Cálculo en Varias Variables

Apunte del curso

(grupos MA2001-01 y MA2001-08)

Aris Daniilidis

Departamento de Ingeniería Matemática

Equipo docente otoño 2021:

Aris Daniilidis	(cátedra)
Sebastián Bustos Atalah	(auxiliar)
Felipe Hernández Castro	(auxiliar)

(Parte II) Diferenciabilidad

La segunda parte de este curso está enfocada en la noción de diferenciabilidad para funciones entre dos espacios normados: tal función se dice diferenciable en un punto, si se puede aproximar, alrededor de dicho punto, por una función lineal continua entre los mismos espacios. Se estudiarán reglas de cálculo para la derivación de la suma, producto (en el caso de funciones a valores reales) y composición de funciones diferenciables. En el caso de dimensión finita, considerando la base canónica, representaremos las derivadas mediante matrices (que no es otra cosa que la representación matricial de las aplicaciones lineales correspondientes). Dichas matrices se llaman matrices Jacobianas y sus coordenadas son las derivadas parciales de la función.

En el caso de que la función tenga valores reales, se definirá la noción del gradiente, que corresponde a la dirección del máximo ascenso de la función, y se definirá el espacio tangente. En esta parte, se verán dos teoremas pilares para este curso, el teorema de función inversa y el teorema de función implicita, este último siendo la base para la geometría diferencial.

Terminaremos esta parte con la derivación de orden superior, el teorema de Schwarz (que servirá en la parte de optimización) y las series de Taylor para aproximar funciones de varias (finitas) variables a valores reales.

Índice de contenidos

1. Aplicaciones lineales

En esta parte, estudiaremos funciones lineales entre espacios normados. Utilizaremos frecuentamente el término *aplicación lineal* o *operador lineal* para referirnos a tal función. En la primera subsección se verá que una aplicación lineal es continua si y sólo si es Lipschitz continua. A partir de esta caracterización, se definirá una norma en el espacio vectorial de las aplicaciones lineales continuas. Luego, en la siguiente subsección nos interesaremos en aplicaciones lineales entre espacios de dimensión finita. Veremos que dichas aplicaciones serán automaticamente (Lipschitz) continuas y recordaremos (del curso anterior del álgebra lineal) como representarlas mediante matrices.

1.1. Norma de una aplicación lineal continua

Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados.

Definición 1.1 (aplicación lineal)**.** Una función $T : X \to Y$ se dice (aplicación o operador) *lineal* si para todo $x, z \in X$, $\lambda \in \mathbb{R}$ se tiene que:

$$T(x + z) = T(x) + T(z) \qquad \text{y} \qquad T(\lambda x) = \lambda T(x).$$

Las aplicaciones lineales son las funciones naturales (morfismos) entre espacios vectoriales, en el sentido que respectan la estructura de los espacios. La siguiente proposición es fundamental e illustra la fuerza atrás de la definición anterior.

Proposición 1.2 (Continuidad para aplicaciones lineales)**.** Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados y $T : X \to Y$ una aplicación lineal. Los siguientes son equivalentes:
(i). Existe $k > 0$ tal que $\|T(x)\|_Y \leq k\,\|x\|_X$ para todo $x \in X$;
(ii). T es Lipschitz continua con constante $k > 0$;
(iii). T es continua ;
(iv). T es continua en 0.

Demostración. Las implicancias **(ii)** $\Longrightarrow$**(iii)**$\Longrightarrow$**(iv)** son obvias.
Mostramos la equivalencia **(i)** $\Leftrightarrow$**(ii).** Para eso, supongamos primero (i) y consideramos dos elementos arbitrarios $y, z \in X$. Por la linealidad de T deducimos que

$$T(y - z) = T(y) + T(-z) = T(y) - T(z).$$

Tomando $x = y - z$ y aplicando (i) obtenemos

$$\|T(y) - T(z)\|_Y \leq k\|y - z\|_X,$$

es decir T es k-Lipschitz. Mostramos ahora el recíproco: por la linealidad de T se tiene $T(0) = 0$. Sea $x \in X$. Como T es k-Lipschitz, deducimos que

$$\|T(x)\|_Y = \|T(x) - T(0)\|_Y \leq k\|x - 0\|_X = k\|x\|_X$$

lo que muestra (i).

Nos queda demostrar que **(iv) $\Longrightarrow$(i).** Para eso, supongamos que T es continua en 0. Entonces (para $\varepsilon = 1$) existe $\delta > 0$ tal que

$$\|x\|_X < \delta \implies \|T(x)\|_Y < 1.$$

Nuestro objetivo es mostrar que (i) se cumple con $k = \delta^{-1}$, es decir

$$\|T(x)\|_Y \leq \frac{1}{\delta}\,\|x\|_X \qquad \text{para todo } x \in X.$$

Supongamos el contrapuesto. Entonces existe $\bar{x} \in X$ tal que

$$\|T(\bar{x})\|_Y > \frac{1}{\delta}\,\|\bar{x}\|_X. \tag{1}$$

(¡Observad que (1) garantiza que $\bar{x} \neq 0$!) Tomando

$$\delta_n = \delta - \frac{1}{n} \underset{n\to\infty}{\nearrow} \delta \qquad \text{e} \qquad z_n := \delta_n \left(\frac{\bar{x}}{\|\bar{x}\|_X}\right)$$

se tiene $\|z_n\|_X = \delta_n < \delta$, y por lo tanto $\|T(z_n)\|_Y < 1$. Entonces, para todo $n \geq 1$ se tiene

$$\|T(z_n)\|_Y = \|T\left(\delta_n \frac{\bar{x}}{\|\bar{x}\|_X}\right)\|_Y = \frac{\delta_n}{\|\bar{x}\|_X}\|T(\bar{x})\|_Y \ < 1 \quad \Longleftrightarrow \quad \delta_n\|T(\bar{x})\|_Y < \|\bar{x}\|_X.$$

Tomando el límite $n \to \infty$ se obtiene $\delta\|T(\bar{x})\|_Y \leq \|\bar{x}\|_X$, lo que contradice (1). La demostración es completa. $\square$

A continuación anotaremos por

$$\mathcal{L}(X,Y) = \{T : X \to Y \text{ lineal, continua }\}$$

el conjunto de las aplicaciones lineales y continuas de X a Y. Es fácil ver que dicho conjunto es un espacio vectorial. Mostramos ahora que $\mathcal{L}(X,Y)$ admite una norma natural que lo transforma en un espacio normado.

Proposición 1.3 (norma de un operador lineal)**.** Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados. Para cada aplicación lineal continua $T : X \to Y$ se tiene:

$$||T||_{\text{op}} := \sup_{\|x\|_X=1} \|T(x)\|_Y \ = \ \text{ínf}\ \{\, k > 0 :\ \|T(x)\|_Y \leq k\,\|x\|_X,\ \ \forall x \in X\}\,. \tag{2}$$

Demostración. Observamos primero que

$$\text{ínf}\ \{\, k > 0 :\ \|T(x)\|_Y \leq k\,\|x\|_X,\ \ \forall x \in X\} = \sup_{x\neq 0}\ \left\{\frac{\|T(x)\|_Y}{\|x\|_X}\right\}. \tag{3}$$

Dado que para cada $\lambda \in \mathbb{R}$, $\lambda \neq 0$ se tiene:

$$\frac{\|T(\lambda x)\|_Y}{\|\lambda x\|_X} = \frac{\|T(x)\|_Y}{\|x\|_X}$$

el supremo en la (3) se puede tomar en la esfera unitaria $S_X := \{x \in X :\ \|x\|_X = 1\}$ y por lo tanto (2) es cierta. $\square$

Ejercicio 1.4 (($(\mathcal{L}(X,Y), ||\cdot||_{\mathrm{op}})$ es un espacio normado)**.** Comprobar que la función $T \mapsto \|T\|_{\mathrm{op}}$ dada por (2) es una norma en $\mathcal{L}(X,Y)$.

La siguiente definición destaca el caso particular donde el espacio normado Y es $\mathbb{R}$. Este caso tiene suma importancia en el análisis funcional.

Definición 1.5 (espacio dual)**.** Sea $(X, \|\cdot\|_X)$ un espacio normado. El espacio normado

$$X^* := \mathcal{L}(X,\mathbb{R}) = \{\, \varphi : X \longrightarrow \mathbb{R} \text{ lineal continua } \}$$

se llama *espacio dual* del espacio normado X y sus elementos se suelen llamar *formas lineales*. La norma $||\cdot||_*$ del espacio dual X^* (que se suele llamar norma dual) esta dada por la formula (proveniente de la formula general (2))

$$||\varphi||_* := ||\varphi||_{\mathrm{op}} = \sup_{\|x\|_X=1} |\varphi(x)| \;=\; \sup_{\|x\|_X=1} \varphi(x).$$

Terminamos la sección con el caso donde X es un espacio de dimensión finita.

Proposición 1.6 (en dimensión finita lineal implica continua)**.** Sea $(Y, \|\cdot\|_Y)$ un espacio normado y $T : \mathbb{R}^d \to Y$ una aplicación lineal. Entonces T es continua, es decir, $T \in \mathcal{L}(\mathbb{R}^d, Y)$.

Demostración. Dado que todas las normas en $\mathbb{R}^d$ son equivalentes, trabajaremos con la norma $||\cdot||_1$. Sea $\{e_1,\ldots,e_d\}$ la base canónica de $\mathbb{R}^d$. Mostraremos que T es Lipschitz continua con constante $k = \max\{||T(e_i)||_Y : i \in \{1,\ldots,d\}\}$. En efecto, sea $x = (x_1,\ldots,x_d) \in \mathbb{R}^d$. Entonces:

$$||T(x)||_Y = \left\| \sum_{i=1}^{d} x_i T(e_i) \right\|_Y \le \sum_{i=1}^{d} |x_i|\, ||T(e_i)||_Y \le\, k \sum_{i=1}^{d} |x_i| = k\, \|x\|_1,$$

lo que muestra lo pedido. □

En dimensión infinita existen aplicaciones lineales que no son continuas. Dimos a continuación un ejemplo.

Ejemplo 1.7 (aplicación lineal discontinua)**.** Sea $X = c_{00}(\mathbb{N})$ el espacio de las sucesiones eventualmente nulas y con norma:

$$||x||_\infty = \sup_{n\ge1} |x_n|, \qquad x = (x_n)_n \in c_{00}(\mathbb{N}).$$

Entonces la aplicación lineal

$$\begin{cases} \varphi : c_{00}(\mathbb{N}) \to \mathbb{R} \\[2ex] \varphi(x) = \sum_{n\ge1} n\, x_n, \quad x = (x_n)_n \in c_{00}(\mathbb{N}) \end{cases}$$

no es continua. En efecto, considerando la base canónica $\{e_n : n \ge 1\}$ de $c_{00}(\mathbb{N})$ se observa que $||e_n||_\infty = 1$, para cada $n \ge 1$, mientras que $\varphi(e_n) = n$.

Ejercicio 1.8 (la función integral como aplicación lineal continua.). **(i).** Sea $L > 0$. Anotemos por $(\mathcal{C}([0,L]), ||\cdot||_\infty)$ el espacio de Banach de funciones continuas en $[0,L]$ a valores reales. Muestre que la función *integral*

$$\begin{cases} I : \mathcal{C}([0,L]) \to \mathbb{R} \\ I(f) = \int_0^L f(s)ds, \qquad \text{para cada } f \in \mathcal{C}([0,L]) \end{cases}$$

es una aplicación lineal continua en $\mathcal{C}([0,L])$, con $||I||_{\text{op}} \leq L$.

(ii). Fijamos ahora $M > 0$ y consideramos el espacio de Banach

$$(\mathcal{C}([0,M] \times [0,L]), ||\cdot||_\infty)$$

de las funciones continuas del subconjunto compacto $[0,M] \times [0,L]$ de $\mathbb{R}^2$ a valores reales. Muestre que la función

$$\begin{cases} \Phi : \mathcal{C}([0,M] \times [0,L]) \to \mathcal{C}([0,M]) \\ \Phi(f)(x) = \int_0^L f(x,s)ds,\ x \in [0,L], \qquad \text{para cada } f \in \mathcal{C}([0,M] \times [0,L]) \end{cases}$$

es una aplicación lineal continua en $\mathcal{C}([0,M] \times [0,L])$ con $||\Phi||_{\text{op}} \leq L$.

1.2. Representación de una aplicación lineal en dimensión finita

En esta sección recordamos del álgebra lineal de que cada aplicacion lineal entre espacios de dimensión finita se puede representar por una matriz.

Empezamos con el caso donde el espacio de llegada es $\mathbb{R}$ y consideramos una forma lineal $\varphi : \mathbb{R}^d \to \mathbb{R}$. (Recordamos que φ debe ser continua, por la Proposición 1.6). Poniendo

$$\varphi_i = \varphi(e_i), \qquad i \in \{1, \ldots, d\}$$

la evaluación de φ en la base canónica $\{e_1, \ldots, e_d\}$ de $\mathbb{R}^d$, se obtiene el vector–fila $(\varphi_1, \ldots, \varphi_d)$ que representa la forma lineal $\varphi \in (\mathbb{R}^d)^*$. En particular, representando cada $x \in \mathbb{R}^d$ por el vector–columna

$$x \underset{\text{representación}}{\longleftrightarrow} \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \qquad \text{donde } x = \sum_{i=1}^d x_i e_i \in \mathbb{R}^d$$

se observa que

$$\varphi(x) = \varphi(\sum_{i=1}^d x_i e_i) = \sum_{i=1}^d x_i \varphi(e_i) = \sum_{i=1}^d x_i \varphi_i = (\varphi_1, \ldots, \varphi_d) \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}.$$

En otras palabras, el resultado de la *actuación* de la forma lineal $\varphi \in (\mathbb{R}^d)^*$ en el vector $x \in \mathbb{R}^d$ *coincide con el producto matricial* del vector $(1 \times d)$ (que representa φ) con el vector $(d \times 1)$ (que representa x). Eso justifica nuestra elección de representar los vectores del espacio primal $\mathbb{R}^d$ por columnas y los vectores del espacio dual $(\mathbb{R}^d)^*$ por filas.

Considerando la norma–2 en $\mathbb{R}^d$, deducimos por la desigualdad Cauchy-Schwarz que

$$|\varphi(x)| = \left|\sum_{i=1}^{d} x_i \varphi_i\right| \leq ||\varphi||_2 \cdot ||x||_2 \qquad \text{para todo } x \in \mathbb{R}^d, \tag{4}$$

y por lo tanto

$$||\varphi||_* = \sup_{\|x\|_2=1} |\varphi(x)| \leq ||\varphi||_2.$$

Tomando $\bar{x} = \sum_{i=1}^{d} \varphi_i e_i$ (el vector de $\mathbb{R}^d$ cuyas coordenadas coinciden con la representación de la forma φ) y

$$\hat{x} = \frac{\bar{x}}{||\bar{x}||_2} = \frac{\bar{x}}{||\varphi||_2}$$

vemos que $||\hat{x}||_2 = 1$ y

$$|\varphi(\hat{x})| = \frac{1}{||\varphi||_2} \left|\sum_{i=1}^{d} \varphi_i^2\right| = ||\varphi||_2,$$

es decir (4) se cumple con igualdad, es decir $||\varphi||_* = ||\varphi||_2$ para cada $\varphi \in (\mathbb{R}^d)^*$, estableciendo que el dual del espacio Euclidiano $(\mathbb{R}^d, ||\cdot||_2)$ es el mismo:

$$(\mathbb{R}^d, ||\cdot||_2)^* = (\mathbb{R}^d, ||\cdot||_2).$$

Ejercicio 1.9 (dual de $\mathbb{R}^d$ con otras normas)**.** Muestre que

$$(\mathbb{R}^d, \|\cdot\|_1)^* = (\mathbb{R}^d, \|\cdot\|_\infty) \qquad \text{y} \qquad (\mathbb{R}^d, \|\cdot\|_\infty)^* = (\mathbb{R}^d, \|\cdot\|_1).$$

Trataremos ahora el caso general de una aplicación lineal $T : \mathbb{R}^d \to \mathbb{R}^\ell$, con $\ell > 1$. Para eso, consideramos las bases canonicas $\{e_1, \ldots, e_d\}$ de $\mathbb{R}^d$ y $\{\tilde{e}_1, \ldots, \tilde{e}_\ell\}$ de $\mathbb{R}^\ell$, luego, para cada vector e_j de la base de $\mathbb{R}^d$, donde $j \in \{1, \ldots, d\}$, anotamos por $(a_{1j}, \ldots a_{dj})$ las coordenadas del vector $T(e_j)$ de $\mathbb{R}^\ell$ en la base canónica de $\mathbb{R}^\ell$, es decir:

$$T(e_j) = \sum_{i=1}^{\ell} a_{ij}\ \tilde{e}_i.$$

Entonces, representando como antes los vectores $x \in \mathbb{R}^d$ por columnas, la matriz con entradas (a_{ij}), con $i \in \{1, \ldots, \ell\}$ y $j \in \{1, \ldots, d\}$ representa[1] la aplicación lineal T.

[1]En esta representación, a_{ij} es la i-coordenada (en la base canónica de $\mathbb{R}^\ell$) de la imagen por T del vector e_j de la base canónica de $\mathbb{R}^d$.

Concretamente, si $x = \sum_{j=1}^{d} x_j e_j \in \mathbb{R}^d$, entonces representando

$$T \underset{\text{representación}}{\longleftrightarrow} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ \vdots & \vdots & & \vdots \\ a_{\ell 1} & a_{\ell 2} & \dots & a_{\ell d} \end{pmatrix} \qquad \text{y} \qquad x \underset{\text{representación}}{\longleftrightarrow} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

tenemos

$$T(x) = T(\sum_{j=1}^{d} x_j e_j) = \sum_{j=1}^{d} x_j T(e_j) = \sum_{i=1}^{\ell} \left(\sum_{j=1}^{d} x_j a_{ij} \right) \tilde{e}_i.$$

En otras palabras, las ℓ–coordenadas del vector $T(x)$ en $\mathbb{R}^\ell$ se obtienen (como vector–columna) por el producto matricial de la matriz $(\ell \times d)$ (que representa $T \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^\ell)$) con el vector–columna $(d \times 1)$ (que representa $x \in \mathbb{R}^d$).

Ejemplo 1.10. (i). $(d = \ell = 1)$ La base canónica de $\mathbb{R}$ (de dimensión 1) es el vector $e = 1$. Por lo tanto, la función lineal

$$\varphi(x) = a\, x, \qquad x \in \mathbb{R}$$

se representa por el número real $a = \varphi(1)$.

(ii). $(d = 2,\ \ell = 1)$ Consideramos la base canónica $\{e_1, e_2\}$ de $\mathbb{R}^2$ y la aplicación lineal

$$T(x_1, x_2) = a_1 x_1 + a_2 x_2, \qquad x = (x_1, x_2) \in \mathbb{R}^2.$$

Entonces T está representada por el vector–fila (a_1, a_2). En efecto, $a_1 = T(1,0) = T(e_1)$ y $a_2 = T(0,1) = T(e_2)$.

(iii). $(d = 3,\ \ell = 2)$ Consideramos la aplicación lineal

$$\begin{cases} T : \mathbb{R}^3 \to \mathbb{R}^2 \\ T(x_1, x_2, x_3) = (x_1 - x_2, 2x_2 + x_3) \end{cases}$$

Observamos que

$$T(e_1) = T(1,0,0) = (1,0), \quad T(e_2) = T(0,1,0) = (-1,2) \quad \text{y} \quad T(e_3) = T(0,0,1) = (0,1)$$

por lo tanto para cada $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ se tiene[2]:

$$T(x) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_3 \\ 2x_2 + x_3 \end{pmatrix}.$$

[2]En la representación matricial, los vectores de $\mathbb{R}^d$ $(d \in \{2,3\})$ se tienen que representar por columnas.

2. Diferenciabilidad

En esta parte introducimos la noción de diferenciabilidad para funciones entre dos espacios normados X e Y. Empezamos por recordar la definición de la derivada de una función $f:\mathbb{R}\to\mathbb{R}$ en un punto $\bar{x}$, luego daremos la definición de una derivada en el caso general y al final consideramos el caso particular donde los espacios X e Y son de dimensión finita.

2.1. Derivada de una función de $\mathbb{R}$ a $\mathbb{R}$ (recordatorio)

Recordamos ahora la definición de la derivada de una función $f:\mathbb{R}\to\mathbb{R}$ en el punto $\bar{x}\in\mathbb{R}$. Decimos que f es diferenciable en $\bar{x}$ con derivada $a=f'(\bar{x})$, si

$$\lim_{x\to\bar{x}}\frac{f(x)-f(\bar{x})}{x-\bar{x}}=a \tag{5}$$

lo que se puede re-escribir como

$$\lim_{x\to\bar{x}}\frac{f(x)-f(\bar{x})-a(x-\bar{x})}{x-\bar{x}}=0 \qquad\Longleftrightarrow\qquad \lim_{u\to 0}\frac{f(\bar{x}+u)-f(\bar{x})-a\cdot u}{u}=0.$$

Observamos que lo anterior es equivalente a pedir que

$$\lim_{|u|\to 0}\frac{|f(\bar{x}+u)-f(\bar{x})-a\cdot u|}{|u|}=0 \tag{6}$$

Contrariamente con la definición (5), la formula anterior se puede extender (y así haremos en la siguiente subsección) para funciones entre espacios normados. Para eso, se tendrá que substituir cada vez el valor absoluto por la norma correspondiente e interpretar la derivada $f'(\bar{x})$ como una función lineal continua:

$$\begin{cases}\varphi:\mathbb{R}\to\mathbb{R}\\ \varphi(u)=a\cdot u, \qquad u\in\mathbb{R}.\end{cases}$$

Con esta interpretación, $f'(\bar{x})\in\mathcal{L}(\mathbb{R},\mathbb{R})$ estará representada por el escalar $\varphi(1)=a\in\mathbb{R}$ (es decir, la evaluación de la forma lineal $\varphi=f'(\bar{x})$ en la base canónica de $\mathbb{R}$ que es el vector $e=1$). Luego (6) se puede escribir como:

$$\lim_{|u|\to 0}\frac{|f(\bar{x}+u)-f(\bar{x})-f'(\bar{x})(u)|}{|u|}=0. \tag{7}$$

Antes de continuar, conviene recordar también la interpretación geométrica de (6) mediante la función afín

$$\ell(x)=f(\bar{x})+\varphi(x-\bar{x})=f(\bar{x})+a(x-\bar{x}), \qquad x\in\mathbb{R}.$$

tLa gráfica de dicha función es una *recta tangente* a la gráfica de la función f en el punto $(\bar{x},f(\bar{x}))$ (véase Figura 1) y la función afín ℓ está aproximando (al orden 1) la función f alrededor del punto $\bar{x}$, en el sentido que

$$\frac{f(\bar{x}+u)-\ell(\bar{x}+u)}{|u|}\underset{|u|\to 0}{\longrightarrow}0.$$

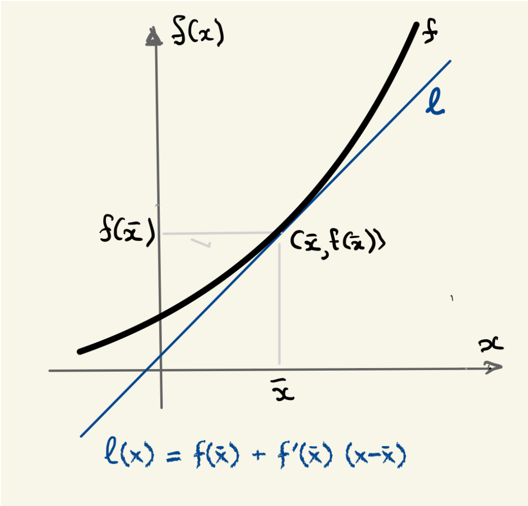


Figura 1: Recta tangente de f en $(\bar{x}, f(\bar{x}))$.

2.2. Derivada de una función entre espacios normados

Consideramos ahora el caso general. Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados.

Definición 2.1 (diferenciabilidad)**.** Decimos que la función $f : X \to Y$ es *diferenciable* en $\bar{x} \in X$, si existe una aplicación lineal continua $T \in \mathcal{L}(X, Y)$ tal que

$$\lim_{||u||_X \to 0} \frac{||f(\bar{x}+u) - f(\bar{x}) - T(u)||_Y}{||u||_X} = 0. \tag{8}$$

En este caso, decimos que T es la *derivada* de f en $\bar{x}$ y la anotamos por $Df(\bar{x})$.

Es importante destacar que la definición de la derivada requiere, de forma implicita, que la aplicación $T \in \mathcal{L}(X, Y)$ que cumple (8) sea única (dado que $Df(\bar{x}) = T$). Eso no es completamente obvio de la definición anterior y se demuestra con la siguiente proposición:

Proposición 2.2 (unicidad de la derivada)**.** Hay a lo máximo una aplicación lineal continua $T \in \mathcal{L}(X, Y)$ que satisface (8).

