Prueba de los principales teoremas

Teorema 1 (Teorema de completitud en dimensión finita) *Todo subespacio M de dimensión finita de un espacio normado es completo. En particular, todo espacio normado de dimensión finita es completo.*

Demostración: Sea $(y_m)_m$ una sucesión de Cauchy arbitraria de M. Sea $\dim M = n$. Entonces para todo $y_m \in M$, existen los números (únicos) $\alpha_k^{(m)}$, $k = 1, 2, \dots, n$, tales que $y_m = \sum_{k=1}^n \alpha_k^{(m)} e_k$. Como $(y_m)_m$ es de Cauchy, entonces, por un lado

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N},\ |\ \forall m > N,\ \forall p \in \mathbb{N}, \quad \|y_m - y_{m+p}\| < \epsilon,$$

y por el otro, por el Lema técnico 1.28, existe un $c > 0$ tal que

$$\epsilon > \|\sum_{k=1}^n (\alpha_k^{(m)} - \alpha_k^{(m+p)}) e_k\| \geq c \sum_{k=1}^n |\alpha_k^{(m)} - \alpha_k^{(m+p)}| \quad \Longrightarrow \sum_{k=1}^n |\alpha_k^{(m)} - \alpha_k^{(m+p)}| \leq \epsilon/c,$$

i.e., las sucesiones $(\alpha_k^{(m)})_m$ son de Cauchy, y por tanto convergen (los $\alpha_k^{(m)}$ son números reales o complejos), a ciertos α_k, $k = 1, 2, \dots, n$. Sea $y = \sum_{k=1}^n \alpha_j e_j \in M$. Entonces

$$\|y_m - y\| = \left\| \sum_{k=1}^n (\alpha_k^{(m)} - \alpha_k) e_k \right\| \leq \sum_{k=1}^n |\alpha_k^{(m)} - \alpha_k| \|e_k\| \stackrel{m\to\infty}{\longrightarrow} 0 \quad \Longrightarrow \lim_{m\to\infty} y_m = y \in M,$$

i.e., M es completo.

Teorema 2 (Equivalencia de las normas en $\mathbb{R}^n$) *Sea $\mathbb{X}$ un espacio vectorial de dimensión finita. Entonces cualquier norma $\|\cdot\|$ en $\mathbb{X}$ es equivalente a cualquier otra norma en $\mathbb{X}$.*

Demostración: Sea $\dim \mathbb{X} = n$ y sea $e_1, \dots, e_n$ una base de $\mathbb{X}$. Entonces para todo $x \in \mathbb{X}$ existen (y son únicos) los números $\alpha_1, \dots, \alpha_n$ tales que $x = \sum_{j=1}^n \alpha_j e_j$. Entonces, por el Lema técnico 1.28

$$\|x\| \geq c(|\alpha_1| + \cdots + |\alpha_n|) \quad \Longrightarrow |\alpha_1| + \cdots + |\alpha_n| \leq \frac{\|x\|}{c}.$$

Pero

$$\|x\|' = \left\| \sum_{j=1}^n \alpha_j e_j \right\|' \leq \sum_{j=1}^n |\alpha_j| \|e_j\|' \leq \max_j \|e_j\|' \sum_{j=1}^n |\alpha_j| \leq \frac{k}{c} \|x\|,$$

donde hemos denotado por k al $\max_j \|e_j\|$. Intercambiando las normas obtenemos $\|x\| \leq \frac{k'}{c'} \|x\|'$, de donde se sigue el resultado. ■

Teorema 3 (Acotación de las aplicaciones lineales) *Toda aplicación lineal $T : \mathbb{X} \mapsto \mathbb{Y}$ de un espacio normado de dimensión finita $\mathbb{X}$ en otro espacio normado cualquiera $\mathbb{Y}$ es acotada.*

Demostración: Sea $\dim \mathbb{X} = n$ y sea $(e_1, \ldots, e_n)$ una base de $\mathbb{X}$. Para todo $x \in \mathbb{X}$, $x = \sum_{k=1}^{n} x_k e_k$. Entonces

$$\|Tx\| = \left\| T \sum_{k=1}^{n} x_k e_k \right\| \leq \sum_{k=1}^{n} |x_k| \|Te_k\| \leq \max_k \|Te_k\| \sum_{k=1}^{n} |x_k|.$$

