Tema 8

# Vector gradiente

Como segundo caso particular de la noción de diferenciabilidad, estudiamos ahora lo que ocurre cuando el espacio normado de partida es $\mathbb{R}^N$ con $N > 1$, y el de llegada es $\mathbb{R}$. Tenemos pues una función real de $N$ variables reales, es decir, un campo escalar en $\mathbb{R}^N$. Su diferencial en un punto, cuando existe, se describe como el producto escalar por un vector de $\mathbb{R}^N$, que será el *vector gradiente*. Las coordenadas de este vector son las *derivadas parciales* de nuestra función, con respecto a cada una de las variables, en el punto considerado. Veremos también la interpretación geométrica y física del gradiente de un campo escalar.

## 8.1. Derivadas direccionales

Volvamos por un momento a la noción general de función diferenciable. Sean pues $X, Y$ espacios normados, $\Omega$ un subconjunto abierto de $X$ y $f : \Omega \to Y$ una función diferenciable en un punto $a \in \Omega$. Al probar la unicidad de la diferencial, obtuvimos de hecho que

$$Df(a)(u) = \lim_{t \to 0} \frac{f(a+tu) - f(a)}{t} \qquad \forall u \in X \tag{1}$$

Por tanto, para estudiar la diferenciabilidad de $f$ en $a$, es natural empezar por la existencia de los límites que aparecen en el segundo miembro de $(1)$. No es necesario trabajar con todos los vectores $u \in X$. Por una parte, el caso $u = 0$ es trivial; por otra, si $v = \lambda u$ con $\lambda \in \mathbb{R}^*$, los cambios de variable $t = \lambda s$ y $s = t/\lambda$ nos dicen que

$$\lim_{t \to 0} \frac{f(a+tu) - f(a)}{t} = y \in Y \iff \lim_{s \to 0} \frac{f(a+sv) - f(a)}{s} = \lambda y \tag{2}$$

Esto permite normalizar cada vector $u \in X \setminus \{0\}$, es decir, suponer que $\|u\| = 1$ sin perder generalidad. Una **dirección** en el espacio normado $X$ es un vector $u \in X$ con $\|u\| = 1$, y denotaremos por $S$ al conjunto de todas las direcciones en $X$, es decir, $S = \{u \in X : \|u\| = 1\}$. Observamos que, para cada $u \in S$, el límite que aparece a la derecha de $(1)$, cuando existe, es el vector derivada de una función de variable real, con valores en el espacio normado $Y$.

Más concretamente, fijamos $r \in \mathbb{R}^+$ tal que $B(a,r) \subset \Omega$ y, para cada $u \in S$, consideramos la función $\varphi_u$ definida de la siguiente forma:

$$\varphi_u : ]-r,r[ \to Y, \qquad \varphi_u(t) = f(a+tu) \quad \forall t \in ]-r,r[$$

Obsérvese que $\varphi_u$ describe el comportamiento de $f$ en la recta que pasa por el punto $a$ y tiene a $u$ como vector de dirección, suficientemente cerca del punto $a$.

Pues bien, dado $u \in S$, decimos que $f$ es **derivable en la dirección** $u$, en el punto $a$, cuando la función $\varphi_u$ es derivable en $0$, en cuyo caso, al vector derivada $\varphi_u'(0)$ lo llamamos **derivada direccional** de $f$ en $a$, según la dirección $u$, y lo denotamos por $f_u'(a)$. Así pues,

$$f_u'(a) = \varphi_u'(0) = \lim_{t \to 0} \frac{\varphi_u(t) - \varphi_u(0)}{t} = \lim_{t \to 0} \frac{f(a+tu) - f(a)}{t}$$

Usando la equivalencia $(2)$, con $\lambda = -1$, vemos que, siempre en el punto $a$, $f$ es derivable en la dirección $u$ si, y sólo si, lo es en la dirección $-u$, en cuyo caso se tiene $f_{-u}'(a) = -f_u'(a)$.

Decimos que $f$ es **direccionalmente derivable** en el punto $a$ cuando es derivable en todas las direcciones $u \in S$. En vista de $(1)$, tenemos la siguiente relación entre la diferenciabilidad y la derivabilidad direccional:

- *Sean $X,Y$ espacios normados, $\Omega$ un abierto de $X$ y $f : \Omega \to Y$ una función. Si $f$ es diferenciable en un punto $a \in \Omega$, entonces $f$ es direccionalmente derivable en $a$ con*

$$f_u{}'(a) = Df(a)(u) \qquad \forall u \in S$$

## 8.2. Derivadas parciales

### 8.2.1. Caso general

A partir de ahora trabajamos en el caso $X = \mathbb{R}^N$, mientras que de momento, $Y$ sigue siendo un espacio normado arbitrario. Fijamos un abierto $\Omega$ de $\mathbb{R}^N$, una función $f : \Omega \to Y$ y un punto $a \in \Omega$. A todos los efectos, de entrada para que las direcciones en $\mathbb{R}^N$ estén definidas sin ambigüedad, usaremos en $\mathbb{R}^N$ la norma euclídea, escribimos $S = \{u \in \mathbb{R}^N : \|u\| = 1\}$ y fijamos $r \in \mathbb{R}^+$ de forma que $B(a,r) \subset \Omega$. La definición de derivada direccional se aplica a las direcciones de los ejes de coordenadas, es decir, las de la base usual $\{e_1, e_2, \ldots, e_N\}$. A partir de este momento, todo lo que hacemos está ligado a dicha base.

Pues bien, dado $k \in \Delta_N$, cuando $f$ es derivable en el punto $a$, en la dirección $e_k$, decimos que $f$ es **parcialmente derivable con respecto a la $k$-ésima variable** en el punto $a$. Entonces, la derivada direccional de $f$ en $a$, según la dirección $e_k$, se denomina **derivada parcial de $f$ con respecto a la $k$-ésima variable** en el punto $a$, y se denota por $\dfrac{\partial f}{\partial x_k}(a)$, es decir,

$$\frac{\partial f}{\partial x_k}(a) = f_{e_k}'(a) = \lim_{t \to 0} \frac{f(a+te_k) - f(a)}{t} \tag{3}$$

Cuando esto ocurre para todo $k \in \Delta_N$, decimos que $f$ es **parcialmente derivable** en $a$, y entonces tenemos $N$ derivadas parciales, $\dfrac{\partial f}{\partial x_k}(a)$ con $k \in \Delta_N$, definidas por la igualdad anterior. Como esta noción es más débil que la derivabilidad direccional, se tiene:

- *Si $f$ es diferenciable en $a$, entonces $f$ es parcialmente derivable en $a$ con:*

$$\frac{\partial f}{\partial x_k}(a) = Df(a)(e_k) \qquad \forall k \in \Delta_N \tag{4}$$

### 8.2.2. Caso de un campo vectorial

Pasemos al caso particular en que $Y = \mathbb{R}^M$, es decir, tenemos un campo vectorial. Como ocurrió con la diferenciabilidad, su derivabilidad parcial equivale a la de todas sus componentes:

- *Sea $\Omega$ un abierto de $\mathbb{R}^N$, $f = (f_1, f_2, \ldots, f_M) : \Omega \to \mathbb{R}^M$ un campo vectorial y $k \in \Delta_N$. Se tiene entonces que $f$ es parcialmente derivable con respecto a la k-ésima variable en un punto $a \in \Omega$, si y sólo si, lo es $f_j$ para todo $j \in \Delta_M$, en cuyo caso:*

$$\frac{\partial f}{\partial x_k}(a) = \left( \frac{\partial f_1}{\partial x_k}(a), \frac{\partial f_2}{\partial x_k}(a), \ldots, \frac{\partial f_M}{\partial x_k}(a) \right)$$

En efecto, basta usar la definición de derivada parcial, caso particular de $(3)$, pues para $a \in \Omega$ y $k \in \Delta_M$ se tiene evidentemente

$$\lim_{t \to 0} \frac{f(a + te_k) - f(a)}{t} = y \in \mathbb{R}^M \iff \lim_{t \to 0} \frac{f_j(a + te_k) - f_j(a)}{t} = y(j) \quad \forall j \in \Delta_M \qquad \blacksquare$$

Así pues, el estudio de la derivabilidad parcial de un campo vectorial se reduce al caso de un campo escalar, en el que nos centramos a partir de ahora.

