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

Recordemos el razonamiento usado, en general, para probar la unicidad de la diferencial. Sean pues X,Y espacios normados, Ω un subconjunto abierto no vacío de X y $f:\Omega \to Y$ una función diferenciable en un punto $a \in \Omega$. Entonces $Df(a) \in L(X,Y)$ verifica que

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

Tomamos $r \in \mathbb{R}^+$ de forma que $B(a,r) \subset \Omega$ y, fijado $v \in X \setminus \{0\}$, consideramos el intervalo abierto $J_v =]-r/\| v \|, r/\| v \| [$. El cambio de variable $x = a + t\,v$, con $t \in J_v$, teniendo en cuenta que $x \in \Omega \setminus \{a\}$ para $t \in J_v \setminus \{0\}$ y que $x \to a$ cuando $t \to 0$, nos da

$$0 = \lim_{t \to 0} \frac{\| f(a+tv) - f(a) - t\,Df(a)(v) \|}{|t|\, \| v \|} = \frac{1}{\| v \|} \lim_{t \to 0} \left\| \frac{f(a+tv) - f(a)}{t} - Df(a)(v) \right\|$$

o lo que es lo mismo,

$$Df(a)(v) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} \tag{1}$$

Hasta aquí nada nuevo, la igualdad anterior, válida para todo $v \in X \setminus \{0\}$, y obviamente también para $v = 0$, es la que asegura la unicidad de la diferencial.

Pero consideremos la función de variable real $\varphi_v : J_v \to Y$ dada por $\varphi_v(t) = f(a+tv)$ para todo $t \in J_v$. La igualdad (1) nos dice que φ_v es derivable en el origen con $\varphi_v'(0) = Df(a)(v)$. Si esto ocurre para todo $v \in X \setminus \{0\}$, diremos que f es **direccionalmente derivable** en a.

Igual que ocurría con el límite de una función en un punto y los límites direccionales, la derivabilidad direccional de f en a es condición necesaria, pero no suficiente, para que f sea diferenciable en a. A la hora de ver si se cumple dicha condición, no es necesario considerar todos los vectores $v \in X \setminus \{0\}$. Si para $u \in X \setminus \{0\}$ sabemos que φ_u es derivable en 0 y tomamos $v = \rho u$ con $\rho \in \mathbb{R}^*$, usando el cambio de variable $t = s/\rho$ tenemos claramente

$$\lim_{t\to 0} \frac{f(a+tv)-f(a)}{t} = \rho \lim_{t\to 0} \frac{f(a+t\rho u)-f(a)}{t\rho} = \rho \lim_{s\to 0} \frac{f(a+su)-f(a)}{s} = \rho\, \varphi_u'(0)$$

luego φ_v también es derivable en 0 con $\varphi_v'(0) = \rho\, \varphi_u'(0)$. Basta por tanto usar vectores $u \in X$ que verifiquen $\|u\| = 1$, que son las *direcciones* del espacio normado X. Veamos pues la definición formal de las derivadas direccionales.

Si $u \in X$ y $\|u\| = 1$, decimos que f es **derivable en la dirección** u, en el punto a, cuando la función $\varphi_u :]-r, r[\to Y$, definida por $\varphi_u(t) = f(a+tu)$ para todo $t \in]-r, r[$, es derivable en 0. Entonces el vector derivada $\varphi_u'(0) \in Y$ es, por definición, la **derivada direccional** de f en el punto a y en la dirección u, que se denota por $f_u'(a)$, es decir,

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

La observación anterior, con $\rho = -1$, nos dice que f es derivable en la dirección u en el punto a si, y sólo si, lo es en la dirección $-u$, en cuyo caso se tiene que $f_{-u}'(a) = -f_u'(a)$. Reiteramos finalmente la relación entre diferenciabilidad y derivabilidad direccional:

- *Sean X, Y espacios normados, Ω un subconjunto abierto no vacío 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 y, para todo $u \in X$ con $\|u\| = 1$, se tiene que $f_u'(a) = Df(a)(u)$.*

Notación. A partir de ahora trabajamos con $X = \mathbb{R}^N$ e $Y = \mathbb{R}$. Para evitar repeticiones, en todo lo que sigue fijamos un abierto Ω de $\mathbb{R}^N$, un campo escalar $f : \Omega \to \mathbb{R}$ 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 $B(a, r) \subset \Omega$.

