Tema 3

Campos escalares y vectoriales derivables. Reglas de derivación.

El estudio de la derivada de Fréchet para campos vectoriales requiere estar familiarizado con el espacio de Banach $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ de los campos vectoriales lineales de $\mathbb{R}^N$ en $\mathbb{R}^M$. Empezamos estableciendo el isomorfismo que existe entre este espacio vectorial y el las matrices $\mathcal{M}_{M\times N}(\mathbb{R})$. Dicho isomorfismo hace corresponder a la composición de campos el producto de las correspondientes matrices.

Es claro que el único campo vectorial lineal acotado es el nulo y que dos campos vectoriales lineales que toman los mismos valores en la esfera unidad coinciden (!Hágase!). A cada $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ le asignamos el número real

$$\|T\| := \max\{\|T(x)\| : \|x\| = 1\}.$$

Probamos que tal aplicación es una norma en $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$, y que para $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ se tiene

$$\|T(x)\| \leq \|T\| \ \|x\|, \ \forall x \in \mathbb{R}^N$$

con lo que todo campo vectorial lineal es lipschitziano y $\|T\|$ es la constante de Lipschitz de T (véase Definición 3.2).

Estudiamos también los isomorfismos topológicos en $\mathbb{R}^N$

$$\text{Iso } (\mathbb{R}^N) = \{T \in \mathcal{L}(\mathbb{R}^N) : \det\ A_T \neq 0\}$$

que serán esenciales a la hora de establecer el Teorema de la función inversa. Probamos que Iso $(\mathbb{R}^N)$ es un abierto de $\mathcal{L}(\mathbb{R}^N)$ y que la aplicación inversión $J : \text{Iso } (\mathbb{R}^N) \to \mathcal{L}(\mathbb{R}^N)$ dada por

$$J(T) := T^{-1}, \ \forall T \in \text{Iso } (\mathbb{R}^N)$$

es continua.

La manera natural de extender a campos vectoriales el concepto de función derivable es el concepto de derivada en el sentido de Fréchet: Un campo vectorial f definido en

$A \subset \mathbb{R}^N$ y con valores en $\mathbb{R}^M$ es derivable en un punto a interior si existe $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ verificando

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

en cuyo caso la aplicación T es única, se denomina la derivada de la función f en el punto a y se nota por $Df(a)$. Los conceptos de derivabilidad y derivada son algebraico-topológicos (no dependen de las normas elegidas en $\mathbb{R}^N$ y $\mathbb{R}^M$).

3.1. El espacio de Banach $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$.

En lo sucesivo, para cada dos naturales N y M, $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ denotará el espacio vectorial de los campos vectoriales lineales de $\mathbb{R}^N$ en $\mathbb{R}^M$, es decir, el conjunto de las aplicaciones $T : \mathbb{R}^N \to \mathbb{R}^M$ tales que

$$T(x+y) = T(x) + T(y), \quad T(\lambda x) = \lambda T(x), \quad \forall x, y \in \mathbb{R}^N, \ \forall \lambda \in \mathbb{R},$$

con las operaciones usuales de suma y producto por escalares. En el caso de que $N = M$ escribiremos simplemente $\mathcal{L}(\mathbb{R}^N)$ en lugar de $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$.

Conviene dejar sentado desde el primer momento que el espacio vectorial $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ y el espacio vectorial $\mathcal{M}_{M\times N}(\mathbb{R})$ de las matrices $M \times N$ de números reales son matemáticamente indistinguibles. Para cada $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ definimos $A_T \in \mathcal{M}_{M\times N}(\mathbb{R})$ por

(3.1.1) $$A_T := \Big(T_i(e_j)\Big)_{\substack{1 \le i \le M \\ 1 \le j \le N}},$$

donde $T_1, \ldots, T_M$ son los campos escalares componentes de T y $\{e_1, \ldots, e_N\}$ es la base canónica de $\mathbb{R}^N$. La aplicación de $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ en $\mathcal{M}_{M\times N}(\mathbb{R})$ definida por $T \to A_T$ es un isomorfismo de $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ sobre $\mathcal{M}_{M\times N}(\mathbb{R})$, cuyo inverso es la aplicación de $\mathcal{M}_{M\times N}(\mathbb{R})$ sobre $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ definida por $A \to T_A$ donde T_A es la aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}^M$ dada por

(3.1.2) $$\Big(T_A(x)\Big)^t := Ax^t, \ \forall x \in \mathbb{R}^N.$$

En particular, si $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R})$, entonces $A_T \in \mathbb{R}^N$ y si $A \in \mathbb{R}^N$, entonces la correspondiente aplicación T_A actúa de la forma siguiente

(3.1.3) $$T_A(x) = (A|x), \ \forall x \in \mathbb{R}^N,$$

donde hemos notado por $(\cdot|\cdot)$ el producto escalar en $\mathbb{R}^N$.

Además esta identificación de $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ con $\mathcal{M}_{M\times N}(\mathbb{R})$ tiene la importante siguiente propiedad. Si $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ y $S \in \mathcal{L}(\mathbb{R}^M, \mathbb{R}^P)$, entonces

(3.1.4) $$A_{(ST)} = A_S A_T,$$

donde, como es usual, notamos por yuxtaposición la composición de los operadores S y T, es decir

$$(ST)(x) = S(T(x)), \forall x \in \mathbb{R}^N.$$

En efecto, si $x \in \mathbb{R}^N$ se tiene que

$$(A_S A_T)x^t = A_S(A_T x^t) = A_S\Big(T(x)\Big)^t = (S(T(x)))^t = ((ST)(x))^t = A_{(ST)}x^t.$$

Sabemos que un campo vectorial es continuo si, y sólo si, lo son sus campos escalares componentes. En el caso particular de campos vectoriales lineales esto es automático, pues en vista de 3.1.3, los campos escalares componentes son polinomios.

Es claro que si $M \neq N$, entonces no existen biyecciones lineales de $\mathbb{R}^N$ sobre $\mathbb{R}^M$. Además, las biyecciones lineales de $\mathbb{R}^N$ sobre $\mathbb{R}^N$ son automáticamente bicontinuas, es decir, son homeomorfismos lineales. Es usual la siguiente nomenclatura:

Definición 3.1 (Isomorfismo topológico). Si X e Y son espacios normados, una aplicación $T : X \longrightarrow Y$ es un *isomorfismo topológico* si T es una aplicación biyectiva, lineal y continua cuya inversa también es continua.

Por el comentario anterior a la definición, toda biyección lineal $T : \mathbb{R}^N \to \mathbb{R}^N$ es un isomorfismo topológico. En lo que sigue, notaremos Iso $(\mathbb{R}^N)$ al conjunto de los isomorfismos topológicos de $\mathbb{R}^N$ sobre $\mathbb{R}^N$.

A partir de 3.1.4 se obtiene que

$$\text{Iso } (\mathbb{R}^N) = \{T \in \mathcal{L}(\mathbb{R}^N) : \det\, A_T \neq 0\}$$

y si $T \in$ Iso $(\mathbb{R}^N)$, entonces

$$A_{(T^{-1})} = (A_T)^{-1}.$$

Veremos enseguida que los campos vectoriales lineales son de hecho algo más que uniformemente continuos.

Definición 3.2 (función lipschitziana). Sean (E, d) y (F, ρ) espacios métricos. Una fuinción $f : E \to F$ es *lipschitziana* si $\exists K \geq 0$ tal que

$$\rho(f(x), f(y)) \leq K d(x, y), \quad \forall x, y \in E. \tag{$*$}$$

A la menor de las constantes que verifican la condición $(*)$ se le llama la *constante de Lipschitz* de f. Evidentemente las funciones lipschitzianas con constante de Lipschitz 0 no son otra cosa que las funciones constantes.

Es inmediato probar que las funciones lipschitzianas son uniformemente continuas. En efecto, dado $\varepsilon > 0$, si f es lipschitziana de razón K, sea $\delta := \frac{\varepsilon}{K+1}$. Se tiene que

$$[d(x, y) < \delta] \ \Rightarrow \ \rho(f(x), f(y)) \leq K\, d(x, y) < \varepsilon.$$

La función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \sqrt{|x|} \qquad (x \in \mathbb{R})$$

es uniformemente continua y, sin embargo, no es lipschitziana. En efecto, si $x, y \in \mathbb{R}$, se tiene que

$$\sqrt{|x|} \leq \sqrt{|x - y| + |y|} \leq \sqrt{|x - y|} + \sqrt{|y|},$$

es decir
$$\sqrt{|x|}-\sqrt{|y|}\leq\sqrt{|x-y|},$$
y en consecuencia, intercambiando los papeles de x e y, que
$$\left|\ \sqrt{|x|}-\sqrt{|y|}\ \right|\leq\sqrt{|x-y|}$$
de donde se deduce inmediatamente que es uniformemente continua (¡Hágase!).

Si fuese lipschitziana de razón $K>0$, tendríamos que
$$\left|\frac{1}{\sqrt{n}}-0\right|\leq K\ \frac{1}{n},\ \forall n\in\mathbb{N}$$
o lo que es lo mismo
$$n\leq K^2,\forall n\in\mathbb{N}$$
y los naturales estarían acotados.

Ejemplos 3.3 (funciones lipschitzianas).

a) Sea $(X,\|\cdot\|)$ un espacio normado. Son lipschitzianas las siguientes aplicaciones:

- La norma en X:
$$|\,\|x\|-\|y\|\,|\leq\|x-y\|.$$
- La suma en X:
$$\|(x_1+y_1)-(x_2+y_2)\|\leq\|x_1-x_2\|+\|y_1-y_2\|\leq 2\|(x_1,y_1)-(x_2,y_2)\|_\infty.$$
- Sabemos que el producto por escalares ni siquiera es uniformemente continuo.

b) Sea (E,d) un espacio métrico. Son lipschitzianas las siguientes aplicaciones:

- La distancia de E:
$$|d(x_1,y_1)-d(x_2,y_2)|\leq|d(x_1,y_1)-d(y_1,x_2)|+|d(y_1,x_2)-d(x_2,y_2|\leq$$
$$d(x_1,x_2)+d(y_1,y_2)\leq 2d((x_1,y_1),(x_2,y_2)).$$
- La distancia a un subconjunto no vacío A de E:
$$|\mathrm{dist}(x,A)-\mathrm{dist}(y,A)|\leq d(x,y).$$
En efecto, para $x,y\in E$, se tiene
$$d(x,a)\leq d(x,y)+d(y,a),\quad\forall a\in A,$$
y en consecuencia
$$\mathrm{dist}(x,A)\leq d(x,y)+\mathrm{dist}(y,A),$$

o lo que es lo mismo

$$\operatorname{dist}(x, A) - \operatorname{dist}(y, A) \leq d(x, y).$$

Ahora, intercambiando los papeles de x e y, se tiene la desigualdad

$$\operatorname{dist}(y, A) - \operatorname{dist}(x, A) \leq d(x, y),$$

que unida a la anterior lleva a

$$\Big| \operatorname{dist}(x, A) - \operatorname{dist}(y, A) \Big| \leq d(x, y).$$

En consecuencia, si (E, d) es un espacio métrico, A un subconjunto no vacío de E y a un número real, entonces los conjuntos

$$U := \{x \in E : \operatorname{dist}(x, A) > a\} \quad \text{y} \quad V := \{x \in E : \operatorname{dist}(x, A) < a\}.$$

son abiertos (ver caracterización topológica de la continuidad global).

La siguiente proposición prueba que una aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}^M$ es lipschitziana, y además da la constante de Lipschitz.

Proposición 3.4. *Sean $N, M \in \mathbb{N}$ y $T : \mathbb{R}^N \to \mathbb{R}^M$ lineal. Entonces se verifican las siguientes afirmaciones:*

i) $\|T\|$ *alcanza el máximo en la esfera unidad y en consecuencia podemos definir*

$$\|T\| := \max\{\|T(x)\| : \|x\| = 1\}.$$

ii) T es lipschitziana, de hecho se verifica que

$$\|T(x)\| \leq \|T\| \|x\|, \forall x \in \mathbb{R}^N. \tag{3.1.5}$$

Además

$$\|T\| = \min\{K \geq 0 : \ \|T(x)\| \leq K\|x\|, \ \forall x \in \mathbb{R}^N\} = \max\{\|T(x)\| : \|x\| \leq 1\}.$$

La primera igualdad nos asegura que $\|T\|$ es la constante de Lipschitz de T.

Demostración:

i) Como T es continua y la esfera unidad es compacta, *i*) es consecuencia de la continuidad de la aplicación norma, de la regla de la cadena para funciones continuas y de la propiedad de compacidad.

ii) Es claro que

$$\Big\| T\Big(\frac{x}{\|x\|}\Big) \Big\| \leq \|T\|, \ \forall x \in \mathbb{R}^N \backslash \{0\},$$

y usando que T es lineal, tenemos

$$\|T(x)\| \leq \|T\| \, \|x\|, \ \forall x \in \mathbb{R}^N \backslash \{0\}.$$

Dado que esta última expresión es también válida para cero, podemos escribirla equivalentemente en la forma

$$\|T(x)\| \leq \|T\| \|x\|, \ \forall x \in \mathbb{R}^N.$$

De la anterior desigualdad se deduce que T es lipschitziana, ya que si $x, y \in \mathbb{R}^N$, obtenemos

$$\|T(x) - T(y)\| = \|T(x-y)\| \leq \|T\| \, \|x-y\|,$$

es decir, T es lipschitziana.