Por otro lado, usando el Lema técnico 1.28 tenemos que existe un $c > 0$ tal que

$$\|x\| = \left\| \sum_{k=1}^{n} x_k e_k \right\| \geq c \sum_{k=1}^{n} |x_k|.$$

Combinando ambas tenemos

$$\|Tx\| \leq \max_k \|Te_k\| \frac{\|x\|}{c} \quad \Longrightarrow \quad \|Tx\| \leq \gamma \|x\|,$$

con $\gamma = \max_k \|Te_k\|/c$. ■

Teorema 4 (Regla de la cadena) *Sean $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ y $g : B \subset \mathbb{R}^m \to \mathbb{R}^k$, A, B abiertos tales que $f(A) \subset B$. Supongamos que f es diferenciable en a y g es diferenciable en $f(a)$. Entonces la función compuesta $g \circ f : A \subset \mathbb{R}^n \to \mathbb{R}^k$ es diferenciable en a y $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$. Lo anterior se puede escribir en coordenadas de la siguiente forma:*

$$D_j(g \circ f)_i(a) = \sum_{k=1}^{m} D_k g_i(f(a)) D_j f_k(a), \quad \frac{\partial (g \circ f)_i(a)}{\partial x_j} = \sum_{l=1}^{m} \frac{\partial g_i(f(a))}{\partial x_l} \frac{\partial f_l(a)}{\partial x_j}$$

donde $i = 1, \ldots, n$, $j = 1, \cdots, k$. Matricialmente lo anterior se escribe como: $D(g \circ f)(a) = Dg(f(a)) \cdot Df(a)$ o $J_{g \circ f}(a) = J_g(f(a)) \cdot J_f(a)$.

Demostración: Calculamos

$$\begin{aligned}(g \circ f)(a + h) - (g \circ f)(a) &= g(f(a + h)) - g(f(a)) \\ &= Dg(f(a))(f(a + h) - f(a)) + o(\|f(a + h) - f(a)\|) \\ &= Dg(f(a))[Df(a)(h) + o(\|h\|)] + o(\|f(a + h) - f(a)\|) \\ &= [Dg(f(a)) \cdot Df(a)](h) + Dg(f(a))o(\|h\|) + o(\|f(a + h) - f(a)\|).\end{aligned}$$

Ahora bien, como $Dg(f(a))$ es una aplicación lineal acotada, entonces $\|Dg(f(a))v\| \leq \|Dg(f(a))\| \|v\|$, de donde se sigue que $Dg(f(a))o(\|h\|) = o(\|h\|)$. Por otro lado $f(a + h) - f(a) = Df(a)(h) + o(\|h\|)$, por tanto

$$\|f(a + h) - f(a)\| \leq \|Df(a)\| \|h\| + o(\|h\|) = \left(\|Df(a)\| + \frac{o(\|h\|)}{\|h\|} \right) \|h\|,$$

pero la expresión entre paréntesis de la última igualdad está acotada pues $Df(a)$ es acotada y $o(\|h\|)/\|h\| \to 0$ si $h \to 0$. De lo anterior se deduce que existe un $M > 0$ tal que $\|f(a + h) - f(a)\| \leq M\|h\|$. Luego $o(\|f(a + h) - f(a)\|) = o(\|h\|)$, de donde se sigue el resultado. ■

Teorema 5 (Condición suficiente de diferenciabilidad I) *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, con A abierto y sea $a \in A$. Supongamos que existen las derivadas parciales de cada una de las componentes de f en un entorno de a con respecto a cada una de las variables y son continuas en a, entonces f es diferenciable en a.*

Demostración: Para probarlo es suficiente considerar el caso de una función escalar.

$$
\begin{aligned}
f(a+h) - f(a) &= f(a_1 + h_1, a_2 + h_2, \cdots, a_{n-1} + h_{n-1}, a_n + h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n) \\
&= [f(a_1+h_1, a_2+h_2, \cdots, a_{n-1}+h_{n-1}, a_n+h_n) - f(a_1, a_2+h_2, \cdots, a_{n-1}+h_{n-1}, a_n+h_n)] \\
&\quad + [f(a_1, a_2+h_2, \cdots, a_{n-1}+h_{n-1}, a_n+h_n) - f(a_1, a_2, a_3+h_3, \cdots, a_{n-1}+h_{n-1}, a_n+h_n)] \\
&\quad + \cdots \\
&\quad + [f(a_1, a_2, \cdots, a_{n-1} + h_{n-1}, a_n + h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n + h_n)] \\
&\quad + [f(a_1, a_2, \cdots, a_{n-1}, a_n + h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n)].
\end{aligned}
$$