Si f es direccionalmente derivable en a, sus derivadas direccionales son:

$$f_u'(a) = \lim_{t\to 0} \frac{f(a+tu)-f(a)}{t} \qquad \forall u \in S$$

Se trata de números reales bien fáciles de calcular, pues son derivadas en el origen de funciones reales de variable real, todas ellas definidas en el intervalo $]-r, r[$. Concretamente, para cada $u \in S$, la derivada direccional $f_u'(a)$ es la derivada en el origen de la función $\varphi_u :]-r, r[\to \mathbb{R}$ definida por $\varphi_u(t) = f(a+tu)$ para todo $t \in]-r, r[$.

Conviene entender el significado intuitivo de las derivadas direccionales, pues las derivadas parciales que enseguida vamos a definir, no son más que casos particulares. Consideremos un desplazamiento desde el punto a, siguiendo la recta que tiene $u \in S$ como vector de dirección y en el sentido indicado por el vector u, hasta el punto $a+tu$ con $0 < t < r$. El campo f experimenta entonces una variación $f(a+tu) - f(a)$, luego el cociente $\dfrac{f(a+tu) - f(a)}{t}$ es la tasa de variación del campo por unidad de longitud. La definición de la derivada direccional nos dice que, cuando t es "pequeño", dicha tasa de variación es "aproximadamente" $f_u'(a)$. Nótese que sólo hemos tenido en cuenta la derivada por la derecha en 0 de la función $t \mapsto f(a+tu)$. Para $t \in \mathbb{R}^-$, en vez de usar la derivada por la izquierda de dicha función, es preferible recordar que $f_{-u}'(a) = -f_u'(a)$ y pensar en la derivada por la derecha en 0 de la función $t \mapsto f(a-tu)$. Vemos entonces que, cuando nos desplazamos en la misma recta pero en el sentido opuesto, el del vector $-u$, la tasa de variación mantiene su valor absoluto pero cambia de signo. Si $f_u'(a) > 0$, el campo aumenta en el sentido del vector u, tanto más rápidamente cuanto mayor sea $f_u'(a)$, y disminuye con la misma rapidez en el sentido del vector $-u$.

8.2. Derivadas parciales

En la definición de las derivadas direccionales de nuestro campo escalar f, podemos usar 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, fijado $j \in I_N$, cuando f es derivable en la dirección e_j, en el punto a, decimos que f es parcialmente derivable con respecto a la j-ésima variable en el punto a. Entonces, la derivada direccional de f en a en la dirección e_j se denomina **derivada parcial** de f con respecto a la j-ésima variable en el punto a, y se denota por $\dfrac{\partial f}{\partial x_j}(a)$, es decir,

$$\frac{\partial f}{\partial x_j}(a) = f_{e_k}'(a) = \lim_{t \to 0} \frac{f(a+te_j) - f(a)}{t}$$

Si esto ocurre para todo $j \in I_N$, decimos que f es **parcialmente derivable** en a, y tenemos N derivadas parciales, $\dfrac{\partial f}{\partial x_j}(a)$ con $j \in I_N$, definidas por la igualdad anterior. Como esta noción es más débil que la derivabilidad direccional, tenemos:

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

$$\frac{\partial f}{\partial x_j}(a) = Df(a)(e_j) \qquad \forall\, j \in I_N \tag{2}$$

En el sentido opuesto, más adelante veremos la condición que debe cumplir una función parcialmente derivable para ser diferenciable, con lo que tendremos una caracterización de la diferenciabilidad en términos de las derivadas parciales. De momento, comprobar que f es parcialmente derivable y calcular sus N derivadas parciales, es el primer paso para probar que f es diferenciable. Conviene por tanto prestar más atención al concepto de derivada parcial.

Usemos las coordenadas del punto a, es decir, pongamos $a = (a_1, a_2, \ldots, a_N)$. Para $j \in I_N$ y $t \in \mathbb{R}$, la j-ésima coordenada de $a + te_j$ es $a_j + t$, mientras que las restantes coinciden con las de a, por lo que escribiremos $a + te_j = (a_1, \ldots a_{j-1},\ a_j + t\ , a_{j+1}, \ldots, a_N)$, sin excluir los casos $j = 1$ y $j = N$. Hemos definido la derivada parcial como $\dfrac{\partial f}{\partial x_j}(a) = \varphi\,'(0)$ donde $\varphi : \left]-r, r\right[\to \mathbb{R}$ es la función dada por

$$\varphi(t) = f(a + te_j) = f(a_1, \ldots a_{j-1},\ a_j + t\ , a_{j+1}, \ldots, a_N) \qquad \forall t \in \left]-r, r\right[$$