### 8.2.3. Caso de un campo escalar

Sea $\Omega$ un abierto de $\mathbb{R}^N$ y $f : \Omega \to \mathbb{R}$ un campo escalar. Fijado un punto $a \in \Omega$ y $k \in \Delta_N$, examinemos con más detalle la derivada parcial de $f$ con respecto a la $k$-ésima variable en el punto $a$. Usamos para ello las coordenadas de $a$, es decir, $a = (a_1, a_2, \ldots, a_N)$.

Para $t \in \mathbb{R}$, la $k$-ésima coordenada de $a + te_k$ es $a_k + t$, mientras las restantes coinciden con las de $a$, por lo que escribiremos $a + te_k = \big(a_1, \ldots a_{k-1},\ a_k + t\ , a_{k+1}, \ldots, a_N\big)$. Esta notación no excluye los casos $k = 1$ y $k = N$, se entiende que $a + te_1 = \big(a_1 + t\ , a_2, \ldots, a_N\big)$ mientras que $a + te_N = \big(a_1, \ldots, a_{N-1},\ a_N + t\big)$. Por definición, la derivada parcial que nos interesa, caso de que exista, viene dada por

$$\frac{\partial f}{\partial x_k}(a) = \lim_{t \to 0} \frac{f\big(a_1, \ldots, a_{k-1}\, , a_k + t\, , a_{k+1}, \ldots, a_N\big) - f(a)}{t}$$

Fijado $r \in \mathbb{R}^+$ tal que $B(a,r) \subset \Omega$, los cambios de variable $t = x_k - a_k$ y $x_k = t + a_k$, teniendo en cuenta que $-r < t < r$ si, y sólo si, $a_k - r < x_k < a_k + r$, mientras que $t \to 0$ si, y sólo si, $x_k \to a_k$, nos dicen que, para cualquier $\alpha \in \mathbb{R}$, se tiene

$$\lim_{t \to 0} \frac{f(a+te_k) - f(a)}{t} = \alpha \iff \lim_{x_k \to a_k} \frac{f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_N) - f(a)}{x_k - a_k} = \alpha$$

Considerando entonces la función $\psi : ]a_k - r, a_k + r[ \to \mathbb{R}$ definida por

$$\psi(x_k) = f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_N) \qquad \forall x_k \in ]a_k - r, a_k + r[$$

la equivalencia anterior nos dice que $f$ es parcialmente derivable con respecto a la $k$-ésima variable en el punto $a$ si, y sólo si, $\psi$ es derivable en el punto $a_k$, en cuyo caso se tiene

$$\frac{\partial f}{\partial x_k}(a) = \psi'(a_k) = \lim_{x_k \to a_k} \frac{f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_N) - f(a)}{x_k - a_k} \tag{5}$$

Nada que no supiéramos ya, el concepto de derivada parcial de un campo escalar coincide con el de derivada de una función real de variable real, pero ahora la relación entre ambas funciones se comprende mejor. Mientras $f$ depende de $N$ variables $x_1, x_2, \ldots, x_N$, vemos que la función $\psi$ de una sola variable, describe “parcialmente” el comportamiento de $f$. Sólo refleja la dependencia de $f$ respecto de la variable $x_k$, cuando por así decirlo, las demás variables se mantienen constantes, pues para obtener $\psi(x_k)$ a partir de $f(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_N)$, tomamos $x_j = a_j$ para todo $j \in \Delta_N \setminus \{k\}$. Por eso decimos que la derivada de $\psi$ en $a_k$ es sólo una “derivada parcial” de $f$ en $a$.

Consideremos por ejemplo el caso $N = 2$. Si $\Omega$ es un abierto de $\mathbb{R}^2$ y $f : \Omega \to \mathbb{R}$ es parcialmente derivable en un punto $(a,b) \in \Omega$, sus dos derivadas parciales vienen dadas por

$$\frac{\partial f}{\partial x}(a,b) = \lim_{x \to a} \frac{f(x,b) - f(a,b)}{x - a} \quad \text{y} \quad \frac{\partial f}{\partial y}(a,b) = \lim_{y \to b} \frac{f(a,y) - f(a,b)}{y - b}$$

Volviendo al caso general, es hora de mover el punto $a \in \Omega$, hasta ahora fijo. Dado $k \in \Delta_N$, decimos que $f$ es **derivable con respecto a la $k$-ésima variable**, cuando lo es en todo $x \in \Omega$, en cuyo caso tenemos la **función derivada parcial** $\dfrac{\partial f}{\partial x_k} : \Omega \to \mathbb{R}$ dada por $x \mapsto \dfrac{\partial f}{\partial x_k}(x)$. Si esto ocurre para todo $k \in \Delta_N$ decimos simplemente que $f$ es **parcialmente derivable**. Las funciones derivadas parciales, $\dfrac{\partial f}{\partial x_k} : \Omega \to \mathbb{R}$ con $k \in \Delta_N$, son entonces campos escalares análogos a $f$, están definidos en el mismo abierto $\Omega$ que $f$.

Si ahora usamos las coordenadas $(x_1, x_2, \ldots, x_N)$ de un punto genérico $x \in \Omega$, la igualdad

$$\frac{\partial f}{\partial x_k}(x) = \frac{\partial f}{\partial x_k}(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_N)$$

plantea una cuestión que conviene aclarar. En el segundo miembro, $x_k$ es la $k$-ésima coordenada del punto $x \in \Omega$ en el que estamos calculando una derivada parcial, pero también aparece en el símbolo $\partial f / \partial x_k$ para indicar cuál de las $N$ derivadas parciales de $f$ en $x$ estamos calculando.

En vez de representar un problema, este doble uso de la variable $x_k$ nos debe ayudar a recordar la regla práctica para calcular derivadas parciales que resuelve la mayoría de los casos, sin necesidad de recurrir a la igualdad $(5)$, como vamos a explicar.

El símbolo $\partial f/\partial x_k$ nos indica la derivada parcial que estamos calculando, así que, en la expresión $f(x_1,\ldots,x_{k-1},\ x_k\ ,x_{k+1},\ldots,x_N)$, típicamente una fórmula en la que aparecen las $N$ variables $x_1,x_2,\ldots,x_N$, debemos pensar que $x_k$ es la única variable, tratar a las demás como constantes, y calcular la derivada de la función de una variable que de esta forma tenemos en mente. Ni que decir tiene, para ello podemos usar todas las reglas de derivación para funciones reales de variable real que podamos necesitar.