A cada uno de los sumandos entre corchetes se le puede aplicar el teorema del valor medio de Lagrange (sólo cambia una única variable y f como función de cada variable es continua y diferenciable en todo un intervalo). Así pues

$$
f(a+h) - f(a) = \sum_{k=1}^{n} \frac{\partial f(y_k)}{\partial x_k} h_k
$$

donde $y_1 = (\xi_1, a_2 + h_2, a_3 + h_3, \ldots, a_n + h_n)$, $y_2 = (a_1, \xi_2, a_3 + h_3, \ldots, a_n + h_n)$, $\ldots$, $y_n = (a_1, a_2, a_3, \ldots, \xi_n)$, con $\xi_k \in (a_k, a_k + h_k)$, $k = 1, 2, \ldots, n$. Calculamos ahora

$$
f(a+h) - f(a) - Df(a)(h) = f(a+h) - f(a) - \sum_{k=1}^{n} \frac{\partial f(a)}{\partial x_k} h_k = \sum_{k=1}^{n} \left[\frac{\partial f(y_k)}{\partial x_k} - \frac{\partial f(a)}{\partial x_k} \right] h_k.
$$

Teniendo en cuenta que $|h_k| \leq \|h\|$ para todo $k = 1, \ldots, n$, tenemos entonces que

$$
\frac{|f(a+h) - f(a) - Df(a)(h)|}{\|h\|} \leq \sum_{k=1}^{n} \left| \frac{\partial f(y_k)}{\partial x_k} - \frac{\partial f(a)}{\partial x_k} \right| \to 0 \text{ cuando } h \to 0,
$$

pues $y_k \to a$, $k = 1, \ldots, n$ cuando $h \to 0$ y las derivadas parciales son continuas por hipótesis. ■

Teorema 6 (del valor medio) *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, diferenciable en A abierto y convexo. Sean $a, b \in A$ y sea s el segmento que los une ($s = \{(1-t)a + tb : t \in [0,1]\}$). Entonces, para cada vector $v \in \mathbb{R}^m$ existe un punto z en el interior del segmento s tal que*

$$
\langle v, f(b) - f(a) \rangle = \langle v, Df(z)(b - a) \rangle,
$$

donde $\langle \cdot, \cdot \rangle$ denota el producto escalar en $\mathbb{R}^m$.

Demostración: Primero hemos de notar que como A es abierto y $a, b \in s \subset A$, entonces el conjunto $U = \{x = a + (b-a)t, t \in [0,1]\} \subset A$. Sea $v \in \mathbb{R}^m$ cualquier vector distinto de cero. Definamos la función $F:[0,1]\mapsto\mathbb{R}$ (por f_k denotamos las componentes de f) definida por

$$F(t)=\langle v, f(a+(b-a)t)\rangle=\sum_{k=1}^{m} v_k f_k(a+(b-a)t).$$

Como f es diferenciable en A, entonces F es continua en $[0,1]$ y diferenciable en $(0,1)$ por lo que podemos aplicar el teorema del valor medio de Lagrange de una variable. Así $F(1)-F(0)=\langle v, f(b)-f(a)\rangle=F'(\xi)$ con $\xi\in(0,1)$. Pero, usando la regla de la cadena y las propiedades de la diferenciación de la suma y multiplicación

$$\frac{d}{dt}f_k(a+(b-a)t)=\sum_{j=1}^{n}\frac{\partial}{\partial x_j}f_k(a+(b-a)t)\cdot(b_j-a_j)=Df_k[a+(b-a)t](b-a).$$

Así que

$$F'(\xi)=\sum_{k=1}^{m}v_k\frac{d}{dt}f_k(a+(b-a)\xi)=\sum_{k=1}^{m}v_k Df_k[a+(b-a)\xi](b-a)=\langle v, Df(z)(b-a)\rangle,$$

donde $z=a+(b-a)\xi$. Como $\xi\in(0,1)$ entonces z pertenece al interior de s. ■

Teorema 7 (Schwarz) *Sea $f: A\subset\mathbb{R}^n\to\mathbb{R}^m$ con A abierto, y sea $x_0\in A$. Si en A existen las derivadas parciales $\dfrac{\partial f(x)}{\partial x_i}$, $\dfrac{\partial f(x)}{\partial x_j}$ y $\dfrac{\partial^2 f(x)}{\partial x_j\partial x_i}$ y la derivada $\dfrac{\partial^2 f(x)}{\partial x_j\partial x_i}$ es continua en x_0, entonces en $x_0\in A$ existe la derivada $\dfrac{\partial^2 f(x_0)}{\partial x_i\partial x_j}$ y $\dfrac{\partial^2 f(x_0)}{\partial x_j\partial x_i}=\dfrac{\partial^2 f(x_0)}{\partial x_i\partial x_j}$.*