Demostración. Supongamos que existen $T_1, T_2 \in \mathcal{L}(X, Y)$ cumpliendo (8). Sea $\varepsilon > 0$ (arbitrario). Entonces para $i \in \{1, 2\}$, existe $\delta_i > 0$ tal que

$$||f(\bar{x}+u) - f(\bar{x}) - T_i(u)||_Y < \frac{\varepsilon}{2}\,||u||_X \qquad \text{para todo } u \in B_X(0, \delta_i).$$

Tomando $\delta = \min\{\delta_1, \delta_2\}$ y $u \in B_X(0, \delta)$ se tiene que

$$\begin{aligned}||T_1(u) - T_2(u)||_Y &\leq ||T_1(u) - (f(\bar{x}+u) - f(\bar{x}))\,||_Y + ||\,(f(\bar{x}+u) - f(\bar{x})) - T_2(u)||_Y \\ &< \frac{\varepsilon}{2}\,||u||_X \;+\; \frac{\varepsilon}{2}\,||u||_X \;=\; \varepsilon\,||u||_X\end{aligned}$$

y por lo tanto

$$\frac{||(T_1 - T_2)(u)||_Y}{||u||_X} \;=\; ||(T_1 - T_2)(\frac{u}{||u||_X})|| \;<\; \varepsilon.$$

Dado que lo anterior es cierto para todo $u \in B_X(0,\delta)$ y que la expresión del lado izquierdo es 0-homogenea con respecto a la variable u se deduce

$$||T_1 - T_2||_{\mathrm{op}} \;=\; \sup_{||v||_X=1} ||(T_1 - T_2)(v)||_Y \;<\; \varepsilon$$

y como $\varepsilon > 0$ es arbitrario, se concluye que $||T_1 - T_2||_{\mathrm{op}} = 0$ y por lo tanto $T_1 = T_2$. $\square$

Mostramos ahora que la diferenciabilidad es una noción más restrictiva (más fuerte) que la continuidad.

Proposición 2.3 (diferenciabilidad implica continuidad)**.** Sea $f : X \to Y$ una función diferenciable en $\bar{x} \in X$. Entonces f es continua en $\bar{x}$.

Demostración. Sea $T = Df(\bar{x}) \in \mathcal{L}(X,Y)$. Mostramos que f es continua en $\bar{x}$. Sea $\varepsilon > 0$ (arbitrario). Entonces existe $\delta > 0$ tal que para todo $u \in B_X(0,\delta)$

$$||f(\bar{x}+u) - f(\bar{x}) - T(u)||_Y < \varepsilon\, ||u||_X.$$

Podemos empequeñecer $\delta > 0$ para asegurar que

$$\delta < \frac{\varepsilon}{||T||_{\mathrm{op}} + \varepsilon}.$$

Entonces para todo $u \in B_X(0,\delta)$ deducimos que

$$\begin{aligned} ||f(\bar{x}+u) - f(\bar{x})||_Y &\leq ||f(\bar{x}+u) - f(\bar{x}) - T(u)||_Y + ||T(u)||_Y \\ &< \varepsilon\, ||u||_X \;+\; ||T||_{\mathrm{op}}\, ||u||_X < (\varepsilon + ||T||_{\mathrm{op}})\,\delta < \varepsilon. \end{aligned}$$

Lo anterior muestra que f es continua en $\bar{x}$. $\square$

El siguiente ejercicio se enfoca en un caso particular de la Definición 2.1.

Ejercicio 2.4 (derivada de una función a valores en $\mathbb{R}^\ell$)**.** Sea $(X, ||\cdot||_X)$ un espacio normado y $f : X \to \mathbb{R}^\ell$ una función. Anotamos por $\{f_i\}_{i=1}^{\ell}$ las coordenadas de la función f en la base canónica $\{\tilde{e}_1, \ldots, \tilde{e}_\ell\}$ de $\mathbb{R}^\ell$, es decir,

$$f(x) = \sum_{i=1}^{\ell} f_i(x)\tilde{e}_i \qquad \text{para todo } x \in X.$$

Muestre que f es diferenciable en $\bar{x} \in X$ *si y sólo si* $f_i : X \to \mathbb{R}$ es diferenciable en $\bar{x}$ para todo $i \in \{1, \ldots, \ell\}$. En este caso, muestre que

$$Df(\bar{x})(u) = \sum_{i=1}^{\ell} Df_i(\bar{x})(u)\tilde{e}_i \qquad \text{para todo } u \in X,$$

es decir, la derivadas $\{Df_i(\bar{x})\}_{i=1}^{\ell}$ son las coordenadas de la aplicación lineal continua $Df(\bar{x}) : X \to \mathbb{R}^\ell$.

2.3. Derivada direccional, derivadas parciales

Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados y sea $f : X \to Y$ una función. Consideramos la restricción de la función f en la recta que pasa por el punto $\bar{x} \in X$ y que tiene dirección $u \in X \setminus \{0\}$:

$$\begin{cases} \varphi : \mathbb{R} \to Y \\ \varphi(t) := f(\bar{x} + tu) \qquad t \in \mathbb{R} \end{cases}$$

Decimos que la función f tiene *derivada direccional* en $\bar{x} \in X$ en la dirección u si la función h es diferenciable en 0. En este caso, anotamos dicha derivada direccional (¡necesariamente única!) por $df(\bar{x}; u)$ y se cumple:

$$df(\bar{x}; u) := \varphi'(0) = \lim_{t \to 0} \frac{f(\bar{x} + tu) - f(\bar{x})}{t}. \tag{9}$$

Proposición 2.5 (derivada direccional vs derivada)**.** Sea $f : X \to Y$ una función diferenciable en $\bar{x} \in X$. Entonces para cada $u \in X \setminus \{0\}$ se tiene:

$$df(\bar{x}; u) = Df(\bar{x})(u),$$

es decir, la derivada direccional de f en $\bar{x}$ y en la dirección u existe y es igual a la evaluación de la derivada $Df(\bar{x}) \in \mathcal{L}(X, Y)$ en u.

Demostración. Anotemos $T := Df(\bar{x})$. Tomamos $\varepsilon > 0$ arbitrario y ponemos $\varepsilon_1 = \varepsilon / ||u||_X$. Entonces existe $\delta_1 > 0$ tal que para todo $v \in B_X(0, \delta_1)$

$$||f(\bar{x} + v) - f(\bar{x}) - T(v)||_Y < \varepsilon_1 \, ||v||_X.$$

Sea $\delta = \delta_1 / ||u||_X$. Entonces para cada $t \in (-\delta, \delta)$, se tiene que $v := tu \in B_X(0, \delta_1)$ y por lo tanto

$$\frac{||f(\bar{x} + tu) - f(\bar{x}) - T(tu)||_Y}{t} < \varepsilon_1 \, ||u||_X$$

o de forma equivalente (dado que $T(tu) = tT(u)$ y $\varepsilon = \varepsilon_1 \, ||u||_X$)

$$\left\| \frac{f(\bar{x} + tu) - f(\bar{x})}{t} - T(u) \right\|_Y < \varepsilon.$$

Lo anterior significa que

$$\lim_{t \to 0} \frac{f(\bar{x} + tu) - f(\bar{x})}{t} = T(u) = Df(\bar{x})(u)$$

lo que muestra lo pedido. $\square$

Es importante notar que la existencia de todas las derivadas direccionales no garantiza la existencia de la derivada, como muestra el siguiente ejemplo.

Ejemplo 2.6 (función no diferenciable con derivadas direccionales)**.** La función $f:\mathbb{R}^2\to\mathbb{R}$ dada por

$$f(x_1,x_2)=\begin{cases}1, & \text{si } \ 0<x_2<x_1^2\\ \\ 0, & \qquad \text{si no}\end{cases}$$

no es continua en el origen $(0,0)\in\mathbb{R}^2$ (y por lo tanto, no es diferenciable en este punto, véase Proposición 2.3). Sin embargo, el lector puede comprobar facilmente que existen todas las derivadas direccionales de f en $(0,0)$ (y son iguales a 0).

Cerramos esta subsección con la siguiente definición para el caso de funciones definidas en un espacio de dimensión finita.

Definición 2.7 (derivadas parciales)**.** Sea $f:\mathbb{R}^d\to Y$ una función y $\{e_1,\ldots,e_d\}$ la base canónica de $\mathbb{R}^d$. Llamamos *j-derivada parcial* de f en $\bar{x}\in\mathbb{R}^d$ y anotamos por $\frac{\partial f}{\partial x_j}(\bar{x})$ o $\partial_j f(\bar{x})$, la derivada direccional de f (en $\bar{x}$) en la dirección e_j. En otras palabras:

$$\partial_j f(\bar{x})\equiv\frac{\partial f}{\partial x_j}(\bar{x}):=df(\bar{x};e_j)$$

Las derivadas parciales son casos particulares de derivada direccional en dimensión finita, donde la dirección considerada es un vector de la base canónica.

2.4. Representación de la derivada en dimensión finita

Como hemos visto, la derivada $Df(\bar{x})$ de una función $f:X\to Y$ en $\bar{x}\in X$ es una aplicación lineal continua y por lo tanto admite una representación matricial, cuando los espacios X e Y son de dimensión finita (véase Subsección 1.2). En efecto, supongamos que $X=\mathbb{R}^d$ e $Y=\mathbb{R}^\ell$. Considerando las bases canonicas en ambos espacios y anotando por $\{f_i\}_{i=1}^{\ell}$ las coordenadas de la función f en la base canónica de $\mathbb{R}^\ell$ obtenemos (*cf.* Ejercício 2.4) que

$$Df(\bar{x})(\cdot)=\sum_{i=1}^{\ell}Df_i(\bar{x})(\cdot)\tilde{e}_i.$$

Luego, la aplicación lineal $T=Df(\bar{x})\in\mathcal{L}(\mathbb{R}^d,\mathbb{R}^\ell)$ se representa por una matriz $(a_{ij})_{i=1}^{\ell}\,{}_{j=1}^{d}$ (*cf.* Subsección 1.2). Los coeficientes de esta matriz se obtienen mediante la formula:

$$a_{ij}=Df_i(\bar{x})(e_j)=df_i(\bar{x};e_j)=\frac{\partial f_i}{\partial x_j}(\bar{x})$$

donde la penúltima igualdad proviene de la Proposición 2.5 aplicada a la i-coordenada $f_i:X\to\mathbb{R}$ y la última igualdad es la Definición 2.7 aplicada a esta misma función. En otras palabras, la entrada a_{ij} de la matriz que representa la derivada $Df(\bar{x})$ es la j-derivada parcial de la coordenada f_i de la función f.

Definición 2.8 (matriz Jacobiana)**.** Sea $f:\mathbb{R}^d\to\mathbb{R}^\ell$ una función diferenciable en $\bar{x}\in\mathbb{R}^d$. Llamamos *(matriz) Jacobiana* de f en $\bar{x}$ —y anotamos por $Jf(\bar{x})$— la matriz $(\ell\times d)$ que representa la derivada $Df(\bar{x})$ en las bases canónicas de $\mathbb{R}^d$ y $\mathbb{R}^\ell$:

$$Df(\bar{x}) \quad \underset{\text{representación}}{\longleftrightarrow} \quad Jf(\bar{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\bar{x}) & \frac{\partial f_1}{\partial x_2}(\bar{x}) & \dots & \frac{\partial f_1}{\partial x_d}(\bar{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_\ell}{\partial x_1}(\bar{x}) & \frac{\partial f_\ell}{\partial x_2}(\bar{x}) & \dots & \frac{\partial f_\ell}{\partial x_d}(\bar{x}) \end{pmatrix} \tag{10}$$

Aunque la matriz de la formula (10) se puede definir formalmente siempre y cuando las derivadas parciales existan, dicha matriz no se puede llamar Jacobiana de f si la función f no es diferenciable en $\bar{x}$. (¡Nótese que la Definición 2.8 requiere la existencia de la derivada $Df(\bar{x})$!).

Recordamos que puede ocurrir que una función no sea diferenciable en un punto, pero que tenga todas sus derivadas parciales (incluso todas las derivadas direccionales) en este punto, véase Ejemplo 2.6. Por otra parte, la Proposición 2.2 (unicidad de la derivada) nos dice que la martiz formada por las derivadas parciales es el único candidato para ser Jacobiana (es decir, para representar la derivada) en dimensión finita. Eso proporciona un método para decidir si una función $f:\mathbb{R}^d\to\mathbb{R}^\ell$ es diferenciable o no en un punto $\bar{x}\in\mathbb{R}^d$ y que en el caso que lo fuera, obtener su derivada[3]:

Paso 1. Calcular todas las derivadas parciales en $\bar{x}$.
(Si hay al menos una derivada parcial que no existe, entonces f no puede ser diferenciable por la Proposición 2.5.)

Paso 2. Formar la matriz de la formula (10) que es el candidato de ser Jacobiana.

Paso 3. Trabajando en coordenadas comprobar directamente si (8) se cumple o no, donde T está substituida por la matriz candidata. La función f será diferenciable (con derivada representada por la matriz candidata) si y sólo si (8) se cumple.

A continuación veremos un ejemplo donde el método anterior se pone en práctica.

Ejemplo 2.9 (decidir sobre la diferenciabilidad)**.** Consideramos la función $f:\mathbb{R}^2\to\mathbb{R}$ dada por

$$f(x_1,x_2)=\begin{cases} \frac{x_1\,|x_2|}{\sqrt{x_1^2+x_2^2}}\,, & \text{si } (x_1,x_2)\neq(0,0) \\ \\ 0\,, & \text{si } (x_1,x_2)=(0,0). \end{cases}$$

Invitamos al lector de comprobar que la función anterior es continua en cada punto. Luego, las dos derivadas parciales existen y son iguales a 0 :

$$\frac{\partial f}{\partial x}(0,0) := df((0,0);e_1) = \lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t} = \lim_{t\to 0}\frac{0-0}{t} = 0 \qquad \text{y}$$

$$\frac{\partial f}{\partial y}(0,0) := df((0,0);e_2) = \lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t} = \lim_{t\to 0}\frac{0-0}{t} = 0.$$

[3]También se podría mirar de primero si f es continua en $\bar{x}$ (si no lo fuera, tampoco podría ser diferenciable ahí, véase Proposición 2.3).

Por lo tanto, si la derivada de f en $(0,0)$ existiera, estaría representada por la matriz (1×2) con entradas 0, es decir, sería necesariamente igual a la aplicación lineal $\mathbf{0}$. Se trata entonces de mirar si el siguiente límite existe y es igual a 0:

$$\lim_{(u_1,u_2)\to(0,0)} \frac{|f((0,0)+(u_1,u_2)) - f(0{,}0) - \mathbf{0}(u_1,u_2)|}{\sqrt{u_1^2+u_2^2}} = \lim_{(u_1,u_2)\to(0,0)} \left|\frac{u_1u_2}{u_1^2+u_2^2}\right| = \quad ?$$

Es fácil ver (tomando $u_1 = u_2 = \frac{1}{n} \to 0$) que el anterior límite no existe, y por lo tanto f no es diferenciable en $(0,0)$.

Ejercicio 2.10. Mostrar que la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x_1,x_2) = \begin{cases} \frac{x_1^2 x_2}{\sqrt{x_1^2+x_2^2}}\,, & \text{si } (x_1,x_2) \neq (0,0) \\ \\ 0\,, & \text{si } (x_1,x_2) = (0,0). \end{cases}$$

es diferenciable en el punto $(0,0)$ y encontrar su derivada.

Terminamos esta subsección introduciendo una terminología importante que se aplica en el caso $\ell = 1$ (es decir, cuando el espacio de llegada es de dimensión 1).

Definición 2.11 (gradiente)**.** Sea $f : \mathbb{R}^d \to \mathbb{R}$ una función diferenciable en $\bar{x} \in \mathbb{R}^d$. Llamamos *gradiente* de f en $\bar{x}$ —y anotamos por $\nabla f(\bar{x})$— el vector de $\mathbb{R}^d$ cuyas coordenadas en la base canónica son las mismas que las de la matriz Jacobiana $Jf(\bar{x})$ que representa la derivada. En otras palabras:

$$\nabla f(\bar{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\bar{x}) \\ \vdots \\ \frac{\partial f}{\partial x_d}(\bar{x}) \end{pmatrix} \tag{11}$$

Es importante destacar que el gradiente $\nabla f(\bar{x})$ (como todos los vectores de $\mathbb{R}^d$) se representa por una matriz $(d \times 1)$ mientras que la Jacobiana $Jf(\bar{x})$ —que representa la derivada $Df(\bar{x}) \in \mathcal{L}(\mathbb{R}^d,\mathbb{R}) = \left(\mathbb{R}^d\right)^* \cong \mathbb{R}^d$ que pertenece al espacio dual $\left(\mathbb{R}^d\right)^*$— es una matrix $(1 \times d)$. Mientras que los espacios $\mathbb{R}^d$ y $\left(\mathbb{R}^d\right)^*$ son obviamente isómorfos (dado que tienen la misma dimensión d), no son identicos. En particular, el espacio dual (conocido también como espacio de formas) actúa sobre el espacio inicial: cada elemento de $\left(\mathbb{R}^d\right)^*$ es una aplicación lineal (continua) de $\mathbb{R}^d$ a valores reales, y dicha actuación se representa por el producto matricial de la matriz $(1 \times d)$ que representa las formas lineales (y en particular la Jacobiana) por la matriz $(d \times 1)$ que representa los elementos del espacio inicial $\mathbb{R}^d$ (en particular el gradiente).

Veremos más adelante la importancia del nuevo objeto introducido (gradiente) y su interpretación geométrica.

2.5. Criterio de diferenciabilidad (dimensión finita).

Veamos ahora un criterio práctico para establecer la diferenciabilidad de una función entre espacios de dimensión finita.

Proposición 2.12 (continuidad de las derivadas partiales)**.** Sea $f:\mathbb{R}^2\to\mathbb{R}$ una función y $x=(x_1,x_2)\in\mathbb{R}^2$. Supongamos que las derivadas parciales $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}$ existen en un abierto convexo[4] $\mathcal{U}$ que contiene el punto x y que son continuas en x. Entonces f es diferenciable en x.

Demostración. Tomamos $u=(u_1,u_2)\in\mathbb{R}^2$ de forma que $x+u\in\mathcal{U}$. Entonces, trabajando en coordenadas tenemos:

$$\begin{aligned} &f(x_1+u_1,x_2+u_2)-f(x_1,x_2)= \\ &\qquad=\underbrace{f(x_1+u_1,x_2+u_2)-f(x_1+u_1,x_2)}_{A}+\underbrace{f(x_1+u_1,x_2)-f(x_1,x_2)}_{B}. \end{aligned} \tag{12}$$

Consideramos las funciones

$$\left\{\begin{array}{l}\varphi_1:[0,u_1]\to\mathbb{R}\\ \varphi_1(s)=f(x_1+s,x_2)\end{array}\right. \qquad \text{y} \qquad \left\{\begin{array}{l}\varphi_2:[0,u_2]\to\mathbb{R}\\ \varphi_2(\tau)=f(x_1+u_1,x_2+\tau).\end{array}\right.$$

Observamos que por nuestra hipotesis, las funciones φ_1 y φ_2 son diferenciables, con derivadas dadas por las formulas (véase (9) y Definición 2.7):

$$\varphi_1'(s)=df\left((x_1+s,x_2);e_1\right):=\frac{\partial f}{\partial x_1}(x_1+s,x_2)\qquad\text{para todo } s\in[0,u_1]\qquad\text{y}$$

$$\varphi_2'(\tau)=df\left((x_1+u_1,x_2+\tau);e_2\right):=\frac{\partial f}{\partial x_2}(x_1+u_1,x_2+\tau)\quad\text{para todo } \tau\in[0,u_2].$$

Por el teorema del valor medio, existen $c_1\in[0,u_1]$ y $c_2\in[0,u_2]$ tales que:

$$B=\varphi_1(u_1)-\varphi_1(0)\equiv\varphi_1'(c_1)\,u_1=\frac{\partial f}{\partial x_1}(x_1+c_1,x_2)\,u_1\qquad\text{y}$$

$$A=\varphi_2(u_2)-\varphi_2(0)\equiv\varphi_2'(c_2)\,u_2=\frac{\partial f}{\partial x_2}(x_1+u_1,x_2+c_2)\,u_2.$$

Substituyendo los valores de A y B en la ecuación (12) obtenemos:

$$f(x_1+u_1,x_2+u_2)-f(x_1,x_2)\;=\;\frac{\partial f}{\partial x_1}(x_1+c_1,x_2)\,u_1\;+\;\frac{\partial f}{\partial x_2}(x_1+u_1,x_2+c_2)\,u_2. \tag{13}$$

Para cada $z=(z_1,z_2)$, $w=(w_1,w_2)\in\mathcal{U}$ anotemos por $T(z,w)$ la aplicación lineal continua de $\mathbb{R}^2$ a $\mathbb{R}$ representada por

$$\left(\frac{\partial f}{\partial x_1}(z),\frac{\partial f}{\partial x_2}(w)\right).$$

[4]Un conjunto $\mathcal{U}$ se dice convexo si para cada $y,z\in\mathcal{U}$ el segmento $[y,z]:=\{ty+(1-t)z:\ t\in[0,1]\}$ esta contentido en $\mathcal{U}$.

La continuidad de las derivadas parciales y la equivalencia de normas en $\mathbb{R}^2$ asegura que la aplicación

$$T : \mathcal{U} \times \mathcal{U} \left(\subset \mathbb{R}^4\right) \to \left(\mathcal{L}(\mathbb{R}^2, \mathbb{R}), ||\cdot||_{\text{op}}\right) \cong \mathbb{R}^2$$

es continua. Obtenemos de (13) que

$$f(x+u) - f(x) = \left(\frac{\partial f}{\partial x_1}(x_1 + c_1, x_2), \frac{\partial f}{\partial x_2}(x_1 + u_1, x_2 + c_2)\right) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = T(x+c, x+\tilde{c})(u)$$

donde $c = (c_1, 0)$ y $\tilde{c} = (u_1, c_2)$. Poniendo

$$T := T(x,x) = \left(\frac{\partial f}{\partial x_1}(x_1, x_2), \frac{\partial f}{\partial x_2}(x_1, x_2)\right)$$

obtenemos

$$|f(x+u) - f(x) - T(u)| = |\left(T(x+c, x+\tilde{c}) - T\right)(u)| \leq ||T(x+c, x+\tilde{c}) - T||_{\text{op}} \cdot ||u||$$

y por lo tanto

$$\frac{|f(x+u) - f(x) - T(u)|}{||u||} \leq ||T(x+c, x+\tilde{c}) - T||_{\text{op}}.$$

Observando que $||u|| \geq \text{máx}\{||c||, ||\tilde{c}||\}$, deducimos que

$$\lim_{||u|| \to 0} ||T(x+c, x+\tilde{c}) - T||_{\text{op}} = 0,$$

lo cual muestra que $T = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x)\right)$ es la derivada de f en x. □

Observación 2.13. La Proposición 2.12 se puede generalizar para funciones $f : \mathbb{R}^d \to \mathbb{R}$ con $d \geq 3$. El lector está invitado a elaborar los detalles imitando la demostración anterior. En particular, en la ecuación (12) tendriamos ahora la suma de d-terminos de forma

$$f(x_1 + u_1, \ldots, \underbrace{x_i + u_i, x_{i+1} + u_{i+1}}, x_{i+2}, \ldots, x_d) - f(x_1 + u_1, \ldots, \underbrace{x_i + u_i, x_{i+1}}, \ldots, x_d),$$

y consideraríamos d-funciones

$$\begin{cases} \varphi_i : [0, u_i] \to \mathbb{R} \\ \varphi_i(s) = (x_1 + u_1, \ldots, \underbrace{x_{i-1} + u_{i-1}, x_i + s}, x_{i+1} \ldots, x_d), \end{cases}$$

donde $i \in \{1, \ldots, d\}$. El resto de la demostración sigue los mismos pasos.

Por último, combinando la Observación 2.13 con el resultado del Ejercício 2.4 obtenemos el siguiente resultado:

Proposición 2.14 (criterio de diferenciabilidad)**.** Sea $f : \mathbb{R}^d \to \mathbb{R}^\ell$ una función. Si para cada $i \in \{1, \ldots \ell\}$ y $j \in \{1, \ldots d\}$ las derivadas parciales $\frac{\partial f_i}{\partial x_j}$ existen en un abierto convexo $\mathcal{U}$ que contiene $\bar{x} \in \mathbb{R}^d$ y son continuas en x, entonces f es diferenciable en $\bar{x}$.

El recíproco de la proposición anterior no es cierto (es decir, la Proposición 2.14 no es una caracterización de la diferenciabilidad). Esto se puede ver incluso en el caso $d = \ell = 1$, donde el resultado se traduce en que si una función diferenciable, su derivada no es necesariamente continua[5].

Ejemplo 2.15. Consideramos la función $f : \mathbb{R} \to \mathbb{R}$ dada por

$$f(t) = \begin{cases} t^2 \sin \frac{1}{t}\,, & \text{si } \ t \neq 0 \\ \\ 0\,, & \text{si } t = 0. \end{cases}$$

Es fácil ver que f es diferenciable en cada punto, con derivada

$$f'(t) = 2t\,\sin\frac{1}{t} - \cos\frac{1}{t}\,, \qquad \text{si } t \neq 0 \qquad \text{y}$$

$$f'(0) = \lim_{t\to 0} \frac{t^2 \sin(t^{-1})}{t} = \lim_{t\to 0} t\sin(t^{-1}) = 0,$$

pero f' no es continua en 0.

Basandonos al ejemplo anterior, el siguiente ejercício illustra un contrejemplo al recíproco de la Proposición 2.12. Invitamos al lector de elaborar los detalles.

Ejercicio 2.16. Mostrar que la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right)\,, & \text{si } \ (x,y) \neq (0,0) \\ \\ 0\,, & \text{si } (x,y) = (0,0). \end{cases}$$

es diferenciable en cada punto (en particular, las funciones $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ existen en todo $\mathbb{R}^2$) pero ambas derivadas parciales son discontinuas en $(0,0)$.