Sea ahora $K \geq 0$ tal que

$$\|T(x)\| \leq K \, \|x\|, \quad \forall x \in \mathbb{R}^N.$$

Se sigue que $\|T\| = \max\{\|T(x)\| : \ \|x\| = 1\} \leq K$, y por tanto

$$\|T\| = \min\{K \geq 0 : \|T(x)\| \leq K \, \|x\|, \ \forall x \in \mathbb{R}^N\},$$

es decir, $\|T\|$ es la constante de Lipschitz de T.

Finalmente, la propiedad de compacidad nos asegura que T alcanza el máximo en la bola unidad cerrada. Es claro que

$$\|T\| \leq \max\{\|T(x)\| : \ \|x\| \leq 1\}$$

y 3.1.5 nos asegura que también

$$\max\{\|T(x)\| : \ \|x\| \leq 1\} \leq \|T\|.$$

Hemos probado que

$$\|T\| = \max\{\|T(x)\| : \ \|x\| \leq 1\}.$$

■

El teorema de Hausdorff nos asegura que todas las normas definibles en el espacio vectorial $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ son equivalentes (es de dimensión $M \times N$). En el siguiente resultado presentamos la forma usual de normar este espacio.

Teorema 3.5 (El espacio de Banach $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$). *La función que a cada aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^M$ le hace corresponder el número real*

$$\|T\| := \max\{\|T(x)\| : \|x\| = 1\}$$

es una norma en $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$, denominada norma de operadores. Además $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ es un espacio de Banach.

Demostración:

Veamos que la función $\|\cdot\| : \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M) \to \mathbb{R}$ es una norma. Para $T, S \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ y $\lambda \in \mathbb{R}$, se tiene:

i)
$$\|T\| = 0 \Rightarrow \left(\|T(x)\| \leq \|T\|\|x\|, \forall x \in \mathbb{R}^N\right) \Rightarrow T = 0.$$

ii)
$$\|\lambda T\| = \max\{\|(\lambda T)(x)\| : \|x\| = 1\} = \max\{|\lambda|\,\|T(x)\| : \|x\| = 1\} = |\lambda|\ \|T\|.$$

iii)
$$\|(T+S)(x)\| = \|T(x)+S(x)\| \leq \|T(x)\|+\|S(x)\| \leq \|T\|\ \|x\|+\|S\|\ \|x\|, \forall x \in \mathbb{R}^N,$$

donde se ha usado 3.1.5. Hemos probado que

$$\|(T+S)(x)\| \leq (\|T\| + \|S\|)\ \|x\|,\ \forall x \in \mathbb{R}^N$$

y en consecuencia, en vista de la proposición anterior, concluimos que

$$\|T+S\| \leq \|T\| + \|S\|.$$

Finalmente $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$, al ser de dimensión finita, es completo para cualquier norma, en particular, para la norma de operadores. ■

Hállense la norma de operadores de $\mathcal{L}((\mathbb{R}^2, \|.\|_1), \mathbb{R})$, $\mathcal{L}((\mathbb{R}^2, \|.\|_2), \mathbb{R})$ (úsese en este caso la desigualdad de Cauchy-Schwarz, Apéndice A) y $\mathcal{L}((\mathbb{R}^2, \|.\|_\infty), \mathbb{R})$.

El siguiente resultado generaliza el hecho de que $\mathbb{R}^*$ sea abierto, ya que Iso $(\mathbb{R})$ se identifica con $\mathbb{R}^*$ (¿Por qué?).

Proposición 3.6. Iso $(\mathbb{R}^N)$ *es un abierto de* $\mathcal{L}(\mathbb{R}^N)$.

Demostración:

La aplicación de $\mathcal{L}(\mathbb{R}^N)$ en $\mathbb{R}$ definida por

$$T \to \det A_T\ (T \in \mathcal{L}(\mathbb{R}^N))$$

es continua (¡Hágase!). En consecuencia

$$\text{Iso } (\mathbb{R}^N) = \{T \in \mathcal{L}(\mathbb{R}^N) :\ \det A_T \neq 0\}$$

es un abierto de $\mathcal{L}(\mathbb{R}^N)$. ■

Veamos para finalizar esta sección que, al igual que la aplicación de $\mathbb{R}^*$ en $\mathbb{R}$ dada por $x \to \frac{1}{x}$ es continua, también la aplicación de Iso $(\mathbb{R}^N)$ en $\mathcal{L}(\mathbb{R}^N)$ dada por $T \to T^{-1}$ es continua.

Proposición 3.7 (Continuidad de la aplicación inversión). *La aplicación inversión* $J : \text{Iso } (\mathbb{R}^N) \to \mathcal{L}(\mathbb{R}^N)$ *definida por*

$$J(T) := T^{-1}, \quad \forall T \in \text{Iso } (\mathbb{R}^N)$$

es continua.

Demostración:

Si $S, T \in \mathcal{L}(\mathbb{R}^N)$, es claro que $ST \in \mathcal{L}(\mathbb{R}^N)$. Veamos ahora la relación entre las normas de los tres operadores. Para $x \in \mathbb{R}^N$ se tiene que

$$\|(ST)(x)\| = \|S(T(x))\| \leq \|S\|\|T(x)\| \leq \|S\|\|T\|\|x\|,$$

y por tanto,

$$\|ST\| \leq \|S\|\|T\|. \tag{3.1.6}$$

Sean $T \in \text{Iso}\,(\mathbb{R}^N)$ y $\{T_n\}$ una sucesión en $\text{Iso}\,(\mathbb{R}^N)$ convergente a T. Para cada natural n se tiene que, en vista de la desigualdad anterior

$$\|J(T_n) - J(T)\| = \|T_n^{-1} - T^{-1}\| = \|T_n^{-1}(T - T_n)T^{-1}\| \leq \|T_n^{-1}\|\ \|T - T_n\|\ \|T^{-1}\|,$$

de donde se deduce la continuidad de J en T sin más que comprobar que la sucesión $\{\|T_n^{-1}\|\}$ está acotada. Para cada natural n tenemos, usando la desigualdad recién obtenida concluimos que

$$\left|\frac{1}{\|T^{-1}\|} - \frac{1}{\|T_n^{-1}\|}\right| = \left|\frac{\|T_n^{-1}\| - \|T^{-1}\|}{\|T^{-1}\|\ \|T_n^{-1}\|}\right| \leq$$

$$\leq \frac{\|T_n^{-1} - T^{-1}\|}{\|T^{-1}\|\ \|T_n^{-1}\|} \leq$$

$$\leq \frac{\|T_n^{-1}\|\ \|T - T_n\|\ \|T^{-1}\|}{\|T^{-1}\|\ \|T_n^{-1}\|} = \|T - T_n\|.$$

Por tanto

$$\lim \frac{1}{\|T_n^{-1}\|} = \frac{1}{\|T^{-1}\|},$$

en consecuencia

$$\lim\ \|T_n^{-1}\| = \|T^{-1}\|,$$

y en particular, la sucesión $\{\|T_n^{-1}\|\}$ está acotada. ∎

3.2. Concepto de derivada.

El concepto de derivada es de los que más han influido en el desarrollo de la matemática. Nuestro objetivo es el estudio de la derivabilidad, así como el cálculo de la derivada cuando ello proceda, de las funciones $f : A \to \mathbb{R}^M$, donde $A \subset \mathbb{R}^N$, siendo M y N dos naturales prefijados.

Esencialmente el estudio de la derivabilidad de un campo vectorial f en un punto a responde al problema de si f es aproximable por una función afín (continua) en el punto a, es decir, por una función g de la forma $g(x) = c + T(x),\ \forall x \in \mathbb{R}^N$, donde $c \in \mathbb{R}^M$ y $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$. La parte lineal de la mejor aproximación afín de la función f en el punto a es la derivada de ésta en a y las propiedades de esta mejor aproximación (equivalentemente de la derivada) repercuten de manera natural sobre las propiedades locales de la función en el punto a. Así, el cálculo diferencial es una potente herramienta para estudiar el <u>comportamiento local de funciones</u>. El cálculo diferencial para aplicaciones entre espacios normados fue iniciado por Fréchet en 1.925, si bien la paternidad ha de ser compartida con muchos otros matemáticos como Stolz, Young, etc.

Empezaremos recordando la noción de derivada para funciones reales de variable real, así como su interpretación geométrica.

Definición 3.8 (derivada de funciones reales de variable real). Sean $A \subset \mathbb{R}$, $a \in A \cap A'$ y $f : A \to \mathbb{R}$ una función. Se dice que f es derivable en el punto a si existe

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

en cuyo caso el valor de tal límite se nota $f'(a)$ y se denomina la <u>*derivada*</u> de la función f en el punto a.

La derivabilidad de una función tiene la siguiente magnífica caracterización en términos de existencia de una mejor aproximación afín, cuya interpretación es clara.

Proposición 3.9 (Caracterización de la derivabilidad). *Sean $A \subset \mathbb{R}, a \in A \cap A'$ y $f : A \to \mathbb{R}$ una función. Equivalen:*

i) f es derivable en a.

ii) Existe una función afín (continua) $g : \mathbb{R} \to \mathbb{R}$ verificando que $g(a) = f(a)$ y que

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

En ese supuesto la función g es única y viene dada por

$$g(x) = f(a) + f'(a)(x - a),\ \forall x \in \mathbb{R}.$$

La Definición 3.8 y la Proposición 3.9 pueden extenderse literalmente a funciones vectoriales de una variable real.

Definición 3.10 (derivada elemental para funciones de $\mathbb{R}$ a $\mathbb{R}^M$). Sean M un natural, $A \subset \mathbb{R}$, $a \in A \cap A'$ y $f : A \to \mathbb{R}^M$ una función. Se dice que f es elementalmente derivable en el punto a si existe

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

en cuyo caso el valor de tal límite (un vector de $\mathbb{R}^M$) se denomina la derivada elemental de f en el punto a y se nota $f'(a)$. Se dice que f es derivable elementalmente en un subconjunto $B \subset A$ si es derivable elementalmente en cada punto de B.

Sea $A_1 \subset A$ el conjunto de puntos donde f es derivable elementalmente. La aplicación $x \to f'(x)$ de A_1 en $\mathbb{R}^M$ se denomina la aplicación derivada elemental de f y se nota f'.

Proposición 3.11 (Caracterización de la derivabilidad elemental). *Sean M un natural, $A \subset \mathbb{R}$, $a \in A \cap A'$ y $f = (f_1, ..., f_M) : A \to \mathbb{R}^M$ una función. Equivalen:*

i) f es derivable elementalmente en a.

ii) $f_1, ..., f_M$ son derivables en a.

iii) Existe una función afín (continua) $g : \mathbb{R} \to \mathbb{R}^M$ verificando que $g(a) = f(a)$ y que

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

En ese supuesto, $f'(a) = (f_1'(a), ..., f_M'(a))$, la función g es única y viene dada por

$$g(x) = f(a) + (x - a)f'(a), \ \forall x \in \mathbb{R}.$$

Demostración:

i) $\Leftrightarrow$ *ii)* Basta observar que para $x \in A\backslash\{a\}$ es

$$\frac{f(x) - f(a)}{x - a} = \left(\frac{f_1(x) - f_1(a)}{x - a}, \ldots, \frac{f_M(x) - f_M(a)}{x - a} \right)$$

y aplicar entonces la Proposición 2.62 sobre la reducción del límite a campos escalares.

ii) $\Leftrightarrow$ *iii)* Es consecuencia inmediata de las Proposiciones 3.9 y 2.62. ∎

El intento de llevar la Definición 3.10 al caso en que el dominio de la aplicación sea un subconjunto de un espacio de dimensión mayor que 1 tropieza con la imposibilidad de dar sentido al límite que en ella aparece. Sin embargo, la caracterización de la derivabilidad de una función en un punto por la existencia de una mejor aproximación afín a la función en el punto (afirmación iii) de la proposición anterior) es perfectamente trasladable al ámbito deseado si se cae en la cuenta de que

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

La anterior propiedad tiene sentido para funciones definidas en un subconjunto A de $\mathbb{R}^N$ y con valores en $\mathbb{R}^M$, sin más que sustituir el valor absoluto por una norma $\|\cdot\|$ en $\mathbb{R}^N$ y la función afín de $\mathbb{R}$ en $\mathbb{R}^M$ por una función afín de $\mathbb{R}^N$ en $\mathbb{R}^M$.