Demostración: Sin pérdida de generalidad vamos a probarlo para una función escalar de dos variables (¿por qué?). Sea la función $u(x)=f(x,b+k)-f(x,b)$, $x_0=(a,b)$, donde h y k son lo suficientes pequeños de forma que el rectángulo $U=\{(x,y),\, x\in(a,a+h), y\in(b,b+k)\}$ esté contenido en A. Definamos la cantidad

$$\Delta f(a,b)=u(a+h)-u(a)=[f(a+h,b+k)-f(a+h,b)]-[f(a,b+k)-f(a,b)].$$

Como f es continua y diferenciable en A (basta que lo sean en un entorno del punto x_0) u también lo es así que podemos aplicar el teorema del valor medio de Lagrange para funciones de una variable

$$\Delta f(a,b)=u(a+h)-u(a)=u'(x)h=\left(\frac{\partial f(x,b+k)}{\partial x}-\frac{\partial f(x,b)}{\partial x}\right)h,\qquad x\in(a,a+h),$$

que aplicando otra vez el teorema del valor medio de Lagrange a la función $\frac{\partial f(x,y)}{\partial x}$ en la variable y nos conduce a

$$\Delta f(a,b)=hk\frac{\partial}{\partial y}\left(\frac{\partial f(x,y)}{\partial x}\right)=hk\frac{\partial^2 f(x,y)}{\partial y\partial x},\quad x\in(a,a+h),\ y\in(b,b+k).$$

Escojamos ahora $(x,y)\in U$ donde $U=\{(x,y),\, x\in(a,a+h), y\in(b,b+k)\}$, siendo h y k las cantidades anteriores, es el rectángulo usado anteriormente. Como $\frac{\partial^2 f(x,y)}{\partial y\partial x}$ es continua en (a,b) entonces para cualquier $\epsilon > 0$ existe una bola $B(x_0,\delta)$ tal que para todos $(x,y) \in B(x_0,\delta)$

$$\left| \frac{\partial^2 f(x,y)}{\partial y \partial x} - \frac{\partial^2 f(a,b)}{\partial y \partial x} \right| < \frac{\epsilon}{2}.$$

Por tanto, si escogemos $U \subset B(x_0,\delta)$ tenemos

$$\left| \frac{\Delta f(a,b)}{hk} - \frac{\partial^2 f(a,b)}{\partial y \partial x} \right| < \frac{\epsilon}{2}. \tag{1}$$

Como existe $\dfrac{\partial f}{\partial y}$, entonces

$$\begin{aligned}\lim_{k\to 0} \frac{\Delta f(a,b)}{hk} &= \lim_{k\to 0} \frac{1}{h} \left[\frac{f(a+h,b+k) - f(a+h,b)}{k} - \frac{f(a,b+k) - f(a,b)}{k} \right] \\ &= \frac{1}{h} \left[\frac{\partial f(a+h,b)}{\partial y} - \frac{\partial f(a,b)}{\partial y} \right],\end{aligned}$$

luego, tomando el límite $k \to 0$ en (1), tenemos

$$\left| \frac{1}{h} \left[\frac{\partial f(a+h,b)}{\partial y} - \frac{\partial f(a,b)}{\partial y} \right] - \frac{\partial^2 f(a,b)}{\partial y \partial x} \right| \leq \frac{\epsilon}{2}.$$

Tomando el límite $h \to 0$ en la expresión anterior se deduce que

$$\lim_{h\to 0} \frac{1}{h} \left[\frac{\partial f(a+h,b)}{\partial y} - \frac{\partial f(a,b)}{\partial y} \right] = \frac{\partial^2 f(a,b)}{\partial y \partial x},$$

pero, por definición, si existe el límite del miembro izquierdo este es igual a $\frac{\partial^2 f(a,b)}{\partial x \partial y}$. ■

Teorema 8 (Heffter-Young) *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, A abierto y sea $a \in A$. Supongamos que existen las derivadas parciales $\dfrac{\partial f(x)}{\partial x_i}$, y $\dfrac{\partial f(x)}{\partial x_j}$ en un entorno de a y son diferenciables en a. Entonces $\dfrac{\partial^2 f(a)}{\partial x_j \partial x_i} = \dfrac{\partial^2 f(a)}{\partial x_i \partial x_j}$.*