Volviendo a la Proposición 2.14, podemos hacer la siguiente observación: si todas las derivadas parciales de una función $f : \mathbb{R}^d \to \mathbb{R}^\ell$ existen y son continuas en un conjunto abierto no vacío $\mathcal{U} \subset \mathbb{R}^d$, entonces f es diferenciable en $\mathcal{U}$ y la función derivada

$$Df : \mathcal{U} \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^\ell) \cong \mathbb{R}^{\ell d}$$

es continua. En efecto, el espacio vectorial $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^\ell)$ es de dimensión finita (es isomorfo al espacio $M_{\ell\times d}$ de matrices que tiene dimensión $\ell \cdot d$). Por lo tanto la convergencia en el espacio normado $\big(\mathcal{L}(\mathbb{R}^d, \mathbb{R}^\ell), ||\cdot||_{\mathrm{op}}\big)$ es equivalente a la convergencia coordenada-por-coordenada de las matrices Jacobianas, lo que se traduce a convergencia de todas las derivadas parciales (véase (10).

$$\underbrace{\lim_{x_n\to x} Df(x_n) = Df(x)}_{\text{convergencia en norma } \|\cdot\|_{\mathrm{op}}} \qquad \Longleftrightarrow \qquad \begin{cases} \lim\limits_{x_n\to x} \frac{\partial f_i}{\partial x_j}(x_n) = \frac{\partial f_i}{\partial x_j}(x_n) \\ \forall i \in \{1,\ldots,\ell\},\ \forall j \in \{1,\ldots,d\} \end{cases}$$

Terminamos esta subsección con la siguiente definición.

[5]Por otra parte, para funciones $f : \mathbb{R} \to \mathbb{R}$ la derivada f' tiene siempre la propiedad de Darboux (es decir, que en cada intervalo (a,b) la derivada f' toma todos los valores entre $f'(a)$ y $f'(b)$).

Definición 2.17 (funciones de clase $\mathcal{C}^1$)**.** Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados y $f : X \to Y$ una función diferenciable (en cada $x \in X$). Decimos que f es de clase $\mathcal{C}^1$, y anotamos $f \in \mathcal{C}^1(X, Y)$, si la función derivada

$$Df : (X, \|\cdot\|_X) \to (\mathcal{L}(X, Y), ||\cdot||_{\mathrm{op}})$$

es continua[6].

Corolario 2.18. *Una función $f : \mathbb{R}^d \to \mathbb{R}^\ell$ es de clase $\mathcal{C}^1$ si y sólo si todas sus derivadas parciales existen y son continuas.*

3. Cálculo con derivadas

En esta sección veremos que la diferenciabilidad se mantiene mediante operaciones básicas: combinaciones lineales, productos (en el caso de funciones a valores reales) y composiciones.

3.1. El espacio vectorial de funciones diferenciables

Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados y $\bar{x} \in X$. Mostramos primero que el conjunto de funciones de X a Y que son diferenciables en $\bar{x}$ es un espacio vectorial.

Proposición 3.1 (diferenciabilidad de la combinación lineal)**.** Sean $f, g : X \to Y$ dos funciones diferenciables en $\bar{x} \in X$. Entonces para cada $\lambda \in \mathbb{R}$ la función:

$$\begin{cases} f + \lambda g : X \to Y \\ (f + \lambda g)(x) := f(x) + \lambda g(x), \qquad \text{para todo } x \in X \end{cases}$$

es diferenciable en $\bar{x}$ con derivada

$$D(f + \lambda g)(\bar{x}) = Df(\bar{x}) + \lambda Dg(\bar{x}).$$

Demostración. Consideramos la aplicación lineal continua

$$T := Df(\bar{x}) + \lambda Dg(\bar{x}) \in \mathcal{L}(X, Y)$$

como candidata para ser derivada en $\bar{x}$ de la función $f + \lambda g$. Entonces para $u \in X \setminus \{0\}$ tenemos:

$$\frac{||(f + \lambda g)(\bar{x} + u) - (f + \lambda g)(\bar{x}) - T(u)||_Y}{||u||_X} \leq \frac{||f(\bar{x} + u) - f(\bar{x}) - Df(\bar{x})(u)||_Y}{||u||_X} + \frac{|\lambda| ||g(\bar{x} + u) - g(\bar{x}) - Dg(\bar{x})(u)||_Y}{||u||_X}.$$

Dado que f y g son diferenciables en $\bar{x}$, el límite de la parte derecha de la desigualdad anterior es igual a 0 cuando $||u||_X \to 0$, de lo que se deduce la diferenciabilidad de la función $\lambda f + g$ con derivada T. □

[6] A veces empleamos una terminología ligeramente ambigua diciendo que la función f es *continuamente diferenciable*.

El resultado anterior dice que las funciones $f : X \to Y$ que son diferenciables en un punto $\bar{x} \in X$ forman un espacio vectorial. Combinando con la Proposición 2.3 deducimos que el espacio

$$\mathcal{D}(X,Y) = \{\ f : X \to Y :\ f \text{ diferentiable } \}$$

es un subespacio vectorial de las funciones continuas $\mathcal{C}(X,Y)$. En particular se tiene:

$$\mathcal{C}^1(X,Y) \le \mathcal{D}(X,Y) \le \mathcal{C}(X,Y) \le Y^X.$$

Consideramos ahora el caso particular donde $Y = \mathbb{R}$.

Proposición 3.2 (diferenciabilidad del producto)**.** Sean $f, g : X \to \mathbb{R}$ dos funciones diferenciables en $\bar{x} \in X$. Entonces la función:

$$\begin{cases} fg : X \to \mathbb{R} \\ (fg)(x) := f(x)g(x), \qquad \text{para todo } x \in X \end{cases}$$

es diferenciable en $\bar{x}$ con derivada

$$D(fg)(\bar{x}) = g(\bar{x})Df(\bar{x}) + f(\bar{x})Dg(\bar{x}). \tag{14}$$

Demostración. La aplicación lineal continua

$$T := g(\bar{x})Df(\bar{x}) + f(\bar{x})Dg(\bar{x}) \in \mathcal{L}(X,\mathbb{R})$$

es ahora nuestro candidato para ser derivada de la función fg en el punto $\bar{x}$. Consideramos $u \in X \setminus \{0\}$ y tenemos:

$$\begin{aligned}(fg)(\bar{x}+u) - (fg)(\bar{x}) - (g(\bar{x})Df(\bar{x}) + f(\bar{x})Dg(\bar{x}))(u) = \\ = f(\bar{x}+u)\,[g(\bar{x}+u) - g(\bar{x}) - Dg(\bar{x})(u)] + \\ + g(\bar{x})\,[f(\bar{x}+u) - f(\bar{x}) - Df(\bar{x})(u)] + [f(\bar{x}+u) - f(\bar{x})]\ Dg(\bar{x})(u).\end{aligned}$$

Recordamos por la Proposición 2.3 que f (siendo diferenciable) es continua en $\bar{x}$. Dividiendo por $||u||_X$ y tomando el límite cuando $||u||_X \to 0$ obtenemos por la (continuidad y) diferenciabilidad de f en $\bar{x}$ que:

$$\frac{f(\bar{x}+u)\,[g(\bar{x}+u) - g(\bar{x}) - Dg(\bar{x})(u)]}{||u||_X} \underset{||u||_X \to 0}{\longrightarrow} f(\bar{x}) \cdot 0 = 0$$

luego por la diferenciabilidad de la función g en $\bar{x}$:

$$\frac{g(\bar{x})\,[f(\bar{x}+u) - f(\bar{x}) - Df(\bar{x})(u)]}{||u||_X} \underset{||u||_X \to 0}{\longrightarrow} 0$$

y por último, por la continuidad de la función f en $\bar{x}$:

$$|f(\bar{x}+u) - f(\bar{x})| \frac{Dg(\bar{x})(u)}{||u||_X} \le |f(\bar{x}+u) - f(\bar{x})|\, ||Dg(\bar{x})||_{\text{op}} \underset{||u||_X \to 0}{\longrightarrow} 0$$

Se deduce que

$$\lim_{||u||_X \to 0} \frac{(fg)(\bar{x}+u) - (fg)(\bar{x}) - (g(\bar{x})Df(\bar{x}) + f(\bar{x})Dg(\bar{x}))(u)}{||u||_X} = 0$$

lo que muestra (14). $\square$

3.2. Derivada de la composición

En esta parte tratamos el caso de la composición de dos funciones. En concreto, fijamos tres espacios normados $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ y $(Z, \|\cdot\|_Z)$. El resultado principal nos dice que la derivada de la composición de dos funciones diferenciables existe y es igual a la composición de las respectivas derivadas. El enunciado preciso es como sigue:

Proposición 3.3 (diferenciabilidad de la composición)**.** Sea $f : X \to Y$ una función diferenciable en $\bar{x} \in X$ y $g : Y \to Z$ una función diferenciable en $\bar{y} = f(\bar{x}) \in Y$. Entonces la función (composición de g con f):

$$\begin{cases} h : X \to Z \\ h(x) := g(f(x)) \qquad \text{para todo } x \in X, \end{cases}$$

es diferenciable en $\bar{x}$ con derivada

$$Dh(\bar{x}) = \underbrace{Dg(f(\bar{x})) \circ Df(\bar{x})}_{\text{composición de opedores lineales continuos}} . \tag{15}$$

Demostración. Ponemos $T_1 := Df(\bar{x}) \in \mathcal{L}(X, Y)$ y $T_2 := Dg(f(\bar{x})) \equiv Dg(\bar{y}) \in \mathcal{L}(Y, Z)$. Apostamos que el operador lineal continuo $T = T_2 \circ T_1 \in \mathcal{L}(X, Z)$ será la derivada de la función h en $\bar{x}$. Para establecer eso, trabajaremos como sigue. Sea $u \in X$. Se tiene por la desigualdad triangular de la norma y la linealidad del operador T_2 que:

$$\begin{aligned} &\|h(\bar{x}+u) - h(\bar{x}) - T_2(T_1(u))\|_Z \leq \\ &\quad \leq \underbrace{\|h(\bar{x}+u) - h(\bar{x}) - T_2(f(\bar{x}+u) - f(\bar{x}))\|_Z}_{A} + \underbrace{\|T_2\left(f(\bar{x}+u) - f(\bar{x}) - T_1(u)\right)\|_Z}_{B} \end{aligned}$$

Sea $\varepsilon > 0$. Dado que f es diferenciable en $\bar{x}$ con derivada T_1, para $\varepsilon_1 = (2||T_2||_{\text{op}})^{-1}\,\varepsilon$, existe $\delta_1 > 0$ tal que para todo $u \in B_X(0, \delta_1)$ se tiene que:

$$\|f(\bar{x}+u) - f(\bar{x}) - T_2(u)\|_Y \leq \frac{\varepsilon}{2\|T_2\|_{\text{op}}} ||u||_X$$

por lo que se deduce que

$$B \leq \|T_2||_{\text{op}} ||f(\bar{x}+u) - f(\bar{x}) - T_1(u)\|_Y \leq \frac{\varepsilon}{2} ||u||_X. \tag{16}$$

En particular, para todo $u \in B_X(0, \delta_1)$ se tiene:

$$|| \underbrace{f(\bar{x}+u) - f(\bar{x})}_{v \in Y} ||_Y \leq \underbrace{\left(\frac{\varepsilon}{2||T_2||_{\text{op}}} + ||T_1||_{\text{op}} \right)}_{k>0} ||u||_X \tag{17}$$

Utilizando ahora que g es diferenciable en $\bar{y} = f(\bar{x})$, para $\varepsilon_2 = (2k)^{-1}\varepsilon$, existe $\delta_2 > 0$ tal que para todo $v \in B_Y(0, \delta_2)$ se tiene que:

$$\|g(\bar{y}+v) - g(\bar{y}) - T_2(v)\|_Z \leq \frac{\varepsilon}{2k} ||v||_Y. \tag{18}$$

Sea $\delta := \text{mín}\{\delta_1, \dfrac{\delta_2}{k}\}$. Tomamos $u \in B_X(0,\delta)$ y ponemos $v := f(\bar{x}+u) - f(\bar{x})$. De la ecuación (17) se obtiene que

$$||v||_Y \leq k||u||_X < k\delta \leq \delta_2 \quad \Longrightarrow \quad v \in B_Y(0,\delta_2).$$

Por lo tanto, dado que $\bar{y}+v = f(\bar{x}+u)$ se deduce de (18) que:

$$A \ \leq \ \|g(f(\bar{x}+u)) - g(f(\bar{x})) - T_2(f(\bar{x}+u) - f(\bar{x}))\|_Z \ \leq \ \frac{\varepsilon}{2k}\,\|v\|_Y \ \leq \ \frac{\varepsilon}{2}\,\|u\|_X\,. \tag{19}$$

Combinando (16) y (19) se concluye que:

$$\|h(\bar{x}+u) - h(\bar{x}) - (T_2 \circ T_1)(u)\|_Z \leq A + B \leq \varepsilon||u||_X$$

para todo $u \in B_X(0,\delta)$, lo cual garantiza que $Dh(\bar{x}) = T_2 \circ T_1$. □

3.3. Regla de la cadena en dimensión finita

Consideramos ahora el caso de dimensión finita, tomando $X = \mathbb{R}^d$, $Y = \mathbb{R}^m$ y $Z = \mathbb{R}^\ell$. Suponiendo que $f : \mathbb{R}^d \to \mathbb{R}^m$ es diferenciable en $\bar{x}$ y $g : \mathbb{R}^m \to \mathbb{R}^\ell$ es diferenciable en $\bar{y} = f(\bar{x})$, entonces podemos representar la derivada $Df(\bar{x})$ por la Jacobiana $Jf(\bar{x})$, que es una matriz $(m \times d)$, luego la derivada $Dg(\bar{y})$ por la Jacobiana $Jg(\bar{y})$ que es una matriz $(\ell \times m)$. En este caso, la composición (15) corresponde a una multiplicación de matrices, y la matriz Jacobiana de h en $\bar{x}$ (que representa la derivada $Dh(\bar{x})$) será la matriz $(\ell \times d)$ que resulta de dicha multiplicación.

$$\underbrace{\begin{pmatrix} \frac{\partial h_1}{\partial x_1}(\bar{x}) & \cdots & \cdots & \frac{\partial h_1}{\partial x_d}(\bar{x}) \\ \vdots & \frac{\partial h_i}{\partial x_j}(\bar{x}) & \cdots & \vdots \\ \frac{\partial h_\ell}{\partial x_1}(\bar{x}) & \cdots & \cdots & \frac{\partial h_\ell}{\partial x_d}(\bar{x}) \end{pmatrix}}_{\text{Jacobiana de } h \text{ en } \bar{x}} = \underbrace{\begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\bar{y}) & \cdots & \frac{\partial g_1}{\partial y_m}(\bar{y}) \\ \vdots & \frac{\partial g_i}{\partial y_k}(\bar{y}) & \vdots \\ \frac{\partial g_\ell}{\partial y_1}(\bar{y}) & \cdots & \frac{\partial g_\ell}{\partial y_m}(\bar{y}) \end{pmatrix}}_{\text{Jacobiana de } g \text{ en } \bar{y} = f(\bar{x})} \underbrace{\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\bar{x}) & \cdots & \cdots & \frac{\partial f_1}{\partial x_d}(\bar{x}) \\ \vdots & \frac{\partial f_k}{\partial x_j}(\bar{x}) & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\bar{x}) & \cdots & \cdots & \frac{\partial f_m}{\partial x_d}(\bar{x}) \end{pmatrix}}_{\text{Jacobiana de } f \text{ en } \bar{x}}$$

Anotando con

$$\frac{\partial f_k}{\partial x_j}(\bar{x}), \quad k \in \{1,\ldots,m\},\ j \in \{1,\ldots,d\},$$

las derivadas parciales de f en $\bar{x}$ y con

$$\frac{\partial g_i}{\partial y_k}(\bar{x}), \quad i \in \{1,\ldots,\ell\},\ k \in \{1,\ldots,m\},$$

las derivadas parciales de g en $\bar{y} = f(\bar{x})$, tenemos la siguiente formula para las derivadas parciales de h en $\bar{x}$ que proviene de la multiplicación de matrices:

$$\frac{\partial h_i}{\partial x_j}(\bar{x}) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(\bar{y})\frac{\partial f_k}{\partial x_j}(\bar{x}) \qquad \text{(regla de la cadena)} \tag{20}$$

Observación 3.4 (conveción de Einstein)**.** Se suelen hacer dos simplificaciones en la formula (20). La primera es que se omite el punto ($\bar{x}$ o respectivamente $\bar{y}$) donde se evaluan las derivadas, escribiendo solo

$$\frac{\partial h_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k} \frac{\partial f_k}{\partial x_j}.$$

La segunda simplificación (conocida como convención de Einstein) consiste en omitir el símbolo de la suma y escribir

$$\frac{\partial h_i}{\partial x_j} = \frac{\partial g_i}{\partial y_k} \frac{\partial f_k}{\partial x_j} \tag{21}$$

con el entendimiendo de que si existe un índice común, se debe sumar en todo el rango de dicho índice (en este caso $k \in \{1, \ldots, m\}$ da lugar a la sumatoria de 1 a k). Cabe mencionar que en las ecuaciones anteriores se suele hacer un abuso de notación que consiste en anotar las derivadas parciales de la función g utilizando como variable del denominador la función f. Con este abuso, la ecuación (21) daría:

$$\frac{\partial h_i}{\partial x_j} = \frac{\partial g_i}{\partial f_k} \frac{\partial f_k}{\partial x_j}$$

(Ese abuso es muy frecuente en formulas de física o de mecánica.)

3.4. Aplicaciones de la regla de la cadena

En esta parte presentaremos unas aplicaciones de la regla de la cadena, que tienen importancia intriseca.

3.4.1. Teorema del valor medio

La primera aplicación que vamos a presentar es el llamado teorema del valor medio, que es un resultado primordial de la teoría de funciones diferenciables. El resultado general es una consecuencia del mismo teorema para funciones de una variable a valores reales combinado con la Proposición 3.3 (derivada de la composición). El enunciado preciso es como sigue:

Proposición 3.5 (teorema del valor medio)**.** Sea $(X, \|\cdot\|_X)$ un espacio normado y $f : X \to \mathbb{R}$ una función diferenciable (a valores reales). Sean $x, y \in X$. Entonces existe $z_* \in [x, y]$ tal que

$$f(y) - f(x) = Df(z_*)(y - x). \tag{22}$$

Demostración. Consideramos la función $h : \mathbb{R} \to X$ con $h(t) = x + t(y - x)$, para todo $t \in \mathbb{R}$. Observamos que la función h es afín, por lo tanto diferenciable, con derivada $h'(t) = y - x$, para todo $t \in \mathbb{R}$. Por la Proposición 3.3 se deduce que la función

$$\begin{cases} \varphi : \mathbb{R} \to \mathbb{R} \\ \varphi(t) = (f \circ h)(t) \end{cases}$$

es diferenciable con derivada

$$\varphi'(t) = Df(h(t))(h'(t)) = Df(h(t))(y-x).$$

Aplicando el teorema del valor medio para la función φ entre los puntos $t=0$ y $t=1$ deducimos la existencia de un punto $t_* \in [0,1]$ tal que $\varphi(1)-\varphi(0)=\varphi'(t_*)$. Poniendo $z_* = h(t_*) = x + t_*(y-x)$ y observando que $\varphi(0) = f(h(0)) = f(x)$ y $\varphi(1) = f(h(1)) = f(y)$ obtenemos (22). $\square$

Observación 3.6 (teorema del valor medio en el caso Euclideano)**.** Si el espacio normado es el espacio Euclideano $(\mathbb{R}^d, \|\cdot\|_2)$, entonces representando la derivada $Df(z_*)$ de f en z_* por la Jaboniana

$$Jf(z_*) = (\frac{\partial f}{\partial x_1}(z_*), \ldots, \frac{\partial f}{\partial x_1}(z_*)) = \nabla f(z_*)^T$$

y usando el producto escalar, se tiene que:

$$Df(z_*)(u) = \nabla f(z_*)^T u = \langle \nabla f(z_*), u\rangle \quad \text{para cada } u \in \mathbb{R}^d,$$

por lo tanto (22) es equivalente a

$$f(y) - f(x) = \langle \nabla f(z_*), y - x\rangle.$$

Ejercicio 3.7. (i) Mostrar que si $f : X \to \mathbb{R}^\ell$ es una función diferenciable, entonces para cada $x, y \in X$ existe $M > 0$ tal que:

$$||f(y) - f(x)||_{\mathbb{R}^\ell} \leq M||x-y||_X.$$

(*Indicación*: Aplicar la Proposición 3.5 para cada coordenada f_i de la función f.)

(ii). Mostrar que si $f : X \to \mathbb{R}^\ell$ es de clase $\mathcal{C}^1$ y $Df(\bar{x}) = 0$, entonces para cada $\varepsilon > 0$ existe $\delta > 0$ tal que:

$$||f(y) - f(x)||_{\mathbb{R}^\ell} \leq \varepsilon||x-y||_X, \qquad \text{para todo } x \in B(\bar{x}, \delta).$$

3.4.2. Derivada de una función a lo largo de una curva

Sea $f : \mathbb{R}^d \to \mathbb{R}$ una función diferenciable y $\gamma : \mathbb{R} \to \mathbb{R}^d$ una curva[7] diferenciable. Anotamos por ϕ la restricción de la función f en $\gamma(\mathbb{R})$, es decir, $\phi(t) = f(\gamma(t))$, para cada $t \in \mathbb{R}$. Entonces por aplicación directa de la Proposición 3.5 se deduce que

$$\phi'(t) = Df(\gamma(t))(\dot{\gamma}(t)) = \left(\frac{\partial f}{\partial x_1}(\gamma(t)), \ldots, \frac{\partial f}{\partial x_1}(\gamma(t))\right) \begin{pmatrix} \dot{\gamma}_1(t) \\ \vdots \\ \dot{\gamma}_d(t) \end{pmatrix} = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\gamma(t))\, \dot{\gamma}_i(t). \tag{23}$$

Veamos un ejemplo concreto:

[7] Si X es un espacio normado, una función de una variable a valores en X se suele llamar curva en X.

Ejemplo 3.8. Consideramos la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x) = ||x||^2 = x_1^2 + x_2^2$ y la curva

$$\gamma(t) = (\gamma_1(t), \gamma_2(t))^T = (\cos t, \sin t)^T, \quad t \in \mathbb{R}.$$

Entonces $Df(x) = (2x_1, 2x_2)$ y $\dot\gamma(t) = (-\sin t, \cos t)^T$. Por lo tanto:

$$Df(\gamma(t))(\dot\gamma(t)) = (2\cos t, 2\sin t)\begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} = -2\cos t \sin t + 2\cos t \sin t = 0,$$

lo que es compatible con el hecho que la restrición de la norma en la esfera unitaria $S^1 := \gamma(\mathbb{R})$ es constante.

Observación 3.9 (formas diferenciables vs vectores tangentes)**.** (★) En la geometría diferencial, las derivadas de las curvas diferenciables sirven para definir los espacios tangentes en una *variedad diferenciable*[8]. En efecto, cada elemento del espacio tangente en un punto $\bar x$ se puede representar como la derivada $\dot\gamma(0)$ de una curva diferenciable γ con $\gamma(0) = \bar x$. En el caso de un espacio normado $(X, \|\cdot\|_X)$, su espacio tangente $\boldsymbol{T}_X(\bar x)$ en un punto $\bar x \in X$ es isomorfo a X, y en este sentido, las derivadas de las curvas de X son elementos de la misma naturaleza como los elementos de X (y se representan por vectores columnas en el caso que $X = \mathbb{R}^d$). Por otro lado, las derivadas de funciones f de X a valores reales son elementos del espacio dual $\boldsymbol{T}_X(\bar x)^* \approx X^*$ al espacio tangente, y se suelen llamar *formas diferenciables.* Las formas diferenciales actúan de forma natural sobre los vectores tangentes. La ecuación (23) es un caso particular, con la forma $Df(\gamma(t)) \in \boldsymbol{T}_{\mathbb{R}^d}(\gamma(t))^* \approx \left(\mathbb{R}^d\right)^*$ actuando sobre $\dot\gamma(t) \in \boldsymbol{T}_{\mathbb{R}^d}(\gamma(t)) \approx \mathbb{R}^d$.

3.4.3. Cambio de coordenadas

En esta parte consideramos funciones diferenciables definidas en el plano $\mathbb{R}^2$ a valores reales. Notemos primero que cada vector no nulo $x \in \mathbb{R}^2\setminus\{(0,0)\}$ se puede representar a la vez mediante sus coordenadas cartesianas $(x_1, x_2)^T$ y mediante las coordenadas polares: el modulo $r = ||x||_2 = \sqrt{x_1^2 + x_2^2}$ y el ángulo $\theta = \arg\tan(x_2/x_1)$. Observamos que la función de cambio de coordenadas (polares hacía cartesianas)

$$\begin{cases} X : (0, +\infty) \times (-\pi, \pi] \to \mathbb{R}^2 \setminus \{(0,0)\} \\ \\ X(r,\theta) = (X_1(r,\theta), X_2(r,\theta))^T = (r\cos\theta, r\sin\theta)^T \end{cases}$$

es una biyección diferenciable[9] con derivada $DX(r,\theta) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ representada por la Jacobiana

$$JX(r,\theta) = \begin{pmatrix} \frac{\partial X_1}{\partial r}(r,\theta) & \frac{\partial X_1}{\partial \theta}(r,\theta) \\ \frac{\partial X_2}{\partial r}(r,\theta) & \frac{\partial X_2}{\partial \theta}(r,\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Si $f : \mathbb{R}^2 \to \mathbb{R}$ es una función diferenciable dada por una formula en base de sus coordenadas cartesianas (por ejemplo, la función del Ejemplo (3.8) $f(x) = x_1^2 + x_2^2 = ||x||_2^2$)

[8] Una variedad diferenciable es un objeto matemático que localmente aparece como espacio vectorial.

