más una traslación de vector $(1,0)$

$$f(\mathbf{x}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Para el segundo caso, tenemos

$$\begin{aligned} f(0,0) &= \begin{pmatrix} 0\cos\theta - 0\,\mathrm{sen}\,\theta + b_0 \\ 0\,\mathrm{sen}\,\theta + 0\cos\theta + b_1 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \\ f(0,1) &= \begin{pmatrix} 0\cos\theta - 1\,\mathrm{sen}\,\theta + b_0 \\ 0\,\mathrm{sen}\,\theta + 1\cos\theta + b_1 \end{pmatrix} = \begin{pmatrix} -1\,\mathrm{sen}\,\theta + b_0 \\ 1\cos\theta + b_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Resulta el sistema:

$$\left.\begin{array}{r} b_0 = 2, \\ b_1 = 0, \\ -\,\mathrm{sen}\,\theta + 2 = 1, \\ \cos\theta + 0 = 0. \end{array}\right\} \Longrightarrow \left.\begin{array}{r} b_0 = 2, \\ b_1 = 0, \\ \mathrm{sen}\,\theta = 1, \\ \cos\theta = 0. \end{array}\right\}$$

Supone una traslación de vector $(2,0)$ y un giro de $\dfrac{\pi}{2}$ radianes. ☻

1.6. Ejercicios

1. Consideramos en el espacio de las funciones continuas en el intervalo $[0,1]$, $C[0,1]$ el producto escalar

$$\langle f, g\rangle = \int_0^1 f(x)\, g(x)\, dx,$$

para $f, g \in C[0,1]$. ¿Se puede definir una norma a partir de él?

Solución: Sí, siempre se puede definir una norma a partir del producto escalar. En este caso es:

$$\|f\| = \left(\int_0^1 f^2(x)\, dx\right)^{1/2}.$$

2. Calcúlense el producto escalar de $(1,2,3,4)$ y $(-1,-2,-3,-4)$ y la norma de estos dos vectores.

   Solución: $(1,2,3,4)\cdot(-1,-2,-3,-4) = 1(-1)+2(-2)+3(-3)+4(-4) = -1-4-9-16 = -30$.
   $\|(1,2,3,4)\| = \sqrt{1^2+2^2+3^2+4^2} = \sqrt{30}$.
   $\|(-1,-2,-3,-4)\| = \sqrt{(-1)^2+(-2)^2+(-3)^2+(-4)^4} = \sqrt{30}$.

3. Sea la norma definida a partir de un producto escalar como

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u}\rangle^{1/2}.$$

   Demuéstrese que cumple las siguientes propiedades, para $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$:

   a) $\|\mathbf{u}\| \geq 0$ y $\|\mathbf{u}\| = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.
   b) $\|\lambda\mathbf{u}\| = |\lambda|\,\|\mathbf{u}\|$.
   c) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

   Solución: La demostración de estas propiedades es sencilla teniendo en cuenta las propiedades del producto escalar. Comenzamos:

   a) $\|\mathbf{u}\| \geq 0$ y $\|\mathbf{u}\| = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$. En efecto:

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u}\rangle^{1/2} \geq 0$$

   porque es una raíz positiva. Además,

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u}\rangle^{1/2} = 0 \Longleftrightarrow \langle \mathbf{u}, \mathbf{u}\rangle = 0 \Longleftrightarrow \mathbf{u} = 0$$

   por la primera propiedad del producto escalar.

   b) $\|\lambda\mathbf{u}\| = |\lambda|\,\|\mathbf{u}\|$, porque

$$\|\lambda\mathbf{u}\|^2 = \langle \lambda\mathbf{u}, \lambda\mathbf{u}\rangle = \lambda\langle \mathbf{u}, \lambda\mathbf{u}\rangle = \lambda^2\langle \mathbf{u}, \mathbf{u}\rangle$$

   y esto es equivalente a

$$\|\lambda\mathbf{u}\| = |\lambda|\langle \mathbf{u}, \mathbf{u}\rangle^{1/2} = |\lambda|\,\|\mathbf{u}\|.$$

c) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$. Tenemos:

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}+\mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}+\mathbf{v}\rangle \\
&= \langle \mathbf{u}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\
&= \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \cdot \|\mathbf{v}\| + \|\mathbf{v}\|^2
\end{aligned}$$

por la desigualdad de Cauchy-Schwarz. Entonces:

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \cdot \|\mathbf{v}\| + \|\mathbf{v}\|^2 \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned}$$

lo que implica que

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

4. Sean E un espacio vectorial normado tal que su norma $\|\cdot\|$ proviene de un producto escalar. Probar que si x, $y \in E$ verifican que $\|x+y\| = \|x\| + \|y\|$, entonces x e y son linealmente dependientes.