Demostración: Nuevamente, y sin pérdida de generalidad, vamos a probarlo para una función escalar de dos variables. Definamos la cantidad

$$\Delta f(a,b) = [f(a+h,b+h) - f(a+h,b)] - [f(a,b+h) - f(a,b)],$$

y sea la función $G(x) = f(x,b+h) - f(x,b)$. Está claro que $\Delta f(a,b) = G(a+h) - G(a)$ y que G es una función de una variable que por condición es continua y diferenciable en un entorno de (a,b). luego existe un $\xi \in (a, a+h)$ tal que

$$\Delta f(a,b) = hG'(\xi) = h \left[\frac{\partial f(\xi,b+h)}{\partial x} - \frac{\partial f(\xi,b)}{\partial x} \right].$$

Como $\frac{\partial f}{\partial x}$ es diferenciable por hipótesis entonces

$$\frac{\partial f(\xi,b+h)}{\partial x} = \frac{\partial f(a,b)}{\partial x} + \frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial x}\right)(\xi - a) + \frac{\partial}{\partial y}\left(\frac{\partial f(a,b)}{\partial x}\right) h + o\left(\sqrt{(\xi-a)^2 + h^2}\right),$$

$$\frac{\partial f(\xi,b)}{\partial x}=\frac{\partial f(a,b)}{\partial x}+\frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial x}\right)(\xi-a)+o\left(\xi-a\right),$$

que, teniendo en cuenta que $o\left(\sqrt{(\xi-a)^2+h^2}\right)=o(h)$ (pues $|\xi-a|<|h|$) nos conduce a

$$\Delta f(a,b)=\frac{\partial}{\partial y}\left(\frac{\partial f(a,b)}{\partial x}\right)h^2+ho(h)$$

de donde se sigue que

$$\lim_{h\to 0}\frac{\Delta f(a,b)}{h^2}=\frac{\partial}{\partial y}\left(\frac{\partial f(a,b)}{\partial x}\right)=\frac{\partial^2 f(a,b)}{\partial y\partial x}$$

Sea ahora la función $H(y)=f(a+h,y)-f(a,y)$. Como $\Delta f(a,b)=H(b+h)-H(b)$ tenemos, de forma análoga que existe un $\xi\in(b,b+h)$ tal que

$$\Delta f(a,b)=hH'(\xi)=h\left[\frac{\partial f(a+h,\xi)}{\partial y}-\frac{\partial f(a,\xi)}{\partial y}\right].$$

Si usamos la diferenciabilidad de $\frac{\partial f}{\partial y}$ obtenemos, de forma completamente análoga

$$\Delta f(a,b)=\frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial y}\right)h^2+ho(h)\quad\Longrightarrow\quad\lim_{h\to 0}\frac{\Delta f(a,b)}{h^2}=\frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial y}\right),$$

de donde se sigue el resultado. ■

Teorema 9 (de Taylor con resto de Lagrange) *Supongamos que $f:A\subset\mathbb{R}^n\mapsto\mathbb{R}$, $f\in C^k(A)$. Sea $a\in A$ y asumamos que el intervalo $[a,a+h]\subset A$ para cierto $h\neq 0$. Entonces*

$$f(a+h)=f(a)+\sum_{l=1}^{k-1}\frac{1}{l!}D^l f(a)(h)+r_k(a,h),$$

donde

$$r_k(a,h)=\frac{1}{k!}D^k f(a+\xi h)(h),\quad \xi\in(0,1).$$

Demostración: Definamos la función $\phi(t)=f(a+ht)$. Como $f\in C^k(A)$ y $l:\mathbb{R}\mapsto\mathbb{R}^n$, $l(t)=x+th$, es $C^\infty(A)$, entonces usando el Teorema de la regla de la cadena la función $\phi:\mathbb{R}\mapsto\mathbb{R}$, $\phi(t)=(f\circ l)(t)$ es $C^k([0,1])$, es decir, ϕ es continuamente diferenciable hasta orden k en $[0,1]$. Además,

$$\phi'(t)=Df(x+ht)\cdot h=\sum_{i=1}^n\frac{\partial f(x+ht)}{\partial x_i}h_i,$$