Este procedimiento se entenderá muy fácilmente con un ejemplo concreto, que exponemos en el caso más sencillo, $N=2$. Entonces $f$ es una función de dos variables que habitualmente se denotan por $x$ e $y$. Por tanto, sus derivadas parciales en un punto $a=(x_0,y_0)\in\Omega$, caso de que existan, vendrán dadas por

$$\frac{\partial f}{\partial x}(x_0,y_0)=\lim_{x\to x_0}\frac{f(x,y_0)-f(x_0,y_0)}{x-x_0} \quad \text{y} \quad \frac{\partial f}{\partial y}(x_0,y_0)=\lim_{y\to y_0}\frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}$$

Cuando $f$ es parcialmente derivable, sus derivadas parciales en un punto genérico $(x,y)\in\Omega$ se denotan entonces por $\dfrac{\partial f}{\partial x}(x,y)$ y $\dfrac{\partial f}{\partial y}(x,y)$, obteniendo dos funciones de dos variables, $\dfrac{\partial f}{\partial x}$ y $\dfrac{\partial f}{\partial y}$, de $\Omega$ en $\mathbb{R}$, que son las dos funciones derivadas parciales de $f$. Concretamente, para todo $(x,y)\in\Omega$, podemos escribir:

$$\frac{\partial f}{\partial x}(x,y)=\lim_{w\to x}\frac{f(w,y)-f(x,y)}{w-x} \quad \text{y} \quad \frac{\partial f}{\partial y}(x,y)=\lim_{w\to y}\frac{f(x,w)-f(x,y)}{w-y}$$

Sea por ejemplo $f:\mathbb{R}^2\to\mathbb{R}$ la función dada por $f(x,y)=e^x\operatorname{sen} y$ para todo $(x,y)\in\mathbb{R}^2$. Si en la expresión $e^x\operatorname{sen} y$, vemos $y$ como constante, pensamos en la función $x\mapsto e^x\operatorname{sen} y$, producto de la exponencial por una constante, cuya función derivada es ella misma, y esto es válido para todo $y\in\mathbb{R}$. Análogamente, para cualquier constante $x\in\mathbb{R}$, la función $y\mapsto e^x\operatorname{sen} y$ es el producto de una constante por el seno, cuya función derivada es el producto de la misma constante por el coseno. Por tanto, $f$ es parcialmente derivable en todo punto de $\mathbb{R}^2$ con

$$\frac{\partial f}{\partial x}(x,y)=e^x\operatorname{sen} y \qquad \text{y} \qquad \frac{\partial f}{\partial y}(x,y)=e^x\cos y \qquad \forall\,(x,y)\in\mathbb{R}^2$$

Nótese que $\dfrac{\partial f}{\partial x}=f$. En general, la definición de $f$ puede ser mucho más complicada, pero el mecanismo de razonamiento es el mismo.

Naturalmente, las dos variables de las que depende nuestra función no siempre se denotan por $x$ e $y$. Incluso es posible que $x$ e $y$ sean dos funciones que estemos estudiando. Por ejemplo, cuando trabajamos con las coordenadas polares en el plano, usamos determinadas restricciones de las funciones $x,y:\mathbb{R}^2\to\mathbb{R}$ definidas, para todo $(\rho,\theta)\in\mathbb{R}^2$, por

$$x(\rho,\theta)=\rho\cos\theta \qquad \text{e} \qquad y(\rho,\theta)=\rho\operatorname{sen}\theta$$

Es claro que $x$ e $y$ son parcialmente derivables en todo punto $(\rho,\theta) \in \mathbb{R}^2$ con

$$\frac{\partial x}{\partial \rho}(\rho,\theta) = \cos\theta\,, \quad \frac{\partial y}{\partial \rho}(\rho,\theta) = \operatorname{sen}\theta\,, \quad \frac{\partial x}{\partial \theta} = -y \quad \text{y} \quad \frac{\partial y}{\partial \theta} = x$$

Nótese que en la tercera y cuarta igualdad hemos podido abreviar, escribiendo una igualdad entre funciones. En este ejemplo y otros similares, para definir las funciones $x$ e $y$, se suele escribir simplemente $x = \rho\cos\theta$ e $y = \rho\operatorname{sen}\theta$, y se entiende que $x,y$ son funciones de dos variables $\rho,\theta$. Dependiendo del uso que queramos hacer de $x$ e $y$, deberemos precisar el conjunto en el que definimos tales funciones, es decir, indicar los valores que pueden tomar las variables $\rho$ y $\theta$. Por ejemplo, si escribimos $x = \rho\cos\theta$ e $y = \rho\operatorname{sen}\theta$ para $\rho > 0$ y $-\pi < \theta < \pi$, está claro que $x$ e $y$ se definen en $\mathbb{R}^+ \times ]-\pi,\pi[$.

## 8.3. Vector gradiente

Calculemos ahora la diferencial de un campo escalar, a partir de las derivadas parciales. Mantenemos la notación: $\Omega$ es un abierto de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}$ un campo escalar y $a \in \Omega$.

Si $f$ es diferenciable en $a$, usando que $Df(a)$ es lineal, junto con la igualdad $(4)$, para todo $x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N$ tenemos

$$Df(a)(x) = Df(a)\left(\sum_{k=1}^{N} x_k\, e_k\right) = \sum_{k=1}^{N} x_k\, Df(a)(e_k) = \sum_{k=1}^{N} x_k\, \frac{\partial f}{\partial x_k}(a)$$

La última suma es el producto escalar de $x$ por el vector cuyas coordenadas en la base usual de $\mathbb{R}^N$ son las $N$ derivadas parciales de $f$ en $a$, que recibe el nombre de vector gradiente, o simplemente gradiente, del campo escalar $f$ en el punto $a$.

Dicho más formalmente, cuando el campo $f$ es parcialmente derivable en $a$, el **gradiente** de $f$ en $a$ es, por definición, el vector $\nabla f(a) \in \mathbb{R}^N$ dado por

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_N}(a)\right) = \sum_{k=1}^{N} \frac{\partial f}{\partial x_k}(a)\, e_k$$

de modo que, para cada $k \in \Delta_N$, la $k$-ésima componente del vector gradiente de $f$ en $a$ es la derivada parcial de $f$ en $a$, con respecto a la $k$-ésima variable.

Acabamos de ver que, cuando $f$ es diferenciable en $a$, su diferencial se obtiene mediante el producto escalar por el vector gradiente de $f$ en $a$.

Recíprocamente, si sólo suponemos que $f$ es parcialmente derivable en $a$, disponemos del vector gradiente $\nabla f(a)$ y, como el producto escalar de $\mathbb{R}^N$ es una forma bilineal simétrica, al hacer el producto escalar por $\nabla f(a)$, obtenemos una aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}$ que es la única posible diferencial de $f$ en $a$. Así pues, la aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}$ dada por

$$T(x) = \big(\nabla f(a) \,\big|\, x\big)$$

es la única posible diferencial de $f$ en $a$, luego $f$ será diferenciable en $a$ si, y sólo si, $T$ cumple la condición que caracteriza a la diferencial. En resumen, tenemos el siguiente resultado.