En la práctica no se usa la función φ, porque es más intuitivo, como vamos a ver enseguida, considerar la función $\psi : \left]a_j - r, a_j + r\right[\to \mathbb{R}$ definida por

$$\psi(x_j) = f(a_1, \ldots a_{j-1},\ x_j\ , a_{j+1}, \ldots, a_N) \qquad \forall x_j \in \left]a_j - r, a_j + r\right[$$

Las funciones φ y ψ se deducen una de otra mediante traslaciones, más concretamente se tiene $\psi(x_j) = \varphi(x_j - a_j)$ para todo $x_j \in \left]a_j - r, a_j + r\right[$ o, lo que es lo mismo, $\varphi(t) = \psi(a_j + t)$ para todo $t \in \left]-r, r\right[$. Deducimos claramente que φ es derivable en 0 si, y sólo si, ψ es derivable en a_j, en cuyo caso se tiene $\varphi\,'(0) = \psi\,'(a_j)$. Así pues, f es parcialmente derivable con respecto a la j-ésima variable en a si, y sólo si, ψ es derivable en a_j, en cuyo caso se tiene

$$\frac{\partial f}{\partial x_j}(a) = \psi\,'(a_j)$$

Mientras f es una función de las N variables $x_1, x_2, \ldots, x_N$, la función ψ de una sola variable, x_j, describe "parcialmente" el comportamiento de f. Sólo refleja la dependencia de f respecto de la variable x_j, para $a_j - r < x_j < a_j + r$, cuando por así decirlo, las demás variables se mantienen constantes, pues para obtener $\psi(x_j)$ a partir de $f(x_1, \ldots, x_{j-1},\ x_j\ , x_{j+1}, \ldots, x_N)$, tomamos $x_k = a_k$ para todo $k \in I_N \setminus \{j\}$. Por eso decimos que la derivada de ψ en a_j es sólo una "derivada parcial" de f en a.

Para el cálculo práctico de derivadas parciales, debemos tener en cuenta que frecuentemente, no sólo calculamos $\dfrac{\partial f}{\partial x_j}(a)$, sino $\dfrac{\partial f}{\partial x_j}(x)$ para todo punto $x \in \Omega$. Lo que en realidad calculamos es por tanto la función $x \mapsto \dfrac{\partial f}{\partial x_j}(x)$, de Ω en $\mathbb{R}$, una función de N variables a la que podemos llamar **función derivada parcial** de f con respecto a la j-ésima variable.

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

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

plantea una cuestión de notación que conviene aclarar. En el paréntesis del segundo miembro, x_j es la j-ésima coordenada del punto $x \in \Omega$ en el que estamos calculando una derivada parcial, pero también aparece en el símbolo $\partial f / \partial x_j$ para indicar cuál de las N derivadas parciales de f en x estamos calculando. Ocurre con esta notación clásica para las derivadas parciales exactamente lo mismo que cuando, en el caso $N = 1$, usamos la notación también clásica para la derivada: $\dfrac{df}{dx}(x)$.

En vez de representar un problema, este doble uso de la variable x_j nos debe ayudar a recordar la regla práctica para calcular derivadas parciales que resuelve la mayoría de los casos. El símbolo $\partial f/\partial x_j$ nos indica la variable x_j respecto de la cual estamos calculando la derivada parcial, luego en la expresión $f(x_1,\ldots,x_{j-1},\ x_j\ ,x_{j+1},\ldots,x_N)$, típicamente una fórmula en la que aparecen las N variables $x_1,x_2,\ldots,x_N$, debemos pensar que en realidad x_j es la única variable, tratar a todas 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 en todo punto de Ω, 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 Ω 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 una constante, obtenemos la función $x\mapsto e^x \operatorname{sen} y$, de $\mathbb{R}$ en $\mathbb{R}$, 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 las 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 la definición de las funciones x e y, se suele escribir simplemente $x = \rho\cos\theta$ e $y = \rho\operatorname{sen}\theta$, y se sobreentiende que x,y son funciones de dos variables ρ,θ. 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 ρ y θ. 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 $\mathbb{R}^+\times\,]-\pi,\pi[$.