Parece pues que en el ambiente siguiente: $A\subset\mathbb{R}^N$, $a\in A\cap A'$, y $f:A\to\mathbb{R}^M$, la función f debe considerarse derivable en a cuando exista una función afín (continua) $g:\mathbb{R}^N\to\mathbb{R}^M$ verificando que $g(a)=f(a)$ y que

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

Como quiera que una tal aplicación afín g ha de tener la forma

$$g(x)=f(a)+T(x-a),\ \forall x\in\mathbb{R}^N$$

para conveniente aplicación lineal $T:\mathbb{R}^N\to\mathbb{R}^M$, la condición anterior es equivalente a

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

Es importante observar que la condición anterior es topológica, es decir involucra solamente las topologías de $\mathbb{R}^N$ y $\mathbb{R}^M$ y no las concretas normas que se tomen. En efecto, fijada una norma en $\mathbb{R}^N$, al ser la noción de límite topológica concluimos que la condición es independiente de la norma elegida en $\mathbb{R}^M$. Ahora si $|\!|\!|\cdot|\!|\!|$ es otra norma en $\mathbb{R}^N$, entonces se tiene para $x\in A\backslash\{a\}$ que

$$\frac{f(x)-f(a)-T(x-a)}{\|x-a\|}=\frac{|\!|\!|x-a|\!|\!|}{\|x-a\|}\,\frac{f(x)-f(a)-T(x-a)}{|\!|\!|x-a|\!|\!|}$$

y como por la equivalencia de $\|\cdot\|$ y $|\!|\!|\cdot|\!|\!|$, existen $0<k,K$ tales que $k\le\dfrac{|\!|\!|x-a|\!|\!|}{\|x-a\|}\le K$, se sigue, teniendo en cuenta de nuevo que la noción de límite es topológica, que

$$\lim_{x\to a}\frac{f(x)-f(a)-T(x-a)}{\|x-a\|}=0 \iff \lim_{x\to a}\frac{f(x)-f(a)-T(x-a)}{|\!|\!|x-a|\!|\!|}=0.$$

Examinamos a continuación la repercusión que tiene la propiedad anterior sobre la continuidad de f en a. De ella deducimos que

$$\lim_{x\to a} f(x)-f(a)-T(x-a)=0,$$

lo que prueba que f es continua al ser T continua.

Para garantizar la unicidad de la aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^M$ verificando la condición requerida, necesitamos poder acercarnos al punto a en cualquier dirección, esto es, si para cada $x \in S(\mathbb{R}^N)$ notamos

$$A_x = \{t \in \mathbb{R} : \ a + tx \in A\},$$

deberá ser $0 \in A'_x$ para todo x en $S(\mathbb{R}^n)$ (véase Ejercicio 3.2). Es importante notar al respecto que la condición topológica más expeditiva para garantizar la anterior propiedad es que a sea un punto interior de A y esta condición será asumida en adelante. En este caso, sea $\delta > 0$ tal que $B(a, \delta) \subset A$. Para $x \in S(\mathbb{R}^N)$ y $t \in \mathbb{R}$ con $0 < |t| < \delta$, se tiene que $a + tx \in A$ y

$$\left\| \frac{f(a+tx) - f(a)}{t} - T(x) \right\| = \frac{\|f(a+tx) - f(a) - tT(x)\|}{|t|} =$$

$$\frac{\|f(a+tx) - f(a) - T(a+tx-a)\|}{\|a+tx-a\|},$$

de donde deducimos que

$$(*) \qquad \lim_{t \to 0} \frac{f(a+tx) - f(a)}{t} = T(x)$$

expresión que nos asegura la unicidad de T dado que una aplicación lineal está determinada por el comportamiento en la esfera unidad (¡Hágase!).

Estamos ya en condiciones de definir de manera coherente el concepto de derivada.

Definición 3.12 (función derivable Fréchet). Sean $M, N \in \mathbb{N}, A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$, y $f : A \to \mathbb{R}^M$ una función. Se dice que f es <u>derivable</u> (en el sentido de Fréchet) o diferenciable en el punto a si existe una aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ verificando

$$(**) \qquad \lim_{x \to a} \frac{f(x) - f(a) - T(x-a)}{\|x-a\|} = 0$$

donde $\|\cdot\|$ es cualquier norma en $\mathbb{R}^N$. En tal caso la aplicación T es única, se denomina la <u>*derivada de una función en un punto*</u> o <u>*diferencial de f en a*</u> y se nota por $Df(a)$. Se dice que f es derivable en un subconjunto $B \subset A$ si es derivable en cada punto de B.

Sea $A_1 \subset A$ el conjunto de puntos donde f es derivable. La aplicación $x \mapsto Df(x)$ de A_1 en $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ se denomina la <u>aplicación derivada</u> de f y se nota Df.

Se dice que f es de <u>*clase $\mathcal{C}^1$*</u> en a, y se nota $f \in C^1(a)$, si f es derivable en un entorno de a y la aplicación Df es continua en a. Se dice que f es de clase C^1 en un subconjunto $B \subset A$ si es de clase C^1 en cada punto de B.

Se dice que f es de clase C^1 cuando lo sea en todos los puntos de su conjunto de definición (que necesariamente será abierto). Notamos por $C^1(A)$ al conjunto de las funciones de clase C^1 en el abierto A.

Queda claro que si $N \geq 2$ se deriva en puntos interiores, lo que por comodidad, no siempre se resaltará en adelante.

Notas 3.13.

1. Resaltamos que los conceptos de derivabilidad y de derivada de una función en un punto involucran sólamente las topologías de $\mathbb{R}^N$ y $\mathbb{R}^M$ y no las normas concretas elegidas. Es decir, son conceptos algebraico-topológicos.

2. Obsérvese que la Proposición 3.11 tiene ahora la siguiente lectura para puntos interiores: Sean $A \subset \mathbb{R}$, $a \in \overset{\circ}{A}$, $M \in \mathbb{N}$ y $f = (f_1, \ldots, f_M) : A \to \mathbb{R}^M$ una función. Entonces las siguientes afirmaciones son equivalentes:

   i) f es derivable elementalmente en a.
   ii) $f_1, \ldots, f_M$ son derivables en a.
   iii) f es derivable en a.

   Además en el caso de que f sea derivable en a, la relación entre ambas derivadas viene dada por
   $$Df(a)(x) = xf'(a), \ \forall x \in \mathbb{R},$$
   mientras que la función afín g viene dada por
   $$g(x) = f'(a)x + (f(a) - af'(a))$$
   (nótese que $Df(a)$ es la aplicación lineal asociada a g). En particular, se obtiene que
   $$Df(a)(1) = f'(a) = (f_1'(a), ..., f_M'(a)).$$

3. Si f es derivable en el punto a, entonces f es continua en a.

De acuerdo con la definición, el estudio de la derivabilidad de una función en un punto conlleva dos problemas: el conocimiento de T dado por (*) y la verificación de (**).

Fijado un vector no nulo x, se dice que la función f es derivable en a en la dirección de x si existe
$$\lim_{t \to 0} \frac{f(a + tx) - f(a)}{t},$$
en cuyo caso el valor de tal límite (que es un vector de $\mathbb{R}^M$) se nota $f'(a; x)$ y se denomina la <u>*derivada de f en a en la dirección de x*</u>. Para $x = 0$ el anterior límite tiene sentido y vale cero, por lo que es natural definir también $f'(a; 0) = 0$. La condición (*) se traduce ahora diciendo que la condición necesaria para que f sea derivable en a es la existencia de $f'(a; .)$, o sea deben existir todas las derivadas direccionales de f en a, siendo en tal caso la candidata a derivada de f en a la aplicación T dada por $T(x) = f'(a; x)$ para todo $x \in \mathbb{R}^N$ (supuesto que sea lineal). Es inmediato (¡Hágase!) que si la función f es derivable en a en las direcciones de los vectores de la esfera unidad, también lo es en la dirección de cualquier vector x, y en tal caso
$$f'(a; x) = \|x\| f'\Big(a, \frac{x}{\|x\|}\Big), \ \forall x \in \mathbb{R}^N \setminus \{0\}.$$

3.3. Campos escalares derivables. Vector gradiente.

Proposición 3.14 (Reducción de la derivabilidad a campos escalares). *Sea $A \subset \mathbb{R}^N$, y $f = (f_1, \dots, f_M) : A \to \mathbb{R}^M$ un campo vectorial. Entonces f es derivable en $a \in \overset{\circ}{A}$ si, y sólo si, cada campo escalar componente es derivable en a, en cuyo caso*

$$Df(a)(x) = (Df_1(a)(x), ..., Df_M(a)(x)),\ \forall x \in \mathbb{R}^n.$$

Además $f \in C^1(a)$ si, y sólo si, $f_i \in C^1(a)$ para $i = 1, \dots, M$.

Demostración:
El enunciado es consecuencia de que para una aplicación lineal $T = (T_1, ..., T_M)$ de $\mathbb{R}^N$ en $\mathbb{R}^M$ se verifica para cada $x \in A \backslash \{a\}$

$$\frac{f(x) - f(a) - T(x-a)}{\|x-a\|} =$$

$$= \left(\frac{f_1(x) - f_1(a) - T_1(x-a)}{\|x-a\|}, ..., \frac{f_M(x) - f_M(a) - T_M(x-a)}{\|x-a\|} \right)$$

y de la Proposición 2.62. De la misma proposición se deduce también que la continuidad en a de la aplicación Df se reduce a la continuidad de las aplicaciones $Df_1, ..., Df_M$ en a, ya que son las componentes de la aplicación Df. ■

En vista del resultado anterior, en el resto de la sección, estudiaremos los campos escalares derivables. Salvo mención expresa en contra, los resultados que se obtengan son válidos también para campos vectoriales.

Definición 3.15 (Derivadas parciales. Vector gradiente). Sea $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Para cada $1 \leq i \leq N$, supongamos que a_i es un punto de acumulación del conjunto

$$A_i := \{x_i \in \mathbb{R} : \ (a_1, \dots a_{i-1}, x_i, a_{i+1}, \dots, a_N) \in A\}.$$

Si la función de A_i en $\mathbb{R}$ definida por $x_i \mapsto f(a_1, \dots a_{i-1}, x_i, a_{i+1}, \dots, a_N)$ es derivable en a_i, se dice que f tiene *derivada parcial* respecto de la variable i-ésima en el punto a. En tal caso el valor del límite

$$\lim_{x_i \to a_i} \frac{f(a_1, \dots a_{i-1}, x_i, a_{i+1}, \dots, a_N) - f(a)}{x_i - a_i}$$

se denomina la derivada parcial de f respecto de la variable i-ésima en el punto a y se nota $D_i f(a)$ (o también $\frac{\partial f}{\partial x_i}(a)$).

Es consecuencia de la definición que el cálculo de la derivada parcial respecto de la variable i-ésima del campo escalar f en un punto genérico $x = (x_1, ..., x_N)$ se ha de llevar a cabo derivando la función real de variable real que resulta al considerar constantes las variables x_j $(j \neq i)$ y por tanto las reglas de derivación ordinarias se podrán utilizar.

Si $A^i \subset A$ es el conjunto de puntos donde f tiene derivada parcial respecto de la variable i-ésima, entonces el campo escalar en A^i definido por $x \mapsto D_i f(x)$ se denomina la *aplicación derivada parcial* de f respecto de la variable i-ésima y se nota $D_i f$ (o también $\frac{\partial f}{\partial x_i}$).

Se dice que el campo escalar f tiene *gradiente* en el punto a si admite derivadas parciales en a con respecto de todas las variables, en cuyo caso definimos el *vector gradiente* de f en a por:

$$\nabla f(a) := (D_1 f(a), ..., D_N f(a)) \in \mathbb{R}^N.$$

Si C es el conjunto de puntos de A donde f tiene gradiente, entonces el campo vectorial en C definido por $x \to \nabla f(x)$ se denomina *aplicación gradiente* de f y se nota ∇f.

Proposición 3.16 (Condición necesaria de deriv. y candidata a derivada). *Sean $A \subset \mathbb{R}^N$ y f un campo escalar en A derivable en un punto a. Entonces f tiene gradiente en a con*

$$D_i f(a) = Df(a)(e_i), \ \forall i \in \{1, \ldots, N\},$$

donde $\{e_1, \ldots, e_N\}$ es la base canónica de $\mathbb{R}^N$.
En consecuencia,

$$Df(a)(x) = (\nabla f(a) \mid x), \ \forall x \in \mathbb{R}^N.$$

Demostración:

Para $i \in \{1, \ldots, N\}$ se tiene que

$$Df(a)(e_i) = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{x_i \to a_i} \frac{f(a + (x_i - a_i)e_i) - f(a)}{x_i - a_i} =$$

$$\lim_{x_i \to a_i} \frac{f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_N) - f(a)}{x_i - a_i} = D_i f(a),$$

donde se ha utilizado (*) de la sección anterior. El resto es consecuencia de la linealidad de $Df(a)$.

■

Notas 3.17.

1. La condición anterior no es suficiente. El campo escalar en $\mathbb{R}^2$

$$f(x, y) = \frac{x^6}{(y - x^2)^2 + x^6} \quad (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

tiene gradiente en $(0, 0)$ igual a $(0, 0)$ y, sin embargo, no es ni tan siquiera continuo en $(0, 0)$ (tómese $y = x^2$).

2. Supongamos que el campo escalar f es derivable en a y tiene derivada no nula. Puesto que

$$Df(a)(x) = (\nabla f(a) \mid x), \quad \forall x \in \mathbb{R}^N$$

se deduce de la desigualdad de Cauchy-Schwarz (Apéndice A) que

$$|Df(a)(x)| \leq \left\|\nabla f(a)\right\|_2 \|x\|_2$$

y que, para $x \in \mathbb{R}^N \backslash \{0\}$, la igualdad se alcanza si, y sólo si, existe $\lambda \in \mathbb{R}^*$ tal que $x = \lambda \nabla f(a)$, de donde se deduce que la derivada de f en el punto a es máxima en la dirección dada por el vector gradiente.

3. Obsérvese que las derivadas parciales no son otra cosa que las derivadas direccionales según los vectores de la base canónica. Es claro que pueden existir las derivadas parciales y no existir una derivada direccional (véase el Ejercicio 3.4).