   Solución: Como $\|x+y\| = \|x\| + \|y\|$ y $\|\cdot\|$ proviene de un producto escalar, entonces $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$. Por otro lado:

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

   De aquí sigue que

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

   y por tanto

   $$\langle x, y\rangle = \|x\|\|y\|.$$

   Razonando por reducción al absurdo, si x e y no son linealmente dependientes entonces se cumple que $x - \lambda y \neq 0$ para todo $\lambda \in \mathbb{R}$. Por consiguiente

   $$\|x - \lambda y\| > 0 \qquad \text{para todo } \lambda \in \mathbb{R}.$$

   Entonces

   $$\langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - 2\lambda\langle x, y\rangle + \lambda^2\langle y, y\rangle > 0 \qquad \text{para todo } \lambda \in \mathbb{R}.$$

   (razonando como en la demostración de la desigualdad de Cauchy-Schwarz) la ecuación anterior en λ no puede tener una raíz real y, por tanto, su discriminante es menor que cero

   $$|\langle x, y\rangle|^2 - \langle y, y\rangle\langle x, x\rangle < 0.$$

Pero como $\langle x, y\rangle = \|x\|\|y\|$ se cumple que

$$|\langle x, y\rangle|^2 - \langle y, y\rangle \langle x, x\rangle = \|x\|^2 |y\|^2 - \langle y, y\rangle \langle x, x\rangle = 0$$

lo cual es una contradicción. Luego existe $\lambda \in \mathbb{R}$ tal que $x = \lambda y$ o equivalentemente x e y son linealmente dependientes.

5. Consideramos en el espacio de las funciones continuas en el intervalo $[0,1]$, $C[0,1]$, el producto escalar

$$\langle f, g\rangle = \int_0^1 f(x) g(x) dx$$

para $f, g \in C[0,1]$. Determine un polinomio de grado 2 que sea ortogonal a $f(x) = x - 3$.

Solución: Si $g(x)$ es un polinomio de grado 2, debe ser de la forma $g(x) = ax^2 + bx + c$. Para que sea ortogonal a $f(x)$ debe cumplirse

$$\begin{aligned}
\langle f, g\rangle &= \int_0^1 f(x) g(x) dx = \int_0^1 (x-3)\left(ax^2 + bx + c\right) dx \\
&= \int_0^1 \left(ax^3 + bx^2 - 3ax^2 + cx - 3bx - 3c\right) dx \\
&= a\int_0^1 x^3 dx + (b-3a)\int_0^1 x^2 dx + (c-3b)\int_0^1 x dx - 3c\int_0^1 dx \\
&= a\frac{1}{4}\, x^4\Big|_0^1 + (b-3a)\,\frac{1}{3}\, x^3\Big|_0^1 + (c-3b)\,\frac{1}{2}\, x^2\Big|_0^1 - 3c\, x\Big|_0^1 \\
&= \frac{a}{4} + \frac{b-3a}{3} + \frac{c-3b}{2} - 3c \\
&= \frac{3a + 4b - 12a + 6c - 18b - 36c}{12} \\
&= \frac{-9a - 14b - 30c}{12} = 0
\end{aligned}$$

Esto ocurre, por ejemplo, si $a = 10, b = 0, c = -3$, siendo en este caso, $g(x) = 10x^2 - 3$

6. Demostrar que la norma de la suma, definida para $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, como

$$\|(x_1, \ldots, x_n)\|_1 = \sum_{i=1}^n |x_i|$$

es, efectivamente, una norma.