[9] en realidad X es un difeomorfismo, es decir, una biyección diferenciable con inversa diferenciable.

entonces $F = f \circ X$ es la misma función expresada en coordenadas polares (en el ejemplo anterior sería $F(r, \theta) = r^2$), se deduce de la Proposición 3.3 que

$$DF(r, \theta) = Df(X(r, \theta)) \circ DX(r, \theta)$$

lo que conduce a siguente representación matricial mediante las Jacobianas:

$$\left(\frac{\partial F}{\partial r}(r,\theta), \frac{\partial F}{\partial \theta}(r,\theta)\right) = \left(\frac{\partial f}{\partial x_1}(X(r,\theta)), \frac{\partial f}{\partial x_2}(X(r,\theta))\right) \begin{pmatrix} \frac{\partial X_1}{\partial r}(r,\theta) & \frac{\partial X_1}{\partial \theta}(r,\theta) \\ \frac{\partial X_2}{\partial r}(r,\theta) & \frac{\partial X_2}{\partial \theta}(r,\theta) \end{pmatrix}. \tag{24}$$

Se deduce de lo anterior que

$$\left\{ \begin{array}{lllll} \frac{\partial F}{\partial r} & = & \frac{\partial f}{\partial x_1}\frac{\partial X_1}{\partial r} + \frac{\partial f}{\partial x_2}\frac{\partial X_2}{\partial r} & = & \left(\frac{\partial f}{\partial x_1}\right)\cos\theta - \left(\frac{\partial f}{\partial x_2}\right)\sin\theta \\ & & & & \\ \frac{\partial F}{\partial \theta} & = & \frac{\partial f}{\partial x_1}\frac{\partial X_1}{\partial \theta} + \frac{\partial f}{\partial x_2}\frac{\partial X_2}{\partial \theta} & = & -r\left(\frac{\partial f}{\partial x_1}\right)\sin\theta + r\left(\frac{\partial f}{\partial x_2}\right)\cos\theta. \end{array} \right. \tag{25}$$

Ejemplo 3.10. En el ejemplo de la función $f(x) = x_1^2 + x_2^2 = ||x||_2^2$ con expresión en coordenadas polares $F(r, \theta) = r^2$ tenemos:

$$Jf(x_1, x_2) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x)\right) = \qquad \text{y} \qquad JF(r,\theta) = (2r, 0)$$

lo que es compatible con (24):

$$(2x_1, 2x_2)\ |_{(x_1,x_2)=X(r,\theta)} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = (2r\cos\theta, 2r\sin\theta) \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = (2r, 0)$$

3.5. Uso práctico de la regla de la cadena.

La regla de la cadena en dimensión finita se traduce por una ecuación matricial (mediante las Jacobianas) la cual nos lleva a una expresión de las derivadas parciales de la composición en términos de las derivadas parciales de las funciones compuestas, véase (24) y (25) respectivamente. En esta subsección veremos una manera informal de usar la regla de la cadena en la práctica, lo que será muy relevante en cursos de física o mecánica. Explicaremos esta técnica mediante dos ejemplos:

Ejemplo 3.11 (derivación de expresiones informales)**. (i).** Consideramos la expresión

$$W = x\exp(y/z)$$

donde las variables x, y e z dependen de otra variable t como sigue:

$$\left\{ \begin{array}{l} x = t^2 \\ y = 1 - t \\ z = 1 + 2t \end{array} \right.$$

Nos interesa calcular la derivada $\dot{W}(t) = \frac{dW}{dt}$. La manera formal (y más lenta) para hacer eso sería considerar las funciones

$$\left\{ \begin{array}{l} F : \mathbb{R}^3 \to \mathbb{R} \\ F(x,y,z) = W \end{array} \right. \qquad \text{y} \qquad \left\{ \begin{array}{l} G : \mathbb{R} \to \mathbb{R}^3 \\ G(t) = (x(t), y(t), z(t))^T \end{array} \right.$$

luego escribir $W(t) = (F \circ G)(t)$ y calcular las Jacobianas respectivas. Otra manera (informal, pero más rápida) sería considerar el diagrama

$$\boxed{\begin{array}{ccc} & \mathbf{W} & \\ / & | & \backslash \\ \mathbf{x} & \mathbf{y} & \boldsymbol{z} \\ | & | & | \\ t & t & t \end{array}}$$

donde las variables aparecen a 3 niveles. En el primer nivel es la variable W, en el segundo nivel son las variables x, y e z y en el tercer nivel es la variable t. Para encontrar la derivada de W con respecto a t, basta derivar W con respecto a cada una de las variables del nivel inferior multiplicando cada vez con la derivada de dichas variables con respecto a t. En concreto:

$$\begin{aligned} \frac{dW}{dt} &= \frac{dW}{dx}\frac{dx}{dt} + \frac{dW}{dy}\frac{dy}{dt} + \frac{dW}{dz}\frac{dz}{dt} = \\ &= 2t\left(e^{y(t)/z(t)}\right) - \left(\frac{x(t)}{z(t)}e^{y(t)/z(t)}\right) - 2\left(\frac{x(t)y(t)}{z(t)^2}e^{y(t)/z(t)}\right) = \\ &= \left(2t - \frac{t^2}{1+2t} - \frac{2t^2(1-t)}{(1+2t)^2}\right)\exp(\frac{1-t}{1+2t}) = \frac{8t^3+5t^2+2t}{(1+2t)^2}\exp(\frac{1-t}{1+2t}). \end{aligned}$$

(ii). Vemos ahora un segundo ejemplo:

$$z = e^r \cos\theta \qquad \text{donde} \quad \begin{cases} r = s \cdot t \\ \theta = \sqrt{s^2+t^2}. \end{cases} \tag{26}$$

Representamos (26) mediante el diagrama:

$$\boxed{\begin{array}{cccccc} & & \mathbf{z} & & & \\ & & / & \backslash & & \\ & \boldsymbol{r} & & & \boldsymbol{\theta} & \\ / & | & & & | & \backslash \\ \mathbf{s} & \mathbf{t} & & & \mathbf{s} & \mathbf{t} \end{array}}$$

Nos interesa calcular la derivada parcial de la variable z con respecto a s. Aplicando la derivación informal en cadena tenemos:

$$\begin{aligned} \frac{dz}{ds} &= \frac{dz}{dr}\frac{dr}{ds} + \frac{dz}{d\theta}\frac{d\theta}{ds} = \left(e^{r(t)}\cos\theta(t)\right)t + \left(-e^{r(t)}\sin\theta(t)\right)\frac{s}{\sqrt{s^2+t^2}} = \\ &= \exp(st)\left(t\cos(\sqrt{s^2+t^2}) - \frac{s\cdot\sin(\sqrt{s^2+t^2})}{\sqrt{s^2+t^2}}\right). \end{aligned}$$

Ejemplo 3.12 (regla de la cadena: más ejemplos)**.** Consideramos la función

$$\varphi(t) = f(t, f(t,t)), \qquad t \in \mathbb{R} \tag{27}$$

donde f es una función diferenciable de $\mathbb{R}^2$ a $\mathbb{R}$. Calculemos su derivada mediante los dos métodos.

Método formal: Introducimos las funciones

$$\begin{cases} I : \mathbb{R} \to \mathbb{R}^2 \\ I(t) = (t,t)^T \end{cases} \qquad \text{y} \qquad \begin{cases} G : \mathbb{R}^2 \to \mathbb{R}^2 \\ G(t,s) = (t, f(t,s))^T. \end{cases}$$

Observamos que $\varphi = f \circ G \circ I$ por lo tanto:

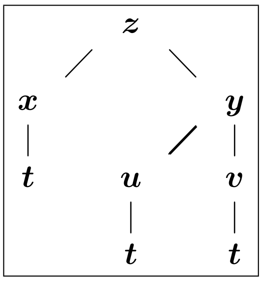
$$\varphi'(t) = Df(G(I(t)) \circ DG(I(t)) \circ DI(t)$$

lo que nos lleva a la expresión:

$$\varphi'(t) = \begin{pmatrix} \frac{\partial f(G(I(t))}{\partial x} & \frac{\partial f(G(I(t))}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ \frac{\partial f(I(t))}{\partial x} & \frac{\partial f(I(t))}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial f(t,f(t,t))}{\partial x} & \frac{\partial f(t,f(t,t))}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \frac{\partial f(t,t)}{\partial x} + \frac{\partial f(t,t)}{\partial y} \end{pmatrix} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\right).$$

Método informal: Representamos (27) mediante el diagrama:



donde $z = f(x,y)$, $x = x(t) = t$, $y = y(u,v)$ con $u = u(t) = t$ y $v = v(t) = t$. Calculamos:

$$\begin{aligned} \varphi'(t) = \frac{dz}{dt} &= \frac{dz}{dx}\frac{dx}{dt} + \frac{dz}{dy}\left(\frac{dy}{dt}\right) = \\ &= \frac{dz}{dx}\frac{dx}{dt} + \frac{dz}{dy}\left(\frac{dy}{du}\frac{du}{dt} + \frac{dy}{dv}\frac{dv}{dt}\right) = \\ &= \frac{\partial f}{\partial x}\cdot 1 + \frac{\partial f}{\partial y}\left(\frac{\partial f}{\partial x}\cdot 1 + \frac{\partial f}{\partial y}\cdot 1\right) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\right) \end{aligned}$$

Ejercicio 3.13. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable. Calcule la derivada de la función

$$\varphi(t) = f(t, f(t, f(t,t))), \qquad t \in \mathbb{R}$$

utilizando las dos maneras que se han visto (la manera formal y la manera informal).

4. Ejemplos más elaborados de derivación

En esta parte veremos unos ejemplos más elaborados de cálculo de derivada que tienen a la vez interés teorico y práctico.

4.1. Derivada de una forma bilineal continua

Sea $(\mathcal{H}, \|\cdot\|)$ un espacio de Hilbert, donde la norma $\|\cdot\|$ proviene de un producto escalar b. Consideramos el espacio producto

$$\mathcal{X} = \mathcal{H} \times \mathcal{H}$$

(que es un espacio vectorial) equipado con la siguiente norma:

$$||(x,y)||_{\mathcal{X}} := \sqrt{||x||^2 + ||y||^2}, \qquad \text{para cada } (x,y) \in \mathcal{X}. \tag{28}$$

Es fácil ver (véase Ejercício 4.1) que la formula anterior es una norma en $\mathcal{X}$ que proviene del producto escalar:

$$\mathcal{B}((x,y),(x',y')) := b(x,x') + b(y,y')$$

Ejercicio 4.1. Mostrar que $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ es un espacio de Hilbert.

Nuestro objetivo es mostrar que el producto escalar b de $\mathcal{H}$, visto como función de $\mathcal{X}$ a valores reales, es diferenciable en cada $(x,y) \in \mathcal{X}$.

Proposición 4.2 (Derivada del producto escalar)**.** Sea $(\mathcal{H}, \|\cdot\|)$ un espacio de Hilbert con producto escalar b. Sea $\mathcal{X} = \mathcal{H} \times \mathcal{H}$ equipado con la norma (28). Entonces la función

$$b : \mathcal{X} \to \mathbb{R}$$

es diferenciable en cada $(x,y) \in \mathcal{X}$ con derivada:

$$Db((x,y))(u_1,u_2) = b(x,u_2) + b(u_1,y), \qquad \text{para cada } (u_1,u_2) \in \mathcal{X}. \tag{29}$$

Demostración. Sea $(x,y) \in \mathcal{X}$. Necesitamos encontrar una aplicación lineal continua $T \in \mathcal{L}(\mathcal{X},\mathbb{R}) := \mathcal{X}^*$ que será un candidato razonable para formar el límite en (8). Para eso, calculamos primero la derivada direccional. Sea $u = (u_1,u_2) \neq (0,0)$ una dirección en $\mathcal{X}$. Entonces:

$$db((x,y);(u_1,u_2)) := \lim_{t\to 0} \frac{b(x+tu_1, y+tu_2) - b(x,y)}{t} = b(x,u_2) + b(u_1,y).$$

El cálculo anterior sugiere que la aplicación

$$\begin{cases} T : \mathcal{X} \to \mathbb{R} \\ T(u_1,u_2) = b(x,u_2) + b(u_1,y)\,, \qquad \text{para cada } (u_1,u_2) \in \mathcal{X}, \end{cases} \tag{30}$$

es el único candidato para ser derivada. Utilizando el hecho que b es una forma bilineal, dejamos al lector de comprobar que la aplicación T definida por (30) es una aplicación lineal en $\mathcal{X}$, es decir, para cada (u_1,u_2), $(v_1,v_2) \in \mathcal{X}$ y $\lambda \in \mathbb{R}$ se tiene:

$$\begin{cases} T((u_1,u_2)+(v_1,v_2)) := T(u_1+v_1, u_2+v_2) = T(u_1,u_2) + T(v_1+v_2) \\ T(\lambda(u_1,u_2)) := T(\lambda u_1, \lambda u_2) = \lambda T(u_1,u_2) \end{cases}$$

Veamos ahora que T es continua: En efecto, para cada $u = (u_1, u_2) \in \mathcal{X}$, usando la desigualdad Cauchy-Schwarz para el producto escalar b en $\mathcal{H}$ deducimos:

$$|T(u_1, u_2)| \leq |b(x, u_2)| + |b(u_1, y)| \leq ||x|| \cdot ||u_2|| + ||y|| \cdot ||u_1||. \tag{31}$$

Poniendo $\alpha_1 = ||x||$, $\alpha_2 = ||y||$ luego $\beta_1 = ||u_2||$ y $\beta_2 = ||u_1||$, aplicando nuevamente la desigualdad Cauchy-Schwarz para el producto escalar canónico $\langle \cdot, \cdot \rangle$ en $\mathbb{R}^2$, obtenemos:

$$||x|| \cdot ||u_2|| + ||y|| \cdot ||u_1|| = \alpha_1\beta_1 + \alpha_2\beta_2 \leq \sqrt{\alpha_1^2 + \alpha_2^2}\sqrt{\beta_1^2 + \beta_2^2} = ||(x, y)||_{\mathcal{X}} \cdot ||(u_1, u_2)||_{\mathcal{X}}. \tag{32}$$

Combinando (31) con (32) concluimos que $T \in \mathcal{L}(\mathcal{X}, \mathbb{R})$ con

$$||T||_{\text{op}} \leq ||(x, y)||_{\mathcal{X}}.$$

Nos queda comprobar que $T = Db((x, y))$. Para eso, consideramos $u = (u_1, u_2) \neq (0, 0)$ y observamos que:

$$\begin{aligned} \frac{|b(x + u_1, y + u_2) - b(x, y) - T(u_1, u_2)|}{||(u_1, u_2)||_{\mathcal{X}}} &= \frac{|b(u_1, u_2)|}{\sqrt{||u_1||^2 + ||u_2||^2}} \\ &\leq \frac{||u_1|| \cdot ||u_2||}{\sqrt{||u_1||^2 + ||u_2||^2}} \;\leq\; \frac{1}{2}||(u_1, u_2)||_{\mathcal{X}} \end{aligned}$$

y por lo tanto

$$\lim_{||u||_{\mathcal{X}} \to 0} \frac{b(x + u_1, y + u_2) - b(x, y) - T(u_1, u_2)}{||(u_1, u_2)||_{\mathcal{X}}} = 0,$$

es decir, $Db((x, y)) = T$. □

Observación 4.3 (derivada de una forma bilineal continua)**.** Invitamos al lector de comprobar que la demostración anterior sigue válida si supongamos que en lugar de ser producto escalar, la función b es solamente una forma bilineal continua en $\mathcal{H}$, es decir: $b(u_1, u_2) \leq M||u_1|| \cdot ||u_2||$, para todo $u_1, u_2 \in \mathcal{H}$. En particular, la demostración anterior no hace uso de la simetría del producto escalar, ni el hecho que está definida positiva. Como consecuencia se obtiene el siguiente resultado (donde usamos la misma notación como en la Proposición 4.2):

Si $b : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ es una forma bilineal continua, entonces b, vista como función en el espacio producto $\mathcal{X} = \mathcal{H} \times \mathcal{H}$, es diferenciable en cada $(x, y) \in \mathcal{X}$ con derivada dada por la formula (29).

4.2. Derivada de una formula integral

Sean $\beta : \mathbb{R} \to \mathbb{R}$ y $f : \mathbb{R}^2 \to \mathbb{R}$ dos funciones de clase $\mathcal{C}^1$. Nos proponemos calcular la derivada

$$\frac{d}{dx}\left[\int_0^{\beta(x)} f(x, s)\, ds\right].$$

Para hacer eso, trabajamos como sigue: definimos las funciones

$$\begin{cases} G : \mathbb{R} \to \mathbb{R}^2 \\ G(x) = (x, \beta(x)) \end{cases} \qquad \text{y} \qquad \begin{cases} F : \mathbb{R}^2 \to \mathbb{R} \\ F(x,y) = \displaystyle\int_0^y f(x,s)\,ds. \end{cases}$$

La función G es diferenciable en todo $x \in \mathbb{R}$ con derivada representada por la Jacobiana $(1, \beta'(x))^T$. Mostramos que F también es diferenciable, ocupando el criterio de la Proposición 2.12 (continuidad de las derivadas parciales). Calculamos las derivadas parciales:

$$\frac{\partial F}{\partial x}(x,y) = \lim_{t\to 0} \int_0^y \left(\frac{f(x+t,s) - f(x,s)}{t} \right) ds.$$

Para cada $s \in \mathbb{R}^2$, la función $x \mapsto f(x,s)$ es diferenciable. Tomando $t \in [-1,1]$ y aplicando el teorema del valor medio existe $\theta(s) \in [0,1]$ tal que

$$f(x+t,s) - f(x,s) = t \cdot \frac{\partial f}{\partial x}(x + \theta(s)t, s).$$

Dado que f es de clase $\mathcal{C}^1$, su derivada parcial $\frac{\partial f}{\partial x}$ es continua y por lo tanto uniformemente continua en el compacto $[x-1, x+1] \times [0,y]$. Deducimos que:

$$\lim_{t\to 0} \max_{s\in[0,L]} \left| \frac{\partial f}{\partial x}(x+\theta(s)t,s) - \frac{\partial f}{\partial x}(x,s) \right| = 0,$$

y por lo tanto:

$$\frac{\partial F}{\partial x}(x,y) = \lim_{t\to 0} \int_0^y \frac{\partial f}{\partial x}(x+\theta(s)t,s)\,ds = \int_0^y \lim_{t\to 0} \left(\frac{\partial f}{\partial x}(x+\theta(s)t,s) \right) ds = \int_0^y \frac{\partial f}{\partial x}(x,s)\,ds.$$

Por otra parte, fijando $x \in \mathbb{R}$ y aplicando el teorema fundamental del cálculo en la función $s \mapsto f(x,s)$ obtenemos:

$$\frac{\partial F}{\partial y}(x,y) = \frac{d}{dy} \int_0^y f(x,s)\,ds = f(x,y).$$

Dejamos al lector de comprobar que ambas derivadas parciales de F

$$\frac{\partial F}{\partial x}(x,y) = \int_0^y \frac{\partial f}{\partial x}(x,s)\,ds \qquad \text{y} \qquad \frac{\partial F}{\partial y}(x,y) = f(x,y)$$

son continuas en $\mathbb{R}^2$ por lo que se deduce que F es derivable con Jacobiana

$$JF(x,y) = \left(\int_0^y \frac{\partial f}{\partial x}(x,s)\,ds, \ f(x,y) \right) \qquad \text{para todo } (x,y) \in \mathbb{R}^2.$$

Concluimos que la función

$$\begin{cases} H : \mathbb{R} \to \mathbb{R} \\ H(x) = (F \circ G)(x) = F(G(x))G(x) = \displaystyle\int_0^{\beta(x)} f(x,s)\, ds \end{cases}$$

es derivable en todo $x \in \mathbb{R}$ con derivada:

$$H'(x) = JF(x, \beta(x)) \cdot JG(x) = \left(\int_0^{\beta(x)} \frac{\partial f}{\partial x}(x,s)\, ds,\ f(x,\beta(x)) \right) \begin{pmatrix} 1 \\ \beta'(x) \end{pmatrix}$$

por lo que se obtiene la formula de derivación:

$$\left(\int_0^{\beta(x)} f(x,s)\, ds \right)' = \int_0^{\beta(x)} \frac{\partial f}{\partial x}(x,s)\, ds + \beta'(x)\, f(x,\beta(x)).$$

5. Gradiente y espacio tangente

En esta parte veremos que el gradiente admite una interesante interpretación geométrica en el espacio Euclideano $(\mathbb{R}^d, \|\cdot\|_2)$. A continuación, consideramos una función diferenciable $f : \mathbb{R}^d \to \mathbb{R}$ y un punto $\bar{x} \in \mathbb{R}^d$ tal que $Df(\bar{x}) \neq 0$ (dicho punto se llama *regular*).

5.1. Dirección del máximo ascenso

Para todo $u \in \mathbb{R}^d$ se tiene:

$$|Df(\bar{x})(u)| = |\langle \nabla f(\bar{x}), u \rangle| \leq ||\nabla f(\bar{x})||_2 \cdot ||u||_2$$

por lo que se deduce que

$$||\nabla f(\bar{x})||_2 \geq \sup_{||u||_2=1} Df(\bar{x})(u) = \sup_{||u||_2=1} df(\bar{x}; u). \tag{33}$$

Ponemos

$$u_* = \frac{\nabla f(\bar{x})}{||\nabla f(\bar{x})||_2} \qquad \text{(dirección unitaria del gradiente)} \tag{34}$$

y observamos que

$$|Df(\bar{x})(u_*)| = |\langle \nabla f(\bar{x}), u_* \rangle| = ||\nabla f(\bar{x})||_2,$$

es decir, el supremo en (33) es máximo y la desigualdad se cumple con igualdad (en particular, $||Df(\bar{x})||_{\mathrm{op}} = ||\nabla f(\bar{x})||_2$). La dirección del gradiente u_* se llama *dirección del* *máximo ascenso* de la función f, en el sentido que dicha dirección maximiza la derivada direccional de f con respecto a todas las posibles direcciones unitarias:

$$df(\bar{x}; u_*) = \sup_{||u||_2=1} df(\bar{x}; u).$$

Por otra parte el modulo $||\nabla f(\bar{x})||_2$ del gradiente corresponde al valor de la derivada direccional en la dirección del máximo ascenso, es decir:

$$df(\bar{x}; u_*) = ||\nabla f(\bar{x})||_2.$$

Ilustraremos la situación con dos ejemplos:

Ejemplo 5.1 (máximo ascenso)**. (i).** Consideramos la función $f : \mathbb{R}^2 \to \mathbb{R}$ con

$$f(x_1, x_2) = x_1^2 + x_2^2 = ||x||_2^2, \qquad x = (x_1, x_2) \in \mathbb{R}^2.$$

El gradiente de f en el punto $x = (x_1, x_2) \in \mathbb{R}^2$ está dado por la formula

$$\nabla f(x_1, x_2) = (2x_1, 2x_2) \quad \Leftrightarrow \quad \nabla f(x) = 2x,$$

es decir, es co-lineal con el vector x. Eso significa que en cada $x \neq 0$ el máximo crecimiento de la función f se alcanza en la dirección definida por x, lo que es normal por la isotropía de f (sus niveles son las esferas centradas en el origen). En particular, la dirección $u_*(1,1)$ del máximo ascenso de f en el punto $(1,1)$ es la que apunte hacia la misma diagonal

$$u_*(1,1) = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}).$$

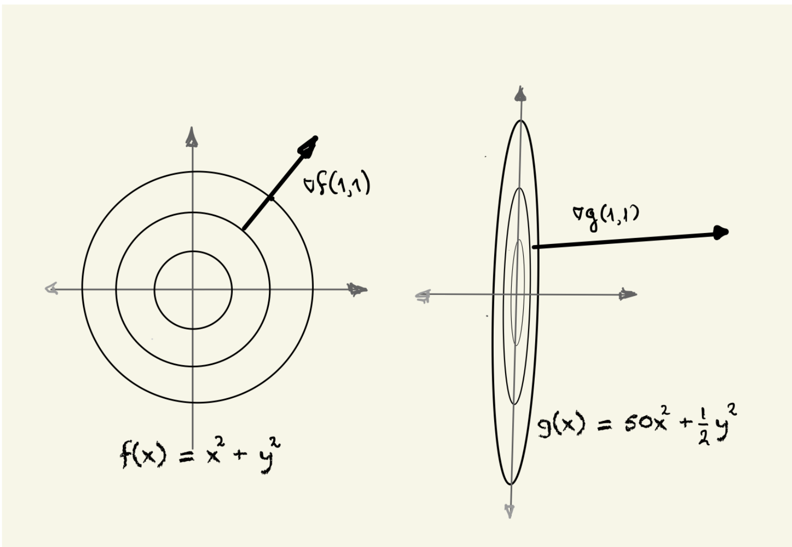


Figura 2: Dirección de máximo ascenso.