Proposición 3.18 (Condición suficiente de derivabilidad). *Sean $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$ y f un campo escalar en A. Supongamos que ∇f existe en un entorno de a y que ∇f es continuo en a. Entonces f es derivable en a.*

Demostración:

Dado $\varepsilon > 0$, por hipótesis, existe $\delta > 0$ tal que

$$\|x - a\|_1 < \delta \Rightarrow \begin{cases} x \in A \\ \exists \nabla f(x) \\ |D_i f(x) - D_i f(a)| \leq \varepsilon, \quad \text{para} \;\; i = 1, \ldots, N. \end{cases}$$

Para $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ tal que $0 < \|x - a\|_1 < \delta$ podemos escribir

$$f(x) - f(a) - (\nabla f(a) \; |x - a) =$$

$$\sum_{i=1}^{N} \big(f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) - D_i f(a)(x_i - a_i) \big)$$

luego, si probamos que para cada $i \in \{1, \ldots, N\}$ se verifica que

$$|f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) - D_i f(a)(x_i - a_i)| \leq$$

$$\leq \varepsilon |x_i - a_i|,$$

entonces tendremos que

$$|f(x) - f(a) - (\nabla f(a)|x - a)| \leq \sum_{i=1}^{N} \varepsilon |x_i - a_i| = \varepsilon \|x - a\|_1$$

y en consecuencia f será derivable en a.

Veamos que son ciertas las citadas desigualdades. Si $x_i = a_i$, la desigualdad es obvia. Supuesto $x_i \neq a_i$, podemos aplicar el Teorema del valor medio para encontrar y_i entre x_i y a_i tal que

$$f(a_1, \ldots, a_{i-1}, x_i, x_{i+1}, \ldots, x_N) - f(a_1, \ldots, a_{i-1}, a_i, x_{i+1}, \ldots, x_N) =$$

$$D_i f(a_1, \ldots, a_{i-1}, y_i, x_{i+1}, \ldots, x_N)(x_i - a_i),$$

de donde se deduce la desigualdad anunciada a partir de las condiciones impuestas a δ.

$\blacksquare$

Ejemplo 3.19. La condición suficiente del anterior resultado no es necesaria. En efecto, el campo escalar $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definido por

$$f(x, y) = (x^2 + y^2) \ \operatorname{sen}\left(\frac{1}{\sqrt{x^2 + y^2}}\right), \ \forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}, \qquad f(0, 0) = 0,$$

es derivable en $(0, 0)$. En efecto, es inmediato que

$$D_1 f(0, 0) = D_2 f(0, 0) = 0$$

y que

$$\frac{f(x, y)}{\sqrt{x^2 + y^2}} \rightarrow 0, \quad \text{cuando} \quad (x, y) \rightarrow (0, 0).$$

Sin embargo ∇f no es continuo en $(0, 0)$. En efecto, como para $(x, y) \neq (0, 0)$ se tiene

$$D_1 f(x, y) = 2x \ \operatorname{sen}\left(\frac{1}{\sqrt{x^2 + y^2}}\right) - \frac{x}{\sqrt{x^2 + y^2}} \ \cos\left(\frac{1}{\sqrt{x^2 + y^2}}\right),$$

basta observar que

$$D_1 f\left(\frac{1}{n\pi}, 0\right) = \frac{2}{n\pi} \ \operatorname{sen}\,(n\pi) - \ \cos\,(n\pi) = (-1)^{n+1}, \ \forall n \in \mathbb{N}$$

para concluir que $D_1 f$ no es continua en $(0, 0)$. Como $f(x, y) = f(y, x)$, para cualquier (x, y), igual le ocurre a $D_2 f$.

Probaremos enseguida que un campo escalar es de clase C^1 si, y sólo si, tiene gradiente continuo. Sin embargo, no hemos obtenido ninguna caracterización de la derivabilidad de un campo escalar en términos del gradiente. A continuación resumimos toda la información para el estudio de la derivabilidad de campos escalares.

Nota 3.20 (Estudio de la derivabilidad de campos escalares). Sean $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$ y f un campo escalar es A. Para el estudio de la derivabilidad de f en a se sugiere seguir los siguientes pasos:

1. Existencia de $\nabla f(a)$. Si no existe $\nabla f(a)$ entonces f no es derivable en a. En caso contrario, la candidata a derivada es la aplicación $x \mapsto (\nabla f(a) \mid x)$.

2. Continuidad de f en a. Si f no es continua en a, entonces no es derivable en a. De ser f continua en a, entonces f *puede ser no derivable* en a (véase la función g del Ejercicio 3.3, que es continua, con gradiente y no es derivable en $(0,0)$).

3. Continuidad de las derivadas parciales. Si ∇f es continuo en a entonces f es derivable en a. No se puede deducir de esto que f es de clase $\mathcal{C}^1$ en a, salvo que se globalice como se prueba en el teorema siguiente.

4. Definición de derivabilidad. Según que la expresión
$$\frac{f(x) - f(a) - (\nabla f(a) \mid x - a)}{\|x - a\|}$$
tienda a cero o no, cuando $x \to a$, f es derivable o no.

Teorema 3.21 (Caracterización de los campos escalares de clase $\mathcal{C}^1$). *Sean A un abierto de $\mathbb{R}^N$ y f un campo escalar en A. Las siguientes afirmaciones son equivalentes:*

i) $f \in \mathcal{C}^1(A)$.

ii) $\nabla f \in \mathcal{C}(A)$, *esto es, f tiene gradiente en cada punto de A y el campo vectorial ∇f es continuo.*

Demostración:

i) $\Rightarrow$ ii) Puesto que f es derivable, por la Proposición 3.16 se sigue que f tiene gradiente en cada punto de A, y además para cada $1 \le i \le N$, se tiene
$$D_i f(a) = Df(a)(e_i), \ \forall a \in A,$$
luego
$$D_i f = E_{e_i} \circ Df,$$
donde $E_{e_i} : \mathcal{L}(\mathbb{R}^N, \mathbb{R}) \longrightarrow \mathbb{R}$ es la evaluación en el vector e_i, esto es,
$$E_{e_i}(T) = T(e_i), \ \forall T \in L(\mathbb{R}^N, \mathbb{R}),$$
aplicación que es lineal y (automáticamente) continua. En vista de la regla de la cadena para funciones continuas, obtenemos que $D_i f$ es continua. Finalmente, como las funciones $D_i f$ son las componentes de ∇f, concluimos que ∇f es continuo a.

ii) $\Rightarrow$ i) Por la Proposición 3.18, f es derivable. Además se tiene que
$$Df(a) = D_1 f(a)\pi_1 + \cdots + D_N f(a)\pi_N,$$
donde π_k es la proyección k-ésima de $\mathbb{R}^N$, para $1 \le k \le N$.
Equivalentemente,
$$Df = \alpha_1 \circ D_1 f + \cdots + \alpha_N \circ D_N f,$$
donde, para cada $i = 1, 2, \cdots, N$, $\alpha_i : \mathbb{R} \longrightarrow \mathcal{L}(\mathbb{R}^N, \mathbb{R})$ es la aplicación lineal (continua) definida por $\alpha_i(t) = t\pi_i$. Basta, en consecuencia, utilizar la regla de la cadena para funciones continuas y que la suma de funciones continuas es continua para concluir que $Df : A \longrightarrow \mathcal{L}(\mathbb{R}^N, \mathbb{R})$ es continua. ∎

3.4. Campos vectoriales derivables. Matriz jacobiana.

Definición 3.22 (matriz jacobiana). Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Se dice que f tiene matriz jacobiana en el punto a si cada uno de los campos escalares componentes tiene gradiente en dicho punto, en cuyo caso se define la *matriz jacobiana* de f en a por

$$J_f(a) := \begin{pmatrix} D_1f_1(a) & ... & D_jf_1(a) & ... & D_Nf_1(a) \\ ... & ... & ... & ... & ... \\ D_1f_i(a) & ... & D_jf_i(a) & ... & D_Nf_i(a) \\ ... & ... & ... & ... & ... \\ D_1f_M(a) & ... & D_jf_M(a) & ... & D_Nf_M(a) \end{pmatrix} = (D_jf_i(a))_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}} .$$

En el caso $M = N$ al determinante de la matriz jacobiana se le denomina determinante jacobiano.

Obsérvese que los N números que componen la fila i-ésima de la matriz jacobiana son las componentes del vector $\nabla f_i(a)$, en consecuencia, en virtud de las Proposición 3.14 y 3.18, si el campo vectorial admite jacobiano continuo en el punto a, entonces es derivable en dicho punto y, en virtud de la Proposición 3.14 y del Teorema 3.21, si A es abierto y el campo escalar admite jacobiano continuo, entonces es de clase $\mathcal{C}^1$. La condición necesaria de derivabilidad y la candidata a derivada se codifican en el siguiente resultado.

Proposición 3.23. *Sean $A \subset \mathbb{R}^N$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Si f es derivable en $a \in \overset{\circ}{A}$, entonces f tiene matriz jacobiana en a y*

$$(Df(a)(x))^t = J_f(a)x^t, \ \forall x \in \mathbb{R}^N.$$

Demostración:

Los campos escalares componentes de f son derivables en a. En consecuencia, en virtud de la Proposición 3.16, dichos campos escalares tienen gradiente en a, es decir f tiene matriz jacobiana en a. Se tiene además que

$$A_{Df(a)} = (Df_i(a)(e_j))_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}} = (D_jf_i(a))_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}} ,$$

donde se ha utilizado la Proposición 3.14. El resto es consecuencia de la igualdad (3.1.2) de la sección 3.1.

∎

Ejemplos 3.24.

1) Toda función constante de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathcal{C}^1$ con

$$Df(a) = 0, \quad \forall a \in \mathbb{R}^N.$$

2) Toda aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathcal{C}^1$ con

$$DT(a) = T, \quad \forall a \in \mathbb{R}^N.$$

3) Toda aplicación bilineal T de $\mathbb{R}^M \times \mathbb{R}^N$ en $\mathbb{R}^P$ (es decir, lineal en ambas variables) es de clase $\mathcal{C}^1$ con derivada dada por

$$DT(a,b)(x,y) = T(x,b) + T(a,y), \ \forall (a,b),(x,y) \in \mathbb{R}^M \times \mathbb{R}^N.$$

Calculemos en primer lugar las derivadas de T en las direcciones de la base canónica. Para ello si $\{e_1, \ldots, e_M\}$ es la base canónica de $\mathbb{R}^M$, y $\{f_1, \ldots, f_N\}$ es la base canónica de $\mathbb{R}^N$, entonces $\{(e_1,0), \ldots, (e_M,0), (0,f_1), \ldots, (0,f_N)\}$ es la base canónica de $\mathbb{R}^M \times \mathbb{R}^N$. Claramente se tiene

$$\frac{T((a,b) + t(e_i,0)) - T(a,b)}{t} = \frac{T((a + te_i, b)) - T(a,b)}{t} =$$
$$\frac{T(a,b) + tT(e_i,b)) - T(a,b)}{t} = T(e_i,b) \quad (i = 1, \ldots, M),$$

es decir,

$$T'((a,b);(e_i,0)) = T(e_i,b) \quad (i = 1, \ldots, M)$$

Haciendo el mismo tipo de cálculo obtenemos que

$$T'((a,b);(0,f_j)) = T(a,f_j) \quad (j = 1, \ldots, N).$$

Hemos probado que

$$J_T(a,b) = \begin{pmatrix} T_1(e_1,b) & \ldots & T_1(e_M,b) & T_1(a,f_1) & \ldots & T_1(a,f_N) \\ T_2(e_1,b) & \ldots & T_2(e_M,b) & T_2(a,f_1) & \ldots & T_2(a,f_N) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ T_P(e_1,b) & \ldots & T_P(e_M,b) & T_P(a,f_1) & \ldots & T_P(a,f_N) \end{pmatrix}.$$

Por tanto de ser T derivable en (a,b), su derivada en (a,b) habría de ser, en virtud de la Proposición 3.23,

$$(x,y) \longmapsto \left(J_T(a,b)(x,y)^t\right)^t = T(x,b) + T(a,y).$$

Como la matriz jacobiana es continua, basta tener en cuenta el comentario que sigue a la Definición 3.22, para concluir que $T \in \mathcal{C}^1(\mathbb{R}^M \times \mathbb{R}^N)$.

Otra forma:

Una aplicación bilineal T verifica que

$$\|T(x,y)\| \leq \|T\|\|x\|\|y\|, \ \forall (x,y) \in \mathbb{R}^M \times \mathbb{R}^N,$$

donde

$$\|T\| = \max\{\|T(x,y)\| : \ \|x\| = \|y\| = 1\}$$

(¡Pruébese!). Así

$$\frac{\|T(x,y) - T(a,b) - T(x-a,b) - T(a,y-b)\|}{\|(x-a,y-b)\|} = \frac{\|T(x-a,y-b)\|}{\|(x-a,y-b)\|} \leq$$
$$\frac{\|T\|\|x-a\|\|y-b\|}{\|(x-a,y-b)\|} \leq \frac{\|T\|\|(x-a,y-b)\|^2}{\|(x-a,y-b)\|} = \|T\|\|(x-a,y-b)\|,$$

donde se ha utilizado la desigualdad comentada. Basta tomar límite en (a,b) para concluir que T es derivable en (a,b) y su derivada es la aplicación que se ha enunciado antes. La aplicación $DT : R^{M+N} \equiv \mathbb{R}^M \times \mathbb{R}^N \longrightarrow \mathcal{L}(\mathbb{R}^M \times \mathbb{R}^N, \mathbb{R}^P)$ es claramente lineal (por la bilinealidad de T), por tanto hemos probado que $T \in \mathcal{C}^1(\mathbb{R}^M \times \mathbb{R}^N)$.