Solución: Tenemos que comprobar que cumple las propiedades:

- No degenerada: $\|\mathbf{u}\| \geq 0$ y $\|\mathbf{u}\| = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.
- Homogénea por homotecias: $\|\lambda\mathbf{u}\| = |\lambda| \|\mathbf{u}\|$.
- Desigualdad triangular: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Comenzamos con la propiedad de no degeneración:

$$\| (x_1, \ldots, x_n) \|_1 = \sum_{i=1}^{n} |x_i| \geq 0$$

obviamente, porque cada $|x_i| \geq 0$. Además,

$$0 = \sum_{i=1}^{n} |x_i| \Longleftrightarrow |x_i| = 0 \text{ para todo } i,$$

lo que implica que cada $x_i = 0$, lo que quiere decir que $\|\mathbf{x}\| = 0$.

Ahora vamos a comprobar ahora que es homogéneo por homotecias:

$$\begin{aligned} \|\lambda\mathbf{x}\|_1 = \|\lambda (x_1, \ldots, x_n) \|_1 &= \sum_{i=1}^{n} |\lambda x_i| = \sum_{i=1}^{n} |\lambda| |x_i| \\ &= |\lambda| \sum_{i=1}^{n} |x_i| = \lambda \|\mathbf{x}\|_1 . \end{aligned}$$

Nos queda por comprobar la desigualdad triangular:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_1 = \|(x_1 + y_1, , \ldots, x_n + y_n)\|_\infty &= \sum_{i=1}^{n} |x_i + y_i| \\ &\leq \sum_{i=1}^{n} (|x_i| + |y_i|) = \sum_{i=1}^{n} = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1 \end{aligned}$$

porque $|x_i + y_i| \leq |x_i| + |y_i|$ para cualquiera números $x_i, y_i \in \mathbb{R}$.

7. Demostrar la siguientes propiedades de la norma:

   a) Para $\mathbf{u}, \mathbf{v} \in V$, se verifica

   $$|\|\mathbf{u}\| - \|\mathbf{v}\|| \leq \|\mathbf{u} \pm \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

   b) Para $\mathbf{u}_i \in V$, $\lambda_i \in \mathbb{R}$, se cumple

   $$\left\| \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \right\| < \sum_{i=1}^{n} |\lambda_i| \|\mathbf{u}_i\| .$$

Solución:

a) Demostramos esta propiedad aplicando la desigualdad triangular y la homogeneidad de la norma. Tenemos:

$$\begin{aligned}\|\mathbf{u}\| &= \|\mathbf{u}+\mathbf{v}-\mathbf{v}\| \le \|\mathbf{u}-\mathbf{v}\| + \|\mathbf{v}\| \Longrightarrow \|\mathbf{u}\| - \|\mathbf{v}\| \le \|\mathbf{u}-\mathbf{v}\|\,,\\ \|\mathbf{v}\| &= \|\mathbf{v}+\mathbf{u}-\mathbf{u}\| \le \|\mathbf{v}-\mathbf{u}\| + \|\mathbf{u}\| \Longrightarrow \|\mathbf{v}\| - \|\mathbf{u}\| \le \|\mathbf{u}-\mathbf{v}\|\,.\end{aligned}$$

De aquí resulta $|\|\mathbf{u}\| - \|\mathbf{v}\|| \le \|\mathbf{u} \pm \mathbf{v}\|$ porque

$$\|\mathbf{u}\| - \|\mathbf{v}\| \le \|\mathbf{u}-\mathbf{v}\|\,, \|\mathbf{v}\| - \|\mathbf{u}\| \le \|\mathbf{u}-\mathbf{v}\|\,.$$

Por otro lado

$$\|\mathbf{u}-\mathbf{v}\| \le \|\mathbf{u}\| + \|-\mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|\,,$$

lo que, junto con la desigualdad triangular, implica

$$\|\mathbf{u} \pm \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|\,.$$

b) Se demuestra teniendo en cuenta las propiedades triangular y homogeneidad por homotecias y aplicando inducción sobre n.

8. Sean E un espacio vectorial normado tal que su norma $\|\cdot\|$ proviene de un producto escalar. Probar que si x, $y \in E$ verifican que $\|x+y\| = \|x\| + \|y\|$, entonces x e y son linealmente dependientes.

