Tema 10

Teorema del valor medio

Podría decirse que hasta ahora sólo hemos sentado las bases para el estudio del cálculo diferencial en varias variables. Hemos introducido el concepto general o abstracto de función diferenciable y comprobado algunas propiedades básicas, como su carácter local o su relación con la continuidad. También hemos visto cómo se concreta dicho concepto en los tres casos particulares que más nos interesan, vector derivada, vector gradiente y matriz jacobiana, viendo al mismo tiempo diversas interpretaciones geométricas y físicas. Por último disponemos de algunos métodos para estudiar la diferenciabilidad de una función y calcular su diferencial, entre los que destaca la regla de la cadena.

Ha llegado ya el momento de estudiar los teoremas fundamentales del cálculo diferencial, empezando como es lógico por intentar generalizar el *teorema del valor medio*, que es sin duda el resultado más importante del cálculo diferencial en una variable. Obtendremos fácilmente una versión del teorema para funciones que parten de un espacio normado arbitrario y toman valores reales, que en particular se aplica a los campos escalares en $\mathbb{R}^N$. Para funciones con valores vectoriales las cosas se complican, la versión literal del teorema es falsa, incluso para funciones de una variable real con valores en $\mathbb{R}^2$. Sin embargo, probaremos una versión más débil, que sí es cierta a plena generalidad y permite deducir algunas consecuencias interesantes. Completamos este tema con una consecuencia del teorema del valor medio para funciones reales de una variable real, útil para comprobar la diferenciabilidad de ciertos campos escalares en $\mathbb{R}^N$.

10.1. Caso escalar

Recordemos el teorema del valor medio para funciones reales de variable real:

- *Sean $a,b \in \mathbb{R}$ con $a < b$ y $f : [a,b] \to \mathbb{R}$ una función continua, que suponemos derivable en el intervalo abierto $]a,b[$.*

  (i) **Teorema del valor medio**: *existe $c \in]a,b[$ tal que $f(b) - f(a) = f'(c)(b-a)$*

  (ii) **Desigualdad del valor medio**: *si existe $M \in \mathbb{R}_0^+$ verificando que $|f'(x)| \leqslant M$ para todo $x \in]a,b[$, entonces $|f(b) - f(a)| \leqslant M(b-a)$.*

Cuando se trabaja con funciones reales de variable real, la desigualdad del valor medio no suele destacarse, siendo como es, consecuencia obvia del teorema. Sin embargo, veremos que, cuando se trabaja con espacios normados, el teorema no siempre es cierto, mientras que la desigualdad es válida a plena generalidad. Por otra parte, la desigualdad es suficiente para conseguir varias consecuencias importantes del teorema. Por ejemplo, permite probar que toda función derivable en un intervalo, con derivada acotada, es lipschitziana y, en el caso de que la derivada sea idénticamente nula, la función es constante.

Con vistas a la generalización del teorema anterior, conviene hacer algunas observaciones. Por una parte, el papel del intervalo $[a,b]$ puede hacerlo cualquier segmento entre dos puntos de un espacio normado, mientras que $]a,b[$ debe ser el mismo segmento, excluidos sus extremos. Concretamente, si X es un espacio normado y $a,b \in X$, escribiremos

$$[a,b] = \big\{a+t(b-a) : t \in [0,1]\big\} \qquad \text{y} \qquad]a,b[= \big\{a+t(b-a) : t \in]0,1[\big\}$$

Obsérvese que, cuando $a \neq b$, único caso que nos interesa, se tiene $]a,b[= [a,b] \setminus \{a,b\}$. Tiene sentido decir que $[a,b]$ es el *segmento cerrado* de extremos a y b, pues se trata de un conjunto cerrado, e incluso compacto. Pero conviene notar que, salvo que X tenga dimensión 1, el conjunto $]a,b[$ no es abierto, de hecho $[a,b]$ tiene interior vacío. Por esta razón, no podemos trabajar con una función f definida solamente en el segmento $[a,b]$, pues no tendría sentido hablar de la diferenciabilidad de f. Una solución aceptable consiste en suponer que f está definida en un abierto que contenga al segmento $[a,b]$. Eso sí, supondremos solamente que f es continua en $[a,b]$ y diferenciable en todo punto de $]a,b[$.

Por otra parte, está claro que la tesis del teorema del valor medio tiene perfecto sentido, otra cosa es que sea cierta, para una función con valores en un espacio normado arbitrario Y. Piénsese que, en el caso $X = Y = \mathbb{R}$ multiplicar $b-a$ por la derivada $f'(c)$ es tanto como aplicar al número real $b-a$ la diferencial $Df(c)$. Por tanto, en general, podemos preguntarnos si existe un punto $c \in]a,b[$ que verifique $f(b)-f(a) = Df(c)(b-a)$, pues ambos miembros de esta igualdad son vectores de Y.

