E

Complementos sobre diferenciabilidad

E.1. Caracterización de las funciones de clase C^1

En la siguiente proposición se caracterizan las funciones de clase C^1 mediante una condición donde no intervienen las derivadas parciales. Esta condición es la que se suele utilizar para definir las funciones de clase C^1 cuando E no es finito dimensional.

Proposición E.1 *Una función* $\mathbf{f} : \Omega \to \mathbf{F}$*, definida en un abierto* $\Omega \subset E = \mathbb{R}^n$*, con valores en un espacio normado* $(F, \| \ \|)$ *es de clase* C^1 *si y sólo si* $\mathbf{f}$ *es diferenciable en cada* $\mathbf{x} \in \Omega$ *y la aplicación* $d\mathbf{f} : \Omega \to \mathcal{L}(E, F)$ *es continua.*

DEM: Si $\mathbf{x} \to d\mathbf{f}(\mathbf{x})$ es continua, para cada $\mathbf{u} \in E$ también lo es $\mathbf{x} \to D_{\mathbf{u}}\mathbf{f}(\mathbf{x}) = d\mathbf{f}(\mathbf{x})\mathbf{u}$, ya que

$$\|d\mathbf{f}(\mathbf{x})\mathbf{u} - d\mathbf{f}(\mathbf{a})\mathbf{u}\| \leq \|d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})\| \, \|\mathbf{u}\|$$

En particular, las derivadas parciales $D_i\mathbf{f}(\mathbf{x})$, $1 \leq i \leq n$ son continuas.

Recíprocamente, si las derivadas parciales $D_i\mathbf{f}(\mathbf{x})$, $1 \leq i \leq n$, son continuas, dado $\mathbf{a} \in \Omega$ y $\epsilon > 0$, existe $\delta > 0$ tal que si $\|\mathbf{x} - \mathbf{a}\| < \delta$, para todo $j \in \{1, \cdots n\}$ se cumple $\|D_j\mathbf{f}(\mathbf{x}) - D_j\mathbf{f}(\mathbf{a})\| < \epsilon$. Entonces, para todo $\mathbf{u} \in \mathbb{R}^n$ con $\|\mathbf{u}\|_2 \leq 1$ se verifica

$$\|[d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})]\mathbf{u}\| = \left\| \sum_{j=1}^{n} [D_j\mathbf{f}(\mathbf{x}) - D_j\mathbf{f}(\mathbf{a})]u_j \right\| \leq$$

$$\leq \sum_{j=1}^{n} \|D_j\mathbf{f}(\mathbf{x}) - D_j\mathbf{f}(\mathbf{a})\| \, |u_j| \leq \epsilon \, \|\mathbf{u}\|_1 \leq \sqrt{n}\epsilon \, \|\mathbf{u}\|_2 = \sqrt{n}\epsilon$$

Considerando el supremo de $\|[d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})]\mathbf{u}\|$ cuando $\|\mathbf{u}\|_2 \leq 1$ se obtiene que $\|d\mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{a})\| \leq \sqrt{n}\epsilon$ si $\|\mathbf{x} - \mathbf{a}\| < \delta$. ■

E.2. La definición general de diferencial segunda

La definición de diferencial segunda que hemos dado en 6.1, en términos de la diferenciabilidad de las derivadas parciales, sólo tiene sentido cuando $E = \mathbb{R}^n$. Con la siguiente proposición se pone de manifiesto que la diferencial segunda $d^2\mathbf{f}(\mathbf{a})$ es realmente la diferencial en $\mathbf{a}$ de la diferencial primera $d\mathbf{f}$. Este hecho es el que permite extender la definición de diferencial segunda al caso en que E sea un espacio normado arbitrario.