   Solución: Como $\|x+y\| = \|x\| + \|y\|$ y $\|\cdot\|$ proviene de un producto escalar, entonces $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$. Por otro lado:

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

   De aquí sigue que

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

   y por tanto
   $$\langle x,y\rangle = \|x\|\|y\|.$$

   Razonando por reducción al absurdo, si x e y no son linealmente dependientes entonces se cumple que $x - \lambda y \ne 0$ para todo $\lambda \in \mathbb{R}$. Por consiguiente
   $$\|x-\lambda y\| > 0 \qquad \text{para todo } \lambda \in \mathbb{R}.$$

Entonces

$$\langle x-\lambda y, x-\lambda y\rangle = \langle x,x\rangle - 2\lambda\langle x,y\rangle + \lambda^2\langle y,y\rangle > 0 \qquad \text{para todo } \lambda\in\mathbb{R}.$$

(razonando como en la demostración de la desigualdad de Cauchy-Schwarz) la ecuación anterior en λ no puede tener una raíz real y, por tanto, su discriminante es menor que cero

$$|\langle x,y\rangle|^2 - \langle y,y\rangle\langle x,x\rangle < 0.$$

Pero como $\langle x,y\rangle = \|x\|\|y\|$ se cumple que

$$|\langle x,y\rangle|^2 - \langle y,y\rangle\langle x,x\rangle = \|x\|^2\|y\|^2 - \langle y,y\rangle\langle x,x\rangle = 0$$

lo cual es una contradicción. Luego existe $\lambda\in\mathbb{R}$ tal que $x=\lambda y$ o equivalentemente x e y son linealmente dependientes.

9. En el espacio de las funciones definidas sobre $\mathbb{R}$ y que son dos veces derivables en el intervalo $[-1,1]$, se define

$$\|f\| = \|f\|_0 + \|f'\|_0\,,$$

donde $\|\cdot\|_0$ es la norma del máximo. Pruébese que $\|\cdot\|$ es una norma en este conjunto.

10. Sea $f:\mathbb{R}^2\to\mathbb{R}^2$ la función definida por

$$f(x,y) = \left(|x|+|y|, \frac{\operatorname{sen} x}{x}\right).$$

Estudiar si existen derivadas parciales de f en el punto $(0,0)$ y determinarlas en caso de que existan tal como está definida la función. Si no es posible con esta definición, estudiar si se puede redefinir en $(0,0)$ de tal forma que f tenga en ese punto derivadas parciales. ¿Es f una función diferenciable en $(0,0)$?

Solución: Observamos que

$$f_2(x,y) = \frac{\operatorname{sen} x}{x}$$

no está definido para $x=0$. Por esto, f no está definida en $(0,y)$. En ese caso no existen derivadas de f en el origen, no siendo tampoco diferenciable.

Veamos si podemos redefinir $f(0,0)$ para que la función tenga derivadas parciales en ese punto. Como

$$\lim_{(x,y)\to(0,0)} f_2(x,y) = \lim_{(x,y)\to(0,0)} \frac{\operatorname{sen} x}{x} = \lim_{(x,y)\to(0,0)} \cos x = 1,$$

podríamos considerar $f_2(0,t) = 1$ y $f(0,t) = (|t|, 1)$ para garantizar la continuidad de f en el origen. En este caso,

$$D_1 f_1(0,0) = \lim_{t\to 0} \frac{f_1(t,0) - f_1(0,0)}{t} = \lim_{t\to 0} \frac{|t| - 0}{t}$$

que tampoco existe. Igualmente, no existe $D_2 f_1(0,0)$. Para f_2

$$D_1 f_2(0,0) = \lim_{t\to 0} \frac{f_2(t,0) - f_2(0,0)}{t} = \lim_{t\to 0} \frac{\frac{\operatorname{sen} t}{t} - 1}{t} = \lim_{t\to 0} \frac{\operatorname{sen} t - t}{t^2} = 0$$
$$D_2 f_2(0,0) = \lim_{t\to 0} \frac{f_2(0,t) - f_2(0,0)}{t} = \lim_{t\to 0} \frac{1-1}{t} = 0$$

por lo que sí que existirían las derivadas parciales de f_2. Pero como no existen $D_1 f_1(0,0)$ y $D_2 f_1(0,0)$, en ninguno de los dos casos es f una función diferenciable.