$$\phi''(t)=\frac{d}{dt}\left[\sum_{i=1}^n\frac{\partial f(x+ht)}{\partial x_i}h_i\right]=\sum_{i=1}^n\left[\sum_{j=1}^n\frac{\partial}{\partial x_j}\left(\frac{\partial f(x+ht)}{\partial x_i}\right)h_j\right]h_i=\sum_{i,j=1}^n\frac{\partial^2 f(x+ht)}{\partial x_j\partial x_i}h_jh_i,$$

y, en general,

$$\phi^{(k)}(t)=\sum_{i_1=1}^n\cdots\sum_{i_k=1}^n\frac{\partial^k f(x+ht)}{\partial x_{i_1}\cdots\partial x_{i_k}}h_{i_1}\cdots h_{i_k}=\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_n\frac{\partial}{\partial x_n}\right)^k f(x+ht).\qquad(2)$$

Nótese que de (2.4) se tiene $\phi^{(k)}(t) = D^k f(a+ht)(h)$.

Entonces podemos usar el Teorema de Taylor para funciones de una variable. Sea $\tau \in (0,1]$ cualquiera, entonces existe un $\xi \in (0,\tau)$ tal que

$$\phi(\tau) = \phi(0) + \phi'(0)\tau + \cdots + \frac{1}{(k-1)!}\phi^{(k-1)}(0)\tau^{k-1} + r_k(\tau),$$

donde tomamos el resto r_k, por ejemplo, en forma de Lagrange

$$r_k(\tau) = \frac{1}{k!}\phi^{(k)}(\xi)\tau^k.$$

Escogiendo $\tau = 1$ y usando (2) con $x = a$ se deduce el teorema. ■

Si tenemos una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, $f \in C^k(A)$, podemos aplicar el teorema de Taylor a cada una de sus componentes.

Teorema 10 (de la función implícita) *Sea $F : A \subset \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}$ definida en un entorno del punto $(x_0, y_0) \in A$, A abierto de $\mathbb{R}^n \times \mathbb{R}$. Supongamos que:*

1. $F(x,y) := F(x_1, x_2, \dots, x_n, y) \in C^{(p)}(A)$, $p \geq 1$,
2. $F(x_0, y_0) := F(x_{01}, x_{02}, \dots, x_{0n}, y_0) = 0$,
3. $F'_y(x_0, y_0) = \dfrac{\partial F(x_{01}, x_{02}, \dots, x_{0n}, y_0)}{\partial y} \neq 0$.

Entonces existe un abierto $I = I_x \times I_y = (x_0 - h, x_0 + h) \times (y_0 - k, y_0 + k)$[10] alrededor del punto (x_0, y_0), $I \subset A$, y una función $f : I_x \subset \mathbb{R}^n \mapsto I_y \subset \mathbb{R}$ tal que:

1. *$F(x,y) = 0$ en I si y sólo si $f(x) = y$,*
2. *$f(x) \in C^{(p)}(I_x)$.*
3. *Para todo $x \in I_x$, las derivadas parciales de $f(x)$ se calculan por la fórmula*

$$\frac{\partial f(x)}{\partial x_i} := \frac{\partial f(x_1, \dots x_n)}{\partial x_i} = -[F'_y(x, f(x))]^{-1} \cdot [F'_{x_i}(x, f(x))], \quad i = 1, 2, \dots, n, \quad (3)$$

donde por F'_{x_i} denotamos la derivada parcial $\frac{\partial F}{\partial x_i}$.

Demostración:

[10] Análogamente al caso de los intervalos definidos justo antes del Teorema 2.8, definiremos el abierto $(x_0 - h, x_0 + h)$ como $(x_{01} - h_2, x_{01} + h_1) \times (x_{02} + -h_2, x_{02} + h_2) \times \cdots \times (x_{0n} - h_n, x_{0n} + h_n)$.

1. Asumamos que $F'_y(x_0, y_0) > 0$, entonces, como $F'_y(x, y)$ es continua en $(x_0, y_0) \in A$, se tiene que existe todo un entorno $I = I_x \times I_y$ de (x_0, y_0) donde $F'_y(x, y) > 0$. Luego, $F(x, y)$ como función de y es estrictamente creciente. Como $F(x_0, y_0) = 0$ entonces tenemos que

$$0 > F(x_0, y_0 - k) < F(x_0, y_0) < F(x_0, y_0 + k) > 0.$$

Como $F(x, y)$ es continua en A entonces los signos de F se mantienen en todo un entorno de cada punto, i.e., existe un entorno I_x (que por simplicidad asumimos igual al de antes) tal que

$$0 > F(x, y_0 - k) < F(x_0, y_0) < F(x, y_0 + k) > 0, \qquad \forall x \in I_x.$$

Pero para cada $x \in I_x$ la función $h(y) := F(x, y)$ cambia de signo en los extremos del intervalo I_y, luego por el Teorema de Bolzano para cada $x \in I_x$ existe un único $y \in I_y$ tal que $F(x, y) = 0$ (la unicidad es consecuencia de la monotonía de $h(y)$). Definiendo la función $f : I_x \mapsto I_y \subset \mathbb{R}$ de forma que $y = f(x)$ obtenemos una función que cumple con que $F(x, f(x)) = 0$.