(ii). Consideramos ahora la función

$$g(x_1, x_2) = 50\, x_1^2 + \frac{1}{2} x_2^2, \qquad x = (x_1, x_2) \in \mathbb{R}^2.$$

En este caso

$$\nabla g(x_1, x_2) = (100\, x_1, x_2)$$

y la dirección $u_*(1,1)$ del máximo ascenso de g en el punto $(1,1)$ es casi horizontal:

$$u_*(1,1) = (\frac{100}{\sqrt{10001}}, \frac{1}{\sqrt{10001}}) \approx (1,0).$$

Ejercicio 5.2 (máximo descenso)**.** Si $\bar{x} \in \mathbb{R}^d$ es un punto regular de una función diferenciable $f : \mathbb{R}^d \to \mathbb{R}$, mostrar que la dirección

$$v_* = -\frac{\nabla f(\bar{x})}{||\nabla f(\bar{x})||_2} = -u_*$$

es la dirección del máximo descenso de f.

5.2. Espacio tangente

En la última parte de la Subsección 2.1 hemos definido la recta tangente al grafo de una función diferenciable $f : \mathbb{R} \to \mathbb{R}$ en el punto $(\bar{x}, f(\bar{x}))$ como el grafo de la función afín $\ell(x) := f(\bar{x}) + f'(\bar{x})(x - \bar{x})$, $x \in \mathbb{R}$, de su aproximación. Poniendo $y = \ell(x)$ e $\bar{y} = f(\bar{x})$ obtenemos la siguiente formula para la recta tangente:

$$y - \bar{y} = f'(\bar{x})(x - \bar{x}) \qquad \Longleftrightarrow \qquad \langle (f'(\bar{x}), -1), (x - \bar{x}, y - \bar{y}) \rangle_{\mathbb{R}^2} = 0 \tag{35}$$

donde $\langle \cdot, \cdot \rangle_{\mathbb{R}^2}$ denota el producto escalar canónico de $\mathbb{R}^2$ entre el vector $(f'(\bar{x}), -1)$ y el vector $(x - \bar{x}, y - \bar{y})$, es decir, (x, y) pertenece a la recta tangente del gráfo de f que pasa por el punto $(\bar{x}, \bar{y})$ si y sólo si

$$(f'(\bar{x}), -1) \perp (x - \bar{x}, y - \bar{y}).$$

Lo anterior admite una generizalización directa en el caso de una función diferenciable $f : \mathbb{R}^d \to \mathbb{R}$. En efecto, para cada $\bar{x} \in \mathbb{R}^d$ se define el *espacio tangente* $\boldsymbol{T}_{\mathrm{gph}(f)}(\bar{x}, f(\bar{x}))$ al grafo de f

$$\mathrm{gph}(f) := \big\{ (x, f(x)) : \; x \in \mathbb{R}^d \big\} \subset \mathbb{R}^{d+1}$$

en el punto $(\bar{x}, f(\bar{x})) \in \mathbb{R}^{d+1}$ como sigue: Para cada $(x, y) = (x_1, \ldots, x_d, y) \in \mathbb{R}^{d+1}$ se tiene:

$$(x, y) \in \boldsymbol{T}_{\mathrm{gph}(f)}(\bar{x}, f(\bar{x})) \qquad \Longleftrightarrow \qquad \langle (\nabla f(\bar{x}), -1), (x - \bar{x}, y - \bar{y}) \rangle_{\mathbb{R}^{d+1}} = 0 \tag{36}$$

En el caso $d = 2$, el espacio tangente se llama *plano tangente* al grafo de una función $f : \mathbb{R}^2 \to \mathbb{R}$ en el punto $(\bar{x}, f(\bar{x})) = (\bar{x}_1, \bar{x}_2, f(\bar{x}_1, \bar{x}_2)) \in \mathbb{R}^3$. En este caso, el plano tangente tiene *normal*

$$\hat{n}_{\bar{x}} = (\nabla f(\bar{x}), -1)$$

y pasa por el punto $(\bar{x}, f(\bar{x}))$.

Ejercicio 5.3 (espacio tangente como núcleo de un gradiente)**.** Sea $f : \mathbb{R}^d \to \mathbb{R}$ una función diferenciable y $\bar{x} \in \mathbb{R}^d$ (arbitrario). Muestre que el punto $(\bar{x}, f(\bar{x})) \in \mathbb{R}^{d+1}$ es un punto regular de la función

$$\begin{cases} \Phi : \mathbb{R}^{d+1} \to \mathbb{R} \\ \Phi(x,y) = f(x) - y \end{cases}$$

y que el espacio tangente $\boldsymbol{T}_{\mathrm{gph}(f)}(\bar{x}, f(\bar{x}))$ coincide con el núcleo de la derivada de Φ en dicho punto.

Veremos ahora algunos ejemplos de como calcular el espacio tangente en situaciones concretas.

Ejemplo 5.4 (rectas tangentes a la esfera S^1)**.** Consideramos la esfera unitaria

$$S^1 := \{(x,y) \in \mathbb{R}^{d+1} : \, x^2 + y^2 = 1\} \subset \mathbb{R}^2$$

y buscamos determinar la ecuación de la recta tangente que pasa por los puntos $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, $(0,1)$ y $(-1,0)$, véase Figura 3.

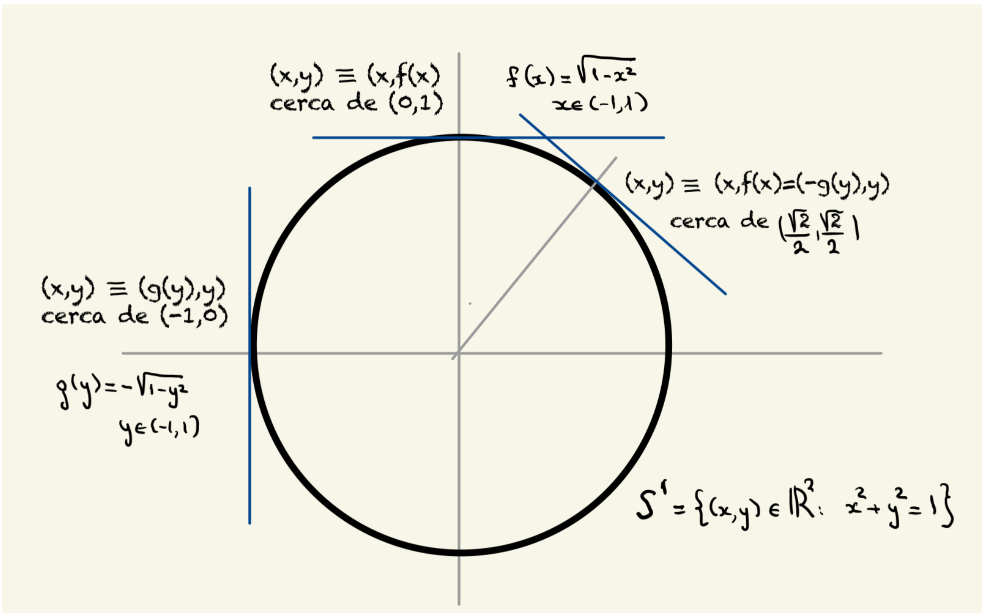


Figura 3: Visualización local de S^1 como grafo de una función.

Observamos que localmente, alrededor de los dos primero puntos, la esfera S^1 coincide con el grafo de la función

$$f(x) = \sqrt{1-x^2}, \quad x \in (-1,1).$$

Dado que $f'(x) = -x(1-x^2)^{-1/2}$, obtenemos $f'(\frac{\sqrt{2}}{2}) = -1$ y $f'(0) = 0$. Por lo tanto, las rectas tangentes a los puntos $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ y $(0,1)$ tienen ecuación:

$$y - \bar{y} = f'(\bar{x})(x - \bar{x}) \quad \Longrightarrow \quad \begin{cases} y - \frac{\sqrt{2}}{2} = (-1)(x - \frac{\sqrt{2}}{2}) \;\Leftrightarrow\; y = -x + \sqrt{2} \\[2ex] y - 1 = 0(x-0) \qquad \Leftrightarrow \qquad y = 1. \end{cases} .$$

Por otra parte, alrededor del punto $(-1,0)$ la esfera no puede ser el grafo de ninguna función $y = f(x)$, pero intercambiando coordenadas y considerando funciones de variable y, vemos que la esfera coincide con el grafo de la función

$$x = g(y) = -\sqrt{1-y^2}, \quad y \in (-1,1),$$

alrededor del punto $(-1,0) \equiv (g(0),0)$. Por lo tanto, en este caso, la ecuación de la recta tangente sera:

$$x - \bar{x} = g'(\bar{y})(y - \bar{y}) \quad \Longrightarrow \quad x - (-1) = 0(y - 0) \Leftrightarrow x = -1.$$

Ejercicio 5.5 (planos tangente a la esfera S^2)**.** Consideramos la esfera unitaria

$$S^2 := \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\} \subset \mathbb{R}^3.$$

Determine la ecuación del plano tangente que pasa por los puntos $(1,0,1)$, $(\sqrt{2},0,0)$ y $(0,0,\sqrt{2})$.

6. Teorema de la Función Inversa

En esta parte veremos el teorema de la función inversa, que constituye el segundo teorema pilar de este curso. Dicho teorema, cuya demostración está basada al teorema del punto fijo de Banach (primer teorema pilar, visto en la Parte I de este apunte) proporciona condiciones suficientes para que una función $f : \mathbb{R}^n \to \mathbb{R}^n$ sea invertible localmente en un entorno de un punto $\bar{x} \in \mathbb{R}^n$ en términos de (la invertibilidad de) su derivada en dicho punto. En otras palabras, es un teorema de existencia local de la función inversa. El teorema se puede generalizar para funciones en un espacio de Banach de dimensión infinita, pero en este apunte nos limitaremos en dimensión finita.

6.1. Resultados preliminares

Anotando por $\mathcal{M}_{d\times d}$ el espacio vectorial de matrices $d \times d$, empezamos por recordar que una matriz $A \in \mathcal{M}_{d\times d}$ es invertible si y sólo si su determinante $\det(A) \neq 0$.

Proposición 6.1 (continuidad de la función "determinante")**.** La aplicación

$$\begin{cases} \det : \mathcal{M}_{d\times d} \to \mathbb{R} \\ A \mapsto \det(A) \end{cases}$$

es continua.

Demostración. Mostramos primero el resultado en el caso $d = 2$. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathcal{M}_{2\times 2}$$

una matriz arbitraria, donde $\{a_{ij} : i, j \in \{1, 2\}\}$ son sus coordenadas en la base canónica[10] de $\mathcal{M}_{2\times 2}$. Entonces expresando la función determinante en coordenadas tenemos que

$$\det(A) = f(\underbrace{a_{11}, a_{12}, a_{21}, a_{22}}_{\text{coordenadas de } A}) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$$

de donde concluye que f (y por lo tanto $\det(\cdot)$) es continua.
Para tratar el caso general de dimensión $d > 2$, consideramos nuevamente una matriz

$$A = \sum_{i,j=1}^{d} a_{ij} E_{ij} \ \in \ \mathcal{M}_{d\times d}$$

(expresada directamente en coordenadas de la base canónica de $\mathcal{M}_{d\times d}$) y usamos la fórmula de Leibniz para expresar su determinante en función de sus coordenadas:

$$\det(A) \ = \ f(\underbrace{a_{11}, \ldots, a_{1d}, a_{21}, \ldots, a_{2d}, \ldots, a_{d1}, \ldots, a_{dd}}_{d\times d \text{ coordenadas de la matriz } A}) \ = \ \sum_{\sigma\in\mathcal{P}_d} \operatorname{sgn}(\sigma) \prod_{i=1}^{d} a_{i,\sigma(i)}$$

donde la suma se calcula sobre el conjunto $\mathcal{P}_d$ de todas las *permutaciones*[11] σ del conjunto $\{1, 2, ..., d\}$ y $\operatorname{sgn}(\sigma) \in \{1, -1\}$ denota la *signatura* de la permutación σ (que es -1 si la permutación es impar y $+1$ si es par). De la fórmula anterior deducimos que:

$$\underbrace{A_n \underset{n\to\infty}{\longrightarrow} A}_{\text{convergencia en norma}} \quad \Longleftrightarrow \quad \left\{ \begin{array}{l} \forall i, j \in \{1, \ldots, d\} \\ a_{ij}(n) \underset{n\to\infty}{\longrightarrow} a_{ij} \end{array} \right. \quad \Longrightarrow \quad f(A_n) \underset{n\to\infty}{\longrightarrow} f(A)$$

es decir, la función determinante es continua. □

Corolario 6.2 (conjunto de matrices invertibles). *El conjunto* $\mathcal{U} := \det^{-1}(\mathbb{R} \setminus \{0\})$ *de las matrices* $d \times d$ *invertibles es un conjunto abierto en* $\mathcal{M}_{d\times d}$.

Demostración. Por la Proposición 6.1 , la función determinante es continua, y el conjunto $\mathcal{U}$ es preimagen del abierto $\mathbb{R} \setminus \{0\}$. □

Corolario 6.3 (la función "inversa de matriz" es continua). *La función*

$$\left\{ \begin{array}{l} \operatorname{Inv} : \mathcal{U} \to \mathcal{U} \\ A \mapsto \operatorname{Inv}(A) := A^{-1} \end{array} \right.$$

es continua.

[10] Recordamos que la base canónica de $\mathcal{M}_{2\times 2}$ consiste de las matrices E_{11}, E_{12}, E_{21} y E_{22}, donde las entradas de la matriz $E_{i_o j_o}$ son todas cero, salvo la entrada (i_0, j_0) que es 1. Por lo tanto, se tiene que $A = a_{11}E_{11} + \ldots + a_{22}E_{22}$

[11] Llamamos *permutación* cualquier biyección de un conjunto finito. Las permutaciones que sólo permutan 2 elementos (manteniendo fijos todos los demás) se llaman *2-ciclos*. Cada permutación se puede escribir como producto (composición) de una cadena de 2-ciclos. De esta forma, una permutación se dice *par* (respectivamente, *impar*) si se descompone en un número par (respectivamente, impar) de 2-ciclos.

Demostración. Sea $A \in \mathcal{M}_{d\times d}$ una matriz invertible (es decir, $\det(A) \neq 0$). Usando coordenadas, tenemos que:

$$\mathrm{Inv}(A) : f(\underbrace{a_{11}, \ldots, a_{1d}, a_{21}, \ldots, a_{2d}, \ldots, a_{d1}, \ldots, a_{dd}}_{d\times d \text{ coordenadas de la matriz } A}) = \frac{1}{\det(A)}\mathrm{Adj}(A)^T,$$

donde $\mathrm{Adj}(A)$ es la matriz adjunta de A cuya entrada (i,j) es igual a $(-1)^{i+j}\det(A_{ij})$, donde A_{ij} es la submatriz $(d-1)\times(d-1)$ que se obtiene por la matriz inicial A eliminando su i-fila y su j-columna. Por la Proposición 6.1, la función determinante es continua (y diferente de 0 en el abierto $\mathcal{U}$) por lo tanto, f es continua $\square$

A continuación consideramos una función diferenciable $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ y llamamos un punto $\bar{x} \in \mathbb{R}^d$ como punto *regular*, si la derivada $Df(\bar{x}) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ es sobreyectiva, lo que equivale, en este caso, en decir que $Df(\bar{x})$ es un automorfismo de $\mathbb{R}^d$ (y por lo tanto, se representa por una matriz Jacobiana $Jf(\bar{x})$ invertible).

Corolario 6.4 (los puntos regulares forman un abierto)**.** *Para cada función $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$, el conjunto de sus puntos regulares es un abierto.*

Demostración. Sea $\bar{x} \in \mathbb{R}^d$ un punto regular de f, es decir, $\det(Jf(\bar{x})) \neq 0$. Por nuestra hipotesis, la función $x \longmapsto Jf(\bar{x})$ es continua, y por la Proposición 6.1, su composición con la función determinante

$$\phi(x) := \det(Jf(x)), \qquad x \in \mathbb{R}^d,$$

también es continua. Dado que $\mathbb{R}^d \setminus \{0\}$ es un subconjunto abierto de $\mathbb{R}^d$, se deduce que existe $\delta > 0$ tal que para todo $x \in B(\bar{x}, \delta)$ se tiene que $\det(Jf(x)) \neq 0$, lo que demuestra lo pedido. $\square$

6.2. Teorema de la función inversa (versión simple)

El teorema de la función inversa, en simples palabras, nos dice que si la derivada $Df(\bar{x})$ en $\bar{x}$ de una función $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ de clase $\mathcal{C}^1$ es invertible[12] (como aplicación lineal continua de $\mathbb{R}^d$ a $\mathbb{R}^d$), entonces lo mismo ocurre con la función f localmente alrededor del punto $\bar{x}$, es decir, se puede definir la función inversa f^{-1}, en un entorno de $f(\bar{x})$, y será también de clase $\mathcal{C}^1$. El teorema de la función inversa es bastante intuitivo para funciones de $\mathbb{R}$ a $\mathbb{R}$. En efecto, si $a = f'(\bar{x}) \neq 0$, el grafo de la función f se aproxima localmente alrededor del punto $(\bar{x}, f(\bar{x}))$ por la recta $y - f(\bar{x}) = a(x - \bar{x})$ y por lo tanto, tiene que ser localmente biyectiva alrededor de $\bar{x}$. En su forma general (véase siguiente subsección), el teorema también nos da la formula de la derivada de la función inversa. En esta subsección demostraremos el teorema en el caso particular donde $\bar{x} = 0 = f(\bar{x})$ y $Df(\bar{x})$ es la aplicación identidad $\mathbb{I}$. La demostración está basada en el teorema del punto fijo de Banach (por lo tanto utiliza la competitud del espacio $\mathbb{R}^d$ (bajo cualquier norma).

Teorema 6.5 (Teorema de la función inversa – Versión simple)**.** *Sea $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$, tal que $f(0) = 0$ y $Df(0) = \mathbb{I}$. Entonces existe un entorno $\mathcal{V}$ del origen del espacio inicial, tal que $f|_{\mathcal{V}}$ es una biyección entre $\mathcal{V}$ y $f(\mathcal{V})$, luego $f^{-1} : f(\mathcal{V}) \longrightarrow \mathcal{V}$ es $\mathcal{C}^1$ con $Df^{-1}(0) = \mathbb{I}$.*

[12] fijamos que es necesario que el espacio inicial y el espacio de llegada sean de la misma dimensión.

Demostración. Dividimos la demostración en 5 pasos.

Paso 1. Consideramos la función:

$$\begin{cases} g : \mathbb{R}^d \to \mathbb{R}^d \\ g(x) = x - f(x) \end{cases}$$

Entonces $g \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$, $g(0) = 0$ y $Dg(0) = 0$. Se deduce (aplicando por ejemplo Ejercício 3.7(ii) para $\varepsilon = 1/4$) que existe $\delta > 0$ tal que:

$$||g(x_1) - g(x_2)|| \le \frac{1}{4}||x_1 - x_1||, \qquad \text{para todo } x_1, x_2 \in \bar{B}(\bar{x}, \delta). \tag{37}$$

Además, tomando posiblemente $\delta > 0$ más pequeño, podemos asegurar (*c.f.* Corolario 6.4) que

$$\det(Jf(x)) \neq 0, \qquad \text{para todo } x \in \bar{B}(\bar{x}, \delta). \tag{38}$$

Paso 2. Sea $y_0 \in \bar{B}(0, \delta/2)$ (arbitrario). Consideramos la función:

$$\begin{cases} \phi_{y_0} : \bar{B}(0, \delta) \to \mathbb{R}^d \\ \phi_{y_0}(x) = \underbrace{x - f(x)}_{g(x)} + y_0 \end{cases} \tag{39}$$

Nuestro objetivo es mostrar que ϕ_{y_0} es una contracción en $\bar{B}(0, \delta)$. Deducimos directamente de (37) que para todo $x_1, x_2 \in \bar{B}(\bar{x}, \delta)$ se tiene que:

$$||\phi_{y_0}(x_1) - \phi_{y_0}(x_2)|| = ||g(x_1) - g(x_2)|| \le \frac{1}{4}||x_1 - x_1||,$$

es decir, ϕ_{y_0} es Lipschitz continua con constante $1/4$. En particular, para todo $x \in \bar{B}(0, \delta)$ tenemos:

$$||\phi_{y_0}(x)|| \le ||\phi_{y_0}(x) - \phi_{y_0}(0)|| + ||\phi_{y_0}(0)|| \le \frac{1}{4}||x - 0|| + ||y_0|| \le \frac{\delta}{2} + \frac{\delta}{2} = \delta,$$

y por lo tanto $\phi_{y_0}(\bar{B}(0, \delta)) \subset \bar{B}(0, \delta)$. Combinando con (39) concluimos que la función ϕ_{y_0} es una contracción, para todo $y_0 \in \bar{B}(0, \delta/2)$. Aplicando el punto fijo de Banach en el espacio métrico completo $\bar{B}(0, \delta)$ deducimos la existencia de un único punto fijo $x_0 \in \bar{B}(0, \delta)$, es decir:

$$x_0 = \phi_{y_0}(x_0) := x_0 - f(x_0) + y_0 \qquad \Longleftrightarrow \qquad f(x_0) = y_0.$$

Eso significa que para todo $y_0 \in \bar{B}(0, \delta/2)$, existe único[13] $x_0 \in \bar{B}(0, \delta)$ tal que $f(x_0) = y_0$. En otras palabras, la función f, restringida en el entorno cerrado

$$\bar{\mathcal{V}} := \bar{B}(0, \delta) \cap f^{-1}(\bar{B}(0, \delta/2)).$$

es una biyección, véase Figura 4.

[13]Hacemos constar que la único de x_0 se puede afirmar sólo con respecto a la bola cerrada $\bar{B}(0, \delta)$. En principio, el punto y_0 puede tener otras preimagenes fuera de dicho conjunto.

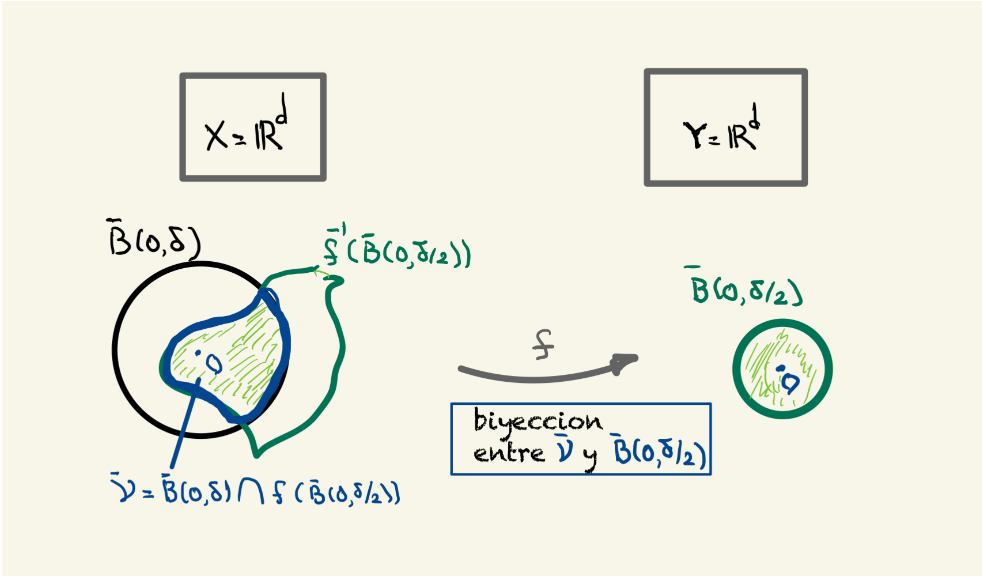


Figura 4: f es biyectiva alrededor del 0.

Paso 3. En esta parte mostramos que la función inversa

$$f^{-1} : \bar{B}(0, \delta/2) \longrightarrow \bar{\mathcal{V}}$$

es Lipschitz continua de constante $4/3$. En efecto, sea $y_1, y_2 \in \bar{B}(0, \delta/2) := f(V)$. Por lo demostrado en el Paso 2, existen únicos puntos $x_1, x_2 \in \bar{\mathcal{V}}$ con

$$f(x_i) = y_i \quad \Longleftrightarrow \quad x_i = f^{-1}(y_i), \qquad i \in \{1, 2\}.$$

En particular, observando que $g(x) = \phi_0(x) = x - f(x)$, es decir $x = f(x) + g(x)$, para todo $x \in \bar{B}(0, \delta)$, deducimos que:

$$\begin{aligned} ||x_1 - x_2|| &= ||(f+g)(x_1) - (f+g)(x_2)|| \leq ||f(x_1) - f(x_2)|| + ||g(x_1) - g(x_2)|| \\ &\leq ||y_1 - y_2|| + \frac{1}{4}||x_1 - x_2|| \qquad \Longrightarrow \qquad \frac{3}{4}||x_1 - x_2|| \leq ||y_1 - y_2|| \end{aligned}$$

de donde se deduce que

$$||x_1 - x_2|| \leq \frac{4}{3}||y_1 - y_2|| \qquad \text{para todo } y_1, y_2 \in \bar{B}(0, \delta/2),$$

lo que muestra lo pedido.
Paso 4. Mostramos que

$$f^{-1} : B(0, \delta/2) \longrightarrow \mathcal{V} := B(0, \delta) \cap f^{-1}(B(0, \delta/2))$$

es diferenciable, con derivada

$$D(f^{-1})(y) = Df(x)^{-1} \qquad (\text{donde } x := f^{-1}(y))$$

para todo $y \in B(0,\delta/2)$. Para eso, sea $r = \delta/2 - ||y|| > 0$ y $v \in B(0,r) \setminus \{0\}$, de forma que $||y+v|| < \delta/2$, es decir, $y+v \in B(0,\delta/2)$. Ponemos

$$u := f^{-1}(y+v) - f^{-1}(y) \qquad \text{y} \qquad x := f^{-1}(y) \in \mathcal{V}$$

de forma que

$$||u|| = ||f^{-1}(y+v) - f^{-1}(y)|| \leq \frac{4}{3}||v|| \qquad \text{(por el Paso 3)} \tag{40}$$

luego

$$x + u = f^{-1}(y+v) \qquad \Longrightarrow \qquad f(x+u) = y+v$$

y por lo tanto

$$v = f(x+u) - y = f(x+u) - f(x).$$

Dado que conocemos el candidato, $Df(x)^{-1}$, para ser derivada de la función f^{-1} en y, formamos la fracción:

$$\frac{||\overbrace{f^{-1}(y+v) - f^{-1}(y)}^{u} - Df(x)^{-1}(\overbrace{\quad v \quad}^{f(x+u)-f(x)})||}{||v||} = \frac{||-Df(x)^{-1}\left[f(x+u) - f(x) - Df(x)(u)\right]||}{||v||}$$

Usando el hecho que f es diferenciable en $x = f^{-1}(y)$ con derivada $Df(x)$ deducimos:

$$\frac{\|f^{-1}(y+v) - f^{-1}(y) - Df(x)^{-1}(v)\|}{||v||} \leq \frac{4}{3}\|Df(x)^{-1}\| \frac{\|f(x+u) - f(x) - Df(x)(u)\|}{||u||}.$$

Deducimos de (40) que $||u|| \underset{||v||\to 0}{\longrightarrow} 0$. Tomando el límite $\|u\| \to 0$ en la anterior desigualdad observamos que el término a derecha (y por lo tanto, también el otro término) converge a 0. Se establece entonces que $Df^{-1}(y) = Df(x)^{-1}$ y la demostración es completa.