11. Razone si es cierto que una función de clase uno, $f : \mathbb{R}^2 \to \mathbb{R}^2$, con $\det f'(x,y) \neq 0$ en cualquier punto de $\mathbb{R}^2$, cumple:

   a) Es invertible en un entorno de cualquier punto de $\mathbb{R}^2$,

   b) Es inyectiva,

   c) Es lineal.

   d) Ninguna de las anteriores.

   Solución: Es cierta la opción **a)** por el Teorema de la función inversa.

   La opción **b)** no es cierta, porque esto significa que tiene inversa global y no siempre ocurre así.

   La opción **c)** no es cierta, se ve con un contraejemplo: $f(x,y) = (e^x \cos y, e^x \operatorname{sen} y)$.

12. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ la función definida como

$$f(x,y) = (ye^x, xy).$$

   ¿En qué puntos podemos asegurar que f tiene inversa local?

Solución: Comprobamos dónde se cumplen las condiciones del teorema de la función inversa. Como cada una de las componentes son infinitamente derivables, f es una función de clase infinito. El determinante de la matriz jacobiana en un punto (x,y) es

$$\det f'(x,y) = \begin{vmatrix} ye^x & e^x \\ y & x \end{vmatrix} = xye^x - ye^x = ye^x(x-1).$$

Podemos asegurar que tiene inversa local si este determinante es distinto de 0 y esto ocurre si

$$y \neq 0, \quad x \neq 1.$$

Así, los puntos en los que podemos asegurar que f tiene inversa local es en los $\{(x,y) \in \mathbb{R}^2 : y \neq 0 \text{ y } x \neq 1\}$.

13. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ la función definida como

$$f(x,y) = (x^2 - y, x^2 + y).$$

Elija los puntos en los que f tiene inversa local en un entorno suyo:

a) $(0,0)$

b) $(1,1)$

c) $(0,1)$

d) Ninguna de las anteriores.

Nota: Nótese que el Teorema de la función inversa da condiciones suficientes, pero no necesarias para que una función tenga inversa en un punto. Por eso, aunque en un punto no se cumplan las condiciones, la función puede tener inversa, y entonces ahí hay que estudiar si aún así, la función es invertible.

Solución: Es correcta la opción b).

Comprobamos que cumple las condiciones del teorema de la función inversa en cada punto. Como cada una de las componentes son infinitamente derivables, f es una función de clase infinito. El determinante de la matriz jacobiana en un punto (x,y) es

$$\det f'(x,y) = \begin{vmatrix} 2x & -1 \\ 2x & 1 \end{vmatrix} = 2x - (-1)2x = 4x.$$

Podemos asegurar que f tiene inversa local en un entorno de cada punto en donde este determinante es distinto de 0. Como se anula en los puntos de la forma $(0, y)$, sabemos que la respuesta b) es correcta.

Pero aunque en los puntos $(0,0)$ y $(0,1)$ el determinante sea 0, no podemos asegurar que no tenga inversa (el Teorema de la función inversa da condiciones suficientes). Por eso, vamos a buscar una expresión de la posible inversa de f y ver si en entornos de estos puntos, existe.

Podemos calcular la inversa de la función despejando:

$$\left.\begin{array}{l} x^2 + y = u \\ x^2 - y = v \end{array}\right\} \Longrightarrow \left.\begin{array}{l} x^2 = \frac{u+v}{2} \\ y = \frac{u-v}{2} \end{array}\right\}$$

Entonces, como obtenemos una expresión de x^2 en función de u y v, si intentamos despejar, para cada valor de u y v obtenemos dos valores de x. Y por esto, observamos que en cualquier entorno de un punto de la forma $(0, y)$, existen dos puntos en él de la forma (ε_1, y) y $(-\varepsilon_1, y)$, tales que la imagen de estos puntos por f es la misma. Por eso, la función no es inyectiva y por tanto, no tiene inversa local en estos puntos. Así podemos concluir que las opciones a) y c) no son correctas.

14. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ la función dada por $f(x, y) = x^3 - y^3 - y^2 x + 1$. Elija la opción correcta.

    a) La función f define una función implícita $y = \varphi(x)$ en un entorno de $(0, 0)$.

    b) La función f define una función implícita $y = \varphi(x)$ en un entorno de $(0, 1)$.

    c) Ninguna de las anteriores.

    Solución: Es correcta la opción **b)**.

    Veamos que se cumplen las hipótesis del teorema de la función implícita en $(0, 1)$. La función f es de clase C^∞ en $\mathbb{R}^2$ por ser polinómica y se tiene que $f(0, 1) = 0$. Puesto que se trata de despejar la variable y, hay que comprobar que $D_2 f(0, 1) \neq 0$:

    $$D_2 f(x, y) = -3y^2 - 2xy \Longrightarrow D_2 f(0, 1) = -3 \neq 0.$$

    Por el teorema de la función implícita, existe un conjunto abierto $U \subset \mathbb{R}^2$ que contiene a (0,1), un intervalo abierto $V \subset \mathbb{R}$ que contiene al 0 y una única función $\varphi : V \to \mathbb{R}$ de clase C^∞ en V tales que $D_2 f(x,y) \neq 0$ para todo $(x,y) \in U$ y

$$\{(x,y) \in U : f(x,y) = 0\} = \{(x,y) \in U : x \in V, y = \varphi(x)\}.$$

Luego en el entorno U de (0,1) f define la función implícita $y = \varphi(x)$.

En el punto $(0,0)$; no se tiene que $f(0,0) = 0$, por tanto, la opción **a)** no es correcta.

15. Se considera la función $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida por

$$f(x,y,z) = 1 - e^{(x^5+y-xz-1)}.$$

Elija la o las opciones correctas:

a) Si $a = 0$, podemos asegurar que la función (x,y,z) define implícitamente una función $z = g(x,y)$ en un entorno de $(1,0,a)$.

b) Si $a = 1$, podemos asegurar que la función (x,y,z) define implícitamente una función $z = g(x,y)$ en un entorno de $(1,0,a)$.

c) Si $a = -1$, podemos asegurar que la función (x,y,z) define implícitamente una función $z = g(x,y)$ en un entorno de $(1,0,a)$.

d) Ninguna de las anteriores.

Solución: Vamos a ver para qué valores de a se cumplen las condiciones del teorema de la función implícita. Para que se verifiquen las hipótesis, en primer lugar f debe ser una función diferenciable con continuidad en un entorno de $(1,0,a)$, lo que se verifica ya que f tiene diferencial, que es

$$f'(x,y,z) = \left(-\left(5x^4 - z\right) e^{(x^5+y-xz-1)}, -e^{(x^5+y-xz-1)}, xe^{(x^5+y-xz-1)}\right).$$

Y a su vez es continua, por ser producto de funciones continuas. Además, debe ser $f(1,0,a) = 0$, que es lo mismo que $1 - e^{-a} = 0$, lo que implica $a = 0$. Finalmente, en el punto $(1,0,0)$ debe ser $D_3 f(1,0,0) \neq 0$. Como

$$D_3 f(x,y,z) = xe^{(x^5+y-xz-1)},$$

entonces $D_3 f(1,0,0) = 1$, por lo que la función f verifica las hipótesis del teorema de la función implícita en este punto.

16. Sea g la función $x = g(y,z)$ definida implícitamente en un entorno del punto $(x_0, y_0, z_0) = (-1,-1,1)$ por la función $F(x,y,z) = x^2 - \ln(xyz) - (x+z)^2 - 5y - 6 = 0$ ¿es diferenciable en $(y_0, z_0) = (-1,1)$? Razone la respuesta.

    Solución: Sí, porque $D_1F(-1,-1,1) \neq 0$ y $F(-1,-1,1) = 0$. El teorema de la función implícita garantiza que g es diferenciable en su entorno de definición.