Previamente conviene recordar que el espacio vectorial $\mathcal{L}(E, F)$, formado por las aplicaciones lineales continuas $L : E \to F$, lo podemos considerar como espacio normado, con la norma

$$\|L\| = \sup\{\|L(\mathbf{x})\|_F : \|\mathbf{x}\|_E \leq 1\}$$

Para lo que sigue, conviene tener presente que para cada $\mathbf{u} \in E$ la evaluación

$$\delta_{\mathbf{u}} : \mathcal{L}(E, F) \to F, \quad \delta_{\mathbf{u}}(L) = L(\mathbf{u})$$

es una aplicación lineal continua, pues, para cada $L \in \mathcal{L}(E, F)$, se cumple

$$\|\delta_{\mathbf{u}}(L)\| \leq M \|L\| \quad \text{con} \quad M = \|\mathbf{u}\|$$

Proposición E.2 *Si $\mathbf{f} : V \to F$ es una aplicación diferenciable definida en un abierto V de un espacio normado E.*

a) Si $\mathbf{g} = d\mathbf{f} : V \to \mathcal{L}(E, F)$ es diferenciable en $\mathbf{a} \in V$ entonces para cada par $(\mathbf{u}, \mathbf{v}) \in E \times E$ existen las derivadas segundas $D_{\mathbf{uv}}\mathbf{f}(\mathbf{a})$, $D_{\mathbf{vu}}\mathbf{f}(\mathbf{a})$ y

$$[d\mathbf{g}(\mathbf{a})(\mathbf{u})]\mathbf{v} = D_{\mathbf{uv}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{vu}}\mathbf{f}(\mathbf{a})$$

b) Cuando $E = \mathbb{R}^n$, una condición necesaria y suficiente para que $\mathbf{g} = d\mathbf{f}$ sea diferenciable en $\mathbf{a} \in V$ es que todas las derivadas parciales $D_i\mathbf{f} : V \to F$, $1 \leq i \leq n$ sean diferenciables en $\mathbf{a}$. En este caso

$$d^2\mathbf{f}(\mathbf{a})(\mathbf{u}, \mathbf{v}) = [d\mathbf{g}(\mathbf{a})(\mathbf{u})](\mathbf{v})$$

DEM: a) Como $\mathbf{f}$ es diferenciable en V, para cada $\mathbf{v} \in E$ y cada $\mathbf{x} \in V$ existe la derivada $D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = d\mathbf{f}(\mathbf{x})\mathbf{v}$. La función $\boldsymbol{\varphi}(\mathbf{x}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{x})$, es el resultado de componer $\mathbf{g} = d\mathbf{f}$ con la evaluación $\delta_{\mathbf{v}} : \mathcal{L}(E, F) \to F$ que es lineal continua. Como la diferencial de una aplicación lineal continua es ella misma, en virtud de la regla de la cadena,

$$\boldsymbol{\varphi}(\mathbf{x}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = \delta_{\mathbf{v}}(d\mathbf{f}(\mathbf{x})) = (\delta_{\mathbf{v}} \circ \mathbf{g})(\mathbf{x})$$

es diferenciable en $\mathbf{a}$, y $d\boldsymbol{\varphi}(\mathbf{a}) = \delta_{\mathbf{v}} \circ d\mathbf{g}(\mathbf{a})$.

Entonces, para cada $\mathbf{u} \in E$ existe la derivada $D_{\mathbf{u}}\boldsymbol{\varphi}(\mathbf{a}) = d\boldsymbol{\varphi}(\mathbf{a})(\mathbf{u})$, igualdad que, en términos de $\mathbf{f}$ y $\mathbf{g} = d\mathbf{f}$, se escribe en la forma

$$D_{\mathbf{u}}(D_{\mathbf{v}}\mathbf{f})(\mathbf{a}) = [\delta_{\mathbf{v}} \circ d\mathbf{g}(\mathbf{a})](\mathbf{u}) = [d\mathbf{g}(\mathbf{a})(\mathbf{u})](\mathbf{v})$$

Obsérvese que la aplicación lineal continua $B = d\mathbf{g}(\mathbf{a}) : E \to \mathcal{L}(E, F)$ asigna a cada $\mathbf{u} \in E$ un elemento $B(\mathbf{u}) \in \mathcal{L}(E, F)$, luego $[B(\mathbf{u})](\mathbf{v}) \in F$ para cada $\mathbf{v} \in E$.