4) La aplicación suma $\sigma : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ dada por

$$\sigma(a,b) = a + b, \quad \forall a,b \in \mathbb{R}^N,$$

y la aplicación producto por escalares dada por

$$\pi(\lambda, a) = \lambda a, \ \ \forall \lambda \in \mathbb{R}, a \in \mathbb{R}^N,$$

son de clase C^1, por ser, en el primer caso lineal, y en el segundo una aplicación bilineal. Por tanto, se tiene

$$D\sigma(a,b) = \sigma, \ \forall a,b \in \mathbb{R}^N, \qquad D\pi(\alpha,a)(\lambda,x) = \alpha x + \lambda a, \ \forall \alpha,\lambda \in \mathbb{R}, \ \forall a,x \in \mathbb{R}^N.$$

5) La aplicación inversión $J : \text{Iso } (\mathbb{R}^N) \longrightarrow \mathcal{L}(\mathbb{R}^N)$ dada por

$$J(T) = T^{-1}, \ \forall T \in \text{Iso } (\mathbb{R}^N)$$

es de clase $\mathcal{C}^1$ con derivada definida por

$$DJ(T)(S) = -T^{-1}ST^{-1}, \ \forall S \in \mathcal{L}(\mathbb{R}^N).$$

En efecto, es claro que la aplicación que se anuncia como $DJ(T)$ es una aplicación lineal. Para $S,T \in \text{Iso } (\mathbb{R}^N)$ se verifica que

$$S^{-1} - T^{-1} - [-T^{-1}(S-T)T^{-1}] = S^{-1} - T^{-1} + T^{-1}(S-T)T^{-1} =$$

$$= S^{-1}(T-S)T^{-1} + T^{-1}(S-T)T^{-1} = (T^{-1} - S^{-1})(S-T)T^{-1}.$$

Así,

$$\frac{\left\|S^{-1}-T^{-1}-\left(-T^{-1}(S-T)T^{-1}\right)\right\|}{\|S-T\|} \leq \|T^{-1}-S^{-1}\|\,\|T^{-1}\| \to 0 \quad (S \to T),$$

donde se ha utilizado la continuidad de J.

Finalmente, veamos que

$$DJ : \text{Iso }(\mathbb{R}^N) \longrightarrow \mathcal{L}(L(\mathbb{R}^N), \mathcal{L}(\mathbb{R}^N))$$

es continua. Escribamos $DJ = \Phi \circ (J, J)$, donde $\Phi : \mathcal{L}(\mathbb{R}^N) \times \mathcal{L}(\mathbb{R}^N) \to \mathcal{L}(\mathcal{L}(\mathbb{R}^N))$ está definida por

$$\Phi(F,G)(S) = -FSG, \ \forall F,G,S \in \mathcal{L}(\mathbb{R}^N).$$

En efecto, para $T \in \text{Iso }(\mathbb{R}^N)$ y $S \in \mathcal{L}(\mathbb{R}^N)$, tenemos que

$$(\Phi \circ (J,J))(T)(S) = \Phi(T^{-1}, T^{-1})(S) = -T^{-1}ST^{-1} = DJ(T)(S).$$

Como Φ es una aplicación bilineal (entre espacios de dimensión finita), Φ es continua. De la continuidad de J se sigue la continuidad de la función vectorial (J, J), luego DJ es continua en virtud de la regla de la cadena para funciones continuas.

6) Cualquier norma $\|.\| : \mathbb{R}^N \longrightarrow \mathbb{R}$ no es derivable en 0 (lo que generaliza la no derivabilidad del valor absoluto en 0). En efecto, si x es un vector no nulo, entonces si $t \in \mathbb{R}^*$ se tiene que

$$\frac{\|tx\| - 0}{t} = \frac{|t|}{t}\|x\|,$$

función que no tiene límite cuando $t \to 0$. Luego no existe ninguna derivada direccional.

3.5. Reglas de derivación.

1) **Linealidad**.
Sean $A \subset \mathbb{R}^N$ y $f, g : A \longrightarrow \mathbb{R}^M$ funciones derivables en a y $\lambda \in \mathbb{R}$. Entonces $f + g$ y λf son derivables en a con

$$D(f+g)(a) = Df(a) + Dg(a)$$

$$D(\lambda f)(a) = \lambda Df(a).$$

Además, si $f, g \in \mathcal{C}^1(a)$, entonces $f + g, \lambda f \in \mathcal{C}^1(a)$.
La comprobación se deja como ejercicio.

2) **Regla de la cadena.**
Sean $A \subset \mathbb{R}^N, B \subset \mathbb{R}^M, f : A \longrightarrow \mathbb{R}^M$ tal que $f(A) \subset B$ y $g : B \longrightarrow \mathbb{R}^P$. Supongamos que f es derivable en a y que g es derivable en $f(a)$. Entonces la composición $h = g \circ f$ es derivable en a con

$$Dh(a) = Dg(f(a)) \circ Df(a)$$

y en consecuencia

$$J_h(a) = J_g(f(a))J_f(a).$$

Además, si $f \in \mathcal{C}^1(a), g \in \mathcal{C}^1(f(a))$, entonces $h \in \mathcal{C}^1(a)$.

Demostración:
Notemos $b = f(a)$. Puesto que f es derivable en a y g es derivable en b, existen funciones

$$r : A \longrightarrow \mathbb{R}, \quad s : B \longrightarrow \mathbb{R},$$

tales que r es continua en a con $r(a) = 0$, s es continua en b con $s(b) = 0$, y

$$\begin{cases} \|f(x) - f(a) - Df(a)(x-a)\| = r(x)\|x-a\|, & \forall x \in A \\ \\ \|g(y) - g(b) - Dg(b)(y-b)\| = s(y)\|y-b\|, & \forall y \in B. \end{cases}$$

Para $x \in A$ se tiene

$$\begin{gathered} \|h(x) - h(a) - (Dg(b) \circ Df(a))(x-a)\| = \\ \|g(f(x)) - g(b) - Dg(b)(f(x) - b) + Dg(b)(f(x) - f(a) - Df(a)(x-a))\| \leq \\ \|g(f(x)) - g(b) - Dg(b)(f(x) - b)\| + \|Dg(b)(f(x) - f(a) - Df(a)(x-a)\| \leq \\ s(f(x))\ \|f(x) - f(a)\| + \|Dg(b)\|\ r(x)\ \|x-a\| = \\ s(f(x))\ \|f(x) - f(a) - Df(a)(x-a) + Df(a)(x-a)\| + \|Dg(b)\|\ r(x)\ \|x-a\| \leq \\ s(f(x))\big[\|f(x) - f(a) - Df(a)(x-a)\| + \|Df(a)\|\|x-a\|\big] + \|Dg(b)\|r(x)\|x-a\| = \\ s(f(x))\big[\ r(x)\ \|x-a\| +\ \|Df(a)\|\|x-a\|\big] + \|Dg(b)\|\ r(x)\ \|x-a\| \end{gathered}$$

de donde se sigue el resultado, pues $\lim_{x \to a} r(x) = 0$, y, como f es continua en a, también $\lim_{x \to a} s(f(x)) = 0$.

Supuesta la continuidad de Df en a y la de Dg en b, probaremos ahora la continuidad de Dh en a. Por ser g de clase $\mathcal{C}^1$ en b, existe $\varepsilon > 0$ tal que g es derivable en $B(b, \varepsilon) \subset B$; usando que $f \in \mathcal{C}^1$ en a, existe $\delta_1 > 0$ tal que f es derivable en $B(a, \delta_1) \subset A$. La continuidad de f en a nos asegura la existencia de $\delta_2 > 0$ tal que $f(B(a, \delta_2)) \subset B(b, \varepsilon)$. Tomando $\delta = \min\{\delta_1, \delta_2\}$ tenemos, aplicando la fórmula que ya hemos probado

$$Dh(x) = Dg(f(x)) \circ Df(x), \ \forall x \in B(a, \delta).$$

En consecuencia, $Dh = \Phi \circ (Dg \circ f, Df)$, donde Φ es la aplicación bilineal de $\mathcal{L}(\mathbb{R}^M \times \mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ en $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^p)$ definida por

$$\Phi(T, S) = T \circ S, \ \ \forall (T, S) \in \mathcal{L}(\mathbb{R}^M, \mathbb{R}^p) \times \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M).$$

Así pues, Dh es continua por la regla de la cadena para funciones continuas y la regla de continuidad de las funciones vectoriales ($Dg \circ f$ es continua en a pues f es continua en a y Dg es continua en $f(a)$ y Df es continua en a). ∎

Notas 3.25.

a) Es interesante observar la forma que adopta la regla de la cadena cuando se involucran derivadas elementales.

i) En el caso $A \subset \mathbb{R} \xrightarrow{f} B \subset \mathbb{R}^M \xrightarrow{g} R^P$ se tiene que

$$(g \circ f)'(a) = Dg(f(a))(f'(a)).$$

En efecto,

$$(g \circ f)'(a) = D(g \circ f)(a)(1) = Dg(f(a))(Df(a)(1)) = Dg(f(a))(f'(a)),$$

donde se ha utilizado la Nota 3.13.2.

ii) En el caso

$$A \subset \mathbb{R} \xrightarrow{f} B \subset \mathbb{R} \xrightarrow{g} \mathbb{R}$$

se tiene que

$$(g \circ f)'(a) = f'(a)g'(f(a)),$$

y por tanto, la regla de la cadena recién enunciada generaliza la conocida para funciones de variable real. En efecto, por el caso anterior

$$(g \circ f)'(a) = Dg(f(a))(f'(a)) = f'(a)Dg(f(a))(1) = f'(a)g'(f(a)),$$

donde se ha vuelto a utilizar la Nota 3.13.2.

b) Veamos ahora la regla de la cadena para las derivadas parciales.

Las entradas de $J_h(a)$ vienen dadas por

$$D_j h_i(a) = \sum_{k=1}^{M} D_k g_i(f(a)) D_j f_k(a), \qquad (1 \leq i \leq P,\ 1 \leq j \leq N),$$

expresión que se conoce como *regla de la cadena para las derivadas parciales.*

Si consideramos el caso $P = 1$, con lo que g y h son campos escalares, la anterior expresión se escribe

$$D_j h(a) = \sum_{k=1}^{M} D_k g(f(a)) D_j f_k(a), \qquad (1 \leq j \leq N),$$

o bien, si notamos por $x_1, \cdots, x_N$ las coordenadas en $\mathbb{R}^N$, y por $y_1, \cdots, y_M$ las coordenadas en $\mathbb{R}^M$, se tendrá

$$\frac{\partial h}{\partial x_j}(a) = \sum_{k=1}^{M} \frac{\partial g}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a), \qquad (1 \leq j \leq N). \qquad (*)$$

Si se identifican las variables con las funciones, es decir,

$$z \equiv z(y_1, \cdots, y_M),$$
$$w \equiv z(y_1(x_1, \cdots, x_N), \cdots, y_M(x_1, \cdots, x_N)) = w(x_1, \cdots, x_N)$$

queda finalmente

$$\frac{\partial w}{\partial x_j}(a) = \sum_{k=1}^{M} \frac{\partial z}{\partial y_k}(f(a)) \frac{\partial y_k}{\partial x_j}(a), \qquad (1 \leq j \leq N),$$

expresión en la que se debe saber reconocer $(*)$.

Ejemplo. Sea $z = z(x, y)$ un campo escalar derivable. Si cambiamos a coordenadas polares, esto es, hacemos

$$\left\{ \begin{array}{l} x = \rho \cos \vartheta \\ y = \rho \operatorname{sen} \vartheta \end{array} \right. ,$$

entonces la función $w(\rho, \vartheta) := z(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta)$ es derivable y sus derivadas parciales vienen dadas por

$$\left\{ \begin{array}{l} \frac{\partial w}{\partial \rho} = \frac{\partial z}{\partial x}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \cos \vartheta + \frac{\partial z}{\partial y}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \operatorname{sen} \vartheta \\ \\ \frac{\partial w}{\partial \vartheta} = -\frac{\partial z}{\partial x}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \rho \operatorname{sen} \vartheta + \frac{\partial z}{\partial y}(\rho \cos \vartheta, \rho \operatorname{sen} \vartheta) \rho \cos \vartheta \end{array} \right.$$

Proposición 3.26 (Carácter local de la derivabilidad y de la derivada). *Sea $A \subset \mathbb{R}^N, a \in \overset{\circ}{A}$ y $f : A \longrightarrow \mathbb{R}^N$. Las siguientes afirmaciones son equivalentes:*

i) f es derivable en a.

ii) $f_{|U}$ es derivable en a para algún entorno $U \subset A$ del punto a.

Además, en caso de que sean ciertas las afirmaciones anteriores, entonces las derivadas de ambas funciones (f y la restricción de f a U) coinciden.