Pues bien, expresado de esta forma, el teorema del valor medio, y la desigualdad que de él se deduce, son ciertos, siempre que el espacio normado de llegada sea $\mathbb{R}$, es decir, para funciones con valores escalares:

Teorema del valor medio escalar. *Sea Ω un abierto de un espacio normado X, $a,b \in \Omega$ tales que $[a,b] \subset \Omega$ y $f : \Omega \to \mathbb{R}$ una función continua en $[a,b]$ y diferenciable en $]a,b[$. Entonces, existe $c \in]a,b[$ tal que*

$$f(b)-f(a) = Df(c)(b-a) \tag{1}$$

Como consecuencia, si $M \in \mathbb{R}_0^+$ verifica que $\|Df(x)\| \leqslant M$ para todo $x \in]a,b[$, se tendrá:

$$|f(b)-f(a)| \leqslant M\|b-a\| \tag{2}$$

Demostración. Definimos una función $\varphi : [0,1] \to \mathbb{R}$, escribiendo $\varphi(t) = f\big(a+t(b-a)\big)$ para todo $t \in [0,1]$. Entonces φ es continua y, por la regla de la cadena, es derivable en $]0,1[$ con derivada dada por

$$\varphi'(t) = Df\big(a+t(b-a)\big)(b-a) \qquad \forall t \in]0,1[$$

El teorema del valor medio para funciones reales de variable real nos da un $t_0 \in]0,1[$ tal que

$$f(b) - f(a) = \varphi(1) - \varphi(0) = \varphi'(t_0) = Df\big(a + t_0\,(b-a)\big)\,(b-a)$$

Tomando $c = a + t_0\,(b-a) \in]a,b[$, obtenemos (1). Para deducir (2) basta usar la definición de la norma en $L(X,\mathbb{R})$:

$$|\,f(b) - f(a)\,| = |\,Df(c)(b-a)\,| \leqslant \|\,Df(c)\,\|\;\|\,b-a\,\| \leqslant M\,\|\,b-a\,\| \qquad \blacksquare$$

En el caso particular de un campo escalar en $\mathbb{R}^N$, usando la desigualdad de Cauchy-Schwartz obtenemos:

- *Sea Ω un abierto de $\mathbb{R}^N$, $a,b \in \Omega$ tales que $[a,b] \subset \Omega$ y $f : \Omega \to \mathbb{R}$ un campo escalar continuo en $[a,b]$ y diferenciable en $]a,b[$. Entonces, existe $c \in]a,b[$ tal que*

  $$f(b) - f(a) = \big(\nabla f(c)\,\big|\,b-a\big)$$

  Por tanto, usando en $\mathbb{R}^N$ la norma euclídea, si existe $M \in \mathbb{R}_0^+$ tal que $\|\,\nabla f(x)\,\| \leqslant M$ para todo $x \in]a,b[$, se tendrá:

  $$|\,f(b) - f(a)\,| \leqslant M\,\|\,b-a\,\|$$

Como se habrá podido observar, la demostración del teorema del valor medio escalar ha consistido en una aplicación muy inmediata del ya conocido para funciones de una variable. De hecho, como sólo se trabaja con la función f en un segmento $[a,b]$, en esencia tenemos una función de una variable. Los resultados anteriores, generalizan formalmente el teorema del valor medio que ya conocíamos, pero en realidad sólo consisten en darse cuenta de que dicho teorema puede aplicarse fácilmente en situaciones más generales. Tiene mayor interés ver lo que ocurre en el caso de un campo vectorial, o incluso para una función entre espacios normados arbitrarios, como enseguida vamos a intentar.

10.2. Caso general

Observamos inmediatamente que la primera afirmación del teorema del valor medio escalar no es cierta para funciones con valores en $\mathbb{R}^2$, incluso aunque sean funciones de una variable real. Concretamente, basta considerar la función $f : \mathbb{R} \to \mathbb{R}^2$ dada por $f(t) = (\cos t, \operatorname{sen} t)$ para todo $t \in \mathbb{R}$, que claramente es derivable en todo punto de $\mathbb{R}$ con

$$f'(t) = \big(-\operatorname{sen} t\,,\, \cos t\big) \neq (0,0) \qquad \forall t \in \mathbb{R}$$

Tomando $a \in \mathbb{R}$ arbitrario y $b = a + 2\pi$ se tiene $f(b) = f(a)$. Si existiese $c \in [a,b]$ verificando que $f(b) - f(a) = Df(c)(b-a) = 2\pi f'(c)$, se tendría $f'(c) = (0,0)$, lo cual es imposible.