- *Para un campo escalar* $f:\Omega\to\mathbb{R}$*, donde* $\Omega$ *es un abierto de* $\mathbb{R}^N$*, y un punto* $a\in\Omega$*, las siguientes afirmaciones son equivalentes:*
  - $(i)$ $f$ *es diferenciable en* $a$*.*
  - $(ii)$ $f$ *es parcialmente derivable en* $a$ *y se verifica que:*

$$\lim_{x\to a}\frac{f(x)-f(a)-\big(\nabla f(a)\,\big|\,x-a\big)}{\|x-a\|}=0 \tag{6}$$

  *En caso de que se cumplan* $(i)$ *y* $(ii)$ *se tiene:*

$$Df(a)(x)=\big(\nabla f(a)\,\big|\,x\big)\qquad \forall x\in\mathbb{R}^N \tag{7}$$

Conviene analizar en general, para cualquier aplicación lineal de $\mathbb{R}^N$ en $\mathbb{R}$, la relación entre diferencial y gradiente que acaba de aparecer. Llegamos así a identificar totalmente los espacios normados $L(\mathbb{R}^N,\mathbb{R})$ y $\mathbb{R}^N$. Para ello es ahora esencial que estemos usando en $\mathbb{R}^N$ la norma euclídea, a todos los efectos. En particular, notando como siempre $S=\{x\in\mathbb{R}^N:\|x\|=1\}$, tenemos

$$\|T\|=\max\{|T(x)|:x\in S\}\qquad \forall T\in L(\mathbb{R}^N,\mathbb{R})$$

- *Sea* $\Phi:\mathbb{R}^N\to L(\mathbb{R}^N,\mathbb{R})$ *la aplicación definida, para todo* $y\in\mathbb{R}^N$*, por* $\Phi(y)=T_y$*, donde* $T_y(x)=\big(y\,\big|\,x\big)$ *para todo* $x\in\mathbb{R}^N$*. Se tiene que* $\Phi$ *es lineal, biyectiva y preserva la norma, luego permite identificar totalmente los espacios normados* $L(\mathbb{R}^N,\mathbb{R})$ *y* $\mathbb{R}^N$*.*

La linealidad del producto escalar en la segunda variable nos dice que $T_y\in L(\mathbb{R}^N,\mathbb{R})$ mientras que, de la linealidad en primera variable deducimos fácilmente que $\Phi$ es lineal.

Para ver que $\Phi$ preserva la norma, fijamos $y\in\mathbb{R}^N\setminus\{0\}$, pues si $y=0$ no hay nada que comprobar. La desigualdad de Cauchy-Schwartz nos dice que $|T_y(x)|\leqslant\|y\|$ para todo $x\in S$, pero tomando $x=y/\|y\|\in S$ obtenemos $T_y(x)=\|y\|$, luego

$$\|\Phi(y)\|=\|T_y\|=\max\big\{|T_y(x)|:x\in S\big\}=\|y\|$$

En particular, $\Phi$ es inyectiva. Finalmente, para ver que es sobreyectiva, fijada $T\in L(\mathbb{R}^N,\mathbb{R})$, tomamos $y=\sum_{k=1}^{N}T(e_k)\,e_k$ y tenemos claramente $T_y(e_k)=T(e_k)$ para todo $k\in\Delta_N$, de donde por ser $T$ y $T_y$ lineales, deducimos que $T=T_y=\Phi(y)$. ■

Recuérdese que $\mathbb{R}^N$ también se identificó en su momento con $L(\mathbb{R},\mathbb{R}^N)$, pero de otra forma. Un vector $y\in\mathbb{R}^N$ se identificó con la aplicación lineal $U_y\in L(\mathbb{R},\mathbb{R}^N)$ dada por $U_y(t)=t\,y$ para todo $t\in\mathbb{R}$, usando el producto por escalares de $\mathbb{R}^N$. En cambio ahora, usamos el producto escalar de $\mathbb{R}^N$ para identificar $y$ con la aplicación lineal $T_y\in L(\mathbb{R}^N,\mathbb{R})$ dada por $T_y(x)=(y|x)$ para todo $x\in\mathbb{R}^N$.

Volviendo a un campo escalar $f:\Omega\to\mathbb{R}$, donde $\Omega$ es un abierto de $\mathbb{R}^N$, diferenciable en $a\in\Omega$, el resultado recién obtenido permite identificar la diferencial $Df(a)$ con el vector gradiente $\nabla f(a)$. Concretamente, la relación entre ambos es:

$$\nabla f(a)=\sum_{k=1}^{N}Df(a)(e_k)\,e_k\qquad\text{y}\qquad Df(a)(x)=\big(\nabla f(a)\,\big|\,x\big)\quad\forall x\in\mathbb{R}^N$$

## 8.4. Condición suficiente de diferenciabilidad

Frecuentemente, se puede probar la diferenciabilidad de un campo escalar, trabajando sólo con sus derivadas parciales. Ello nos permitirá obtener una caracterización muy cómoda de los campos escalares de clase $C^1$. Ambos resultados son también útiles para campos vectoriales, sin más que aplicarlos a sus componentes, que son campos escalares.

Sea pues $\Omega$ un abierto de $\mathbb{R}^N$ y supongamos que queremos probar la diferenciabilidad de un campo escalar $f : \Omega \to \mathbb{R}$ en un punto $a \in \Omega$. Sabemos que ello equivale a probar que $f$ es parcialmente derivable en el punto $a$ y que

$$\lim_{x \to a} \frac{f(x) - f(a) - \big(\nabla f(a) \,\big|\, x - a\big)}{\|x - a\|} = 0 \tag{8}$$

La idea es evitar el cálculo de este límite, a cambio de trabajar un poco más con las derivadas parciales.

En la práctica $f$ suele ser parcialmente derivable, no sólo en $a$, sino en un entorno abierto de $a$. En tal caso, por el carácter local de la diferenciabilidad, podemos sustituir $f$ por su restricción a dicho entorno, así que no perdemos generalidad suponiendo que $f$ es parcialmente derivable en $\Omega$. Entonces, en vez de comprobar la igualdad $(8)$, puede ser más fácil comprobar la continuidad en el punto $a$ de las derivadas parciales, es decir, comprobar que

$$\lim_{x \to a} \frac{\partial f}{\partial x_k}(x) = \frac{\partial f}{\partial x_k}(a) \qquad \forall k \in \Delta_N \tag{9}$$

Pues bien, vamos a probar que $(8)$ es consecuencia de $(9)$, y de hecho, que una condición ligeramente más débil que $(9)$ ya es suficiente para deducir $(8)$. Concretamente, fijado $k \in \Delta_N$, podemos suponer solamente la existencia de la derivada parcial de $f$ con respecto a la $k$-ésima variable en el punto $a$, mientras que las demás derivadas parciales sí se suponen definidas en todo punto de $\Omega$ y continuas en $a$. Esta hipótesis, obviamente más débil que $(9)$, es ya suficiente para la diferenciabilidad de $f$ en $a$.

- *Sea $\Omega$ un abierto de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}$ un campo escalar, $a \in \Omega$ y $k \in \Delta_N$. Supongamos que se verifican las dos condiciones siguientes:*

  $(i)$ *$f$ es parcialmente derivable con respecto a la $k$-ésima variable en el punto $a$*

  $(ii)$ *Para $j \in \Delta_N \setminus \{k\}$, $f$ es parcialmente derivable con respecto a la $j$-ésima variable en todo punto $x \in \Omega$ y la función derivada parcial $\dfrac{\partial f}{\partial x_j} : \Omega \to \mathbb{R}$ es continua en $a$.*

  *Entonces $f$ es diferenciable en el punto $a$.*

Salvo un obvio intercambio de variables, que no afecta a la diferenciabilidad de $f$, podemos suponer que $k = N$, es decir, que la derivada parcial para la que sólo se supone su existencia en el punto $a$, es la última. Tomamos $r \in \mathbb{R}^+$ tal que $B(a,r) \subset \Omega$, y se trata de probar $(8)$. Para que se comprenda mejor el razonamiento, conviene destacar una igualdad evidente.