2. Que $f(x)$ así definida es continua es inmediato de la construcción anterior pues de esta se sigue que $\forall \epsilon > 0$ (suficientemente pequeño de forma que $y_0 \pm \epsilon \in I_y$) siempre podemos encontrar un $\delta > 0$ tal que $\|x - x_0\| < \delta$ y $x \in I_x$ de forma que $F(x, y)$ cambie de signo en los extremos de $[y_0 - \epsilon, y_0 + \epsilon]$, luego $f(x) \in (y_0 - \epsilon, y_0 + \epsilon)$, i..e., $\forall \epsilon > 0$, $\exists \delta > 0$, tal que $|f(x) - f(x_0)| < \epsilon$, que es justo lo que se quería probar.

3. Para probar que $f \in C^{(1)}(I_x)$ hacemos lo siguiente: Elegimos $\Delta x = h e_i$, e_i i-ésimo vector de la base canónica de $\mathbb{R}^n$. Como f es continua en I_y, entonces podemos escribir $f(x + \Delta x) = f(x) + \Delta y = y + \Delta y$, donde $\Delta y = f(x + \Delta x) - f(x) \to 0$ si $\Delta x \to 0$. Como $F(x, y) = F(x + \Delta x, y + \Delta y) = 0$ entonces, aplicando el teorema del valor medio 6 para funciones escalares sabemos que existe un $\theta \in (0, 1)$ tal que

$$\begin{aligned} 0 =& F(x + \Delta x, y + \Delta y) - F(x, y) = DF(x + \theta h e_i, y + \theta \Delta y) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ =& \frac{\partial F(x + \theta h e_i, y + \theta \Delta y)}{\partial x_i} h + \frac{\partial F(x + \theta h e_i, y + \theta \Delta y)}{\partial y} \Delta y. \end{aligned}$$

Como $F(x, y) \in C^{(1)}(I)$ entonces tomando límites $\Delta x \to 0$ obtenemos

$$0 = \frac{\partial F(x, y)}{\partial x_i} + \frac{\partial F(x, y)}{\partial y} \lim_{h \to 0} \frac{\Delta y}{h}$$

de donde, teniendo en cuenta que $\lim_{h \to 0} \Delta y / h = \frac{\partial f(x)}{\partial x_i}$, y que como hemos visto en todo un entorno de (x_0, y_0) $\dfrac{\partial F(x, y)}{\partial y} \neq 0$, se sigue el resultado (3).

4. Finalmente, para probar que $f(x) \in C^{(p)}(I_x)$ basta notar que en la parte derecha fórmula (3) solo aparece la función f y F. Del apartado anterior se sigue que ambas son diferenciables, entonces derivando consecutivamente (3) se obtiene el resultado ya que cada vez que se deriva en la parte derecha aparecen las derivadas parciales sucesivas de f de un orden menor que las que aparecen en la parte izquierda. ∎

Teorema 11 (Condición necesaria de extremo relativo) *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A abierto, $a \in A$. Supongamos que f tiene en a un extremo relativo. Entonces, si existen las derivadas parciales $\frac{\partial f}{\partial x_k}$, $k = 1, \dots, n$ éstas son iguales a cero en a, i.e., $\frac{\partial f}{\partial x_k}(a) = 0$, $k = 1, \dots, n$. En particular si f es diferenciable en a, entonces $Df(a) = 0$.*

Demostración: Sea $a = (a_1, \cdots, a_n) \in A$, A abierto, un extremo de la función $f : A \subset \mathbb{R}^n \to \mathbb{R}$ diferenciable en a. Definamos la función $\phi_k(x_k) := f(a_1, \dots, a_{k-1}, x_k, a_{k+1}, \dots, a_n)$. Es obvio que ϕ_k es diferenciable en su única variable y tiene un extremo en $x_k = a_k$, luego el Lema de Fermat para funciones escalares de una variable nos asegura que $\phi_k'(a_k) = 0$, de donde se sigue el resultado. ■