La aplicación lineal $B : E \to \mathcal{L}(E, F)$ se identifica con la aplicación $E \times E \to F$ que asocia al par $(\mathbf{u}, \mathbf{v}) \in E \times E$ el vector $[B(\mathbf{u})](\mathbf{v}) \in F$, y se escribe $B(\mathbf{u}, \mathbf{v})$ en vez de $[B(\mathbf{u})](\mathbf{v})$. Con este convenio $(\mathbf{u}, \mathbf{v}) \to B(\mathbf{u}, \mathbf{v})$ es una aplicación bilineal y si M es la norma de B, como aplicación lineal continua de E en $\mathcal{L}(E, F)$, es fácil ver que: $\|B(\mathbf{u}, \mathbf{v})\| \leq M \|\mathbf{u}\| \|\mathbf{v}\|$ para todo $(\mathbf{u}, \mathbf{v}) \in E \times E$. En lo que sigue $d\mathbf{g}(\mathbf{a}) = d(d\mathbf{f})(\mathbf{a})$ se considera como una aplicación bilineal $B : E \times E \to F$.

Cuando $E = \mathbb{R}^n$ sabemos que esta aplicación bilineal es simétrica, pero la demostración dada 6.7, basada en la simetría de las derivadas parciales, no sirve en la situación actual y debemos adaptarla al caso general:

La función de dos variables reales $\boldsymbol{\psi}(s, t) = \mathbf{f}(\mathbf{a} + s\mathbf{u} + t\mathbf{v})$ está definida y es diferenciable en un cierto entorno de $(0, 0)$, $U = \{(s, t) : |s| < r, |t| < r\}$. Claramente

$$D_1\boldsymbol{\psi}(s, t) = d\mathbf{f}(\mathbf{a} + s\mathbf{u} + t\mathbf{v})\mathbf{u}, \quad D_2\boldsymbol{\psi}(s, t) = d\mathbf{f}(\mathbf{a} + s\mathbf{u} + t\mathbf{v})\mathbf{v}$$

y en virtud de la regla de la cadena $D_1\boldsymbol{\psi}$ y $D_2\boldsymbol{\psi}$ son diferenciables en $(0, 0)$. Teniendo en cuenta que

$$D_2\boldsymbol{\psi}(s, 0) = d\mathbf{f}(\mathbf{a} + s\mathbf{u})\mathbf{v} = D_{\mathbf{v}}\mathbf{f}(\mathbf{a} + s\mathbf{u}), \quad D_1\boldsymbol{\psi}(0, t) = d\mathbf{f}(\mathbf{a} + t\mathbf{v})\mathbf{u} = D_{\mathbf{u}}\mathbf{f}(\mathbf{a} + t\mathbf{v})$$

y usando la definición de derivada parcial se obtiene

$$D_{\mathbf{u}}D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = D_{12}\boldsymbol{\psi}(0, 0), \quad D_{21}\boldsymbol{\psi}(0, 0) = D_{\mathbf{v}}D_{\mathbf{u}}\mathbf{f}(\mathbf{a})$$

Aplicando el teorema 6.7 a la función $\boldsymbol{\psi}$ se concluye que $D_{\mathbf{uv}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{vu}}\mathbf{f}(\mathbf{a})$.
b) Si $E = \mathbb{R}^n$, la función $D_j\mathbf{f}(\mathbf{x})$, es el resultado de componer $d\mathbf{f} : V \to \mathcal{L}(\mathbb{R}^n, F)$ con la evaluación $\delta_{\mathbf{e}_j} : \mathcal{L}(\mathbb{R}^n, F) \to F$ (que es lineal y continua), luego, según la regla de la cadena, $D_j\mathbf{f}(\mathbf{x})$ es diferenciable en $\mathbf{a}$ para todo $j \in \{1 \cdots n\}$.