Si escribimos $a = (a_1, a_2, \ldots, a_N)$, para $x = (x_1, x_2, \ldots, x_N) \in B(a,r)$ se tiene:

$$\begin{aligned} f(x) - f(a) &= \big( f(x) - f(a_1, x_2, \ldots, x_N) \big) \\ &+ \big( f(a_1, x_2, \ldots, x_N) - f(a_1, a_2, x_3, \ldots, x_N) \big) \\ &+ \ldots + \big( f(a_1, \ldots, a_{N-1}, x_N) - f(a) \big) \\ &= \sum_{j=1}^{N} \big( f(a_1, \ldots, a_{j-1}, x_j, \ldots, x_N) - f(a_1, \ldots, a_j, x_{j+1}, \ldots, x_N) \big) \end{aligned}$$

En la anterior igualdad se debe entender que $(a_1, \ldots, a_{j-1}, x_j, \ldots, x_N) = x$ para $j = 1$, así como que $(a_1, \ldots, a_j, x_{j+1}, \ldots, x_N) = a$ para $j = N$. Deducimos claramente que

$$f(x) - f(a) - \big( \nabla f(a) \,\big|\, x - a \big) = \sum_{j=1}^{N} R_j(x) \qquad \forall x \in B(a,r) \tag{10}$$

donde, para cada $j \in \Delta_N$ hemos escrito:

$$R_j(x) = f(a_1, \ldots, a_{j-1}, x_j, \ldots, x_N) - f(a_1, \ldots, a_j, x_{j+1}, \ldots, x_N) - (x_j - a_j) \frac{\partial f}{\partial x_j}(a) \tag{11}$$

Trabajaremos ahora por separado con cada uno de los sumandos que han aparecido.

Dado $\varepsilon > 0$, de $(i)$ obtenemos $\delta \in ]0, r[$ tal que, para $x_N \in \mathbb{R}$ con $|x_N - a_N| < \delta$, se tiene

$$\left| f(a_1, \ldots, a_{N-1}, x_N) - f(a) - (x_N - a_N) \frac{\partial f}{\partial x_N}(a) \right| \leqslant \frac{\varepsilon}{N} |x_N - a_N| \tag{12}$$

Por otra parte, la hipótesis $(ii)$ nos permite conseguir que el mismo $\delta$ verifique además que, para todo $w \in B(a, \delta)$ se tenga

$$\left| \frac{\partial f}{\partial x_j}(w) - \frac{\partial f}{\partial x_j}(a) \right| \leqslant \frac{\varepsilon}{N} \qquad \forall j \in \Delta_{N-1} \tag{13}$$

Para comprobar $(8)$ y concluir así la demostración, fijamos $x = (x_1, x_2, \ldots x_N) \in B(x, \delta)$, y es importante tener en cuenta que, a partir de este momento, el vector $x$, y por tanto todas sus componentes, estarán fijos.

Por una parte tenemos que $|x_N - a_N| \leqslant \|x - a\| < \delta$ luego $(12)$ nos dice que

$$|R_N(x)| \leqslant (\varepsilon/N) |x_N - a_N| \leqslant (\varepsilon/N) \|x - a\| \tag{14}$$

Por otra parte, fijamos $j \in \Delta_{N-1}$ para trabajar en el intervalo $J = ]a_j - \delta, a_j + \delta[$ con la función $\psi : J \to \mathbb{R}$ definida por

$$\psi(t) = f(a_1, \ldots, a_{j-1}, t, x_{j+1}, \ldots, x_N) \qquad \forall t \in J$$

Para cada $t \in J$, que $f$ sea parcialmente derivable con respecto a la $j$-ésima variable en el punto $(a_1, \ldots, a_{j-1}, t, x_{j+1}, \ldots, x_N) \in \Omega$, significa que $\psi$ es derivable en el punto $t$ con

$$\psi'(t) = \frac{\partial f}{\partial x_j}(a_1, \ldots, a_{j-1}, t, x_{j+1}, \ldots, x_N)$$

Usando el teorema del valor medio, obtenemos $u \in \mathbb{R}$, con $|u - a_j| \leqslant |x_j - a_j|$, tal que

$$\psi(x_j) - \psi(a_j) = (x_j - a_j)\,\psi'(u) = (x_j - a_j)\,\frac{\partial f}{\partial x_j}(a_1, \ldots, a_{j-1}, u, x_{j+1}, \ldots, x_N)$$

Tomando $w = (a_1, \ldots, a_{j-1}, u, x_{j+1}, \ldots, x_N)$, de la igualdad anterior deducimos que

$$R_j(x) = \psi(x_j) - \psi(a_j) - (x_j - a_j)\,\frac{\partial f}{\partial x_j}(a) = (x_j - a_j)\left(\frac{\partial f}{\partial x_j}(w) - \frac{\partial f}{\partial x_j}(a)\right)$$

Es claro que $\| w - a \| \leqslant \| x - a \| < \delta$, luego podemos usar $(13)$ para obtener que

$$|R_j(x)| \leqslant (\varepsilon/N)\,|x_j - a_j| \leqslant (\varepsilon/N)\,\|x - a\| \qquad \forall\, j \in \Delta_{N-1} \tag{15}$$

Teniendo en cuenta $(10)$, de $(14)$ y $(15)$, deducimos finalmente que

$$\big|\, f(x) - f(a) - \big(\nabla f(a) \,\big|\, (x - a)\big) \big| \leqslant \sum_{j=1}^{N} |R_j(x_j, x_{j+1}, \ldots, x_N)| \leqslant \varepsilon\, \|x - a\|$$

Esto prueba que se verifica $(8)$, así que $f$ es diferenciable en el punto $a$. ■

El resultado anterior nos permitirá obtener una útil caracterización de los campos escalares de clase $C^1$ en términos de sus derivadas parciales.

Si $\Omega$ es de nuevo un abierto de $\mathbb{R}^N$ y $f \in D(\Omega)$, con vistas a estudiar la continuidad de la función diferencial $Df : \Omega \to L(\mathbb{R}^N, \mathbb{R})$, para cada $x \in \Omega$, podemos identificar $Df(x)$ con $\nabla f(x)$, y considerar la **función gradiente** $\nabla f : \mathbb{R}^N \to \mathbb{R}^N$, que a cada punto $x \in \Omega$ hace corresponder el gradiente de $f$ en $x$, que es un campo vectorial. Entonces $Df$ será continua en un punto $a \in \Omega$ si, y sólo si, lo es $\nabla f$, pues sabemos que

$$\| Df(x) - Df(a) \| = \| \nabla f(x) - \nabla f(a) \| \qquad \forall\, x, a \in \Omega$$

Enunciamos ya la caracterización que buscamos:

- *Si $\Omega$ es un abierto de $\mathbb{R}^N$ y $f : \Omega \to \mathbb{R}$ un campo escalar, las siguientes afirmaciones son equivalentes:*
  - $(i)$ $f \in C^1(\Omega)$
  - $(ii)$ *$f$ es parcialmente derivable en todo punto de $\Omega$ y su función gradiente es continua, es decir, $\nabla f \in \mathcal{C}(\Omega, \mathbb{R}^N)$.*
  - $(iii)$ *$f$ es parcialmente derivable en todo punto de $\Omega$ y todas sus derivadas parciales son funciones continuas, es decir, $\partial f / \partial x_k \in \mathcal{C}(\Omega)$ para todo $k \in \Delta_N$.*

La equivalencia entre $(ii)$ y $(iii)$ es evidente, pues las derivadas parciales son las componentes de la función gradiente. La afirmación $(i)$ implica obviamente que $f$ es diferenciable y del resultado anterior se deduce que $(ii)$ también lo implica. Una vez que $f \in D(\Omega)$, la equivalencia entre $(i)$ y $(ii)$ ya se ha comentado. ■

## 8.5. Interpretación física del gradiente

El uso en Física del gradiente de un campo escalar se explica fácilmente a partir de una interpretación de las derivadas direccionales, que pasamos a explicar. Trabajamos por tanto con un abierto $\Omega \subset \mathbb{R}^N$ y un campo escalar $f : \Omega \to \mathbb{R}$, direccionalmente derivable en $a \in \Omega$.