Queda claro que el cálculo de derivadas parciales consiste en calcular derivadas de funciones reales de variable real, dominarlo es cuestión de práctica. Por ello, aunque sea repetitivo, veamos algún ejemplo en el caso $N=3$, con lo que tenemos tres variables, que habitualmente se denotan por x,y,z. Si Ω es un abierto de $\mathbb{R}^3$ y $f:\Omega\to\mathbb{R}$, las derivadas parciales de f en un punto $(x_0,y_0,z_0)\in\Omega$, cuando existan, vendrán dadas por

$$\begin{aligned}
\frac{\partial f}{\partial x}(x_0,y_0,z_0) &= \lim_{x\to x_0}\frac{f(x,y_0,z_0)-f(x_0,y_0,z_0)}{x-x_0}\\
\frac{\partial f}{\partial y}(x_0,y_0,z_0) &= \lim_{y\to y_0}\frac{f(x_0,y,z_0)-f(x_0,y_0,z_0)}{y-y_0}\\
\frac{\partial f}{\partial z}(x_0,y_0,z_0) &= \lim_{z\to z_0}\frac{f(x_0,y_0,z)-f(x_0,y_0,z_0)}{z-z_0}
\end{aligned}$$

No hay problema en trabajar en un punto genérico $(x,y,z)\in\Omega$ y escribir, por ejemplo,

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

Consideremos por ejemplo las funciones $x,y,z:\mathbb{R}^3\to\mathbb{R}$ cuyas restricciones se usan para trabajar con las coordenadas esféricas. Vienen dadas, para $(r,\theta,\varphi)\in\mathbb{R}^3$ por

$$x = r\cos\theta\,\cos\varphi\,, \qquad y = r\operatorname{sen}\theta\,\cos\varphi \qquad \text{y} \quad z = r\operatorname{sen}\varphi$$

Claramente las tres funciones son parcialmente derivables en todo punto de $\mathbb{R}^3$. Para simplificar la escritura de sus derivadas parciales, podemos omitir el punto en el que las calculamos, se sobreentiende que se trata de un punto genérico $(r,\theta,\varphi)\in\mathbb{R}^3$, con lo que tenemos:

$$\begin{aligned}
&\frac{\partial x}{\partial r} = \cos\theta\,\cos\varphi\,, && \frac{\partial x}{\partial\theta} = -r\operatorname{sen}\theta\,\cos\varphi\,, && \frac{\partial x}{\partial\varphi} = -r\cos\theta\,\operatorname{sen}\varphi\\
&\frac{\partial y}{\partial r} = \operatorname{sen}\theta\,\cos\varphi\,, && \frac{\partial y}{\partial\theta} = r\cos\theta\,\cos\varphi\,, && \frac{\partial y}{\partial\varphi} = -r\operatorname{sen}\theta\operatorname{sen}\varphi\\
&\frac{\partial z}{\partial r} = \operatorname{sen}\varphi\,, && \frac{\partial z}{\partial\theta} = 0\,, && \frac{\partial z}{\partial\varphi} = r\cos\varphi
\end{aligned}$$

8.3. Vector gradiente

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

Si f es diferenciable en a, para todo $x \in \mathbb{R}^N$, usando que $Df(a)$ es lineal, junto con la igualdad (2), tenemos

$$Df(a)(x) = Df(a)\left(\sum_{j=1}^{N} x_j e_j\right) = \sum_{j=1}^{N} x_j Df(a)(e_j) = \sum_{k=1}^{N} x_j \frac{\partial f}{\partial x_j}(a)$$

La última suma es el producto escalar de $x \in \mathbb{R}^N$ 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.

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

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_N}(a)\right) = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(a)\, e_j$$

de modo que la j-ésima coordenada del vector gradiente de f en a es la derivada parcial con respecto a la j-ésima variable de f en a, para todo $j \in I_N$.