Sin embargo, probaremos a plena generalidad la segunda de las afirmaciones del teorema del valor medio escalar, es decir, la desigualdad del valor medio. Ciertamente es más débil que el teorema, pero será suficiente para deducir interesantes consecuencias.

La dificultad para probarlo se concentra en el caso de funciones de una variable real, que es el que resolvemos previamente.

Lema. *Sea* Y *un espacio normado y sean* $g : [0,1] \to Y$ *y* $\alpha : [0,1] \to \mathbb{R}$ *dos funciones continuas en* $[0,1]$ *y derivables en* $]0,1[$, *verificando que*

$$\| g'(t) \| \leqslant \alpha'(t) \qquad \forall t \in]0,1[\tag{3}$$

Se tiene entonces la siguiente desigualdad:

$$\| g(1) - g(0) \| \leqslant \alpha(1) - \alpha(0) \tag{4}$$

Demostración. Fijado $\varepsilon > 0$, consideramos el conjunto

$$\Lambda = \big\{ t \in [0,1] \, : \, \| g(t) - g(0) \| \leqslant \alpha(t) - \alpha(0) + \varepsilon \, t + \varepsilon \big\}$$

y a poco que se piense, la demostración estará casi concluida si probamos que $1 \in \Lambda$.

Por ser g y α continuas, la función $\varphi : [0,1] \to \mathbb{R}$ definida por

$$\varphi(t) = \| g(t) - g(0) \| - \big(\alpha(t) - \alpha(0) \big) - \varepsilon \, t \qquad \forall t \in [0,1]$$

también es continua, de lo que deduciremos dos consecuencias:

En primer lugar, como $\varphi(0) = 0$, la continuidad de φ en 0 nos permite encontrar $\eta \in]0,1[$ tal que, para $t \in [0,\eta]$ se tenga $\varphi(t) < \varepsilon$, con lo que $[0,\eta] \subset \Lambda$.

Por otra parte, como $\Lambda = \{ t \in [0,1] \, : \, \varphi(t) \leqslant \varepsilon \}$, la continuidad de φ nos dice que Λ es un subconjunto cerrado de $[0,1]$, luego es compacto, y en particular tiene máximo. Sea pues $t_0 = \text{máx}\, \Lambda$ y anotemos que $t_0 \geqslant \eta > 0$. Nuestro objetivo es probar que $t_0 = 1$, así que supondremos que $t_0 < 1$ para llegar a una contradicción.

Al ser $0 < t_0 < 1$, tenemos que g y α son derivables en t_0, luego existe $\delta > 0$ tal que, para todo $t \in [0,1]$ con $|t - t_0| \leqslant \delta$, se tiene

$$\begin{aligned} \| g(t) - g(t_0) - g'(t_0)(t - t_0) \| &\leqslant \frac{\varepsilon}{2} |t - t_0| \qquad \text{y también} \\ |\alpha(t) - \alpha(t_0) - \alpha'(t_0)(t - t_0)| &\leqslant \frac{\varepsilon}{2} |t - t_0| \end{aligned}$$

Obviamente podemos suponer que $t_0 + \delta < 1$, para tomar $t = t_0 + \delta$ y obtener

$$\begin{aligned} \| g(t_0 + \delta) - g(t_0) - \delta g'(t_0) \| &\leqslant \frac{\varepsilon}{2} \delta \qquad \text{así como} \\ |\alpha(t_0 + \delta) - \alpha(t_0) - \delta \alpha'(t_0)| &\leqslant \frac{\varepsilon}{2} \delta \end{aligned} \tag{5}$$

Llegaremos a contradicción viendo que $t_0 + \delta \in \Lambda$. Para ello, usando la primera desigualdad de (5), la hipótesis (3) con $t = t_0$, y el hecho de que $t_0 \in \Lambda$, tenemos:

$$\begin{aligned} \| g(t_0 + \delta) - g(0) \| &\leqslant \| g(t_0 + \delta) - g(t_0) - \delta g'(t_0) \| + \delta \| g'(t_0) \| + \| g(t_0) - g(0) \| \\ &\leqslant \frac{\varepsilon}{2} \delta + \delta \alpha'(t_0) + \alpha(t_0) - \alpha(0) + \varepsilon t_0 + \varepsilon \end{aligned} \tag{6}$$