Paso 5. Mostramos ahora que f^{-1} es de clase $\mathcal{C}^1$ en el abierto $B(0,\delta/2)$. En el Paso 4 hemos mostrado que se puede definir la función derivada:

$$\begin{cases} D(f^{-1}) : B(0,\delta/2) \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d) \\ D(f^{-1})(y) := \left[Df(f^{-1}(y))\right]^{-1}. \end{cases}$$

Observamos que

$$D(f^{-1}) := \mathrm{Inv} \circ Df \circ f^{-1}$$

donde la función Inv es continua por el Corolario 6.3, la función Df es continua por hipotesis y la función f^{-1} es continua por el Paso 3. Concluimos que $D(f^{-1})$ es continua como composición de funciones continuas y por lo tanto f^{-1} es de clase $\mathcal{C}^1$. $\square$

Observación 6.6 (comentarios sobre las hipotesis del Teorema 6.5)**.** La hipotesis que f es de clase $\mathcal{C}^1$ asegura que alrededor del punto $\bar{x} = 0$ las derivadas de f quedan invertibles, véase (38), lo que se utilizó explicitamente en el Paso 4 para definir (el candidato de ser) la derivada de la función f^{-1} en los puntos $y = f(x)$, con x cerca de $\bar{x} = 0$. El siguiente ejemplo muestra que eso es necesario incluso cuando $d = 1$. En efecto, la función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} x + x^2 \sin \frac{1}{x}, & \text{si } x \neq 0 \\ \\ 0, & \text{si } x = 0 \end{cases}$$

es diferenciable en cada punto con derivada

$$f(x) = \begin{cases} 1 + 2x \sin \frac{1}{x} - \cos \frac{1}{x}, & \text{si } x \neq 0 \\ \\ 1, & \text{si } x = 0, \end{cases}$$

luego satisface $f(0) = 0$ y $f'(0) = 1$, pero la derivada f' no es continua en $\bar{x} = 0$. Por lo tanto el Teorema 6.5 no se puede aplicar, y de hecho f no es localmente biyectiva en ningún entorno de 0.

Antes de presentar la versión general del teorema de la función inversa, veamos una ilustración de su uso en la versión simple.

Ejemplo 6.7 (ilustración del teorema de la función inversa)**.** Consideramos la función

$$\begin{cases} f : \mathbb{R}^2 \to \mathbb{R}^2 \\ f(x, y) = (x_1 + x_2^2, x_1^2 + x_2), \qquad (x_1, x_2) \in \mathbb{R}^2. \end{cases}$$

Observamos que $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R}^2)$, $f(0, 0) = (0, 0)$ y

$$Jf(x_1, x_2) = \begin{pmatrix} 1 & 2x_2 \\ 2x_1 & 1 \end{pmatrix}, \qquad (x_1, x_2) \in \mathbb{R}^2,$$

por lo tanto $Jf(0, 0) = \mathbb{I}$. Por el Teorema 6.5 deducimos la existencia de dos funciones $\psi_1, \psi_2 \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ de forma que para todo (x_1, x_2) en un entorno de $(0, 0)$ se tiene que:

$$\left.\begin{array}{r} x_1 + x_2^2 = y_1 \\ x_1^2 + x_2 = y_2 \end{array}\right\} \qquad \Longleftrightarrow \qquad \left\{\begin{array}{l} x_1 = \psi_1(y_1, y_2) \\ x_2 = \psi_2(y_1, y_2) \end{array}\right.$$

es decir, que el sistema al lado izquierda se puede resolver de forma explicita. (Las funciones ψ_1, ψ_2 son las componentes de la inversa local f^{-1} de f alrededor del $(0, 0)$.

6.3. Teorema de la función inversa (versión general)

Veamos ahora el teorema de la función inversa en su forma general (en dimensión finita).

Teorema 6.8 (Teorema de la función inversa)**.** *Sea $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ y $\bar{x} \in \mathbb{R}^d$ un punto regular (es decir,* $\det(Jf(\bar{x})) \neq 0$*). Entonces existe un entorno $\tilde{\mathcal{V}}$ de $\bar{x}$ tal que la función $f|_{\tilde{\mathcal{V}}} : \tilde{\mathcal{V}} \to f(\tilde{\mathcal{V}})$ es una biyección entre $\tilde{\mathcal{V}}$ y $f(\tilde{\mathcal{V}})$, la función $f^{-1} : f(\tilde{\mathcal{V}}) \longrightarrow \tilde{\mathcal{V}}$ es de clase $\mathcal{C}^1$ y*

$$Df^{-1}(y) = \left[Df(f^{-1}(y))\right]^{-1}, \qquad \textit{para todo } y \in f(\tilde{\mathcal{V}}).$$

Demostración. Sea $T = Df(\bar{x}) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ (invertible) e $\bar{y} = f(\bar{x})$. Definimos la función

$$\varphi(u) := T^{-1}(f(\bar{x}+u) - f(\bar{x})), \qquad u \in \mathbb{R}^d. \tag{41}$$

Observamos que $\varphi \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ con $\varphi(0) = 0$. Aplicando la regla de la cadena (Proposición 3.3) obtenemos:

$$D\varphi(u) = T^{-1} \circ Df(\bar{x}+u), \qquad u \in \mathbb{R}^d, \tag{42}$$

y en particular

$$D\varphi(0) = T^{-1} \circ Df(\bar{x}) = T^{-1} \circ T = \mathbb{I}.$$

Concluimos que la función φ cumple las hipotesis del Teorema 6.5 y admite localmente alrededor de 0 una función inversa $\varphi^{-1} : \varphi(\mathcal{V}) \longrightarrow \mathcal{V}$ de clase $\mathcal{C}^1$, donde siguiendo la notación empleada en la demostración del Teorema 6.5 los conjuntos

$$\mathcal{V} := B(0,\delta) \cap \varphi^{-1}(B(0,\delta/2)) \quad \text{y} \quad \varphi(\mathcal{V}) = B(0,\delta/2)$$

son entornos del origen (en el espacio inicial y respectivamente en el espacio de llegada) y están en biyección mediante la función φ. En particular, para cada $u \in \mathcal{V}$ deducimos de (41) que

$$\bar{x} + u = \bar{x} + \varphi^{-1}(T^{-1}(\overbrace{f(\bar{x}+u)}^{y} - \overbrace{f(\bar{x})}^{\bar{y}})). \tag{43}$$

Poniendo

$$\mathcal{W} := \bar{y} + T(\varphi(\mathcal{V})) = \bar{y} + T(B(0,\delta/2))$$

tenemos que

$$y \in \mathcal{W} \;\Leftrightarrow\; y - \bar{y} \in T(\overbrace{B(0,\delta/2)}^{\varphi(\mathcal{V})}) \;\Leftrightarrow\; \exists! u \in \mathcal{V} : \;\varphi(u) = T^{-1}(y - \bar{y}). \tag{44}$$

Combinando con (41) obtenemos que para cada $y \in \mathcal{W}$ existe un único elemento

$$x := \bar{x} + u \in \bar{x} + \mathcal{V} := \tilde{\mathcal{V}}$$

tal que

$$\varphi(u) = T^{-1}(f(x) - \bar{y}) = T^{-1}(y - \bar{y}) \quad \Longrightarrow y = f(x).$$

Por lo tanto f es una biyección entre el entorno $\tilde{\mathcal{V}}$ de $\bar{x}$ y el entorno $\mathcal{W} = f(\tilde{\mathcal{V}})$ de $\bar{y} = f(\bar{x})$. Además, por (44) obtenemos:

$$f^{-1}(y) = x = \bar{x} + u = \bar{x} + \varphi^{-1}\left(T^{-1}(y - \bar{y})\right), \qquad \text{para todo } y \in \mathcal{W}$$

de lo que se deduce que

$$\begin{aligned} D(f^{-1})(y) &= D(\varphi^{-1})(T^{-1}(y-\bar{y})) \circ T^{-1} = D(\varphi^{-1})(\varphi(u)) \circ T^{-1} = D\varphi(u)^{-1} \circ T^{-1} \\ &= [T \circ D\varphi(u)]^{-1} = Df(\bar{x}+u)^{-1} = Df(x)^{-1} = Df(f^{-1}(y))^{-1}. \end{aligned}$$

Esto muestra que f^{-1} es diferenciable con derivada continua, como composición de funciones continuas:

$$Df^{-1} = \mathrm{Inv} \circ Df \circ f^{-1}.$$

La demostración es completa. $\square$

7. Teorema de la Función Implicita

En esta parte veremos el famoso teorema de la función implicita, que es el tercer teorema pilar de este curso. Dicho teorema está basado en el teorema de la función inversa que acabamos de ver. Recordamos que este último, en su turno, era una consecuencia del teorema del punto fijo de Banach. Por lo tanto, esta cadena de teoremas (el teorema del punto fijo de Banach, el teorema de la función inversa y el teorema de la función implicita) es válida sólo en espacios completos. El teorema de la función implicita se considera como la base de la geometría diferencial.

Antes de enunciar el teorema, y con el fin de entender y apreciar el enunciado, veremos un ejemplo concreto.

Ejemplo 7.1 (ilustración del teorema de la función implicita)**.** Consideramos la función diferenciable

$$\begin{cases} \Phi : \mathbb{R}^2 \to \mathbb{R} \\ \Phi(x,y) = x^2 + y^2 - 1, \qquad (x,y) \in \mathbb{R}^2 \end{cases} \tag{45}$$

Observamos que la circunferencia

$$\mathcal{C} = \left\{(x,y) \in \mathbb{R}^2 : \ x^2 + y^2 = 1\right\}$$

coincide con la preimagen $\Phi^{-1}(\{0\})$ y que para cualquier $(\bar{x},\bar{y}) \in \mathcal{C}$ la ecuación de la recta tangente $T_{\mathcal{C}}(\bar{x},\bar{y})$ a $\mathcal{C}$ coincide con el transado al $(\bar{x},\bar{y})$ del nucleo

$$\mathrm{Ker}(D\Phi) = \left\{(u,v) \in \mathbb{R}^2 : \ D\Phi(\bar{x},\bar{y})(u,v) = 0\right\} = \left\{(u,v) \in \mathbb{R}^2 : \ \bar{x}u + \bar{y}v = 0\right\}.$$

En otras palabras,

$$T_{\mathcal{C}}(\bar{x},\bar{y}) = (\bar{x},\bar{y}) + \mathrm{Ker}(D\Phi) := \left\{(x,y) \in \mathbb{R}^2 : \ \bar{x}(x-\bar{x}) + \bar{y}(y-\bar{y}) = 0\right\}. \tag{46}$$

También observamos que si $\bar{x} \in (-1,1)$, entonces $\mathcal{C}$ está localmente representada con el grafo de la función $y = f(x)$ alrededor del punto $(\bar{x},\bar{y})$. Más precisamente, la función $f(x) = \sqrt{1-x^2}$, si $\bar{y} > 0$, y $f(x) = -\sqrt{1-x^2}$, si $\bar{y} < 0$. De hecho, en este caso, la tangente $T_{\mathcal{C}}(\bar{x},\bar{y})$ dada por (46) también coincide con el grafo de la función afín

$$y = f(\bar{x}) + f'(\bar{x})(x-\bar{x}). \tag{47}$$

(Invitamos al lector de comprobar esta asersión.) Luego, observamos que si $\bar{x} \in \{-1,1\}$, entonces $\mathcal{C}$ ya no se podrá representar mediante una función $y = f(x)$ alrededor del punto $(\bar{x},\bar{y})$, pero si que se podrá representar mediante una función $x = g(y)$, y precisamente con la función $g(y) = \sqrt{1-y^2}$, si $\bar{x} > 0$, y $g(y) = -\sqrt{1-y^2}$, si $\bar{x} < 0$. Y también en este caso, la tangente $T_{\mathcal{C}}(\bar{x},\bar{y})$ coincidiría con el grafo de la función afín

$$x = g(\bar{y}) + g'(\bar{y})(y-\bar{y}). \tag{48}$$

En realidad, podriamos representar la circunferencia $\mathcal{C}$ por una función $x = g(y)$ alrededor del punto $(\bar{x},\bar{y})$, y su tangente $T_{\mathcal{C}}(\bar{x},\bar{y})$ por (48), siempre y cuando $\bar{y} \in (-1,1)$.

En el ejemplo anterior vimos que a partir de una formula $\Phi(x,y)=0$ podriamos despejar localmente una variable como función de la otra (es decir, $y=f(x)$ o $x=g(y)$). Además vimos que la ecuación de la tangente al conjunto $\mathcal{C}=\Phi^{-1}(\{0\})$ en cualquier punto $(\bar{x},\bar{y})\in\mathcal{C}$ se puede determinar de manera intriseca, *cf.* (46), o mediante la función representante, *cf.* (47)–(48). El objetivo del teorema de la función implicita es formalizar —y extender a varias variables— la situación anterior.

Empezamos con una versión simple y luego presentamos la versión general.

Teorema 7.2 (Teorema de la función implicita – versión simple)**.** *Sea $\Phi\in\mathcal{C}^1(\mathbb{R}^d,\mathbb{R})$ una función de clase $\mathcal{C}^1$ y $\bar{x}\in\mathbb{R}^d$ tal que $\Phi(\bar{x})=0$ y*

$$\frac{\partial\Phi}{\partial x_d}(\bar{x})\neq 0 \qquad \textit{(en particular } \bar{x} \textit{ es un punto regular).} \tag{49}$$

Entonces considerando la decomposición

$$\bar{x}=(\bar{x}_1,\ldots,\bar{x}_d)\equiv(\bar{u},\bar{v})\in\mathbb{R}^{d-1}\times\mathbb{R}$$

existen entornos $\mathcal{U}$ de $\bar{u}\in\mathbb{R}^{d-1}$ y $\mathcal{V}$ de $\bar{v}\in\mathbb{R}$ y una función

$$f:\mathcal{U}\to\mathbb{R}$$

de clase $\mathcal{C}^1$ tal que para todo $(u,v)\in\mathcal{U}\times\mathcal{V}$ se tiene[14]*:*

$$\Phi(u,v)=0 \qquad\Longleftrightarrow\qquad v=f(u).$$

Demostración. Por simplicidad, trabajaremos en el espacio normado $(\mathbb{R}^d,||\cdot||_\infty)$. Consideramos la siguiente función:

$$\begin{cases} H:\mathbb{R}^d\to\mathbb{R}^d \\ H(u,v):=(u,\Phi(u,v)), \qquad (u,v)\in\mathbb{R}^{d-1}\times\mathbb{R}. \end{cases}$$

Es facil ver que H es de clase $\mathcal{C}^1$ y

$$JH(\bar{x})=JH(\bar{u},\bar{v})=\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & \vdots & \vdots \\ \vdots & & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \\ \frac{\partial\Phi}{\partial x_1}(\bar{x}) & \cdots & & \cdots & \underbrace{\frac{\partial\Phi}{\partial x_d}(\bar{x})}_{\neq 0} \end{pmatrix}$$

es invertible, debido a nuestra hipotesis (49). Anotemos a continuación $\bar{c}=\Phi(\bar{u},\bar{v})$. Aplicando el Teorema 6.8 deducimos que existen entornos $\mathcal{U}\times\mathcal{V}$ de $\bar{x}=(\bar{u},\bar{v})\in\mathbb{R}^{d-1}\times\mathbb{R}$

[14] En otras palabras, el conjunto $\mathcal{C}=\Phi^{-1}(\{0\})$ se representa localmente por el grafo de f, es decir:

$$\mathcal{C}\cap(\mathcal{U}\times\mathcal{V})=\mathrm{gph}(f)\cap(\mathcal{U}\times\mathcal{V}).$$

y $\mathcal{U}'\times\mathcal{W}$ de $H(\bar{x}) = (\bar{u}, \Phi(\bar{u}, \bar{v})) \equiv (\bar{u}, \bar{c}) \in \mathbb{R}^{d-1} \times \mathbb{R}$ tal que H es una biyección entre $\mathcal{U} \times \mathcal{V}$ y $\mathcal{U}' \times \mathcal{W}$ con inversa $H^{-1} : \mathcal{U}'\times\mathcal{W} \to \mathcal{U} \times \mathcal{V}$ de clase $\mathcal{C}^1$. En particular, existen $\phi_1 \in \mathcal{C}^1(\mathcal{U}', \mathbb{R}^{d-1})$ y $\phi_2 \in \mathcal{C}^1(\mathcal{U}', \mathbb{R})$ tales que:

$$H^{-1}(u, c) = (\phi_1(u, c), \phi_2(u, c)), \qquad \text{para todo } (u, c) \in \mathcal{U}'\times\mathcal{W}.$$

Dado que

$$(u, c) = \overbrace{\big(H \circ H^{-1}\big)}^{\mathbb{I}_d}(u, c) = H(\overbrace{\phi_1(u, c)}^{u}, \overbrace{\phi_2(u, c)}^{v}) \equiv (\overbrace{\phi_1(u, c)}^{u}, \overbrace{\Phi(\phi_1(u, c), \phi_2(u, c))}^{\Phi(u,v)}),$$

se deduce que $\phi_1(u, c) = u$ (y por lo tanto $\mathcal{U} = \mathcal{U}'$). Ponemos

$$v := \phi_2(u, c) \qquad \text{y} \qquad c := \Phi(\overbrace{\phi_1(u, c)}^{u}, \overbrace{\phi_2(u, c)}^{v}) = \Phi(u, v). \tag{50}$$

Dado que (50) es válida para todo $u \in \mathcal{U}$ y $c \in \mathcal{W}$, fijando el valor $c = \bar{c}$ y considerando la función de clase $\mathcal{C}^1$

$$\begin{cases} f : \mathcal{U} \to \mathbb{R} \\ f(u) := \phi_2(u, \bar{c}), \end{cases}$$

se decude que

$$\bar{c} = \Phi(\phi_1(u, \bar{c}), \phi_2(u, \bar{c})) = \Phi(u, f(u)), \qquad \text{para todo } u \in \mathcal{U}.$$

En particular, $(u, v) \in \mathcal{U} \times \mathcal{W}$ cumple $\Phi(u, v) = \bar{c}$ si y sólo si $v := f(u)$.

La demostración es completa. $\square$

Observación 7.3 (fórmula de la derivada de f)**.** Se deduce del Teorema 7.2 que si

$$\Phi \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}), \qquad \Phi(\bar{u}, \bar{v}) = 0 \quad \text{y} \quad \frac{\partial \Phi}{\partial x_d}(\bar{u}, \bar{v}) \neq 0,$$

entonces existe $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R})$ tal que

$$\Phi(u, f(u)) = 0, \qquad \text{para todo } u \in \mathcal{U}.$$

Aplicando la regla de la cadena obtenemos la siguiente fórmula para la derivada de la función f:

$$f'(u) = -\frac{\partial \Phi}{\partial u}(u, v) \,/\, \frac{\partial \Phi}{\partial v}(u, v).$$

Anotando por $\mathcal{C} := \Phi^{-1}(\{0\})$, la ecuación de la tangente $T_{\mathcal{C}}(\bar{u}, \bar{v})$ está dada por la fórmula:

$$T_{\mathcal{C}}(\bar{u}, \bar{v}) = (\bar{u}, \bar{v}) + \mathrm{Ker}(D\Phi)) := \left\{(u', v') \in \mathbb{R}^d : \; D\Phi(\bar{u}, \bar{v})(u' - \bar{u}, v' - \bar{v}) = 0\right\}.$$

Presentamos ahora la versión general:

Teorema 7.4 (Teorema de la función implicita – versión general)**.** *Sea* $\Phi \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^\ell)$ *una función de clase* $\mathcal{C}^1$. *Consideramos la decomposición*

$$x = (u, v) \in \mathbb{R}^{d-\ell} \times \mathbb{R}^\ell$$

y supongamos que existe $\bar{x} = (\bar{u}, \bar{v}) \in \mathbb{R}^{d-\ell} \times \mathbb{R}^\ell$ *tal que* $\Phi(\bar{u}, \bar{v}) = 0$ *y que la derivada parcial* $D_v\Phi(\bar{x})$ *es invertible*[15]. *Entonces existen* $\mathcal{U} \subset \mathbb{R}^{d-\ell}$ *entorno de* $\bar{u}$, $\mathcal{V} \subset \mathbb{R}^\ell$ *entorno de* $\bar{v} \in \mathbb{R}^\ell$ *y una función*

$$f : \mathcal{U} \to \mathbb{R}^\ell$$

de clase $\mathcal{C}^1$ *tal que para todo* $(u, v) \in \mathcal{U} \times \mathcal{V}$ *se tiene:*

$$\Phi(u, v) = 0 \qquad \Longleftrightarrow \qquad v = f(u).$$

Demostración. La demostración es muy similar a la del Teorema 7.2. Ponemos

$$m := d - \ell$$

definimos la función de clase $\mathcal{C}^1$

$$\begin{cases} H : \mathbb{R}^{m+\ell} \to \mathbb{R}^{m+\ell} \\ H(u, v) = (u, \Phi(u, v)), \qquad (u, v) \in \mathbb{R}^m \times \mathbb{R}^\ell \end{cases}$$

y observanos que

$$JH(\bar{x}) = JH(\bar{u}, \bar{v}) = \begin{pmatrix} \mathbb{I}_m & \mathbf{0}_{m\times\ell} \\ & \\ D_u\Phi(\bar{x}) & \underbrace{D_v\Phi(\bar{x})}_{\text{invertible}} \end{pmatrix}$$

es invertible por lo tanto podemos deducir la existencia local de la inversa $H^{-1} := (\phi_1, \phi_2)$ donde $\phi_1 \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^m)$ y $\phi_2 \in \mathcal{C}^1(\mathcal{U}, \mathbb{R}^\ell)$ y $\mathcal{U} \subset \mathbb{R}^m$. Invitamos al lector de completar la demostración siguiendo los pasos de la demostración del Teorema 7.2. □

Terminamos esta sección con un ejemplo de uso del teorema de la función implicita.