Recíprocamente, supongamos que todas las derivadas parciales $D_j\mathbf{f}(\mathbf{x})$, son diferenciables en $\mathbf{a}$. Si consideremos F^n dotado de la norma

$$\|(\mathbf{y}_1, \mathbf{y}_2, \cdots \mathbf{y}_n)\|_\infty = \text{máx}\{\|\mathbf{y}_j\| : 1 \leq j \leq n\}$$

la aplicación $\mathbf{x} \to (D_1\mathbf{f}(\mathbf{x}), \cdots D_n\mathbf{f}(\mathbf{x}))$, definida en V con valores en en F^n, es diferenciable en $\mathbf{a}$ (basta razonar como en 5.18).

Para cada $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n) \in F^n$ la aplicación lineal $T_{\mathbf{y}} : \mathbb{R}^n \to F$, definida por $T_{\mathbf{y}}(\mathbf{x}) = \sum_{i=1}^n x_i\mathbf{y}_i$, es continua y verifica $\|T_{\mathbf{y}}(\mathbf{x})\| \leq \|\mathbf{y}\|_\infty \|\mathbf{x}\|_1$, luego $\|T_{\mathbf{y}}\| \leq \|\mathbf{y}\|_\infty$. Esto significa que la aplicación lineal $T : F^n \to \mathcal{L}(\mathbb{R}^n, F)$, $\mathbf{y} \to T_{\mathbf{y}}$ es continua. Como $d\mathbf{f} : U \to \mathcal{L}(\mathbb{R}^n, F)$ es el resultado de componer $\mathbf{x} \to (D_1\mathbf{f}(\mathbf{x}), \cdots D_n\mathbf{f}(\mathbf{x}))$ con la aplicación lineal continua T, en virtud de la regla de la cadena, $d\mathbf{f}$ es diferenciable en $\mathbf{a}$. ■

E.3. Teorema de Schwarz sobre la igualdad de las derivadas mixtas

Aunque las derivadas $D_{11}f(a,b)$, $D_{22}f(a,b)$ no intervienen en la conclusión del teorema de Young y sus corolarios 6.5, 6.6, sin embargo las hipótesis de estos resultados llevan implícita la existencia de estas derivadas. Para el caso de funciones de dos variables hay otras hipótesis, donde no intervienen las derivadas $D_{11}f$, $D_{22}f$, que garantizan la igualdad de las derivadas mixtas $D_{12}\mathbf{f}(\mathbf{a}) = D_{21}\mathbf{f}(\mathbf{a})$. Uno de ellos, que se obtendrá más adelante como consecuencia del teorema de Green afirma que si en un entorno U de (a,b) existen y son continuas las derivadas parciales $D_1f(x,y)$, $D_2f(x,y)$, $D_{12}f(x,y)$ y $D_{21}f(x,y)$ entonces $D_{21}f(x,y) = D_{12}f(x,y)$ para todo $(x,y) \in U$. Una mejora sustancial de este resultado y del corolario 6.6 es el siguiente teorema de Schwarz

Teorema E.3 [Schwarz] *Sea $\mathbf{f} : \Omega \to E$ una aplicación definida en un abierto $\Omega \subset \mathbb{R}^2$ con valores en un espacio normado $(E, \| \|)$ tal que en un entorno de $(a,b) \in \Omega$ existen las derivadas parciales $D_1\mathbf{f}(x,y)$, $D_2\mathbf{f}(x,y)$, $D_{21}\mathbf{f}(x,y)$. Si $D_{21}\mathbf{f}(x,y)$ es continua en (a,b) entonces existe $D_{12}\mathbf{f}(a,b)$ y se cumple $D_{12}\mathbf{f}(a,b) = D_{21}\mathbf{f}(a,b)$.*

DEM: Basta hacer la demostración cuando $(a,b) = (0,0)$ y $D_{21}\mathbf{f}(0,0) = \mathbf{0}$ pues el caso general se reduce a este considerando $\mathbf{g}(x,y) = \mathbf{f}(a+x, b+y) - xyD_{21}\mathbf{f}(a,b)$. En este caso el objetivo es demostrar que existe y vale $\mathbf{0}$ el límite

$$D_{12}\mathbf{f}(0,0) = \lim_{h \to 0} \frac{D_2\mathbf{f}(h,0) - D_2\mathbf{f}(0,0)}{h}$$