Demostración:

Es consecuencia del carácter local del límite. ■

Nota. Supongamos que el conjunto A es unión disjunta de dos subconjuntos A_1 y A_2 y $f_1 : A_1 \longrightarrow \mathbb{R}^M, f_2 : A_2 \longrightarrow \mathbb{R}^M$ son campos vectoriales tales que

$$f(x) = \begin{cases} f_1(x) & \text{si} \quad x \in A_1 \\ f_2(x) & \text{si} \quad x \in A_2 \end{cases}$$

Por el resultado anterior, la derivabilidad de f en $\overset{\circ}{A}_1$ y en $\overset{\circ}{A}_2$ no es más que la derivabilidad de f_1 en $\overset{\circ}{A}_1$ y de f_2 en $\overset{\circ}{A}_2$, respectivamente (cuyo estudio posiblemente se puede llevar a cabo mediante el uso de las reglas de derivación). En los puntos de $(\mathrm{Fr}(A_1) \cup \mathrm{Fr}(A_2)) \cap A$, posiblemente haya que hacer un estudio particular.

Proposición 3.27 (Derivación de la función inversa).

a) Sea $A \subset \mathbb{R}^N, a \in \overset{\circ}{A}$ y $f : A \longrightarrow \mathbb{R}^N$ un campo vectorial. Supongamos que f es inyectiva, derivable en a y que $f(a) \in f(\overset{\circ}{A})$, entonces

$$f^{-1} \textit{ es derivable en } f(a) \Leftrightarrow \begin{cases} \det J_f(a) \neq 0 \\ f^{-1} \textit{ es continua en } f(a) \end{cases}$$

Además, si se verifica lo anterior, se tiene

$$Df^{-1}(f(a)) = Df(a)^{-1},$$

y en consecuencia,

$$J_{f^{-1}}(f(a)) = J_f(a)^{-1}.$$

b) Sean A y B subconjuntos abiertos de $\mathbb{R}^N$ y f un homeomorfismo de A sobre B. Si $f \in \mathcal{C}^1(a)$ para algún $a \in A$ y $\det J_f(a) \neq 0$, entonces $f^{-1} \in \mathcal{C}^1(f(a))$.

Demostración:

a) $\Rightarrow$ f^{-1} es continua en $f(a)$ por ser derivable. Como

$$f \circ f^{-1} = Id_{f(A)}, \quad f^{-1} \circ f = Id_A,$$

la regla de la cadena nos da

$$Df(a) \circ Df^{-1}(f(a)) = Id_{\mathbb{R}^N}, \qquad Df^{-1}(f(a)) \circ Df(a) = Id_{\mathbb{R}^N}$$

con lo que $Df(a) \in \mathrm{Iso}\ (\mathbb{R}^N)$ y $Df^{-1}(f(a)) = Df(a)^{-1}$, por tanto $\det J_f(a) \neq 0$.

$\Leftarrow$ Al ser f derivable en a, existe una función $r : A \longrightarrow \mathbb{R}$ continua en a con $r(a) = 0$ que verifica además

$$\|f(x) - f(a) - Df(a)(x - a)\| = r(x)\ \|x - a\|,\ \forall x \in A.$$

Se tiene que

$$
\begin{aligned}
x-a &= [Df(a)]^{-1}(Df(a)(x-a)) \\
&= Df(a)^{-1}\Big(Df(a)(x-a)-(f(x)-f(a))+(f(x)-f(a)\Big) \qquad (3.5.1)\\
&= -Df(a)^{-1}\Big(f(x)-f(a)-Df(a)(x-a))\Big)+Df(a)^{-1}(f(x)-f(a).
\end{aligned}
$$

En consecuencia,

$$x-a-Df(a)^{-1}(f(x)-f(a)) = -Df(a)^{-1}(f(x)-f(a)-Df(a)(x-a))$$

y por tanto

$$\|x-a-Df(a)^{-1}(f(x)-f(a))\| \leq \|Df(a)^{-1}\|\ r(x)\ \|x-a\|. \qquad (3.5.2)$$

También se sigue de 3.5.1 que

$$\|x-a\| \leq \|Df(a)^{-1}\|\ r(x)\ \|x-a\| + \|Df(a)^{-1}\|\ \|f(x)-f(a)\| \Rightarrow$$

$$\Rightarrow\ \Big(1-r(x)\ \|Df(a)^{-1}\|\ \Big)\ \|x-a\| \leq \|Df(a)^{-1}\|\ \|f(x)-f(a)\|$$

y cuando $r(x) < \dfrac{1}{\|Df(a)^{-1}\|}$ tendremos

$$\|x-a\| \leq \frac{\|Df(a)^{-1}\|}{1-r(x)\ \|Df(a)^{-1}\|}\ \|f(x)-f(a)\|$$

y por tanto, en vista de 3.5.2,

$$\|x-a-Df(a)^{-1}(f(x)-f(a))\| \leq r(x)\frac{\|Df(a)^{-1}\|^2}{1-r(x)\ \|Df(a)^{-1}\|}\ \|f(x)-f(a)\|$$

de donde se sigue el resultado, pues $\lim_{x\to a} r(x) = 0$ (¿Por qué?).

b) Como Iso $(\mathbb{R}^N)$ es un abierto por la Proposición 3.6, la continuidad de la aplicación $x \mapsto Df(x)$ en el punto a garantiza la existencia de un abierto U de X tal que

$$a \in U \subset A \quad \text{y} \quad Df(x) \in \text{Iso } (\mathbb{R}^N),\ \forall x \in U.$$

El apartado a) asegura que f^{-1} es derivable en cada punto de $V := f(U)$ (que es un abierto de $\mathbb{R}^N$ tal que $f(a) \in V \subset B$) y

$$Df^{-1}(y) = (Df(f^{-1}(y)))^{-1},\ \forall y \in V.$$

En consecuencia, $Df^{-1} = J \circ Df \circ f^{-1}$, lo que prueba la continuidad de Df^{-1} en $f(a)$ y, por tanto, $f^{-1} \in \mathcal{C}^1(f(a))$. ∎

Notas 3.28.

1. Este resultado generaliza para puntos interiores el visto para funciones reales de una variable real. En efecto, si $I \subset \mathbb{R}$ es un intervalo y $f : I \longrightarrow \mathbb{R}$ es una función (real de variable real), $a \in \overset{\circ}{I}$ y $b = f(a)$.

   Supongamos que f es continua, inyectiva y derivable en a. Entonces b es un punto interior de $f(A)$ y las siguientes afirmaciones son equivalentes:

   i) $f'(a) \neq 0$ y f^{-1} es continua en b.

   ii) f^{-1} es derivable en b.

   Además, en caso de que se cumplan i) y ii) se tiene: $(f^{-1})'(b) = \frac{1}{f'(a)}$. En efecto,

$$Df(a) \in \text{Iso } (\mathbb{R}) \Leftrightarrow f'(a) \neq 0 \;\; \text{y} \;\; Df^{-1}(b) = Df(a)^{-1} \Leftrightarrow (f^{-1})'(b) = \frac{1}{f'(a)}.$$

2. Es bueno resaltar que la regla antes establecida no es el Teorema de la función inversa.

3.6. Interpretación geométrica del concepto de derivada. Hiperplano tangente.

Definición 3.29. Una _variedad afín_ es el trasladado por un vector de un subespacio vectorial.

Sean $A \subset \mathbb{R}^N$ y $f : A \to \mathbb{R}^M$. Si f es derivable en el punto a, entonces la variedad afín de $\mathbb{R}^N \times \mathbb{R}^M$ que pasa por el punto $(a, f(a))$ con variedad de dirección Graf $(Df(a))$, esto es

$$(a, f(a)) + \text{Graf }(Df(a)),$$

la cual coincide con la gráfica de la aproximación afín

$$g(x) = f(a) + Df(a)(x - a).$$

En efecto,

$$\begin{aligned}\{(x, g(x)) : x \in \mathbb{R}^N\} &= \{(a + (x - a), f(a) + Df(a)(x - a)) : x \in \mathbb{R}^N\} \\ &= (a, f(a)) + \text{Graf }(Df(a)).\end{aligned}$$

Dicha variedad afín de $\mathbb{R}^N \times \mathbb{R}^M$ es la "más próxima" a Graf (f) en el punto $(a, f(a))$. Se denomina _variedad afín tangente a la gráfica de f en el punto $(a, f(a))$_. Obsérvese que la aplicación $x \mapsto (x, Df(a)(x))$ de $\mathbb{R}^N$ en Graf $(Df(a))$ es un isomorfismo y en consecuencia la variedad afín tangente a la gráfica de f en el punto $(a, f(a))$ es afínmente isomorfa a $\mathbb{R}^N$.

Un subconjunto H de un espacio vectorial X es un _hiperplano vectorial_ si es un subespacio maximal, es decir, si $H \neq X$ y el único subespacio vectorial que contiene estrictamente a H es X. El Teorema de extensión de la base nos permite afirmar que los hiperplanos vectoriales son los subespacios de codimensión uno. El trasladado A por un vector a de un hiperplano vectorial H se llama un _hiperplano_ (afín), es decir, $A = a + H$. Para la caracterización de los hiperplanos vectoriales y de los hiperplanos (afines) véase el apéndice C.

Supuesto que el campo escalar f es derivable en a, la variedad afín tangente a la gráfica de f en el punto $(a, f(a))$ es el hiperplano de $\mathbb{R}^{N+1}$ dado por

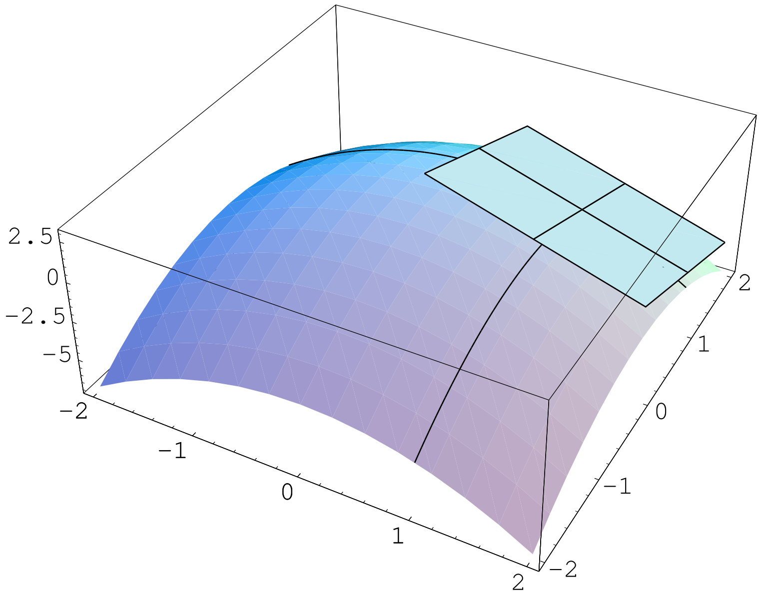
$$\{(a, f(a)) + \big(x, Df(a)(x)\big) : \ x \in \mathbb{R}^N\}.$$

En virtud de la Proposición 3.16 este hiperplano se escribe ahora

$$\{(a, f(a)) + \big(x, (\nabla f(a)|x)\big) : x \in \mathbb{R}^N\} = \Big\{\Big(x, f(a) + \big(\nabla f(a)|x - a\big)\Big) : x \in \mathbb{R}^N\Big\}.$$

Por tanto, el _hiperplano tangente_ a la gráfica del campo escalar f en el punto $(a, f(a)) = (a_1, \ldots, a_N, f(a_1, \ldots, a_N))$ es el hiperplano de ecuación

$$x_{N+1} = f(a) + D_1 f(a)(x_1 - a_1) + ... + D_N f(a)(x_N - a_N).$$



A continuación caracterizamos la variedad afín tangente a la gráfica de una función en un punto, en particular, los hiperplanos tangentes. Para ello recordamos, en primer lugar, que una *curva en $\mathbb{R}^N$* es una aplicación continua γ de un intervalo I de $\mathbb{R}$ en $\mathbb{R}^N$. Normalmente la forma más cómoda de visualizar una curva es considerar su *imagen* (en $\mathbb{R}^N$, al menos cuando $N \leq 3$) en lugar de su *gráfica* en $\mathbb{R}^{N+1}$. Supongamos que γ es derivable en un punto $t_0 \in I$ con $\gamma'(t_0) \neq 0$, entonces la recta que pasa por el punto $\gamma(t_0)$ con dirección $\gamma'(t_0)$, esto es,

$$\{\gamma(t_0) + t\gamma'(t_0) : t \in \mathbb{R}\}$$

es la *recta tangente a γ en el punto $\gamma(t_0)$.*

Claramente $\mathrm{Im}\,(\gamma) = \pi_{\mathbb{R}^N}(\mathrm{Graf}\,(\gamma))$ (donde hemos identificado de forma natural $\mathbb{R}^{N+1} = \mathbb{R} \times \mathbb{R}^N$). Si la curva se visualiza en $\mathbb{R}^N$, es natural visualizar también en $\mathbb{R}^N$ la tangente en un punto como la proyección de la recta tangente a la gráfica de γ en el punto $(t_0, \gamma(t_0))$. Esta viene dada por

$$\{(t_0, \gamma(t_0)) + (t, t\gamma'(t_0)) : t \in \mathbb{R}\} = \{(t_0, \gamma(t_0)) + t(1, \gamma'(t_0)) : t \in \mathbb{R}\}$$

y su proyección es

$$\{\gamma(t_0) + t\gamma'(t_0) : t \in \mathbb{R}\},$$

que es la recta tangente a γ en $\gamma(t_0)$ cuando $\gamma'(t_0) \neq 0$.

Seguidamente mostramos que para un campo vectorial derivable f, la variedad afín tangente a la gráfica de f en el punto $(a, f(a))$ coincide con el conjunto de todas las rectas tangentes a las curvas cuya imagen está contenida en $\mathrm{Graf}\,(f)$ y que pasan por el punto $(a, f(a))$.