Por otra parte, usando la segunda desigualdad de (5) tenemos también

$$\begin{aligned}\delta\alpha'(t_0)+\alpha(t_0)&=\alpha(t_0+\delta)-\big(\alpha(t_0+\delta)-\alpha(t_0)-\delta\alpha'(t_0)\big)\\&\leqslant\alpha(t_0+\delta)+\frac{\varepsilon}{2}\delta\end{aligned}\tag{7}$$

De (6) y (7) deducimos claramente que

$$\begin{aligned}\|g(t_0+\delta)-g(0)\|&\leqslant\frac{\varepsilon}{2}\delta+\alpha(t_0+\delta)+\frac{\varepsilon}{2}\delta-\alpha(0)+\varepsilon t_0+\varepsilon\\&=\alpha(t_0+\delta)-\alpha(0)+\varepsilon(t_0+\delta)+\varepsilon\end{aligned}$$

es decir, que $t_0+\delta\in\Lambda$. Esto es una clara contradicción, ya que $t_0+\delta>t_0=\max\Lambda$.

Así pues hemos comprobado que $t_0=1$ y en particular $1\in\Lambda$, es decir

$$\|g(1)-g(0)\|\leqslant\alpha(1)-\alpha(0)+2\varepsilon$$

Como esto es válido para todo $\varepsilon\in\mathbb{R}^+$, tenemos (4), como queríamos. ■

En el caso $Y=\mathbb{R}^2$, aunque también se podría hacer en general, el resultado anterior tiene una interpretación física que lo hace muy plausible. Pensemos que g describe un movimiento curvilíneo en el plano y α un movimiento rectilíneo, durante el mismo intervalo de tiempo. La hipótesis (3) significa que la celeridad del primer móvil nunca supera a la del segundo. Entonces $\|g(1)-g(0)\|$ es la distancia entre las posiciones inicial y final del móvil más lento, que será menor o igual que la longitud de la trayectoria recorrida, pero esta longitud ha de ser a su vez menor o igual que $\alpha(1)-\alpha(0)$, que es la distancia recorrida por el móvil más rápido.

Desigualdad del valor medio. *Sean X e Y espacios normados, Ω abierto de X, $a,b\in X$ tales que $[a,b]\subset\Omega$ y $f:\Omega\to Y$ una función. Supongamos que f es continua en $[a,b]$ y diferenciable en $]a,b[$, y que existe $M\in\mathbb{R}_0^+$ tal que $\|Df(x)\|\leqslant M$ para todo $x\in]a,b[$. Se tiene entonces:*

$$\|f(b)-f(a)\|\leqslant M\|b-a\|$$

Demostración. Basta aplicar el lema anterior a las funciones $g:[0,1]\to Y$ y $\alpha:[0,1]\to\mathbb{R}$ definidas, para todo $t\in[0,1]$, por

$$g(t)=f\big((1-t)a+tb\big)\qquad\text{y}\qquad\alpha(t)=M\|b-a\|t$$

Es claro que g y α son continuas en $[0,1]$ y derivables en $]0,1[$ con

$$\begin{aligned}\|g'(t)\|&=\|Df\big((1-t)a+tb\big)(b-a)\|\\&\leqslant\|Df\big((1-t)a+tb\big)\|\,\|(b-a)\|\leqslant M\|b-a\|=\alpha'(t)\qquad\forall t\in]0,1[\end{aligned}$$

Aplicando pues el lema anterior, obtenemos la desigualdad buscada:

$$\|f(b)-f(a)\|=\|g(1)-g(0)\|\leqslant\alpha(1)-\alpha(0)=M\|b-a\|$$ ■

Como primera aplicación de este resultado, es natural ponerse en una situación que permita usarlo con cualquier par de puntos $a,b \in \Omega$, lo cual es bien fácil:

Corolario 1. *Sean X e Y espacios normados, Ω un abierto convexo de X y $f:\Omega \to Y$ una función diferenciable. Supongamos que existe $M \in \mathbb{R}_0^+$ verificando que $\| Df(x) \| \leqslant M$ para todo $x \in \Omega$. Entonces f es lipschitziana con constante M, es decir:*

$$\| f(b) - f(a) \| \leqslant M \| b - a \| \qquad \forall a,b \in \Omega$$

Obsérvese que, en caso de que podamos tomar $M = 0$, esto es, cuando la diferencial Df es idénticamente nula, obtenemos que $f(b) = f(a)$ para cualesquiera $a,b \in \Omega$, es decir, que f es constante. Pero podemos conseguir el mismo resultado con una hipótesis más débil que la convexidad de Ω:

Corolario 2. *Sean X e Y espacios normados, Ω un subconjunto abierto y conexo de X y $f:\Omega \to Y$ una función diferenciable tal que $Df(x) = 0$ para todo $x \in \Omega$. Entonces f es constante.*

Demostración. Fijado $a \in \Omega$, consideramos el conjunto $A = \{x \in \Omega : f(x) = f(a)\}$ y se trata de probar que $A = \Omega$. Por ser f continua, A es un subconjunto cerrado de Ω, pero vamos a comprobar que también es abierto. Dado $x \in A$, como Ω es abierto, existe $r \in \mathbb{R}^+$ tal que $B(x,r) \subset \Omega$. Podemos entonces aplicar el corolario anterior, a la restricción de f al abierto convexo $B(x,r)$, cuya función diferencial es idénticamente nula, obteniendo que f es constante en $B(x,r)$. Por tanto tenemos que $f(y) = f(x) = f(a)$ para todo $y \in B(x,r)$, es decir, $B(x,r) \subset A$. Como $x \in A$ era arbitrario, hemos probado que A es abierto. Puesto que Ω es conexo y $A \neq \emptyset$, porque $a \in A$, concluimos que $A = \Omega$ como se quería. ■

Destacamos una consecuencia obvia: con las mismas hipótesis sobre Ω, si f y g son dos funciones diferenciables en Ω tales que $Df = Dg$, entonces $f - g$ es constante. Dicho de otra forma, una función diferenciable en un subconjunto abierto y conexo de un espacio normado, queda determinada por su función diferencial, salvo adición de una constante.

10.3. Una condición suficiente para la diferenciabilidad

Completamos este tema con una útil aplicación del teorema del valor medio, pero no de las versiones aquí probadas, sino de la que ya conocíamos para funciones reales de variable real. Se trata de un resultado que permite probar la diferenciabilidad de un campo escalar, trabajando solamente con sus derivadas parciales. Deduciremos una cómoda caracterización de los campos escalares de clase C^1. Este resultado también es útil para campos vectoriales, sin más que aplicarlo a sus componentes, que son campos escalares.

Sea pues Ω 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 Ω. 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

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

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 $k \in I_N$, 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 Ω y continuas en a. Esta hipótesis, obviamente más débil que (9), es ya suficiente para la diferenciabilidad de f en a:

- *Sea Ω un abierto de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}$ un campo escalar, $a \in \Omega$ y $k \in I_N$. 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 $j \in I_N \setminus \{k\}$, 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.

Aunque sea repetitivo, merece la pena probar primero el caso $N = 2$, para después hacer el razonamiento análogo en el caso general.

Caso $\mathbf{N = 2}$. Salvo intercambio de las variables, que no afecta a la diferenciabilidad de f, podemos suponer $k = 2$. Por tanto, escribiendo $a = (x_0, y_0)$ la hipótesis (i) nos dice que existe la derivada parcial $\dfrac{\partial f}{\partial y}(x_0, y_0)$, mientras en (ii) estamos suponiendo la existencia de $\dfrac{\partial f}{\partial x}(x, y)$ para todo $(x, y) \in \Omega$ y que la función derivada parcial $\dfrac{\partial f}{\partial x} : \Omega \to \mathbb{R}$ es continua en (x_0, y_0).

Usando en $\mathbb{R}^2$ la norma del máximo, tomamos $r \in \mathbb{R}^+$ de forma que la bola abierta de centro (x_0, y_0) y radio r, que denotaremos simplemente por B, esté contenida en Ω, y se trata de probar (8) en nuestro caso.

Para abreviar, escribimos

$$R(x,y) = f(x,y) - f(x_0,y_0) - \big(\nabla f(x_0,y_0) \,\big|\, (x,y) - (x_0,y_0)\big) \qquad \forall (x,y) \in B$$

Fijado $\varepsilon > 0$ debemos encontrar $\delta \in]0, r[$ tal que, denotando por B_δ a la bola abierta de centro (x_0, y_0) y radio δ, se tenga

$$|R(x,y)| \leqslant \varepsilon \, \| (x,y) - (x_0,y_0) \| \qquad \forall (x,y) \in B_\delta \tag{10}$$

Para todo $(x,y) \in B$, teniendo en cuenta que, obviamente

$$f(x,y) - f(x_0,y_0) = \big(f(x,y) - f(x_0,y) \big) + \big(f(x_0,y) - f(x_0,y_0) \big) \tag{11}$$

podemos escribir $R(x,y) = R_1(x,y) + R_2(y)$ donde

$$R_1(x,y) = f(x,y) - f(x_0,y) - (x-x_0) \frac{\partial f}{\partial x}(x_0,y_0) \qquad \forall (x,y) \in B$$
$$R_2(y) = f(x_0,y) - f((x_0,y_0) - (y-y_0) \frac{\partial f}{\partial y}(x_0,y_0) \qquad \forall y \in]\, y_0 - r \, , \, y_0 + r \, [$$

así que trabajaremos por separado con los dos sumandos que han aparecido.