Teorema 12 (Condición suficiente de extremo) *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ dos veces diferenciable en $a \in A$, A abierto, y sea $x = a$ un punto crítico de f, i.e., $Df(a) = 0$. Entonces*

1. *Si la segunda diferencial $D^2 f(a)(x)$ es definida positiva en a, entonces f tiene un mínimo relativo en a.*

2. *Si la segunda diferencial $D^2 f(a)(x)$ es definida negativa, entonces f tiene un máximo relativo en a.*

3. *Si la segunda diferencial $D^2 f(a)(x)$ es indefinida, i.e., si existen $x, y \in \mathbb{R}^n$ tales que $D^2 f(a)(x) > 0 > D^2 f(a)(y)$, entonces f tiene un punto de silla en a.*

Demostración: Usamos la fórmula local de Taylor (¿por qué?)

$$f(x) = f(a) + Df(a)(h) + \frac{1}{2} D^2 f(a)(h) + o(\|h\|^2),$$

donde $x = a + h$. Teniendo en cuenta que $Df(a) = 0$ (se asume que $h \neq 0$) tenemos

$$f(x) - f(a) = \frac{1}{2} D^2 f(a)(h) + o(\|h\|^2) = \frac{1}{2}\|h\|^2 \left(\sum_{i,j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} \frac{h_i}{\|h\|} \frac{h_j}{\|h\|} + o(1) \right),$$

donde $o(1)$ es tal que $\lim_{x\to a} o(1) = \lim_{h\to 0} o(1) = 0$. De la expresión anterior se sigue que el signo de $f(x) - f(a)$ coincide con el signo de la expresión dentro de los paréntesis. Denotemos por $Q(h', a)$ la forma cuadrática

$$Q(h', a) = \sum_{i,j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} \frac{h_i}{\|h\|} \frac{h_j}{\|h\|}, \quad h' = \frac{h}{\|h\|}, \quad \|h'\| = 1.$$

Nótese que Q está definida sobre el vector $h' = h/\|h\|$ con norma uno, i.e., Q es una función continua (polinomio) sobre un cerrado y acotado (la esfera $S = \{x \in \mathbb{R}^n | \|x - a\| = 1\}$), luego por el Teorema de Weierstrass Q alcanza su máximo y mínimo absolutos, que denotaremos por M y m respectivamente. Luego para todo $h' \in S$ tenemos $m \leq Q(h', a) \leq M$. Supongamos ahora que Q es definida positiva, entonces $Q(h', a) > 0$ para todo h', y por tanto $m > 0$. Ahora bien como $\lim_{h\to 0} o(1) = 0$, eso indica que para todo $\epsilon > 0$ existe un $\delta > 0$ tal que si $\|x - a\| = \|h\| < \delta$ entonces $|o(1)| < \epsilon$. Si tomamos $\epsilon = m/2$

entonces existe un entorno de $x=a$ tal que $Q(h',a)+o(1)>0$ (¿por qué), y por tanto en dicho entorno de a $f(x)-f(a)>0$, i.e., f tiene un mínimo local. Si por el contrario Q es definida negativa entonces $M<0$. Razonando igual que antes y tomando $\epsilon<|M|/2$, tenemos que $f(x)-f(a)<0$, i.e., f tiene un máximo. Lo anterior prueba 1 y 2.

Probemos ahora 3. Sean a_m y a_M los puntos donde Q alcanza su mínimo y máximo absolutos sobre S (nótese que $\|a_m\|=\|a_M\|=1$), respectivamente. Como Q es indefinida entonces $m<0<M$ (¿por qué?). Por tanto tenemos $0>m\leq Q(h',a)\leq M>0$. Si elegimos ahora $h=x-a=ta_m$ con t suficientemente pequeño tendremos (¿por qué?)

$$f(x)-f(a)=\frac{1}{2}D^2f(a)(h)+o(\|h\|^2)=\frac{1}{2}D^2f(a_mt)(h)+o(\|a_mt\|^2)=\frac{t^2}{2}(m+o(1)),$$

y para ciertos x en un entorno de a el signo de $f(x)-f(a)$ es igual al signo de m, o sea, negativo. Repitiendo el razonamiento para $h=x-a=ta_M$ con t suficientemente pequeño obtendremos que para ciertos x en un entorno de a $f(x)-f(a)>0$, luego en a no puede haber ningún extremo de f. ■