Proposición 3.30. *Sea $A \subset \mathbb{R}^N$ un abierto y $f : A \longrightarrow \mathbb{R}^M$ una función derivable. Si $a \in A$ y $u \in \mathbb{R}^N \times \mathbb{R}^M$, equivalen:*

i) $u \in \text{Graf } Df(a)$.

ii) Existen $\delta > 0$ y $\gamma : [-\delta, \delta] \longrightarrow \mathbb{R}^N \times \mathbb{R}^M$ continua tal que

$$\gamma([-\delta, \delta]) \subset \text{Graf } (f), \quad \gamma(0) = (a, f(a)) \ \text{ y } \ \gamma'(0) = u.$$

Demostración:
i) $\Rightarrow$ ii) Notaremos por π_1 y π_2 a las proyecciones de $\mathbb{R}^N \times \mathbb{R}^M$ sobre $\mathbb{R}^N$ y $\mathbb{R}^M$, respectivamente. Si $u \in \text{Graf } Df(a)$, sea $\delta > 0$ tal que

$$[a - \delta\pi_1(u), a + \delta\pi_1(u)] \subset A$$

y consideremos la curva $\gamma : [-\delta, \delta] \longrightarrow \mathbb{R}^N \times \mathbb{R}^M$ definida por

$$\gamma(t) = (a + t\pi_1(u), f(a + t\pi_1(u))).$$

Es claro que $\gamma([-\delta, \delta]) \subset \text{Graf } (f)$, $\gamma(0) = (a, f(a))$ y

$$\gamma'(0) = (\pi_1(u), Df(a)(\pi_1(u))),$$

donde se han utilizado las reglas elementales de derivación. Por otra parte, como $u \in \text{Graf } (Df(a))$, entonces $u = \big(\pi_1(u),\ Df(a)(\pi_1(u))\big)$ y en consecuencia $\gamma'(0) = u$.

ii) $\Rightarrow$ i) Supongamos ahora que γ es una curva verificando las condiciones requeridas en ii). Entonces, como $\text{Im } \gamma \subset \text{Graf } f$, se verifica que

$$f \circ \pi_1 \circ \gamma = \pi_2 \circ \gamma$$

y las reglas elementales de derivación nos aseguran que

$$Df(a)(\pi_1(\gamma'(0))) = \pi_2(\gamma'(0)),$$

esto es, $u = \gamma'(0) \in \text{Graf } (Df(a))$. ∎

3.7. Apéndice A) Desigualdad de Cauchy-Schwarz.

Proposición.

$$|(x|y)| \leq \|x\|_2 \; \|y\|_2, \; \forall x, y \in \mathbb{R}^N,$$

donde $(\cdot|\cdot)$ *denota el producto escalar en* $\mathbb{R}^N$*, esto es*

$$(x|y) := \sum_{k=1}^{N} x_k y_k,$$

y la igualdad ocurre si, y sólo si, los vectores son linealmente dependientes.

Demostración:

Sean $x, y \in \mathbb{R}^N$. Si $y = 0$ se da la igualdad y la condición (¿Por qué?). En caso contrario, la desigualdad de Cauchy-Schwarz equivale a probar

$$0 \leq \frac{(x|x)(y|y) - (x|y)^2}{(y|y)},$$

desigualdad que probamos a continuación. En efecto

$$\begin{aligned}
\frac{(x|x)(y|y) - (x|y)^2}{(y|y)} &= (x|x) - \frac{(x|y)}{(y|y)}(x|y) \\
&= (x|x) - 2\frac{(x|y)}{(y|y)}(x|y) + \frac{(x|y)}{(y|y)}(x|y) \\
&= (x|x) - 2\frac{(x|y)}{(y|y)}(x|y) + \left(\frac{(x|y)}{(y|y)}\right)^2 (y|y) \\
&= \left(x - \frac{(x|y)}{(y|y)}y \;\middle|\; x - \frac{(x|y)}{(y|y)}y\right) \geq 0
\end{aligned}$$

Finalmente, si por ejemplo $y = \alpha x$ para conveniente $\alpha \in \mathbb{R}$, entonces

$$(x|y)^2 = (x|\alpha x)^2 = \alpha^2 \|x\|_2^4 = \|x\|_2^2 \; \|\alpha x\|_2^2 = \|x\|_2^2 \; \|y\|_2^2.$$

■

3.8. Apéndice B) Normas duales.

Proposición. *Si consideramos en $\mathbb{R}^N$ la norma $\|.\|_1$ (resp. $\|.\|_2$, $\|.\|_\infty$), entonces la norma de operadores en $\mathcal{L}(\mathbb{R}^N,\mathbb{R}) \equiv \mathbb{R}^N$ es $\|.\|_\infty$ (resp. $\|.\|_2$, $\|.\|_1$).*

Demostración:

Sabemos que los espacios vectoriales $\mathcal{L}(\mathbb{R}^N,\mathbb{R})$ y $\mathbb{R}^N$ son matemáticamente indistinguibles vía la correspondencia que a cada $a \in \mathbb{R}^N$ le asigna la forma lineal

$$x \longmapsto (a|x).$$

1. El dual de $\left(\mathbb{R}^N, \|.\|_1\right)$ es isométricamente isomorfo a $\left(\mathbb{R}^N, \|.\|_\infty\right)$.

   Se tiene que
   $$|(a|x)| \leq \|a\|_\infty \; \|x\|_1 \;\; (a, x \in \mathbb{R}^N),$$
   y en consecuencia
   $$\|a\| \leq \|a\|_\infty.$$
   Por otra parte si $k \in \{1, ..., N\}$ es tal que $\|a\|_\infty = |a_k|$, tomado $x_0 = e_k$ se tiene también que
   $$\|a\| \geq |(a|x_0)| = |a_k| = \|a\|_\infty.$$
   Hemos probado que $\|a\| = \|a\|_\infty$ y que la norma dual se alcanza en x_0.

2. El dual de $\left(\mathbb{R}^N, \|.\|_2\right)$ es isométricamente isomorfo a $\left(\mathbb{R}^N, \|.\|_2\right)$.

   La prueba es análoga al caso anterior. La primera desigualdad es en este caso la desigualdad de Cauchy-Schwarz y el punto donde se alcanza la norma dual es por ejemplo
   $$x_0 = \frac{a}{\|a\|_2}.$$

3. El dual de $\left(\mathbb{R}^N, \|.\|_\infty\right)$ es isométricamente isomorfo a $\left(\mathbb{R}^N, \|.\|_1\right)$.

   La prueba es también análoga al primer caso siendo ahora x_0 (vector donde se alcanza la norma dual) cualquier vector que verifique
   $$\big|x_0(k)\big| = 1, \quad a_k \; x_0(k) \geq 0 \;, \; k \in \{1, ..., N\}.$$

■

3.9. Apéndice C) Hiperplanos.

Proposición. *Sea H un subconjunto de $\mathbb{R}^N$. Equivalen las siguientes afirmaciones:*

i) H es un hiperplano vectorial.

ii) H es el núcleo de una forma lineal no nula.

Además, dos formas lineales no nulas determinan el mismo hiperplano vectorial, si y sólo si, son proporcionales.

Demostración:

$i) \Rightarrow ii)$ Sea $e \in \mathbb{R}^N \setminus H$. Se tiene que $\mathbb{R}^N = H \bigoplus < e >$, es decir:

$$\forall x \in \mathbb{R}^N, \exists ! h \in H,\ \exists ! \lambda \in \mathbb{R} :\ x = h + \lambda e.$$

La aplicación $f : \mathbb{R}^N \to \mathbb{R}$ definida por

$$f(x) = \lambda \ \text{ si } \ x = h + \lambda e \ \ (h \in H, \lambda \in \mathbb{R})$$

es un funcional lineal no nulo cuyo núcleo es H.

$ii) \Rightarrow i)$ Sea f un funcional no nulo tal que $H = Ker(f)$. Se tiene que H es un subespacio vectorial de $\mathbb{R}^N$. Probemos que H es maximal. Supongamos que S es un subespacio vectorial de $\mathbb{R}^N$ que contenga estrictamente a H. Fijado $y \in S \setminus H$, se tiene para cada vector $x \in \mathbb{R}^N$ que

$$x = \left(x - f(x) \frac{y}{f(y)} \right) + f(x) \frac{y}{f(y)} \ \in S$$

ya que $x - f(x)\frac{y}{f(y)} \ \in H$ y $f(x)\frac{y}{f(y)} \ \in S$. Hemos probado que $S = \mathbb{R}^N$.

Por último, sean f, g funcionales lineales no nulos. Si existe un real λ tal que $g = \lambda f$, es claro que $Ker(f) = Ker(g)$. Recíprocamente, si $Ker(f) = Ker(g)$, dado un vector $a \in \mathbb{R}^N \setminus Ker(f)$ se tiene que $g = \frac{g(a)}{f(a)}\ f$ ya que $x - \frac{f(x)}{f(a)}\ a \ \in Ker(f), \forall x \in \mathbb{R}^N$, y por tanto

$$0 = g(x) - \frac{f(x)}{f(a)} g(a),\ \forall x \in \mathbb{R}^N.$$

∎

Corolario. *Un subconjunto A de $\mathbb{R}^N$ es un hiperplano, si y sólo si, existen un funcional no nulo f y un número real α tales que*

$$A = \{x \in \mathbb{R}^N :\ f(x) = \alpha\},$$

es decir, existen $A_1, ..., A_N, \alpha \in \mathbb{R}$ tales que

$$A = \{x \in \mathbb{R}^N :\ A_1 x_1 + ... + A_N x_N = \alpha\}.$$

Demostración:

Supongamos que $A = a + H$ para convenientes $a \in \mathbb{R}^N$ y H hiperplano vectorial. Si f es un funcional no nulo tal que $H = Ker(f)$, entonces se verifica que

$$A = \{x \in \mathbb{R}^N : \ f(x) = f(a)\}.$$

Recíprocamente, supongamos que $A = \{x \in \mathbb{R}^N : \ f(x) = \alpha\}$ para conveniente $\alpha \in \mathbb{R}$. Puesto que f es sobreyectiva (¡Hágase!), podemos elegir $a \in \mathbb{R}^N$ tal que $f(a) = \alpha$. Se tiene entonces que

$$A = f^{-1}(\alpha) = \{x \in \mathbb{R}^N : \ f(x) = \alpha\} = \{x \in \mathbb{R}^N : \ x - a \in Ker(f)\} = a + Ker(f).$$

■

Proposición. *Sean f un funcional lineal no nulo y x_0 un vector de $\mathbb{R}^N$, entonces*

$$\operatorname{dist}\Big(x_0, f^{-1}(\alpha)\Big) = \frac{|f(x_0) - \alpha|}{\|f\|}.$$

En consecuencia, si el funcional f viene dado por

$$f(x) = A_1 x_1 + ... + A_N x_N, \ \forall x \in \mathbb{R}^N, (A_1^2 + ... + A_N^2 > 0)$$

y se considera en $\mathbb{R}^N$ la norma euclídea, entonces la fórmula anterior es

$$\operatorname{dist}\big(x_0, A\big) = \frac{|A_1 x_0(1) + ... + A_N x_0(N) - \alpha|}{\sqrt{A_1^2 + ... + A_N^2}}.$$

Demostración:

Basta probar que $\operatorname{dist}\big(x_0, Ker(f)\big) = \frac{|f(x_0)|}{\|f\|}$. En efecto, sea a tal que $f(a) = \alpha$, entonces, en virtud de la linealidad de f, se tiene que

$$\operatorname{dist}\Big(x_0, f^{-1}(\alpha)\Big) = \operatorname{dist}\big(x_0 - a, Ker(f)\big).$$

Al ser f lipschitziana de razón $\|f\|$, se tiene para cada $x \in Ker(f)$ que

$$|f(x_0)| = |f(x_0 - x)| \leq \|f\| \|x_0 - x\| \qquad (*)$$

de donde se deduce que

$$\frac{|f(x_0|}{\|f\|} \leq \operatorname{dist}\big(x_0, Ker(f)\big).$$

Por definición de norma de un aplicación lineal, existe una sucesión $\{x_n\}$ en la esfera unidad de $\big(\mathbb{R}^N, \|\cdot\|\big)$ tal que $\|f\| < \frac{n+1}{n} \ |f(x_n)|, \forall n \in \mathbb{N}$. Al ser $\Big\{x_0 - f(x_0) \ \frac{x_n}{f(x_n)}\Big\}$ una sucesión en $Ker(f)$, se tiene para cada n natural que

$$\operatorname{dist}\big(x_0, Ker(f)\big) \leq \Big\|x_0 - \Big(x_0 - f(x_0) \ \frac{x_n}{f(x_n)}\Big)\Big\| = \Big|\frac{f(x_0)}{f(x_n)}\Big| < \ \frac{n+1}{n} \ \frac{|f(x_0)|}{\|f\|},$$

de donde deducimos tomando límites que $\text{dist}\big(x_0, Ker(f)\big) \leq \frac{|f(x_0)|}{\|f\|}$. Hemos probado que

$$\text{dist}\big(x_0, Ker(f)\big) = \frac{|f(x_0)|}{\|f\|}.$$

Nótese que también se puede razonar que la desigualdad (*) es, de hecho, una igualdad por alcanzarse la norma de operadores.