Ejemplo 7.5 (uso del teorema de la función implicita)**.** Consideramos el sistema no lineal subdeterminado (2×3):

$$\begin{cases} x^2y = xyz \\ \\ xy^2 - 2xy^2z^2 + z^4 = 0. \end{cases} \tag{51}$$

Definimos la función

$$\begin{cases} \Phi : \mathbb{R}^3 \to \mathbb{R}^2 \\ \\ \Phi(x, y, z) = (x^2y - xyz,\ xy^2 - 2xy^2z^2 + z^4), \end{cases}$$

[15]En particular, $D\Phi(\bar{x}) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^\ell)$ es sobreyectiva y $\bar{x}$ es un punto regular.

y observamos que $\Phi \in \mathcal{C}^1(\mathbb{R}^3, \mathbb{R}^2)$, luego que para todo $(x, y, z) \in \mathbb{R}^3$

$$J\Phi(x,y,z) = \begin{pmatrix} 2xy - yz & x^2 - xz & -xy \\ y^2 - 2y^2z^2 & 2xy - 4xz^2 & -4xy^2z + 4z^3 \end{pmatrix}.$$

En particular, $\Phi(1,1,1) = 0$ y

$$J\Phi(1,1,1) = \begin{pmatrix} 1 & 0 & -1 \\ -1 & -2 & 0 \end{pmatrix}.$$

En particular, considerando la decomposición

$$(x,y,z) = ((x,y),z) \in \mathbb{R}^2 \times \mathbb{R}$$

observamos que $D\Phi_v(1,1,1) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ es invertible[16]. Ponemos $z = t$ y aplicamos el Teorema 7.4 para deducir la existencia de dos funciones $x = x(t)$ y $y = y(t)$ definidas en $(1-\delta, 1+\delta)$ de clase $\mathcal{C}^1$ y con valores cerca de 1 tales que:

$$\begin{cases} x(t)^2 \cdot y(t) = x(t) \cdot y(t) \cdot t \\ \\ x(t) \cdot y(t)^2 - 2x(t) \cdot y(t)^2 \cdot t^2 + t^4 = 0. \end{cases}$$

Dado que $x(t) \neq 0$ e $y(t) \neq 0$ cuando $t \in (1-\delta, 1+\delta)$ deducimos de la primera ecuación que $x(t) = t$. Substituyendo a la segunda ecuación $x(t) = t$ obtenemos, despues de un cálculo elemental (y suponiendo $\delta < 1 - \sqrt{2}/2$), que

$$y(t) = \sqrt{\frac{t^3}{2t^2 - 1}}.$$

Por lo tanto, hemos logrado a resolver explicitamente el sistema no lineal alrededor de la solución $(\bar{x}, \bar{y}, \bar{z}) = (1,1,1)$, obteniendo las soluciones:

$$\left\{ \left(t, \sqrt{\frac{t^3}{2t^2-1}}, t \right) : \quad t \in (1-\delta, 1+\delta) \right\}.$$

8. Derivadas de orden superior

Sea $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ dos espacios normados y sea $f : X \to Y$ una función diferenciable en cada $x \in X$. Se define entonces la función derivada

$$\begin{cases} Df : X \to \mathcal{L}(X,Y) \\ \\ x \mapsto Df(x), \qquad x \in X \end{cases}$$

que también es una función entre dos espacios normados, el espacio $(X, \|\cdot\|_X)$ y el espacio $(\mathcal{L}(X,Y), \|\cdot\|_{\text{op}})$. Si Df es continua (es decir, $Df \in \mathcal{C}(X, \mathcal{L}(X,Y))$), entonces decimos

[16]en efecto: $\det(J\Phi_v(1,1,1)) = \begin{vmatrix} 1 & 0 \\ -1 & -2 \end{vmatrix} = -2 \neq 0$

que f es de clase $\mathcal{C}^1$, y anotamos $f \in \mathcal{C}^1(X,Y)$, véase Definición 2.17. Si ahora Df es diferenciable en cada $x \in X$, entonces se puede definir la segunda derivada

$$\begin{cases} D^2 f : X \to \mathcal{L}(X, \mathcal{L}(X,Y)) \\ x \mapsto D(Df)(x), \qquad x \in X \end{cases}$$

entre los espacios normados $(X, \|\cdot\|_X)$ y $(\mathcal{L}(X,\mathcal{L}(X,Y)), \|\cdot\|_{\mathrm{op}})$. Este último es el espacio de las aplicaciones lineales continuas entre el espacio normado $(X, \|\cdot\|_X)$ y el espacio normado $(\mathcal{L}(X,Y)), \|\cdot\|_{\mathrm{op}})$ de las aplicaciones lineales continuas de X a Y. Obviamente este proceso podría continuar. Si $D^2 f$ es continua, entonces decimos que f es de clase $\mathcal{C}^2$ y anotamos $f \in \mathcal{C}^2(X,Y)$. Si además es diferenciable, entonces se define la función $D^3 f$ (tercera derivada de f) y el proceso se puede continuar hasta el infinito (dando lugar a las funciones de clase $\mathcal{C}^\infty$.

En lo que sigue estudiaremos funciones entre espacios de dimensión finita, y representaremos las derivadas de orden superior por tensores. En la Subsección 8.1 nos ocuparemos del caso de funciones de clase $\mathcal{C}^2$ y veremos que en el caso $X = \mathbb{R}^d$ e $Y = \mathbb{R}$ la segunda derivada $D^2 f(x)$ se representa por una matriz $(d \times d)$ simetrica. Este resultado sigue cierto para cada coordenada f_i, en el caso de funciones $f = (f_1, \dots f_\ell)$ a valores en $\mathbb{R}^\ell$. Luego, en la Subsección 8.2 veremos como aproximar una función de clase $\mathcal{C}^k$ por su polinomio de Taylor de orden k por analogía de lo que ocurre para funciones de $\mathbb{R}$ a $\mathbb{R}$.

8.1. Derivadas de orden 2. Teorema de Schwarz

El principal resultado de esta subsección es el llamado teorema de Schwarz que afirma que si $f : \mathbb{R}^d \to \mathbb{R}^\ell$ es una función de clase $\mathcal{C}^2$ (entre espacios normados de dimensión finita), entonces el orden en que tomemos las segundas derivadas parciales no es relevante.

Antes de continuar, explicamos la notación que será empleada a partir de ahora: si $f : \mathbb{R}^d \to \mathbb{R}^\ell$ es de clase $\mathcal{C}^2$, entonces para cada $i \in \{1, \dots, \ell\}$ (coordenada de la función f) y cada $j_1 \in \{1, \dots, d\}$ (variable de la función f) la derivada parcial

$$\frac{\partial f_i}{\partial x_{j_1}} : \mathbb{R}^d \to \mathbb{R}$$

existe y es una función de clase $\mathcal{C}^1$. Por lo tanto, la derivada $D\left[\frac{\partial f_i}{\partial x_{j_1}}\right](x)$ de dicha derivada parcial existe para cada $x \in \mathbb{R}^d$ y se representa mediante sus derivadas parciales, que son las derivadas parciales en orden 2 de la función f, es decir[17]:

$$\frac{\partial^2 f_i(x)}{\partial x_{j_2} \partial x_{j_1}} := \frac{\partial}{\partial x_{j_2}}[\frac{\partial f_i}{\partial x_{j_1}}](x), \qquad j_2 \in \{1, \dots, d\}.$$

En el caso donde la variable que se vuelva a tomar la derivada parcial coincide con la variable inicial (es decir, $j_1 = j_2 \equiv j$), entonces solemos anotar:

$$\frac{\partial^2 f_i(x)}{\partial x_j^2} := \frac{\partial^2 f_i(x)}{\partial x_j \partial x_j}.$$

[17]La Proposición 2.14 también funciona aquí: la función $f : \mathbb{R}^d \to \mathbb{R}^\ell$ es de clase $\mathcal{C}^2$ si y sólo si las todas sus derivadas parciales del orden 2 existen y son continuas.

Estamos listos para enunciar el resultado principal de esta subseción.

Teorema 8.1 (Teorema de Schwarz). *Sea $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^\ell)$. Entonces para cada $x \in \mathbb{R}^d$, $i \in \{1, \ldots, \ell\}$ y $j_1, j_2 \in \{1, \ldots, d\}$ se tiene que:*

$$\frac{\partial^2 f_i(x)}{\partial x_{j_1} \partial x_{j_2}} = \frac{\partial^2 f_i(x)}{\partial x_{j_2} \partial x_{j_1}}.$$

Demostración. Es suficiente demostrar el resultado en el caso $\ell = 1$ y $d = 2$. Consideramos entonces una función $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$, $x = (x_1, x_2) \in \mathbb{R}^2$ y $u = (u_1, u_2) \in \mathbb{R}^2 \setminus \{(0,0)$ y anotamos por $\Delta^2(u_1, u_2)$ la variación:

$$\Delta^2(u_1, u_2) := [f(x_1 + u_1, x_2 + u_2) - f(x_1, x_2 + u_2)] - [f(x_1 + u_1, x_2) - f(x_1, x_2)] \quad (52)$$

Definimos ahora la función

$$\phi_1(t) := [f(x_1 + t, x_2 + u_2) - f(x_1, x_2 + u_2)] - [f(x_1 + t, x_2) - f(x_1, x_2)] .$$

Se deduce de nuestra hipotesis que $\phi_1 \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$, $\phi_1(0) = 0$ y $\phi_1(u_1) = \Delta^2(u_1, u_2)$. Aplicando el teorema del valor medio en el intervalo $[0, u_1]$ obtenemos la existencia de un elemento $s_1 \in [0, u_1]$ tal que

$$\Delta^2(u_1, u_2) = \phi_1(u_1) - \phi_1(0) = \phi_1'(s_1) u_1 := \left[\frac{\partial f}{\partial x_1}(x_1 + s_1, x_2 + u_2) - \frac{\partial f}{\partial x_1}(x_1 + s_1, x_2) \right] u_1.$$

Consideramos ahora la función

$$\phi_2(t) := \frac{\partial f}{\partial x_1}(x_1 + s_1, x_2 + t) - \frac{\partial f}{\partial x_1}(x_1 + s_1, x_2),$$

y observamos que $\phi_2 \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$, $\phi_2(0) = 0$ y

$$\phi_2(u_2) = \frac{\Delta^2(u_1, u_2)}{u_1}.$$

Por el teorema del valor medio, obtenemos $s_2 \in [0, u_2]$ tal que

$$\frac{\Delta^2(u_1, u_2)}{u_1} = \phi_2(u_2) - \phi_2(0) = \phi_2'(s_2) u_2 := \frac{\partial}{\partial x_2} \left[\frac{\partial f}{\partial x_1} \right] (x_1 + s_1, x_2 + s_2)\, u_2.$$

por lo que se deduce que

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1 + s_1, x_2 + s_2) = \frac{\Delta^2(u_1, u_2)}{u_1\, u_2}. \quad (53)$$

Observamos ahora que (52) se puede reescribir como sigue:

$$\Delta^2(u_1, u_2) := [f(x_1 + u_1, x_2 + u_2) - f(x_1 + u_1, x_2)] - [f(x_1, x_2 + u_2) - f(x_1, x_2)] . \quad (54)$$

Trabajando como antes, introducimos ahora la función ψ_2 de clase $\mathcal{C}^2$

$$\psi_2(t) := [f(x_1 + u_1, x_2 + t) - f(x_1 + u_1, x_2)] - [f(x_1, x_2 + t) - f(x_1, x_2)] .$$

Aplicando el teorema del valor medio en el intervalo $[0, u_2]$ obtenemos $\hat{s}_2 \in [0, u_2]$ tal que:

$$\Delta^2(u_1, u_2) = \psi_2(u_2) - \psi_2(0) = \psi_2'(\hat{s}_2)u_1 := \left[\frac{\partial f}{\partial x_2}(x_1 + u_1, x_2 + \hat{s}_2) - \frac{\partial f}{\partial x_1}(x_1, x_2 + \hat{s}_2)\right] u_2.$$

Luego, definiendo la función ψ_1 de clase $\mathcal{C}^1$

$$\psi_1(t) := \frac{\partial f}{\partial x_2}(x_1 + t, x_2 + \hat{s}_2 t) - \frac{\partial f}{\partial x_2}(x_1 + t, x_2 + \hat{s}_2)$$

y aplicando el teorema del valor medio en el intervalo $[0, u_1]$ obtenemos $\hat{s}_1 \in [0, u_1]$ tal que:

$$\frac{\Delta^2(u_1, u_2)}{u_2} = \psi_1(u_1) - \psi_1(0) = \psi_1'(\hat{s}_1)u_1 := \frac{\partial}{\partial x_1}\left[\frac{\partial f}{\partial x_2}\right](x_1 + \hat{s}_1, x_2 + \hat{s}_1)\, u_1.$$

Resumiendo, para cada $(u_1, u_2) \neq (0, 0)$ existen s_1, $\hat{s}_1 \in [0, u_1]$ y s_2, $\hat{s}_2 \in [0, u_2]$ tal que:

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1 + s_1, x_2 + s_2) = \frac{\Delta^2(u_1, u_2)}{u_1\, u_2} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 + \hat{s}_1, x_2 + \hat{s}_2). \tag{55}$$

Teniendo en cuenta que

$$\lim_{(u_1,u_2)\to 0}(s_1, s_2) = 0 \qquad \text{y} \qquad \lim_{(u_1,u_2)\to 0}(\hat{s}_1, \hat{s}_2) = 0,$$

y que las derivadas parciales de orden 2 son continuas en x, se deduce que todos los límites en (55) existen y

$$\lim_{(u_1,u_2)\to 0} \frac{\Delta^2(u_1, u_2)}{u_1\, u_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2),$$

lo que demuestra lo solicitado. □

Corolario 8.2 (funciones de clase $\mathcal{C}^k$)**.** *Sea $f \in \mathcal{C}^k(\mathbb{R}^d, \mathbb{R}^\ell)$, $k \geq 2$. Entonces para cada $x \in \mathbb{R}^d$, cada coordenada $i \in \{1, \ldots, \ell\}$ y cada permutación $\sigma \in \mathcal{P}_k$ del conjunto $\{j_1, \ldots, j_k\} \subset \{1, \ldots, d\}$ se tiene que:*

$$\frac{\partial^k f_i(x)}{\partial x_{j_1} \ldots \partial x_{j_k}} = \frac{\partial^k f_i(x)}{\partial x_{\sigma(j_1)} \ldots \partial x_{\sigma(j_k)}}.$$

Demostración. El resultado se obtiene aplicando sucesivamente el Teorema 8.1 a la función f_i y sus derivadas parciales. □

En el caso de funciones a valores reales ($\ell = 1$), el Teorema 8.1 garantiza que la matriz que representa la segunda derivada en x de la función $f : \mathbb{R}^d \to \mathbb{R}$ (que coincide con la matriz Jacobiana que representa la derivada en x de la función $Df : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}) := \left(\mathbb{R}^d\right)^*$) sea una matriz simétrica. Introducimos la siguiente definición.

Definición 8.3 (Hessiana)**.** Sea $f : \mathbb{R}^d \to \mathbb{R}$ una función dos veces diferenciable. Llamamos matriz *Hessiana* en $x \in \mathbb{R}^d$ y anotamos por $\mathcal{H}_f(x)$, la matriz Jacobiana $J(Df)(x)$, es decir la matriz que representa la segunda derivada $D^2 f(x)$.

$$\mathcal{H}_f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_d \partial x_1} \\ \vdots & \frac{\partial^2 f(x)}{\partial x_j^2} & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_d} & \cdots & \frac{\partial^2 f(x)}{\partial x_d^2} \end{pmatrix}$$

Quisieramos anotar que, igual con el caso de la matriz Jacobiana, la matriz de las derivadas parciales de orden 2 se podría formar siempre y cuando estas últimas existan, pero la matriz formada no será la Hessiana al menos que Df es diferenciable (es decir, que f es dos veces diferenciable). Por el Teorema 8.1, si además f es clase $\mathcal{C}^2$, entonces la matriz Hessiana es simétrica.

Miramos ahora un ejemplo donde el teorema de Schwarz se aplica:

Ejemplo 8.4 (ilustración del teorema de Schwarz)**.** Consideramos la función

$$\begin{cases} f : \mathbb{R}^3 \to \mathbb{R} \\ f(x,y,z) = z \exp(xy^2) \cos y. \end{cases}$$

La función f es de clase $\mathcal{C}^\infty$ porque sus derivadas parciales (de cualquier orden) existen y son nuevamente diferenciables. En particular:

$$\left[\begin{array}{l} \frac{\partial f}{\partial x}(x,y,z) = y^2 z \exp(xy^2) \cos(y) \\ \frac{\partial f}{\partial y}(x,y,z) = 2xyz \exp(xy^2) \cos(y) - z \exp(xy^2) \sin y \\ \frac{\partial f}{\partial z}(x,y,z) = \exp(xy^2) \cos(y) \end{array} \right.$$

El Teorema 8.1 afirma que la derivada

$$\frac{\partial}{\partial y}[\frac{\partial f_i}{\partial x}](x,y,z) = \frac{\partial}{\partial y}[y^2 z \exp(xy^2) \cos(y)]$$

coincide con la derivada

$$\frac{\partial}{\partial x}[\frac{\partial f_i}{\partial y}](x,y,z) = \frac{\partial}{\partial x}[2xyz \exp(xy^2) \cos(y) - z \exp(xy^2) \sin y]$$

(lo que no es obvio en un principio, en vista de las fórmulas involucradas). Dejamos al lector de comprobar que todas las derivadas cruzadas de orden 2 son iguales y formar la matriz Hessiana de f.

Ejercicio 8.5. Considere la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} \dfrac{xy(x^2 - y^2)}{x^2 + y^2}, & \text{si } (x,y) \neq (0,0) \\ 0, & \text{si } (x,y) = (0,0). \end{cases}$$

(i). Muestre que f es diferenciable en $\mathbb{R}^2$ y define la función derivada $(x,y) \mapsto Df(x,y)$ en cada punto.
(ii). Muestre que la función Df es diferenciable en $\mathbb{R}^2$, pero que la función

$$(x,y) \mapsto D^2 f(x,y)$$

no es continua en $(0,0)$ (es decir, f es dos veces derivable en $\mathbb{R}^2$, pero no es de clase $\mathcal{C}^2$).
(iii). Muestre que

$$\frac{\partial^2 f(0{,}0)}{\partial x \partial y} = 1 \qquad \text{y} \qquad \frac{\partial^2 f(0,0)}{\partial y \partial x} = -1,$$

es decir, el teorema de Schwarz no se cumple en el punto $(0,0)$.

8.2. Aproximación mediante polinomios de Taylor

Recordamos primero que una función $\phi : \mathbb{R} \to \mathbb{R}$ de clase $\mathcal{C}^k$ se puede aproximar alrededor del cero mediante su polinomio de Taylor de orden k dado por la formula:

$$p_\phi^k(t) := \phi(0) + \phi'(0)\, t + \frac{1}{2!}\phi''(0)\, t^2 + \ldots + \frac{1}{k!}\phi^{(k)}(0)\, t^k$$

en el sentigo que el error que se comete en considerar el valor $p_\phi^k(t)$ de la aproximación polinomial —en lugar del valor real $\phi(t)$— es del orden[18] $\mathrm{o}_k(t)$. Más precisamente, existe $t_* \in [0,t]$ tal que

$$\phi(t) = p_\phi^{k-1}(t) + \frac{1}{k!}\phi^{(k)}(t_*)\, t^k \qquad \text{(igualdad exacta}^{19}\text{)}$$

de lo que se deduce que

$$\mathrm{o}_k(t) = \phi(t) - p_\phi^k(t) = \left[\frac{\phi^{(k)}(t_*) - \phi^{(k)}(0)}{k!}\right] t^k.$$

Observación 8.6 (serie de Taylor)**.** El caso $k = \infty$ (funciones infinitamente derivables o de clase $\mathcal{C}^\infty$) permite considerar formalmente una serie, la llamada serie de Taylor:

$$\sum_{n=1}^{\infty} \frac{1}{n!}\phi^{(n)}(0)\, t^n$$

La función ϕ se dice analítica real (lo que se suele anotar por $\mathcal{C}^\omega$) si dicha serie tiene radio de convergencia no trivial y es una representación exacta de la función ϕ. Eso suele ser el caso de las funciones definidas directamente por una fórmula, como por ejemplo la función exponencial:

$$\exp t = \sum_{n=0}^{\infty} \frac{t^n}{n!}, \qquad \text{para todo } t \in \mathbb{R},$$

o las funciones $\sin t$, $\cos t$, $\arctan t$. En estos ejemplos, el radio de convergencia de la serie es todo $\mathbb{R}$. También se pueden considerar funciones con radio de convergencia finito, como por ejemplo la función $\ln(1+t)$, $t > -1$, con radio de convergencia igual a 1.

[18] Recordamos que la notación $\mathrm{o}_k(t)$ se refiere a una expresión funcional de t que cumple $\lim_{t \to 0} \frac{\mathrm{o}_k(t)}{t^k} = 0$.
[19] El caso $k = 1$ corresponde al teorema del valor medio que hemos evocado en varias ocasiones.

Observación 8.7 (jerarquización de funciones)**.** Podriamos considerar una jerarquización de las funciones, empezando de las "más suaves" a las menos regulares: las mejores funciones se consideran las funciones afines, que son polinomios de grado menor o igual a uno. Luego vienen las funciones polinomiales (que coinciden con su polinomio de Taylor del mismo grado), las funciones analíticas reales (que coinciden con la serie de Taylor en el radio de convergencia de esta última), las funciones $\mathcal{C}^\infty$ (que admiten aproximación polinomial de cualquier orden), las funciones $\mathcal{C}^k$, $k \geq 1$ (que admiten aproximación polinomial de cualquier orden menor o igual a k), las funciones diferenciables (que solo se aproximan alrededor de un punto por su tangente), luego las funciones continuas que no admiten ningun tipo de aproximación por funciones suaves.

Ejercicio 8.8. Muestre que la función $f : \mathbb{R} \to \mathbb{R}$

$$f(t) = \begin{cases} \exp(-1/t^2)\,, & \text{si } t \neq 0 \\ \\ 0\,, & \text{si } t = 0 \end{cases}$$

es infinitamente diferenciable, pero no analítica.

(*Hint*: Calcule la serie de Taylor de f alrededor del punto $t = 0$.)

8.2.1. Polinomios de Taylor de varias variables

En lo que sigue, generalizamos la teoría anterior para funciones de varias variables a valores reales. En efecto, sea $f \in \mathcal{C}^k(\mathbb{R}^d, \mathbb{R})$, $k \geq 1$. Fijamos $x \in \mathbb{R}^d$ y $u \in \mathbb{R}^d \setminus \{0\}$ y definimos la función

$$\phi(\tau) = f(x + \tau u), \qquad \tau \in \mathbb{R},$$

que es la restricción de la función f en la recta que pasa por x y tiene dirección u. Observamos que $\phi(0) = f(x)$, luego $\phi \in \mathcal{C}^k(\mathbb{R}, \mathbb{R})$ y mediante la regla de la cadena:

$$\phi'(t) = Df(x + tu)(u), \qquad \text{luego} \qquad \phi''(t) = D^2 f(x + tu)(u, u)$$

y de manera general:

$$\phi^{(n)}(t) = D^{(n)} f(x + tu)(\underbrace{u, \ldots, u}_{n \text{ veces}}), \qquad \text{donde } n \in \{1, \ldots, k\}.$$

Trabajando en coordenadas, obtenemos:

$$\phi'(t) = \nabla f(x + tu)^T \cdot u = \sum_{j_1=1}^{d} \frac{\partial f(x + tu)}{\partial x_{j_1}} u_{j_1}$$

luego

$$\phi''(t) = u^T \cdot \mathcal{H}_f(x + tu) \cdot u = \sum_{j_2=1}^{d} \sum_{j_1=1}^{d} \frac{\partial^2 f(x + tu)}{\partial x_{j_2} \partial x_{j_1}} u_{j_1} u_{j_2}$$

y para $n \in \{1, \ldots, k\}$

$$\phi^{(n)}(t) = \sum_{j_n=1}^{d} \cdots \sum_{j_1=1}^{d} \frac{\partial^{(n)} f(x+tu)}{\partial x_{j_n} \cdots \partial x_{j_1}} u_{j_1} \cdots u_{j_n}.$$

Evaluando las derivadas $\phi^{(n)}$ en $t = 0$ y aplicando el teorema de Taylor para la función ϕ obtenemos:

$$\begin{aligned} p_\phi^k(\tau) &:= \phi(0) + \phi'(0)\,\tau + \frac{1}{2!}\phi''(0)\,\tau^2 + \ldots + \frac{1}{k!}\phi^{(k)}(0)\,\tau^k = \\ &= f(x) + Df(x)(u)\,\tau + \ldots + \frac{1}{k!} D^{(k)} f(x)(u, \ldots, u)\,\tau^k. \end{aligned}$$

En particular, para $\tau = 1$ obtenemos una aproximación polinomial[20] valor $\phi(1) = f(x+u)$ como sigue:

$$f(x+u) \simeq f(x) + Df(x)(u) \; + \ldots + \frac{1}{k!} D^{(k)} f(x)(\underbrace{u, \ldots, u}_{k \text{ veces}}), \tag{56}$$

o en coordenadas:

$$f(x+u) \simeq f(x) + \sum_{j_1=1}^{d} \frac{\partial f(x)}{\partial x_{j_1}} u_{j_1} + \ldots + \frac{1}{k!} \sum_{j_k=1}^{d} \cdots \sum_{j_1=1}^{d} \frac{\partial^{(k)} f(x)}{\partial x_{j_k} \cdots \partial x_{j_1}} \; u_{j_1} \cdots u_{j_k} \tag{57}$$

8.2.2. Algebra multi-lineal, tensores.

Hemos visto en la Subsección 1.2 que una aplicación lineal $T : \mathbb{R}^d \to \mathbb{R}$ se representa mediante un vector, cuyas coordenadas se obtienen evaluando la aplicación lineal T en la base canónica $\{e_1, \ldots e_d\}$ de $\mathbb{R}^d$. Luego, en la Subsección 2.4 vimos que la derivada $Df(x)$ de una función $f : \mathbb{R}^d \to \mathbb{R}$ es un caso particular de aplicación lineal y se representa por la Jacobiana $\nabla f(x)^T$. Dimos ahora la siguiente definición.

Definición 8.9 (forma multi-lineal)**.** Una aplicación

$$T^{(n)} : \underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{n \text{ veces}} \to \mathbb{R}$$

se dice forma multi-lineal (o n-forma lineal), si para cada $i \in \{1, \ldots, d\}$, fijando las remantentes $n-1$ coordenadas a valores $\{\bar{x}_j : j \neq i\}$, la aplicación

$$u \mapsto T^{(n)}(\bar{x}_1, \ldots \bar{x}_{i-1}, u, \bar{x}_{i+1}, \ldots, \bar{x}_d)$$

es lineal. Una n-forma lineal se dice continua, si existe $c > 0$ tal que

$$|T^{(n)}(u_1, \ldots, u_d)| \leq c\,||u_1|| \cdots ||u_n||$$

[20]Tal y como en el caso de una variable, la función $f \in \mathcal{C}^k(\mathbb{R}^d)$ también admite una representación exacta. Dejamos al lector la tarea de escribir la igualdad exacta, substituyendo el último término de la expresión en (56) por $\frac{1}{k!} D^{(k)} f(x + t_* u)(u, \ldots, u$ donde $t_* \in [0, t]$.