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, el producto escalar por el vector gradiente es una aplicación lineal y continua de $\mathbb{R}^N$ en $\mathbb{R}$ que es la única posible diferencial de f en a, luego f será diferenciable en a si, y sólo si, dicha aplicación lineal cumple la condición que caracteriza a la diferencial. En resumen:

- *Para un campo escalar $f : \Omega \to \mathbb{R}$ donde Ω 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:*

$$\lim_{x \to a} \frac{f(x) - f(a) - \big(\nabla f(a) \,\big|\, x - a\big)}{\| x - a \|} = 0 \tag{3}$$

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

$$Df(a)(x) = \big(\nabla f(a) \,\big|\, x\big) \qquad \forall x \in \mathbb{R}^N \tag{4}$$

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:

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

- *Sea* $\Phi : L(\mathbb{R}^N,\mathbb{R}) \to \mathbb{R}^N$ *la aplicación definida por* $\Phi(T) = \sum_{j=1}^{N} T(e_j)\,e_j$ *para toda* $T \in L(\mathbb{R}^N,\mathbb{R})$. *Se tiene que* Φ *es lineal, biyectiva y conserva la norma, luego permite identificar totalmente los espacios normados* $L(\mathbb{R}^N,\mathbb{R})$ *y* $\mathbb{R}^N$.

La linealidad de Φ es evidente. Si $T \in L(\mathbb{R}^N,\mathbb{R})$ y $\Phi(T) = 0$, tendremos $T(e_j) = 0$ para todo $j \in I_N$, pero siendo $\{e_1, e_2, \ldots, e_N\}$ una base, de la linealidad de T deducimos claramente que $T = 0$, luego Φ es inyectiva. Para ver que también es sobreyectiva, fijado $y \in \mathbb{R}^N$ definimos $T(x) = (y|x)$ para todo $x \in \mathbb{R}^N$. Como el producto escalar es lineal en la segunda variable, tenemos $T \in L(\mathbb{R}^N,\mathbb{R})$, pero es claro que $\Phi(T) = \sum_{j=1}^{N} (y|e_j)\,e_j = y$.

Para ver que Φ conserva la norma, fijamos $T \in L(\mathbb{R}^N,\mathbb{R}) \setminus \{0\}$, pues si $T = 0$ no hay nada que comprobar. Tomando $y = \Phi(T) \neq 0$, acabamos de ver que $T(x) = (y|x)$ para todo $x \in \mathbb{R}^N$, luego

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

y bastará probar que este último máximo coincide con $\|y\|$. Pero eso es exactamente lo que nos dice la desigualdad de Cauchy-Schwartz: $|(y|x)| \leqslant \|y\|\,\|x\| = \|y\|$ para todo $x \in S$, y la igualdad se consigue tomando $x = y/\|y\| \in S$, que depende linealmente de y. ■

Recuérdese que $\mathbb{R}^N$ también se identificó en su momento con $L(\mathbb{R},\mathbb{R}^N)$, pero de forma diferente. Para ver un vector $y \in \mathbb{R}^N$ como aplicación lineal $T : \mathbb{R} \to \mathbb{R}^N$, definíamos $T(t) = t\,y$ para todo $t \in \mathbb{R}$, luego usábamos el producto por escalares de $\mathbb{R}^N$. Ahora, para ver y como una aplicación lineal $T : \mathbb{R}^N \to \mathbb{R}$ tomamos $T(x) = (y|x)$ para todo $x \in \mathbb{R}^N$, usando el producto escalar de $\mathbb{R}^N$.

Volviendo a nuestro campo escalar f, definido en un abierto Ω de $\mathbb{R}^N$, y diferenciable en un punto $a \in \Omega$, el resultado recién obtenido permite identificar la aplicación lineal $Df(a)$ con el vector gradiente $\nabla f(a)$, pero esta identificación es diferente de la que hacíamos con el vector derivada, cuando f era una función de variable real con valores en $\mathbb{R}^N$. Como vimos en (4), ahora $Df(a)$ se identifica con la forma lineal en $\mathbb{R}^N$ que consiste en hacer el producto escalar por el vector gradiente.

El uso en Física del gradiente de un campo escalar se explica fácilmente a partir de la interpretación que hicimos de las derivadas direccionales. Si el campo escalar f es diferenciable en un punto $a \in \Omega$, sus derivadas direccionales en el punto a vienen dadas por

$$f_u{}'(a) = Df(a)(u) = \big(\nabla f(a)\,|\,u\big) \qquad \forall u \in S$$

Supongamos ahora que $\nabla f(a) \neq 0$ y consideremos la dirección $v = \nabla f(a)/\|\nabla f(a)\|$. Para toda dirección $u \in S$ tenemos entonces