La última expresión se deduce de la anterior sin más que recordar que la norma euclídea es autodual. ■

3.10. Referencias recomendadas.

[BoRoVe], [Cra], [FeVicd], [FeVie], [MaHo], [Jur] y [MaHo].

3.11. Resumen del resultados del Tema 3

$\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M) \equiv \mathcal{M}_{M\times N}(\mathbb{R})$.

El espacio vectorial $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ y el espacio vectorial $\mathcal{M}_{M\times N}(\mathbb{R})$ de las matrices $M \times N$ de números reales son matemáticamente indistinguibles. Para cada T en $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ definimos $A_T \in \mathcal{M}_{M\times N}(\mathbb{R})$ por

$$A_T := \Big(T_i(e_j)\Big)_{\substack{1 \le i \le M \\ 1 \le j \le N}},$$

donde $T_1, \cdots, T_M$ son los campos escalares componentes de T y $\{e_1, ..., e_N\}$ es la base canónica de $\mathbb{R}^N$. La aplicación de $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ en $\mathcal{M}_{M\times N}(\mathbb{R})$ definida por $T \to A_T$ es un isomorfismo de $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ sobre $\mathcal{M}_{M\times N}(\mathbb{R})$, cuyo inverso es la aplicación de $\mathcal{M}_{M\times N}(\mathbb{R})$ sobre $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ definida por $A \to T_A$ donde T_A es la aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}^M$ dada por

$$\Big(T_A(x)\Big)^t := Ax^t, \forall x \in \mathbb{R}^N.$$

Isomorfismo topológico. Si X e Y son espacios normados, una aplicación $T : X \longrightarrow Y$ es un <u>*isomorfismo topológico*</u> si T es una aplicación biyectiva, lineal y continua cuya inversa también es continua. Si $X = Y = \mathbb{R}^N$, toda aplicación lineal y biyectiva es un isomorfismo topológico. Notaremos $\mathrm{Iso}(\mathbb{R}^N)$ al conjunto de los isomorfismos de $\mathbb{R}^N$ en $\mathbb{R}^N$.

$$\mathrm{Iso}(\mathbb{R}^N) := \{T \in \mathcal{L}(\mathbb{R}^N) : \det A_T \neq 0\}.$$

Función lipschitziana. Sean (E, d) y (F, ρ) espacios métricos. Se dice que $f : E \to F$ es lipschitziana si existe $K \geq 0$ verificando

$$\rho(f(x), f(y)) \leq Kd(x, y), \quad \forall x, y \in E.$$

La menor constante que verifica esta desigualdad se denomina la constante de Lipschitz de T.

El espacio de Banach $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$. Norma del espacio $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$.

La función que a cada aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}^M$ le hace corresponder el número real

$$\|T\| := \max\{\|T(x)\| : \|x\| = 1\}$$

es una norma en $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$, denominada <u>norma de operadores</u>. $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ es un espacio de Banach. Además se verifica que

$$\|T(x)\| \leq \|T\|\|x\|, \forall x \in \mathbb{R}^N$$

$$\|T\| = \min\{K \geq 0 : \ \|T(x)\| \leq K\|x\|, \ \forall x \in \mathbb{R}^N\} = \max\{\|T(x)\| : \|x\| \leq 1\}.$$

La primera igualdad nos asegura que $\|T\|$ es la constante de Lipschitz de T.

Proposición. Iso $(\mathbb{R}^N)$, *el conjunto de los isomorfismos de* $\mathbb{R}^N$ *en* $\mathbb{R}^N$, *es un abierto de* $\mathcal{L}(\mathbb{R}^N)$ *y la aplicación inversión* $J : \text{Iso }(\mathbb{R}^N) \to \mathcal{L}(\mathbb{R}^N)$ *definida por*

$$J(T) := T^{-1}, \quad \forall T \in \text{Iso }(\mathbb{R}^N)$$

es continua.

Función derivable Fréchet. Sean $M, N \in \mathbb{N}$, $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$, y $f : A \to \mathbb{R}^M$ una función. Se dice que f es derivable (en el sentido de Fréchet) o diferenciable en el punto a si existe una aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ verificando

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

donde se consideran normas cualesquiera en $\mathbb{R}^N$ y $\mathbb{R}^M$. En tal caso la aplicación T es única, se denomina la derivada de la función f en el punto a o diferencial de f en a y se nota por $Df(a)$. Se dice que f es derivable en un subconjunto $B \subset A$ si es derivable en cada punto de B.

Sea $A_1 \subset A$ el conjunto de puntos donde f es derivable. La aplicación $x \to Df(x)$ de A_1 en $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ se denomina la aplicación derivada de f o la diferencial de f y se nota Df.

Se dice que f es de clase C^1 en a, y se nota $f \in C^1(a)$, si f es derivable en un entorno de a y la aplicación Df es continua en a. Se dice que f es de clase C^1 en un subconjunto $B \subset A$ si es de clase C^1 en cada punto de B.

Se dice que f es de clase C^1 cuando lo sea en todos los puntos de su conjunto de definición (que necesariamente será abierto). Notamos por $C^1(A)$ al conjunto de las funciones de clase C^1 en el abierto A.

Reducción de la derivabilidad a campos escalares. *Sean* A *un subconjunto de* $\mathbb{R}^N$ *y* $f = (f_1, \ldots, f_M) : A \to \mathbb{R}^M$ *un campo vectorial. Entonces* f *es derivable en* $a \in \overset{\circ}{A}$ *si, y sólo si, cada campo escalar componente es derivable en* a*, en cuyo caso*

$$Df(a)(x) = (Df_1(a)(x), ..., Df_M(a)(x)), \ \forall x \in \mathbb{R}^n.$$

Además $f \in C^1(a)$ *si, y sólo si,* $f_i \in C^1(a)$ *para* $i = 1, \ldots, M$.

Derivadas parciales. Vector gradiente. Sea $A \subset \mathbb{R}^N$, $a \in A$ y f un campo escalar en A. Para cada $1 \leq i \leq N$, supongamos que a_i es un punto de acumulación del conjunto

$$A_i := \{x_i \in \mathbb{R} : \ (a_1, \ldots a_{i-1}, x_i, a_{i+1}, \ldots, a_N) \in A\}.$$

Si la función de A_i en $\mathbb{R}$ definida por $x_i \to f(a_1, \ldots a_{i-1}, x_i, a_{i+1}, \ldots, a_N)$ es derivable en a_i, se dice que f tiene derivada parcial respecto de la variable i-ésima en el punto a. En tal caso el valor del límite

$$\lim_{x_i \to a_i} \frac{f(a_1, \ldots a_{i-1}, x_i, a_{i+1}, \ldots, a_N) - f(a)}{x_i - a_i}$$

se denomina la derivada parcial de f respecto de la variable i-ésima en el punto a y se nota $D_i f(a)$ (o también $\frac{\partial f}{\partial x_i}(a)$).

Es consecuencia de la definición que el cálculo de la derivada parcial respecto de la variable i-ésima del campo escalar f en un punto genérico $x = (x_1, ..., x_N)$ se ha de llevar a cabo derivando la función real de variable real que resulta al considerar constantes las variables x_j $(j \neq i)$ y por tanto las reglas de derivación ordinarias se podrán utilizar.

Si $A^i \subset A$ es el conjunto de puntos donde f tiene derivada parcial respecto de la variable i-ésima, entonces el campo escalar en A^i definido por $x \mapsto D_i f(x)$ se denomina la aplicación derivada parcial de f respecto de la variable i-ésima y se nota $D_i f$ (o también $\frac{\partial f}{\partial x_i}$).

Se dice que el campo escalar f tiene gradiente en el punto a si admite derivadas parciales en a con respecto de todas las variables, en cuyo caso definimos el vector gradiente de f en a por:

$$\nabla f(a) := (D_1 f(a), ..., D_N f(a)) \in \mathbb{R}^N.$$

Si C es el conjunto de puntos de A donde f tiene gradiente, entonces el campo vectorial en C definido por $x \to \nabla f(x)$ se denomina aplicación gradiente de f y se nota ∇f.

Estudio de la derivabilidad de campos escalares.

Sean $A \subset \mathbb{R}^N$, $a \in \overset{\circ}{A}$ y f un campo escalar en A. Para el estudio de la derivabilidad de f en a se sugiere seguir los siguientes pasos:

1. Continuidad de f en a. Si f no es continua en a, entonces no es derivable en a. De ser f continua en a, entonces f *puede no ser derivable* en a (véase la función g del Ejercicio 3.4, que es continua, tiene gradiente y no es derivable en $(0,0)$).

2. Existencia de $\nabla f(a)$. Si no existe $\nabla f(a)$, entonces f no es derivable en a. En caso contrario, la candidata a derivada es la aplicación $x \to \Big(\nabla f(a) \mid x\Big)$.

3. Continuidad de las derivadas parciales. Si ∇f es continuo en a entonces f es derivable en a. No se puede deducir de esto que f es de clase $\mathcal{C}^1$ en a, salvo que se globalice.

4. Definición de derivabilidad. Según que la expresión

$$\frac{f(x) - f(a) - \Big(\nabla f(a) \mid x - a\Big)}{\|x - a\|}$$

tienda a cero o no, cuando $x \to a$, f es derivable o no.

Matriz jacobiana. Sean $A \subset \mathbb{R}^N$, $a \in A$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Se dice que f tiene matriz jacobiana en el punto a si cada uno de los campos escalares componentes tiene gradiente en dicho punto, en cuyo caso se define la matriz jacobiana de f en a por

$$J_f(a) := \begin{pmatrix} D_1 f_1(a) & ... & D_j f_1(a) & ... & D_N f_1(a) \\ ... & ... & ... & ... & ... \\ D_1 f_i(a) & ... & D_j f_i(a) & ... & D_N f_i(a) \\ ... & ... & ... & ... & ... \\ D_1 f_M(a) & ... & D_j f_M(a) & ... & D_N f_M(a) \end{pmatrix} = (D_j f_i(a))_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}}.$$

En el caso $M = N$ al determinante de la matriz jacobiana se le denomina determinante jacobiano.

Proposición. *Sean $A \subset \mathbb{R}^N$ y $f : A \to \mathbb{R}^M$ un campo vectorial. Si f es derivable en $a \in \overset{\circ}{A}$, entonces f tiene matriz jacobiana en a y*

$$(Df(a)(x))^t = J_f(a)x^t, \ \forall x \in \mathbb{R}^N.$$

Conviene recordar:

- La derivabilidad es un concepto local.

- Toda aplicación lineal T de $\mathbb{R}^N$ en $\mathbb{R}^M$ es de clase $\mathcal{C}^1$ con

$$DT(a) = T, \quad \forall a \in \mathbb{R}^N.$$

- **Linealidad**.
  Sean $A \subset \mathbb{R}^N$ y $f, g : A \longrightarrow \mathbb{R}^M$ funciones derivables en a y $\lambda \in \mathbb{R}$. Entonces $f + g$ y λf son derivables en a con

$$D(f+g)(a) = Df(a) + Dg(a)$$
$$D(\lambda f)(a) = \lambda Df(a).$$

  Además, si $f, g \in \mathcal{C}^1(a)$, entonces $f + g, \lambda f \in \mathcal{C}^1(a)$.

- **Regla de la cadena.**
  Sean $A \subset \mathbb{R}^N, B \subset \mathbb{R}^M, f : A \longrightarrow \mathbb{R}^M$ tal que $f(A) \subset B$ y $g : B \longrightarrow \mathbb{R}^P$. Supongamos que f es derivable en a y que g es derivable en $f(a)$. Entonces la composición $h = g \circ f$ es derivable en a con

$$Dh(a) = Dg(f(a)) \circ Df(a)$$

  y en consecuencia

$$J_h(a) = J_g(f(a))J_f(a).$$

  Además, si $f \in \mathcal{C}^1(a), g \in \mathcal{C}^1(f(a))$, entonces $h \in \mathcal{C}^1(a)$.

- **Derivada de la función inversa.**
  Sea $A \subset \mathbb{R}^N, a \in \overset{\circ}{A}$ y $f : A \longrightarrow \mathbb{R}^N$ un campo vectorial. Supongamos que f es inyectiva, derivable en a y que $f(a) \in f(\overset{\circ}{A})$, entonces

$$f^{-1} \text{ es derivable en } f(a) \Leftrightarrow \begin{cases} \det J_f(a) \neq 0 \\ f^{-1} \text{ es continua en } f(a) \end{cases}$$

  Además, si se verifica lo anterior, se tiene

$$Df^{-1}(f(a)) = Df(a)^{-1},$$

  y en consecuencia,

$$J_{f^{-1}}(f(a)) = J_f(a)^{-1}.$$

  Si además A y B son subconjuntos abiertos de $\mathbb{R}^N$, f es un homeomorfismo de A sobre B y $f \in \mathcal{C}^1(a)$ para algún $a \in A$ con $\det J_f(a) \neq 0$, entonces $f^{-1} \in \mathcal{C}^1(f(a))$.

Interpretación geométrica. *Si un campo escalar f es derivable en un punto a, entonces la gráfica de f tiene un hiperplano tangente en el punto $(a, f(a))$ que viene dado por la expresión*

$$x_{N+1} = f(a) + D_1 f(a)(x_1 - a_1) + \ldots + D_N f(a)(x_N - a_N).$$

Dicho hiperplano coincide con el conjunto de todas las rectas tangentes en $(a, f(a))$ a las curvas cuya imagen está contenida en la gráfica de f y que pasan por el punto $(a, f(a))$.