Por definición de derivada parcial, podemos tomar $\delta \in]\, 0, r\, [$ de forma que

$$\left| f(x_0,y) - f(x_0,y_0) - (y-y_0) \frac{\partial f}{\partial y}(x_0,y_0) \right| \leqslant \frac{\varepsilon}{2} \, |y-y_0| \quad \forall y \in]y_0 - \delta, y_0 + \delta[\tag{12}$$

Además, la continuidad de $\partial f / \partial x$ en (x_0, y_0) permite conseguir que el mismo δ verifique también que

$$\left| \frac{\partial f}{\partial x}(u,v) - \frac{\partial f}{\partial x}(x_0,y_0) \right| \leqslant \frac{\varepsilon}{2} \qquad \forall (u,v) \in B_\delta \tag{13}$$

Para comprobar (10) y concluir así la demostración, fijamos $(x,y) \in B_\delta$. Es importante tener presente que, en todo lo que sigue, x e y están fijos. En virtud de (12) tenemos directamente

$$|R_2(y)| \leqslant \frac{\varepsilon}{2} |y-y_0| \leqslant \frac{\varepsilon}{2} \, \| (x,y) - (x_0,y_0) \| \tag{14}$$

Para la desigualdad análoga con R_1, trabajamos en el intervalo $J =]x_0 - \delta, x_0 + \delta\,[$ con la función $\psi : J \to \mathbb{R}$ definida por $\psi(t) = f(t,y)$ para todo $t \in J$.

Para cada $t \in J$, que f sea parcialmente derivable con respecto a la primera variable en el punto $(t,y) \in \Omega$, significa precisamente que ψ es derivable en el punto t con $\psi'(t) = \dfrac{\partial f}{\partial x}(t,y)$. Aplicamos a ψ el teorema del valor medio obteniendo un punto $u \in J$ tal que

$$f(x,y) - f(x_0,y) = \psi(x) - \psi(x_0) = (x-x_0)\, \psi'(u) = (x-x_0) \frac{\partial f}{\partial x}(u,y)$$

de donde deducimos que

$$R_1(x,y) = (x-x_0) \left(\frac{\partial f}{\partial x}(u,y) - \frac{\partial f}{\partial x}(x_0,y_0) \right)$$

Por ser $|u-x_0|<\delta$ y $|y-y_0|<\delta$, podemos usar (13) con $v=y$ obteniendo:

$$|R_1(x,y)| \leqslant \frac{\varepsilon}{2}\,|x-x_0| \leqslant \frac{\varepsilon}{2}\,\|(x,y)-(x_0,y_0)\| \tag{15}$$

De las desigualdades (14) y (15), deducimos directamente (10):

$$|R(x,y)| \leqslant |R_1(x,y)| + |R_2(y)| \leqslant \varepsilon\,\|(x,y)-(x_0,y_0)\|$$

Resuelto el caso $N=2$, para que se entienda la forma de tratar el caso $N>2$, resaltamos la diferente naturaleza de los sumandos R_1 y R_2, pues R_2 sólo depende de la variable y, mientras que R_1 depende también de x. Ello hace que la acotación de R_2 pueda deducirse directamente de la definición de la derivada parcial de f respecto de y en el punto (x_0,y_0). Para R_1 hemos tenido que usar la existencia de la derivada parcial de f con respecto a x en todos los puntos de un segmento, para poder aplicar el teorema del valor medio, y la acotación de R_1 se ha deducido de la continuidad de dicha derivada parcial en el punto (x_0,y_0).

Caso general. Se supone de nuevo sin perder generalidad que $k=N$, es decir, que la derivada parcial de la que sólo se sabe su existencia en el punto a, es la última. La definición de B y de la función $R:B\to\mathbb{R}$ son análogas al caso $N=2$. Usando en $\mathbb{R}^N$ la norma del máximo, B es una bola abierta de centro a y radio $r\in\mathbb{R}^+$ tal que $B\subset\Omega$. Definimos $R:B\to\mathbb{R}$ por:

$$R(x) = f(x) - f(a) - \big(\nabla f(a)\,\big|\,x-a\big) \qquad \forall x\in B$$

Fijado $\varepsilon>0$, debemos encontrar $\delta\in]0,r[$ tal que, llamando B_δ a la bola abierta de centro a y radio δ, se tenga la desigualdad análoga a (10):

$$|R(x)| \leqslant \varepsilon\,\|x-a\| \qquad \forall x\in B_\delta \tag{16}$$

Aparece aquí la diferencia con el caso $N=2$, La función R se descompondrá en N sumandos.