El caso $n = 1$ corresponde a una forma lineal. Si $f : \mathbb{R}^d \to \mathbb{R}$ es una función k-veces diferenciable, entonces para todo $n \in \{1, \ldots, k\}$, la n-derivada $D^{(n)}f(x)$ en $x \in \mathbb{R}^d$ de f es una n-forma lineal. Por la Proposición 2.3 la función f es de clase $\mathcal{C}^n$ para todo $n < k$ y por lo tanto, se deduce del teorema de Schwarz (*cf.* Teorema 8.1) que $T^{(n)} := D^{(n)}f(x)$ es simetrica (es decir, $T^{(n)}(\sigma X) = T^{(n)}(X)$ para todo $X \in \mathbb{R}^d \times \ldots \times \mathbb{R}^d$ y permutación $\sigma \in \mathcal{P}_n$ de $\{1, \ldots, d\}$). Lo mismo ocurre para la k-derivada, si f es de clase $\mathcal{C}^k$.

Las formas 2-lineales se representan por matrices $(d \times d)$ cuyas entradas $T^{(2)}_{j_1 j_2}$ se obtienen evaluando $T^{(2)}(e_{j_1}, e_{j_2})$, donde $j_1, j_2 \in \{1, \ldots, d\}$[21]. En el caso $T^(2) := D^2 f(x)$, la matriz obtenida es la matriz Hessiana $\mathcal{H}_f(x)$ de f en $x \in \mathbb{R}^d$. Luego, para $n \geq 2$, la n-forma lineal $T^(n) := D^n f(x)$ se representa por la matriz

$$T^{(n)}_{j_1 \cdots j_n} = \frac{\partial^{(n)} f(x)}{\partial x_{jn} \cdots \partial x_{j_1}}.$$

El lector puede comparar la formula (56) con su representación en coordenadas (57).

8.2.3. Notación multi-índice

Con el fin de simplificar la fórmula (57) introducimos la siguiente notación: para el multi-índice

$$\overrightarrow{\alpha} := (\alpha_1, \ldots, a_d) \in \mathbb{N}^d$$

anotemos

$$|\overrightarrow{\alpha}| = \alpha_1 + \ldots + a_d = ||\overrightarrow{\alpha}||_1 \qquad \text{y} \qquad \overrightarrow{\alpha}! := \alpha_1! \cdots a_d!$$

luego para cada vector $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$ anotemos

$$u^{\overrightarrow{\alpha}} := u_1^{\alpha_1} \cdots u_d^{\alpha_d}$$

Miremos ahora la formula (57) en el caso $d = 2$, es decir, cuando $u = (u_1, u_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Explicitando las sumatorias obtenemos la expresión

$$f(x+u) \simeq f(x) + \underbrace{\left[\frac{\partial f(x)}{\partial x_1} u_1 + \frac{\partial f(x)}{\partial x_2} u_2\right]}_{\nabla f(x)^T \cdot u} + \frac{1}{2} \underbrace{\left[\frac{\partial^2 f(x)}{\partial x_1^2} u_1^2 + 2\frac{\partial^2 f(x)}{\partial x_1 \partial x_2} u_1 u_2 + \frac{\partial^2 f(x)}{\partial x_2^2} u_2^2\right]}_{u^T \cdot \mathcal{H}_f(x) \cdot u}$$
$$+ \frac{1}{3!}\left[\frac{\partial^{(3)} f(x)}{\partial x_1^3} u_1^3 + 3\frac{\partial^{(3)} f(x)}{\partial x_1^2 \partial x_2} u_1^2 u_2 + 3\frac{\partial^{(3)} f(x)}{\partial x_1 \partial x_2^2} u_1 u_2^2 + \frac{\partial^{(3)} f(x)}{\partial x_2^3} u_2^3\right] + \ldots$$

La expresión anteriorl se puede abreviar mediante el polinomio de Newton como sigue:

$$f(x+u) \simeq f(x) + \sum_{i=0}^{1} \binom{1}{i} \frac{\partial^{(1)} f(x)}{\partial x_1^{1-i} \partial x_2^i} u_1^{1-i} u_2^i + \frac{1}{2} \sum_{i=0}^{2} \binom{2}{i} \frac{\partial^{(2)} f(x)}{\partial x_1^{2-i} \partial x_2^i} u_1^{2-i} u_2^i +$$
$$+ \frac{1}{3!} \sum_{i=0}^{3} \binom{3}{i} \frac{\partial^{(3)} f(x)}{\partial x_1^{3-i} \partial x_2^i} u_1^{3-i} u_2^i + \ldots = \sum_{n=0}^{k} \frac{1}{k!} \sum_{i=0}^{n} \binom{n}{i} \frac{\partial^{(n)} f(x)}{\partial x_1^{n-i} \partial x_2^i} u_1^{n-i} u_2^i$$

[21]Se puede observar que $T^{(2)}$ es simétrica, si y sólo si para todo $j_1, j_2 \in \{1, \ldots, d\}$ se tiene $T^{(2)}_{j_1 j_2} = T^{(2)}_{j_2 j_1}$.

donde se utilizan las convenciones:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} := 1, \qquad 0! = 1, \qquad \text{y} \qquad \frac{\partial^{(0)} f(x)}{\partial x_1^0 \partial x_2^0} \equiv f(x).$$

Dado que

$$\frac{1}{n!}\begin{pmatrix} n \\ i \end{pmatrix} = \frac{1}{i!(n-i)!}$$

obtenemos:

$$\begin{aligned} f(x+u) \;&\simeq\; f(x) + \sum_{n=1}^{k}\sum_{i=0}^{n} \frac{1}{i!(n-i)!}\,\frac{\partial^{(n)} f(x)}{\partial x_1^{n-i}\partial x_2^{i}}\,u_1^{n-i}u_2^{i} \\ &= f(x) + \sum_{|\overrightarrow{\alpha}|=1}^{k} \frac{1}{\overrightarrow{\alpha}!}\frac{\partial^{(|\overrightarrow{\alpha}|)} f(x)}{\partial x^{\overrightarrow{\alpha}}} u^{\overrightarrow{\alpha}} \end{aligned}$$

donde $\overrightarrow{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{N}^2$ y para $|\overrightarrow{\alpha}| = \alpha_1 + \alpha_2 = n$ anotemos:

$$\frac{\partial^{(|\overrightarrow{\alpha}|)} f(x)}{\partial x^{\overrightarrow{\alpha}}} := \frac{\partial^{(n)} f(x)}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}}.$$

La formula anterior es directamente generalizable para funciones $f : \mathbb{R}^d \to \mathbb{R}$. En este caso, consideramos multi-índices $\overrightarrow{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$ y para $|\overrightarrow{\alpha}| = \alpha_1 + \ldots + \alpha_d = n$ anotemos:

$$\frac{\partial^{(|\overrightarrow{\alpha}|)} f(x)}{\partial x^{\overrightarrow{\alpha}}} := \frac{\partial^{(n)} f(x)}{\partial x_1^{\alpha_1}\ldots\partial x_2^{\alpha_d}}.$$

La fórmula (57) se escribe de forma equivalente mediante la siguiente fórmula compacta:

$$\begin{aligned} f(x+u) \;&\simeq f(x) + \sum_{|\overrightarrow{\alpha}|=1}^{k} \frac{1}{\overrightarrow{\alpha}!}\;\frac{\partial^{(|\overrightarrow{\alpha}|)} f(x)}{\partial x^{\overrightarrow{\alpha}}} u^{\overrightarrow{\alpha}} \\ &:= f(x) + \sum_{n=1}^{k}\sum_{\alpha_1+\ldots+\alpha_d=n}^{k} \left(\frac{1}{\alpha_1!\cdots\alpha_d!}\right)\frac{\partial^{(n)} f(x)}{\partial x_1^{\alpha_1}\cdots\partial x_2^{\alpha_d}} u_1^{\alpha_1}\cdots u_d^{\alpha_d} \end{aligned}$$

9. Problemas propuestos

Incluimos en esta sección una lista de problemas, clasificados por temas y de dificultad variada, que están basados en la teoría desarrolada en esta segunda parte del curso.

9.1. Ejercicios sobre diferenciabilidad y derivadas parciales

D1. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por:

$$f(x,y) = \begin{cases} x^{4/3} \sin(y/x)\,, & \text{si } x \neq 0, \\ 0\,, & \text{si } x = 0. \end{cases}$$

a) Estudie la continuidad de f.

b) Calcule, donde exista, la derivada de f, y en el caso que no exista justifique su respuesta.

c) Decida si la función f es diferenciable en $\mathbb{R}^2$.

D2. Considere la función

$$f(x,y) = \begin{cases} \left(\frac{xy}{x^2+y^2}\right) \sin\left(\frac{x^2y}{x^2+y^2}\right), & \text{si } (x,y) \neq (0,0), \\ 0\,, & \text{si } (x,y) = (0,0). \end{cases}$$

a) Estudie la continuidad de la función en $\mathbb{R}^2$.

b) Calcule, si es posible, las derivadas parciales de f en $(0,0)$.

c) Estudie la diferenciabilidad de la función en $(0,0)$.

D3. Considere la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por:

$$f(x,y) = \begin{cases} \dfrac{x^3 y \sin(xy^2)}{x^6 + y^6}\,, & \text{si } (x,y) \neq (0,0), \\ 0\,, & \text{si } (x,y) = (0,0). \end{cases}$$

a) Para cada punto $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, calcule las derivadas parciales de la función f. Demuestre que es diferenciable en este punto.

b) Determine si existen las derivadas parciales en el punto $(0,0)$ y las derivadas direccionales en la dirección $(e_1, e_2) \in \mathbb{R}$ con $(e_1, e_2) \neq (0,0)$.

c) Determine si f es diferenciable en $(0,0)$.

D4. Considere la función $f:\mathbb{R}^2 \to \mathbb{R}$ definida por:

$$f(x,y)=\begin{cases}\dfrac{1-\cos\left(\sqrt{x^2+y^2}\right)}{x^2+y^2}\,, & \text{si } (x,y)\neq(0,0),\\[2ex] 1/2\,, & \text{si } (x,y)=(0,0).\end{cases}$$

a) Demuestre que f es diferenciable para todo punto $(x,y)\neq(0,0)$.

b) Calcule, si existen, las derivadas parciales en el punto $(0,0)$.

c) Determine si f es diferenciable en $(0,0)$.

D5. Considere la función $f:\mathbb{R}^2 \to \mathbb{R}$ definida como

$$f(x,y)=\begin{cases}\dfrac{|x|\sin(y)}{\sqrt{x^2+y^2}}\,, & \text{si } (x,y)\neq(0,0),\\[2ex] 0\,, & \text{si } (x,y)=(0,0).\end{cases}$$

a) Demuestre que f es continua en $(0,0)$.

b) Determine si existen las derivadas parciales en el $(0,0)$.

c) Determine si f es diferenciable en el $(0,0)$.

D6. Sea $f:\mathbb{R}^2 \to \mathbb{R}$ la función definida por:

$$f(x,y)=\begin{cases}\dfrac{|xy|^{\alpha}}{(x^2+y^2)^{3/2}}\,, & \text{si } (x,y)\neq(0,0),\\[2ex] 0\,, & \text{si } (x,y)=(0,0).\end{cases}$$

a) Muestre que f es continua en $\mathbb{R}^2\backslash\{(0,0)\}$.

b) Pruebe que f es continua en $(0,0)$ si y solamente si $\alpha>3/2$.

c) Calcule las derivadas parciales en todo punto de $\mathbb{R}^2\backslash\{(0,0)\}$.

d) Calcule las derivadas parciales en $(0,0)$.

e) Encuentre los valores de $\alpha\in\mathbb{R}$ tales que f es diferenciable en $(0,0)$.

D7. Considere la función $f:\mathbb{R}^2 \to \mathbb{R}$ definida de la siguiente manera:

$$f(x,y)=\begin{cases}\dfrac{x|y|^{\alpha}}{x^4+y^4+x^2}\,, & \text{si } (x,y)\neq(0,0),\\[2ex] 0\,, & \text{si } (x,y)=(0,0).\end{cases}$$

(i). Demostrar si $\alpha>3$, entonces f es continua.

Para lo que sigue, fijamos el valor $\alpha = 2$.

(ii). Sea θ tal que $\cos(\theta) \neq 0$, defina $\phi(t) = f(t\cos(\theta), t\sin(\theta))$ y muestre que

$$\phi'(0) = \frac{\sin^2\theta}{\cos(\theta)}.$$

(iii). Encuentre todas las direcciones $e \in \mathbb{R}^2 \backslash \{(0,0)\}$ para las cuales existe la derivada direccional de f en el punto $(0,0)$, y calcula su valor.

(iv). ¿Es f diferenciable en $(0,0)$ para $\alpha = 2$? Justifique.

9.2. Ejercicios sobre la regla de la cadena

R1. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función homogénea de grado p, es decir:

$$f(tx, ty) = t^p f(x,y), \qquad \forall (x,y) \in \mathbb{R}^2, \quad \forall t \in \mathbb{R} \setminus \{0\}.$$

a) Demuestre que las derivadas parciales son funciones homogéneas de grado $p-1$.

b) Demuestre que $\Big\langle (x,y), \nabla f(x,y) \Big\rangle = pf(x,y)$.

R2. Considere la función $F : \mathbb{R}^2 \to \mathbb{R}$ definida en cada $(x,y) \in \mathbb{R}^2$ por

$$F(x,y) = f(g(xy), f(y^2, x))$$

donde $f : \mathbb{R}^2 \to \mathbb{R}$ y $g : \mathbb{R} \to \mathbb{R}$ son funciones diferenciables. Calcule las derivadas parciales de la función F.

R3. Considere $f : \mathbb{R}^2 \to \mathbb{R}$ diferenciable. Se define $g : \mathbb{R}^2 \to \mathbb{R}$ por

$$g(x,y) = f(u(x,y), v(x,y))$$

donde

$$u(x,y) = x + y \qquad \text{y} \qquad v(x,y) = \sin(x-y).$$

a) Demuestre que:

$$\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 = 2\left(\frac{\partial f}{\partial u}\right)^2 + 2\cos^2(x-y)\left(\frac{\partial f}{\partial v}\right)^2$$

y explicite en que puntos están evaluadas cada una de las derivadas parciales.

b) Verifique lo anterior para $f(u,v) = v^2 - u\sin^{-1}(v)$.

R4. Sea $f : \mathbb{R} \to \mathbb{R}$ y $g : \mathbb{R} \to \mathbb{R}$ dos funciones dos veces derivables. Considere la función $z : \mathbb{R}^2 \to \mathbb{R}$ definida por:

$$z(x, y) = xf(x + y) + yg(x + y).$$

Pruebe que:

$$\frac{\partial^2 z}{\partial x^2} - 2\frac{\partial^2 z}{\partial \partial y} + \frac{\partial^2 z}{\partial y^2} = 0.$$

R5. Sean $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R}^2 \to \mathbb{R}$ dos funciones diferenciables tales que

$$g(x, y) = xyf(\frac{x + y}{xy}).$$

Demuestre que:

$$x^2\frac{\partial g}{\partial x} - y^2\frac{\partial g}{\partial y} = G(x, y)g(x, y),$$

y encuentre explícitamente a $G(x, y)$.

9.3. Ejercicios sobre el teorema de la función inversa

N1. Considere el sistema de ecuaciones

$$\begin{aligned} u &= x^2 - xy \\ v &= y - x \end{aligned}$$

Muestre que en un entorno de cada punto $(x_0, y_0) \in \mathbb{R}^2$, con $x_0 \neq y_0$, es posible expresar las variables x, y como funciones de las variables (u, v). Calcule $\frac{\partial y}{\partial u}(u_0, v_0)$.

N2. Determine todos los puntos donde $F(x, y) = (xy^2 - \sin \pi x, y^2 - 25x^2 + 1)$ es invertible localmente y encuentre las derivadas parciales de la inversa en esos puntos.

N3. Sea $g_i : \mathbb{R}^n \to \mathbb{R}$ funciones de clase $\mathcal{C}^1(\mathbb{R}^n)$, para todo $i \in \{1, 2, ..., n\}$. Suponga además que

$$\text{máx}\{||\nabla g_i(x)||_\infty : x \in \mathbb{R}^n\} < \frac{1}{2n}, \forall i \in \{1, ..., n\}.$$

Pruebe que el siguiente sistema

$$\begin{aligned} x_1 &= g_1(x) \\ x_2 &= g_2(x) \\ &\vdots \\ x_n &= g_n(x) \end{aligned}$$

posee una única solución en $\mathbb{R}^n$, donde $x = (x_1, ..., x_n)$.

N4. Sea $f : \mathbb{R} \to \mathbb{R}$ de clase $\mathcal{C}^1$ tal que existe $\alpha \in (0,1)$, $\forall x \in \mathbb{R}$, $|f'(x)| \leq \alpha$ defina

$$F(x,y) = (x + f(y), y + f(x))^t.$$

a) Demuestre que F es de clase $\mathcal{C}^1$.

b) Ahora se demostrará que F es biyectiva. Para esto, dado $(a,b) \in \mathbb{R}^2$ defina $g(x,y) = (a - f(y), b - f(x))$. Demuestre que g es contractante y concluya.

c) Demuestre que $\forall (x,y) \in \mathbb{R}^2$, $JF(x,y)$ es invertible.

d) Concluya que F y F^{-1} son diferenciables.

9.4. Ejercicios sobre espacios tangentes y el teorema de la función implícita

M1. Considere la superficie $\mathbb{S}$ in $\mathbb{R}^3$ dada por la ecuación

$$x^2 + y^2 + z^2 = 1.$$

Sea P un plano tangente a $\mathbb{S}$ con ecuación normal:

$$\Big\langle\, (x,y,z) - (x_0,y_0,z_0),\ (1,-1,2) \,\Big\rangle \;=\; 0\,,$$

donde (x_0,y_0,z_0) es un punto de tangencia. Determine todos los valores posibles para el punto $(x_0,y_0,z_0) \in \mathbb{R}^3$.

M2. Probar que la ecuación

$$x^2y - xy^2 + z^2 \cos xz = 1,$$

define una función implícita $z = z(x,y)$ en un entorno del punto $(0, \sqrt{2}, 1)$.

M3. Muestre que las ecuaciones

$$\begin{aligned} x^2 - y^2 - u^3 + v^2 + 4 &= 0, \\ 2xy + y^2 - 2u^2 + 3v^2 + 8 &= 0, \end{aligned}$$

determinan localmente funciones $u(x,y), v(x,y)$ en un entorno del punto $(x,y) = (2,-1)$ con $u(2,-1) = 2$, $v(2,-1) = 1$. Determine además el valor de las derivadas $Du(2,-1)$ y $Dv(2,-1)$.

M4. Considere la superficie dada por $z = \ln(2x+y)$ y el paraboloide de ecuación

$$z = 5 + (x+1)^2 + (y+2)^2.$$

Determine la ecuación de la recta dada por la intersección de los planos tangentes a las superficies en los puntos $(-1,3,0)$ y $(2,9,10)$ respectivamente.

M5. Sea $f : \mathbb{R} \to \mathbb{R}$ una función de clase $\mathcal{C}^1$. Sea

$$C = \{(x, y, z) \in \mathbb{R}^3 : z = xf(\frac{x}{y}), y \neq 0\}.$$

Muestre que cada plan tangente a la superficie C pasa por el origen $(0, 0, 0) \in \mathbb{R}^3$.

M6. Suponga que $F : \mathbb{R}^2 \to \mathbb{R}^2$ es una función $\mathcal{C}^1$, y que $F(0, 0) = 0$. Que condiciones sobre F garantizan que la ecuación $F(F(x, y), y) = 0$ puede resolverse localmente alrededor del punto $(0, 0)$ despejando la variable $y = y(x)$ como una función de x de clase $\mathcal{C}^1$.

9.5. Ejercicios sobre series de Taylor

T1. Encuentre el polinomio de Taylor de orden 2 de la función

$$f(x, y) = e^{-(x^2+y^2)}$$

en el punto $(0, 0) \in \mathbb{R}^2$.

T2. *a*) Encuentre el polinomio de Taylor de orden 2 de la función

$$f(x, y) = (\sin(x + y))^2 + x^2 y,$$

en torno a $(1, -1)$.

b) Pruebe que $|f(1 + h_1, -1 + h_2) - \mathcal{P}_2(h_1, h_2)| \leq 8||(h_1, h_2)||^2$ donde $\mathcal{P}_2$ es el polinomio de orden 2 de f en torno de $(1, -1)$.

T3. Considere la función $f(x, y) = \sqrt{x} \cdot \sqrt[3]{y}$, definida en el subconjunto abierto de $\mathbb{R}^2$

$$\mathcal{U} = \{(x, y) \in \mathbb{R}^2 : \, x > 0, \, y > 0\}.$$

a) Obtenga el polinomio de Taylor de segundo orden entrono al punto $(x_0, y_0) \in \mathcal{U}$.

b) Obtenga un valor aproximado de la cantidad $\sqrt{1,03} \cdot \sqrt[3]{0,98}$.

9.6. Otros ejercicios

O1. Si $f(x, y, z) = (x, y, z)$, para cada $(x, y, z) \in \mathbb{R}^3$, demostrar que la matriz jacobiana $Jf(x, y, z)$ es la matriz identidad de dimensión 3. Luego:

a) Hallar todos las funciones diferenciables $f : \mathbb{R}^3 \to \mathbb{R}^3$ para las cuales la matriz jacobiana sea la identidad de dimensión 3.

b) Sea p, q, r funciones continuas de $\mathbb{R} \to \mathbb{R}$ dadas. Hallar todas las funciones diferenciables $f : \mathbb{R}^3 \to \mathbb{R}^3$ tales que la matriz jacobiana es una matriz diagonal, donde las entradas en su diagonal tienen como valores las funciones $p(x), q(y), r(z)$.

O2. Sea $f(x, y)$ una función clase $\mathcal{C}^1$ tal que:

$$x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y) = 0$$

Pruebe que f es una función constante.

O3. Sean X, Y espacios de Banach. Sea $f : X \to Y$ una función de clase $\mathcal{C}^1$. Sea $a \in X$ tal que la derivada $Df(a) \in \mathcal{L}(X, Y)$ es una función sobreyectiva. Pruebe que, si existe espacio normado X_1 tal que $X = X_1 \bigoplus \mathrm{Ker}(Df(a))$, entonces la función g definida por

$$\begin{aligned} g\colon E_1 \times \mathrm{Ker}(Df(a)) &\to Y \\ (x, y) &\mapsto g(x, y) = f(x + y), \end{aligned}$$

es una función con diferenciales parciales continuas y $D_2 g(a)$ es un isomorfismo (función lineal biyectiva).

Indicación: Recuerde que $\mathrm{Ker}(Df(a)) = \{x \in X \mid Df(a)(x) = 0\}$.

O4. Considere la función $J : \mathcal{C}([0, 1], \mathbb{R}) \to \mathbb{R}$ definida para cada $f \in \mathcal{C}([0, 1], \mathbb{R})$ por

$$J(f) = \int_0^1 (f(x) - x^2)f(x)dx.$$

Considere en $\mathcal{C}([0, 1], \mathbb{R})$ la norma $\|\cdot\|_\infty$.

a) Calcule la derivada direccional de J en el punto f, en la dirección h, es decir, calcule $dJ(f; h)$.

b) Demuestre que la forma lineal $\ell : \mathcal{C}([0, 1], \mathbb{R}) \to \mathbb{R}$ definida por

$$\ell(h) = \int_0^1 (2f(x) - x^2)h(x)dx$$

es la derivada de J en f, es decir $DJ(f) = \ell$.

O5. Para una función $F : \mathbb{R}^2 \to \mathbb{R}$ considere la ecuación en derivadas parciales

$$\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 = e^{4x}\operatorname{sen}^2(y).$$

Encuentre una solución a este problema, definido en todo $\mathbb{R}^2$.
Indicación: Considere una solución de forma $F(x,y) = \phi(e^x\cos(y), e^x\operatorname{sen}(y))$.

O6. Sea $u(x,t)$ una función de $\mathbb{R}^2$ a valores en $\mathbb{R}$ que satisface la ecuación

$$\frac{\partial u}{\partial t}(x,t) + u(x,t)\frac{\partial u}{\partial x}(x,t) = 0 \quad \forall (x,t) \in \mathbb{R}^2.$$

Además, sea $x = f(t)$ con $f : \mathbb{R} \to \mathbb{R}$ tal que

$$f'(t) = u(x,t).$$

Pruebe que la función $g(t) = u(f(t),t)$, $t \in \mathbb{R}$ es constante.

O7. Sea $f(x,y) = x\cos(y/x) + \tan(y/x)$. Muestre que $x^2 f_{xx} + 2xyf_{xy} + y^2 f_{zz} = 0$ en los puntos donde f es de clase $\mathcal{C}^2$.

O8. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ de clase $\mathcal{C}^2$, definimos el operador laplaciano como

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Considere las coordenadas polares $x = \rho\cos(\theta)$ $y = \rho\sin(\theta)$, donde $\rho \geq 0$, $\theta \in [0, 2\pi)$. Muestre que para $g(\rho,\theta) = f(\rho\cos(\theta), \rho\sin(\theta))$ y $\rho \neq 0$ se tiene

$$\Delta f = \frac{\partial^2 g}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial g}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 g}{\partial \theta^2}$$

y diga donde están evaluadas las derivadas.

O9. Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ una función continuamente diferenciable. Pruebe que f no puede ser inyectiva si $n > m$.