Fijada una dirección $u \in S$, al desplazarnos una distancia $t > 0$ desde el punto $a$, en la dirección y el sentido del vector $u$, el campo $f$ experimenta una variación de $f(a+tu) - f(a)$ unidades. Podemos decir por tanto que la derivada direccional $f'_u(a)$ es la *tasa de variación* del campo en el punto $a$ y en la dirección $u$. En un "pequeño" desplazamiento desde $a$, en la dirección y sentido del vector $u$, el campo varía "aproximadamente" a razón de $f'_u(a)$ unidades de campo por unidad de longitud.

Si $f$ es diferenciable en $a \in \Omega$, se tiene $f_u{}'(a) = \big(\nabla f(a) \,|\, u\big)$ para todo $u \in S$. Supongamos que $\nabla f(a) \neq 0$ y consideremos la dirección $v = \nabla f(a) / \|\nabla f(a)\|$. Para toda dirección $u \in S$, la desigualdad de Cauchy-Schwartz nos dice que

$$f_u{}'(a) = (\nabla f(a) \,|\, u) \leqslant \|\nabla f(a)\| = (\nabla f(a) \,|\, v) = f'_v(a)$$

de donde deducimos que

$$f_v{}'(a) = \text{máx}\,\big\{ f_u{}'(a) : u \in S \big\} > 0$$

Además, sabemos que la desigualdad de Cauchy-Schwartz sólo es una igualdad cuando en ella aparece el producto escalar de dos vectores linealmente dependientes. Por tanto, si $u \in S$ verifica que $f'_u(a) = f'_v(a)$, se tendrá $u = \lambda v$ con $\lambda \in \mathbb{R}$, pero entonces $f'_u(a) = \lambda f'_v(a)$ y deducimos que $\lambda = 1$, es decir, $u = v$.

Tenemos así una caracterización del vector gradiente normalizado: $v = \nabla f(a) \,/\, \|\nabla f(a)\|$ es la única dirección $v \in S$ que hace que la derivada direccional $f'_v(a)$ sea máxima. Obsérvese que esta caracterización del vector gradiente es independiente de la base que estemos usando en $\mathbb{R}^N$. Si cambiamos de base en $\mathbb{R}^N$, las derivadas direccionales no cambian, luego el vector gradiente seguirá siendo el mismo, aunque obviamente cambien sus coordenadas.

El significado físico de esta caracterización es claro, y se facilita si tenemos en cuenta que el vector $v = \nabla f(a) \,/\, \|\nabla f(a)\|$ tiene la misma dirección y sentido que $\nabla f(a)$. Por tanto, al desplazarnos desde el punto $a$ en la dirección y el sentido del vector gradiente, conseguimos la máxima tasa de aumento del campo por unidad de longitud, o si se quiere, hacemos que el campo aumente lo más rápidamente posible. Son además la única dirección y el único sentido que producen dicha máxima tasa de aumento, que viene dada por la norma euclídea del vector gradiente, $f_v{}'(a) = \|\nabla f(a)\|$. Dicho más intuitivamente, un "pequeño" desplazamiento en la dirección y sentido del vector gradiente, hace que el campo aumente "aproximadamente" a razón de $\|\nabla f(a)\|$ unidades, de las que usemos para medirlo, por unidad de longitud.

Por otra parte, como $f'_{-v}(a) = -f'_v(a)$, al desplazarnos por la misma recta, pero en el sentido opuesto, el del vector $-\nabla f(a)$, conseguimos que el valor del campo disminuya lo más rápidamente posible,"aproximadamente" a razón de $\|\nabla f(a)\|$ unidades de campo por unidad de longitud.

Resaltamos que lo dicho es válido cuando $f$ es diferenciable en $a$, no basta con que $f$ sea parcialmente derivable en $a$, suponiendo además que $\nabla f(a) \neq 0$. Cuando $\nabla f(a) = 0$, todas las derivadas direccionales de $f$ en $a$ se anulan, luego el campo varía muy lentamente cerca del punto $a$, su tasa de variación es "prácticamente" nula en todas las direcciones. Se dice entonces que $a$ es un *punto crítico* o un *punto estacionario* del campo $f$.

## 8.6. Plano tangente a una superficie explícita

Recordemos que una curva explícita en $\mathbb{R}^2$ es la gráfica de una función continua, definida en un intervalo y con valores reales, el tipo más sencillo de curva paramétrica. Análogamente, consideramos ahora el tipo más sencillo de superficie en $\mathbb{R}^3$.

Llamamos **superficie explícita** en $\mathbb{R}^3$ a la gráfica de una función continua $f : \Omega \to \mathbb{R}$, donde $\Omega$ es un subconjunto no vacío, abierto y conexo, de $\mathbb{R}^2$. Se trata por tanto del conjunto

$$\Sigma = \operatorname{Gr} f = \Big\{ \big( x\,,\, y\,,\, f(x,y) \big) \,:\, (x,y) \in \Omega \Big\} \subset \mathbb{R}^3$$

Nótese que $\Sigma$ determina de manera única a $f$, pues para $(x,y) \in \mathbb{R}^2$, es claro que $(x,y) \in \Omega$ si, y sólo si, existe $z \in \mathbb{R}$ tal que $(x,y,z) \in \Sigma$, en cuyo caso $f(x,y)$ es el único $z \in \mathbb{R}$ que verifica dicha condición. Se dice que la igualdad

$$z = f(x,y) \qquad \big( (x,y) \in \Omega \big)$$

es la **ecuación explícita** de la superficie $\Sigma$.

Geométricamente, es también natural considerar superficies explícitas de otros dos tipos, que se obtienen intercambiando los ejes de coordenadas. Concretamente, la misma función $f$ también da lugar a las superficies explícitas

$$\Sigma_1 = \Big\{ \big( f(y,z)\,,\, y\,,\, z \big) \,:\, (y,z) \in \Omega \Big\} \qquad \text{y} \qquad \Sigma_2 = \Big\{ \big( x\,,\, f(x,z)\,,\, z \big) \,:\, (x,z) \in \Omega \Big\}$$

cuyas ecuaciones explícitas serán $x = f(y,z)$ con $(y,z) \in \Omega$ e $y = f(x,z)$ con $(x,z) \in \Omega$, respectivamente. Trabajaremos solamente con una superficie $\Sigma$ del primer tipo, la gráfica de $f$, pues toda la discusión que haremos se traduce fácilmente para aplicarla a los otros dos tipos.