En vez de (11), si $a=(a_1,a_2,\ldots,a_N)$, para $x=(x_1,x_2,\ldots,x_N)\in B$ usaremos la igualdad evidente

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

Por tanto, descomponemos R en la forma

$$R(x) = \sum_{j=1}^{N} R_j(x_j,\ldots x_N) \qquad \forall x\in B \tag{17}$$

donde, si $j\in I_N$ y $x_i\in]a_i-r,\,a_i+r[$ para $i=j,j+1,\ldots,N$, estamos escribiendo

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

Por hipótesis, existe $\delta \in]0,r[$ tal que, para $x_N \in \mathbb{R}$ con $|x_N - a_N| < \delta$, se tiene

$$\left| f(a_1,\ldots,a_{N-1},x_N) - f(a) - (x_N - a_N)\frac{\partial f}{\partial x_N}(a) \right| \leqslant \frac{\varepsilon}{N}\,|x_N - a_N| \tag{18}$$

que es la afirmación análoga a (12) en nuestro caso.

Además, para $j \in I_{N-1}$, la continuidad de $\partial f/\partial x_j$ en a permite conseguir que el mismo δ verifique también que, para todo $w \in B_\delta$ se tenga

$$\left| \frac{\partial f}{\partial x_j}(w) - \frac{\partial f}{\partial x_j}(a) \right| \leqslant \frac{\varepsilon}{N} \quad \forall\, j \in I_{N-1} \tag{19}$$

que son $N-1$ desigualdades análogas a la única que teníamos en (13).

Para comprobar (16) y concluir así la demostración, fijamos $x = (x_1,x_2,\ldots x_N) \in B_\delta$. En virtud de (18) tenemos directamente

$$|R_N(x_N)| \leqslant \frac{\varepsilon}{N}|x_N - a_N| \leqslant \frac{\varepsilon}{N}\,\|x-a\| \tag{20}$$

Para acotar los otros $N-1$ sumandos, fijamos $j \in I_{N-1}$ y trabajamos en $J =]a_j - \delta, a_j + \delta[$ con la función $\psi : J \to \mathbb{R}$ definida por

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

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 ψ es derivable en el punto t con

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

Aplicamos a ψ el teorema del valor medio, obteniendo un punto $u \in J$ tal que

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

de donde, usando la definición de R_j, deducimos claramente que

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

Es claro que $w = (a_1,\ldots,a_{j-1}, u, x_{j+1},\ldots,x_N) \in B_\delta$ luego (19) nos dice que

$$|R_j(x_j,\ldots,x_N)| \leqslant \frac{\varepsilon}{N}\,|x_j - a_j| \leqslant \frac{\varepsilon}{N}\,\|x-a\| \tag{21}$$

y esta desigualdad es válida para todo $j \in I_{N-1}$. Teniendo en cuenta (17), de (20) y (21), deducimos finalmente (16):

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

Esto prueba que f es diferenciable en el punto a, como queríamos. ■

El principal interés del resultado anterior estriba en que nos proporciona una caracterización de los campos escalares de clase C^1 en términos de sus derivadas parciales:

- *Si* Ω *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* $x\in\Omega$ *y, para todo* $k\in I_N$, *la función derivada parcial* $\partial f/\partial x_k$ *es continua en* Ω.

La afirmación (i) implica obviamente que f es diferenciable, y acabamos de ver que (ii) también lo implica. Una vez que f es diferenciable, sabemos que la continuidad de la función diferencial Df equivale a la de todas las derivadas parciales de f. ■

Como aplicación del resultado anterior, recordemos por ejemplo el razonamiento usado para probar que un producto de funciones diferenciables es diferenciable. Nos basábamos en el hecho de que la función polinómica $P:\mathbb{R}^2\to\mathbb{R}$ definida por $P(x,y)=xy$ para todo $x,y\in\mathbb{R}^2$, es de clase C^1 en $\mathbb{R}^2$. Comprobar este hecho fue laborioso, pues sólo podíamos usar la definición de diferenciabilidad. Es claro que P es parcialmente derivable en todo punto $(x,y)\in\mathbb{R}^2$ con

$$\frac{\partial P}{\partial x}(x,y)=y \qquad \text{y} \qquad \frac{\partial f}{\partial y}(x,y)=x$$

luego ambas funciones derivadas parciales son continuas en $\mathbb{R}^2$ y el resultado anterior nos dice directamente que $P\in C^1(\mathbb{R}^2)$.