17. Sean los puntos $\mathbf{p}_0 = (0,1)$, $\mathbf{p}_1 = (1,1)$, $\mathbf{p}_2 = (2,-1)$. ¿Son una referencia afín? En caso de que lo sean, determínenense las coordenadas baricéntricas $\lambda_0, \lambda_1, \lambda_2$ de un punto (x,y) respecto a esta referencia.

    Solución: Comprobamos que son una referencia afín. Veamos que los siguientes vectores son linealmente independientes:

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{p}_1 - \mathbf{p}_0 = (1,1) - (0,1) = (1,0)\,,\\ \mathbf{v}_2 &= \mathbf{p}_2 - \mathbf{p}_0 = (2,-1) - (0,1) = (2,-2)\,.\end{aligned}$$

Podemos comprobar que son linealmente independientes viendo que no es 0 el determinante de la matriz

$$\begin{vmatrix} 1 & 0 \\ 2 & -2 \end{vmatrix} = -2 \neq 0.$$

Por tanto, base de $\mathbb{R}^2$, ya que son dos vectores linealmente independientes y las bases de $\mathbb{R}^2$ están formadas por dos vectores.

Las coordenadas baricéntricas de un punto (x,y) son

$$(x,y) = \lambda_0(0,1) + \lambda_1(1,1) + \lambda_2(2,-1),$$

con $1 = \lambda_0 + \lambda_1 + \lambda_2$. Entonces

$$\begin{aligned}(x,y) &= \lambda_0(0,1) + \lambda_1(1,1) + (1 - \lambda_0 - \lambda_1)(2,-1)\\ &= (-2\lambda_0 - \lambda_1 + 2, 2\lambda_0 + 2\lambda_1 - 1)\,.\end{aligned}$$

Entonces obtenemos las ecuaciones

$$\left.\begin{aligned} x &= -2\lambda_0 - \lambda_1 + 2, \\ y &= 2\lambda_0 + 2\lambda_1 - 1. \end{aligned}\right\}$$

y restándolas, llegamos a:

$$x + y = \lambda_1 + 1 \Longrightarrow \lambda_1 = x + y - 1.$$

Sustituyendo esta igualdad en la segunda ecuación, tenemos el valor de λ_0:

$$\begin{aligned} 2\lambda_0 &= -x - \lambda_1 + 2 = -x - x - y + 1 + 2 = -2x - y + 3 \\ &\Longrightarrow \lambda_0 = -x - \frac{1}{2}y + \frac{3}{2}. \end{aligned}$$

Y ya sólo nos queda utilizar la relación entre las coordenadas baricéntricas para tener λ_2:

$$\lambda_2 = 1 - \lambda_0 - \lambda_1 = 1 + x + \frac{1}{2}y - \frac{3}{2} - x - y + 1 = \frac{1}{2} - \frac{1}{2}y.$$

Estas son las coordenadas baricéntricas.

18. Dada la afinidad

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

determínese la imagen de la recta $y = x + 2$ por esta aplicación afín.

Solución: Los puntos de la recta $y = x + 2$ son de la forma $(x, x + 2)^t$. Su imagen por f es

$$f(x, x+2) = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ x+2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3x+5 \\ -x-2 \end{pmatrix}.$$

Obtenemos los puntos de la forma

$$\begin{pmatrix} 3x+5 \\ -x-2 \end{pmatrix}.$$

Esto es lo mismo que escribir que son los puntos (x, y) que cumplen:

$$x = 3\lambda + 5, \quad y = -\lambda - 2,$$

o, considerando $\lambda = -y - 2$, la recta

$$x = 3(-y-2) + 5 = -3y - 1.$$

Bibliografía

[1] Farin, G., 2002. Curves and Surfaces for CAGD. 5a ed. Academic Press, San Diego.

[2] Rodríguez Marín, L. 2008. Ampliación de Cálculo, primera parte. 3º Edición. Ed. UNED.

[3] Struik, D. J. ; 1973. Geometría diferencial clásica, Aguilar, Madrid.

[4] Valdés, A.; 2014. Notas de Geometría diferencial con aplicaciones. Consultado el 29 de junio de 2014 en http://www.mat.ucm.es/ avaldes/GDA.pdf, difundido bajo una licencia Creative Commons.