Cuando $f$ es diferenciable en un punto $(x_0, y_0) \in \Omega$, veamos la relación entre el vector gradiente $\nabla f(x_0, y_0)$ y la superficie $\Sigma$. Para abreviar, escribimos

$$z_0 = f(x_0, y_0)\,, \quad \alpha_0 = \frac{\partial f}{\partial x}(x_0, y_0) \quad \text{y} \quad \beta_0 = \frac{\partial f}{\partial y}(x_0, y_0)$$

y consideramos el plano afín $\Pi$ definido por la ecuación

$$z - z_0 = \alpha_0\,(x - x_0) + \beta_0\,(y - y_0)$$

que también es una superficie explícita, concretamente $\Pi = \operatorname{Gr} g$ donde $g : \mathbb{R}^2 \to \mathbb{R}$ es la función definida por

$$g(x,y) = z_0 + \alpha_0\,(x - x_0) + \beta_0\,(y - y_0) \qquad \forall\,(x,y) \in \mathbb{R}^2 \tag{16}$$

Usando la relación entre la diferencial de $f$ y su gradiente, vemos que $g$ es una función afín que ya hemos manejado anteriormente:

$$g(x,y) = f(x_0,y_0) + Df(x_0,y_0)\big((x,y)-(x_0,y_0)\big) \qquad \forall\,(x,y)\in\mathbb{R}^2$$

Al comentar el significado analítico de la diferencial, vimos que $g$ es la función afín que nos da una buena aproximación de $f$ cerca del punto $(x_0,y_0)$, es decir,

$$\lim_{(x,y)\to(x_0,y_0)} \frac{|\,f(x,y)-g(x,y)\,|}{\|\,(x,y)-(x_0,y_0)\,\|} = 0 \tag{17}$$

Geométricamente, esto significa que la distancia (vertical) entre el punto $\big(x,y,f(x,y)\big)$ de la superficie $\Sigma$ y el correspondiente punto $\big(x,y,g(x,y)\big)$ del plano $\Pi$, tiende a cero cuando ambos puntos se acercan a $(x_0,y_0,z_0)$, "mucho más rápidamente" que $\|\,(x,y)-(x_0,y_0)\,\|$, así que la siguiente definición está plenamente justificada.

Sea $\Sigma = \operatorname{Gr} f$ una superficie explícita en $\mathbb{R}^3$, donde $f:\Omega\to\mathbb{R}$ es una función continua en un abierto conexo $\Omega\subset\mathbb{R}^2$. Si $f$ es diferenciable en un punto $(x_0,y_0)$ y $z_0 = f(x_0,y_0)$, se dice que el plano $\Pi$ de ecuación explícita

$$z-z_0 = (x-x_0)\,\frac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\,\frac{\partial f}{\partial y}(x_0,y_0) \tag{18}$$

es el **plano tangente** a la superficie $\Sigma$ en el punto $(x_0,y_0,z_0)$.

Decimos también que el vector

$$\left(\frac{\partial f}{\partial x}(x_0,y_0)\;,\;\frac{\partial f}{\partial y}(x_0,y_0)\;,\;-1\right)\in\mathbb{R}^3 \tag{19}$$

es un **vector normal** a la superficie $\Sigma$ en el punto $(x_0,y_0,z_0)$. Ello se explica porque, en vista de $(18)$, se trata de un vector normal al plano tangente $\Pi$, es claramente ortogonal a todos los vectores de la forma $v-u$ con $u,v\in\Pi$ es decir, a todas las rectas contenidas en $\Pi$. Tenemos así la interpretación geométrica de la diferencial $Df(x_0,y_0)$ o del vector gradiente $\nabla f(x_0,y_0)$.

Obsérvese que, para definir el vector que aparece en $(19)$, normal al plano dado por $(18)$, o lo que es lo mismo, considerar la función afín $g$ dada por $(16)$, bastaría disponer del gradiente de $f$ en el punto $(x_0,y_0)$, es decir, que $f$ fuese sólo parcialmente derivable en $(x_0,y_0)$. Sin embargo, el plano así obtenido sólo es una buena aproximación de la superficie $\Sigma$ cerca del punto $(x_0,y_0,z_0)$ cuando se cumple $(17)$, que equivale a la diferenciabilidad de $f$ en $(x_0,y_0)$.

Esta es la razón por la que sólo hemos definido el plano tangente y el vector normal a una superficie $\Sigma = \operatorname{Gr} f$ en el punto $(x_0,y_0,z_0)\in\Sigma$, cuando $f$ es diferenciable en $(x_0,y_0)$. Así pues, la noción geométricamente relevante, la que justifica la definición de plano tangente y vector normal, es la diferenciabilidad, no basta la derivabilidad parcial.

Más adelante, la regla de la cadena para las derivadas parciales, nos permitirá reafirmar aún más la interpretación geométrica que acabamos de hacer.

## 8.7. Extremos absolutos y relativos

Recordemos ciertas nociones elementales, que conocemos bien en el caso de una función real de variable real, pero tienen perfecto sentido en contextos mucho más generales.

Si $A$ es un conjunto no vacío arbitrario y $f : A \to \mathbb{R}$ una función, decimos que $f$ tiene o alcanza un **máximo absoluto** en un punto $a \in A$, cuando $f(a) = \max f(A)$, es decir,

$$f(a) \geqslant f(x) \qquad \forall x \in A$$

Obviamente, no debemos confundir el valor máximo de $f$, que es $f(a)$, con el punto $a$ donde se alcanza dicho valor. Análogamente, decimos que $f$ tiene un **mínimo absoluto** en $a \in A$ cuando $f(a) = \min f(A)$ es decir,

$$f(a) \leqslant f(x) \qquad \forall x \in A$$

Obviamente, $f$ tiene un mínimo absoluto en $a$ si, y sólo si, la función $-f$ tiene un máximo absoluto en $a$, de modo que ambas nociones se reducen una a la otra. Decimos que $f$ tiene un **extremo absoluto** en un punto $a \in A$ cuando tiene un máximo absoluto en $a$ o un mínimo absoluto en $a$.

El problema de averiguar si $f$ tiene un máximo o un mínimo absoluto en algún punto de $A$, y en su caso encontrarlos, es lo que, en términos muy generales, se conoce como un *problema de optimización* o un problema de extremos. Claramente, sin hipótesis sobre el conjunto $A$ y la función $f$, este problema es intratable.

Conocemos una respuesta a este problema en un contexto muy general. Concretamente, sabemos que si $A$ es un espacio métrico compacto y $f$ es continua, entonces $f$ tiene un máximo absoluto y un mínimo absoluto en sendos puntos de $A$, aunque nada sabemos sobre cómo encontrarlos. En casos más concretos, suele ser útil la siguiente noción de extremo relativo.

Supongamos que $A$ es un subconjunto de un espacio métrico $E$. Decimos que $f : A \to \mathbb{R}$ tiene un **máximo relativo** en un punto $a \in A$, cuando existe $r \in \mathbb{R}^+$ verificando que

$$B(a,r) \subset A \qquad \text{y} \qquad f(a) \leqslant f(x) \quad \forall x \in B(a,r)$$

Obsérvese que en particular se ha de tener $a \in A^\circ$.

Análogamente, $f$ tiene un **mínimo relativo** en $a \in A$, cuando existe $r \in \mathbb{R}^+$ tal que

$$B(a,r) \subset A \qquad \text{y} \qquad f(x) \leqslant f(a) \quad \forall x \in B(a,r)$$

De nuevo es claro que $f$ tiene un mínimo relativo en $a$ si, y sólo si, $-f$ tiene un máximo relativo en $a$. Finalmente, decimos que $f$ tiene un **extremo relativo** en un punto $a \in A$, cuando tiene un máximo relativo o un mínimo relativo en $a$.

Observemos que las nociones de extremo absoluto y relativo no están relacionadas. Cierto que, cuando $f$ tiene un extremo relativo en $a$, la restricción de $f$ a una bola $B(a,r) \subset A$ tiene un máximo absoluto en $a$, pero fuera de dicha bola, $f$ puede tomar valores estrictamente mayores y menores que $f(a)$, en cuyo caso $f$ no tendrá un extremo absoluto en $a$.