Teniendo en cuenta que

$$D_2\mathbf{f}(h,0) = \lim_{k \to 0} \frac{\mathbf{f}(h,k) - \mathbf{f}(h,0)}{k}; \quad D_2\mathbf{f}(0,0) = \lim_{k \to 0} \frac{\mathbf{f}(0,k) - \mathbf{f}(0,0)}{k}$$

todo se reduce a demostrar que existe y vale $\mathbf{0}$ el límite iterado

$$D_{12}\mathbf{f}(0,0) = \lim_{h \to 0} \left(\lim_{k \to 0} \frac{\Delta(h,k)}{hk} \right)$$

donde

$$\Delta(h,k) = \mathbf{f}(h,k) - \mathbf{f}(h,0) - \mathbf{f}(0,k) + \mathbf{f}(0,0)$$

Por hipótesis, para algún $r > 0$, las derivadas parciales $D_1\mathbf{f}$, $D_2\mathbf{f}$, $D_{21}\mathbf{f}$ existen en todo punto de $B(r) = \{(x,y) : |x| < r, |y| < r\}$.
Dado $\epsilon > 0$, en virtud de la continuidad de $D_{21}\mathbf{f}$ en $(0,0)$, existe $\delta \in (0,r)$ tal que

$$|s| < \delta, |t| < \delta \Rightarrow \|D_{21}\mathbf{f}(s,t)\| < \epsilon$$

En lo que sigue se supone $|h| < \delta$ y $|k| < \delta$. Para cada $s \in (-\delta, \delta)$, la función $t \to D_1\mathbf{f}(s,t)$ es derivables en el intervalo $(-\delta, \delta)$, donde su derivada cumple

$\|D_{21}\mathbf{f}(s,t)\| < \epsilon$. Aplicando a la función $t \to D_1\mathbf{f}(s,t)$ el teorema del incremento finito en el intervalo de extremos 0, k, se obtiene

$$|s| < \delta \Rightarrow \|D_1\mathbf{f}(s,k) - D_1\mathbf{f}(s,0)\| \leq \epsilon|k|$$

Consideremos ahora la función $s \to \mathbf{F}(s,k) = \mathbf{f}(s,k) - \mathbf{f}(s,0)$, que está definida para $|s| < r$. Cuando $|s| < \delta$ su derivada cumple

$$\|D_1\mathbf{F}(s,k)\| = \|D_1\mathbf{f}(s,k) - D_1\mathbf{f}(s,0)\| < \epsilon|k|$$

luego, en virtud del teorema del incremento finito $\|\mathbf{F}(h,k) - \mathbf{F}(0,k)\| \leq \epsilon|h||k|$. Como $\mathbf{F}(h,k) - \mathbf{F}(0,k) = \Delta(h,k)$ queda establecido que

$$|h| < \delta,\ |k| < \delta \Rightarrow \|\Delta(h,k)\| \leq \epsilon|h||k|$$

lo que significa que existe el límite doble

$$\lim_{(h,k)\,\to\,(0,0)} \frac{\Delta(h,k)}{hk} = 0$$

Por otra parte, en virtud de las hipótesis, para cada $h \in (-r,r)$ existe el límite

$$\lim_{k\,\to\,0} \frac{\Delta(h,k)}{hk} = \frac{1}{h}(D_2\mathbf{f}(h,0) - D_2 f(0,0))$$

luego, según 3.2, (véase también C.4) debe existir el límite iterado

$$\lim_{h\,\to\,0} \left(\lim_{k\,\to\,0} \frac{\Delta(h,k)}{hk} \right) = 0$$

lo que significa que existe y vale $\mathbf{0}$ la derivada

$$D_{12}\mathbf{f}(0,0) = \lim_{h\,\to\,0} \frac{1}{h}(D_2\mathbf{f}(h,0) - D_2\mathbf{f}(0,0))$$

■