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

Equivalentemente, usando que $-f_v{}'(a) = f_{-v}{}'(a)$, tenemos

$$f_{-v}{}'(a) \leqslant f_u{}'(a) \leqslant f_v{}'(a) \qquad \forall u \in S$$

Concluimos por tanto que

$$f_v{}'(a) = \text{máx}\left\{ f_u{}'(a) : u \in S \right\} > 0 \qquad \text{y} \qquad f'_{-v}(a) = \text{mín}\left\{ f_u{}'(a) : u \in S \right\} < 0$$

La interpretación física de estas igualdades es clara. Por una parte, al desplazarnos desde el punto a en la dirección v, que es la del vector gradiente $\nabla f(a)$, conseguimos la máxima tasa de aumento del campo por unidad de longitud, es decir, hacemos que el campo aumente lo más rápidamente posible. Esa máxima tasa de aumento es la norma euclídea del vector gradiente, $f_v{}'(a) = \|\nabla f(a)\|$, luego un "pequeño" desplazamiento en la dirección 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, en la dirección $-v$, opuesta a la del vector gradiente, el valor del campo disminuye lo más rápidamente posible, "aproximadamente" a razón de $\|\nabla f(a)\|$ unidades de campo por unidad de longitud.

Resaltamos que todo 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 que a es un *punto crítico* o un *punto estacionario* del campo f.

8.4. 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 Ω es un subconjunto no vacío, abierto y conexo, de $\mathbb{R}^2$. Se trata por tanto del conjunto

$$S = \text{Gr}\, f = \left\{ \big(x, y, f(x,y)\big) : (x,y) \in \Omega \right\} \subset \mathbb{R}^3$$

Nótese que S determina de forma única a la función 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 $(x,y,z) \in S$, en cuyo caso $f(x,y)$ es el único $z \in \mathbb{R}$ que verifica dicha condición. Se dice que la igualdad $z = f(x,y)$ con $(x,y) \in \Omega$ es la *ecuación explícita* de la superficie S.

Cuando f es diferenciable en un punto $(x_0, y_0) \in \Omega$, vamos a interpretar geométricamente el vector gradiente $\nabla f(x_0, y_0)$ en términos de la superficie S. Para abreviar, escribimos

$$z_0 = f(x_0, y_0), \qquad \alpha = \frac{\partial f}{\partial x}(x_0, y_0) \quad \text{y} \quad \beta = \frac{\partial f}{\partial y}(x_0, y_0)$$

y consideramos el plano afín Π definido por la ecuación

$$\alpha(x - x_0) + \beta(y - y_0) - (z - z_0) = 0$$

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

$$g(x,y) = z_0 + \alpha(x - x_0) + \beta(y - y_0) \qquad \forall (x,y) \in \mathbb{R}^2 \tag{5}$$

Usando (4), 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), lo cual significa que

$$\lim_{(x,y)\to(x_0,y_0)} \frac{|\,f(x,y)-g(x,y)\,|}{\|\,(x,y)-(x_0,y_0)\,\|} = 0 \tag{6}$$

Geométricamente, la distancia (vertical) entre el punto $\big(x,y,f(x,y)\big)$ de la superficie S y el correspondiente punto $\big(x,y,g(x,y)\big)$ del plano Π, tiende a cero "mucho más rápidamente" que $\|\,(x,y)-(x_0,y_0)\,\|$, así que la siguiente definición está plenamente justificada.

Sea $S=\operatorname{Gr} f$ 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 Π de ecuación

$$\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) - (z-z_0) = 0 \tag{7}$$

es el **plano tangente** a la superficie S en el punto (x_0,y_0,z_0). Decimos también que el vector

$$\left(\frac{\partial f}{\partial x}(x_0,y_0)\,,\;\frac{\partial f}{\partial y}(x_0,y_0)\,,\;-1\right)\in\mathbb{R}^3 \tag{8}$$

es un **vector normal** a la superficie S en el punto (x_0,y_0,z_0). Ello se explica porque, en vista de (7), se trata de un vector normal al plano tangente Π, es claramente ortogonal a todos los vectores de la forma $u-v$ con $u,v\in\Pi$ es decir, a todas las rectas contenidas en Π.

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 (8), normal al plano dado por (7), o lo que es lo mismo, considerar la función afín g dada por (5), 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 S cerca del punto (x_0,y_0,z_0) cuando se cumple (6), lo cual equivale a que f sea diferenciable en (x_0,y_0), como vimos en (3). 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, servirá también para reafirmar la interpretación geométrica que acabamos de hacer.