En sentido opuesto, es claro que si $f$ tiene en $a$ un extremo absoluto, entonces tendrá un extremo relativo si, y sólo si, $a \in A^\circ$. Por tanto, salvo que $A$ sea abierto, $f$ puede perfectamente tener un extremo absoluto en un punto $a \in A \setminus A^\circ$, que nunca puede ser un extremo relativo.

Pues bien, a la hora de encontrar los puntos en los que un campo escalar puede tener un extremo relativo, entra en juego el cálculo diferencial, igual que ocurría con funciones reales de variable real. Supongamos pues que $A$ es un subconjunto no vacío de $\mathbb{R}^N$ y que $f : A \to \mathbb{R}$ es un campo escalar diferenciable en un punto $a \in A^\circ$. Tanto el punto de vista físico como la intuición geométrica sugieren que, si $f$ tiene un extremo relativo en $a$, se debe tener $\nabla f(a) = 0$. En efecto, si fuese $\nabla f(a) \neq 0$, el campo aumentaría en la dirección y el sentido del vector $\nabla f(a)$ y disminuiría en sentido opuesto, luego $f$ no puede tener un extremo relativo en $a$. Igualmente, en el caso $N = 2$, cabe esperar que, si $f$ tiene un extremo relativo en el punto $a = (a_1, a_2)$, el plano tangente a la superficie $\operatorname{Gr} f$ en el punto $\big(a_1, a_2, f(a_1, a_2)\big)$ tenga que ser horizontal, es decir, $\nabla f(a_1, a_2) = 0$.

Comprobamos fácilmente que las ideas anteriores son correctas, e incluso, esta vez no es necesario que $f$ sea diferenciable en el punto $a$, basta la derivabilidad parcial:

- **Condición necesaria de extremo relativo.** *Sea $A$ un subconjunto no vacío de $\mathbb{R}^N$, y sea $f : A \to \mathbb{R}$ un campo escalar. Si $f$ tiene un extremo relativo en un punto $a \in A^\circ$ y es parcialmente derivable en dicho punto, entonces $\nabla f(a) = 0$.*

Supongamos que $f$ tiene en $a$ un máximo relativo, pues en el otro caso, sustituimos $f$ por $-f$. Sea pues $r \in \mathbb{R}^+$ tal que $B(a,r) \subset A$ y $f(x) \leqslant f(a)$ para todo $x \in B(a,r)$. Fijado $k \in \Delta_N$, para todo $t \in ]0, r[$ se tiene entonces

$$\frac{f(a+te_k)-f(a)}{t} \leqslant 0 \quad \text{luego} \quad \frac{\partial f}{\partial x_k}(a) = \lim_{t \to 0^+} \frac{f(a+te_k)-f(a)}{t} \leqslant 0$$

Pero análogamente, para $-r < t < 0$ se tiene también

$$\frac{f(a+te_k)-f(a)}{t} \geqslant 0 \quad \text{luego} \quad \frac{\partial f}{\partial x_k}(a) = \lim_{t \to 0^-} \frac{f(a+te_k)-f(a)}{t} \geqslant 0$$

Esto prueba que $\dfrac{\partial f}{\partial x_k}(a) = 0$ para todo $k \in \Delta_N$, es decir, $\nabla f(a) = 0$. ■

Como se hizo en la interpretación física del gradiente, cuando un campo escalar $f : A \to \mathbb{R}$ es parcialmente derivable en un punto $a \in A^\circ$, y verifica que $\nabla f(a) = 0$, se dice que $a$ es un **punto crítico** de $f$. Asi pues, los puntos en los que $f$ es parcialmente derivable y tiene un extremo relativo, han de ser puntos críticos. Esto ayuda con frecuencia a encontrarlos.

## 8.8. Ejercicios

1. Calcular todas las derivadas direccionales en el punto $(-1,0,0)$ de la función $f:\mathbb{R}^3\to\mathbb{R}$ definida por
$$f(x,y,z) = x^3 - 3xy + z^3 \qquad \forall\,(x,y,z)\in\mathbb{R}^3$$

2. Sea $J$ un intervalo abierto en $\mathbb{R}$ y $\Omega$ un subconjunto abierto de $\mathbb{R}^N$. Si $f:J\to\Omega$ es una función derivable en un punto $a\in J$ y $g:\Omega\to\mathbb{R}$ es diferenciable en el punto $b=f(a)$, probar que la función $h=g\circ f:J\to\mathbb{R}$ es derivable en el punto $a$ con
$$h'(a) = \big(\nabla g(b)\,\big|\,f'(a)\big)$$

3. Sea $\Omega$ un subconjunto abierto de $\mathbb{R}^N$ y $f,g:\Omega\to\mathbb{R}$ funciones parcialmente derivables en un punto $a\in\Omega$. Probar que las funciones $f+g$ y $fg$ son parcialmente derivables en $a$ con
$$\nabla(f+g)(a) = \nabla f(a) + \nabla g(a) \qquad\qquad \text{y}$$
$$\nabla(fg)(a) = g(a)\,\nabla f(a) + f(a)\,\nabla g(a)$$
Suponiendo que $g(x)\neq 0$ para todo $x\in\Omega$, probar también que $f/g$ es parcialmente derivable en $a$ con
$$\nabla(f/g)(a) = \frac{g(a)\,\nabla f(a) - f(a)\,\nabla g(a)}{g(a)^2}$$

4. Fijado $p\in\mathbb{R}^*$, se considera la función $f:\mathbb{R}^N\setminus\{0\}\to\mathbb{R}$ definida por $f(x) = \|x\|^p$ para todo $x\in\mathbb{R}^N\setminus\{0\}$, donde $\|\cdot\|$ es la norma euclídea. Probar que $f\in C^1\big(\mathbb{R}^N\setminus\{0\}\big)$ con
$$\nabla f(x) = p\,\|x\|^{p-2}x \qquad \forall\,x\in\mathbb{R}^N\setminus\{0\}$$
Como consecuencia, encontrar una función $g\in C^1(\mathbb{R}^N,\mathbb{R}^N)$ que verifique
$$\nabla g(x) = x \qquad \forall\,x\in\mathbb{R}^N$$

5. Calcular la ecuación del plano tangente a la superficie explícita de ecuación $z = x + y^3$ con $(x,y)\in\mathbb{R}^2$, en el punto $(1,1,2)$.

6. Sea $\Omega$ un subconjunto abierto y conexo de $\mathbb{R}^2$ y $f:\Omega\to\mathbb{R}$ una función diferenciable. Se considera la superficie explícita $S\subset\mathbb{R}^3$ dada por
$$S = \big\{\,\big(x,\,f(x,z)\,,\,z\big)\;:\;(x,z)\in\Omega\,\big\}$$
Calcular la ecuación del plano tangente a $S$ en un punto arbitrario $(x_0,y_0,z_0)\in S$.

7. Sea $\Omega$ un subconjunto abierto de $\mathbb{R}^2$ y $f:\Omega\to\mathbb{R}^2$ una función parcialmente derivable en todo punto de $\Omega$. Probar que, si la función $\nabla f:\Omega\to\mathbb{R}^2$ está acotada, entonces $f$ es continua. Usando la función $f:\mathbb{R}^2\to\mathbb{R}$ definida por
$$f(x,y) = \frac{xy}{\sqrt{x^2+y^2}} \quad \forall\,(x,y)\in\mathbb{R}^2\setminus\{(0,0)\} \quad \text{y} \quad f(0,0) = 0$$
comprobar que, con las mismas hipótesis, no se puede asegurar que $f$ sea diferenciable.