Destacamos una consecuencia del resultado anterior, que facilita el uso del Corolario 2, deducido de la desigualdad del valor medio:

- *Sea* Ω *un abierto conexo de* $\mathbb{R}^N$ *y* $f:\Omega\to\mathbb{R}$ *un campo escalar parcialmente derivable. Si* $\dfrac{\partial f}{\partial x_k}(x)=0$ *para todo* $x\in\Omega$ *y todo* $k\in I_N$, *entonces* f *es constante.*

El resultado anterior nos dice que $f\in C^1(\Omega)$ y en particular f es diferenciable, pero es claro que Df es idénticamente nula en Ω, luego basta usar el Corolario 2. ■

Terminamos este tema con una última aplicación del teorema del valor medio para funciones reales de variable real. Lo hacemos para funciones de dos variables, pero es fácil adivinar y demostrar la versión para N variables con $N>2$. Intuitivamente, si Ω es un abierto de $\mathbb{R}^2$ y un campo escalar parcialmente derivable $f:\Omega\to\mathbb{R}$ verifica que una de sus derivadas parciales es idénticamente nula en Ω, parece natural pensar que f sólo depende de la otra variable. Esto es obviamente falso cuando Ω no es conexo, e incluso no es difícil probar que puede ser falso aunque Ω sea conexo. Lo probaremos cuando Ω es *convexo*, y es fácil adivinar una condición suficiente más general.

- *Sea* Ω *un abierto convexo de* $\mathbb{R}^2$ *y* $f:\Omega\to\mathbb{R}$ *un campo escalar, parcialmente derivable con respecto a la primera variable, con* $\dfrac{\partial f}{\partial x}(x,y)=0$ *para todo* $(x,y)\in\Omega$. *Entonces existe un intervalo* $J\subset\mathbb{R}$ *y una función* $\varphi:J\to\mathbb{R}$ *verificando:*

$$(x,y)\in\Omega \quad\Longrightarrow\quad y\in J \;\;\text{y}\;\; f(x,y)=\varphi(y)$$

Sea π la segunda proyección coordenada en $\mathbb{R}^2$, es decir, $\pi(x,y) = y$ para todo $(x,y) \in \mathbb{R}^2$. Está bien claro que debemos tomar $J = \pi(\Omega)$, pues entonces, para todo $(x,y) \in \Omega$ se tiene obviamente que $y \in J$. Puesto que Ω es conexo π es continua, vemos que J es un subconjunto conexo de $\mathbb{R}$, es decir, un intervalo.

Pasamos ahora a definir la función $\varphi : J \to \mathbb{R}$ que buscamos, para lo cual fijamos $y \in J$ y consideramos el conjunto $H_y = \{x \in \mathbb{R} : (x,y) \in \Omega\}$. Para cualesquiera $a,b \in H_y$ y $t \in [0,1]$, tenemos que $(a,y) \in \Omega$ y $(b,y) \in \Omega$, luego

$$\big((1-t)a + tb\,,\,y\big) = (1-t)(a,y) + t(b,y) \in \Omega$$

ya que Ω es convexo. Por tanto, H_y es un subconjunto convexo de $\mathbb{R}$, es decir, un intervalo. Consideramos la función $h_y : H_y \to \mathbb{R}$ dada por $h_y(x) = f(x,y)$ para todo $x \in H_y$. Fijado $x \in H_y$, que f sea parcialmente derivable con respecto a la primera variable en $(x,y) \in \Omega$, significa precisamente que h_y es derivable en el punto x, y tenemos

$$h_y'(x) = \frac{\partial f}{\partial x}(x,y) = 0 \qquad \forall\, x \in H_y$$

Deducimos que la función h_y es constante, una constante que llamaremos $\varphi(y)$.

Tenemos así una función bien definida $\varphi : J \to \mathbb{R}$. Para todo $(x,y) \in \Omega$ vemos claramente que $y \in J$ y que $x \in H_y$, luego $f(x,y) = h_y(x) = \varphi(y)$. $\blacksquare$

Es fácil adivinar la hipótesis más general sobre Ω que permite obtener el mismo resultado. Consiste en suponer que Ω es conexo y que, para cada $y \in \mathbb{R}$, el conjunto $\{x \in \mathbb{R} : (x,y) \in \Omega\}$ es un intervalo, que puede ser vacío. Por otra parte, si queremos que la función φ sea derivable, basta suponer que f es parcialmente derivable, también con respecto a la segunda variable.