Diferenciabilidad en espacios vectoriales normados

Pedro Gajardo A.

Universidad Técnica Federico Santa María
E-mail address: pedro.gajardo@usm.cl

Índice general

CAPíTULO 1

Diferenciabilidad en espacios vectoriales normados

1.1. Introducción

En el presente documento se aborda el cálculo diferencial de funciones definidas en un espacio vectorial normado (e.v.n.) con valores en otro e.v.n.

Las siguientes dos secciones definen las nociones fundamentales de este apunte. Primero, se comienza definiendo la derivada parcial de una función f en un punto x_0 de su dominio con respecto a un vector v, que denotamos $Df(x_0; v)$. Luego, se definirá el diferencial de una función en un punto x_0, que denotamos $Df(x_0)$. El diferencial $Df(x_0)$ es una función lineal continua que nos da una aproximación de primer orden de la función en el punto x_0 (ver Nota 1.3.2) y la derivada parcial $Df(x_0; v)$ es un vector en el espacio donde f toma sus valores, que nos permite calcular en forma efectiva el diferencial mediante la fórmula $Df(x_0)(v) = Df(x_0; v)$ para todo v (fórmula (1.12)). En estas secciones se entregan también algunas fórmulas para el cálculo de la derivada parcial y de el diferencial de la suma de dos funciones, del producto de una función por un escalar, junto con otras reglas de cálculo. El cálculo para la composición de funciones (regla de la cadena) es presentado en la Sección 1.5.

Por otro lado, se presenta el Teorema del valor medio que es fundamental en cálculo diferencial, como por ejemplo, para poder caracterizar propiedades de funciones (monotonía, Lipschitzianidad, convexidad, etc) con respecto a la información que entregan sus diferenciales. Este teorema sólo será válido para funciones con valores en $\mathbb{R}$. Una variante de este teorema, para funciones a valores en un espacio vectorial, esta dado por el Teorema 1.3.7 llamado de los incrementos finitos.

En la Sección 1.4 se introducen las funciones de clase C^1 que constituyen la clase más importante de funciones diferenciables.

Dos teoremas fundamentales se desarrollan en la Sección 1.7: el Teorema de la función inversa y el de la función implícita.

Haciendo uso de la noción de derivada de orden superior presentado en la Sección 1.8, se define la noción de desarrollo limitado de orden N de una función en un punto de su dominio. El resultado fundamental de esta última sección está dado por el Teorema 1.9.3 donde se calcula explícitamente el desarrollo limitado de orden dos de una función de clase C^2.

1.2. Derivada parcial con respecto a un vector

Definición 1.2.1. Dada una función f definida en un abierto A de un e.v.n. X con valores en un e.v.n. Y, se llama derivada parcial de f en $x_0 \in A$, con respecto al vector $v \in X$, al elemento de Y definido por:

$$Df(x_0; v) := \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(x_0 + tv) - f(x_0)}{t} \tag{1.1}$$

cuando el límite existe.

Nota 1.2.1. Si definimos la función $\phi :]-\varepsilon, \varepsilon[\to Y$ por $\phi(t) = f(x_0 + tv)$, es fácil ver que se tiene la fórmula $Df(x_0; v) = \phi'(0)$ (la derivada de ϕ en 0) según el cálculo de funciones de una variable real. Con esto vemos que cuando $Y = \mathbb{R}$ y $\|v\| = 1$, la cantidad $Df(x_0; v)$ se interpreta como la pendiente de f en x_0 en la dirección v.

Definición 1.2.2. La función f de la Definición 1.2.1 se dirá parcialmente derivable en x_0, si $Df(x_0; v)$ existe para todo $v \in X$.

Definición 1.2.3. Si en la Definición 1.2.1 consideramos que $X = \mathbb{R}^n$ y si denotamos por $e_1, ..., e_n$ la base canónica de $\mathbb{R}^n$, entonces la derivada parcial $Df(x_0; e_i)$, cuando existe, la denotaremos $\frac{\partial f}{\partial x_i}(x_0) \in Y$ o bien $\partial_i f(x_0)$ y la llamaremos derivada parcial de f en x_0 con respecto a x_i.

Nota 1.2.2. De la fórmula (1.1) observamos que la derivada parcial de f en x_0 con respecto a x_i

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x_0) &= \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(x_0 + te_i) - f(x_0)}{t} \\ &= \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a_1, ..., a_i + t, ..., a_n) - f(a_1, ..., a_i, ..., a_n)}{t} \end{aligned}$$

corresponde a la derivada de la función de una variable $f(a_1, ..., a_{i-1}, \cdot\, , a_{i+1}, ..., a_n)$ en a_i.

Nota 1.2.3. Cuando se habla de la derivada parcial de una función f con respecto a x_i, se entiende que se trata de la función $\frac{\partial f}{\partial x_i} : A \to Y$ que a cada $x \in A$ le hace corresponder $\frac{\partial f}{\partial x_i}(x) \in Y$.

Teorema 1.2.1. *Si $Df(x_0; v)$ existe para la función f, de la Definición 1.2.1, $Df(x_0; \lambda v)$ también existirá para todo $\lambda \in \mathbb{R}$ y se tiene la igualdad*

$$Df(x_0; \lambda v) = \lambda Df(x_0; v). \tag{1.2}$$

Demostración. Si $\lambda = 0$, la fórmula es evidente. Si $\lambda \neq 0$, hacemos el cambio de variable $s = \lambda t$, y obtenemos

$$Df(x_0; \lambda v) = \lim_{\substack{\frac{s}{\lambda} \to 0 \\ \frac{s}{\lambda} \neq 0}} \lambda \frac{f(x_0 + sv) - f(x_0)}{s} = \lambda \lim_{\substack{s \to 0 \\ s \neq 0}} \frac{f(x_0 + sv) - f(x_0)}{s} = \lambda Df(x_0; v).$$

□

Proposición 1.2.1. *Dadas dos funciones f y g definidas en un abierto A de un e.v.n. X, con valores en un e.v.n. Y y, dos elementos $x_0 \in A$ y $v \in X$ tales que $Df(x_0; v)$ y $Dg(x_0; v)$ existen, entonces*

(i) $D[f+g](x_0;v)$ *existe y se tiene*

$$D[f+g](x_0;v) = Df(x_0;v) + Dg(x_0;v); \tag{1.3}$$

(ii) *Para* $\lambda \in \mathbb{R}, D[\lambda f](x_0;v)$ *existe y se tiene*

$$D[\lambda f](x_0;v) = \lambda Df(x_0;v); \tag{1.4}$$

(iii) *Si* $Y = \mathbb{R}, D[fg](x_0;v)$ *existe y se tiene*

$$D[fg](x_0;v) = g(x_0)Df(x_0;v) + f(x_0)Dg(x_0;v); \tag{1.5}$$

(iv) *Si* $Y = \mathbb{R}$ *y* $f(x_0) \neq 0, D[1/f](x_0;v)$ *existe y se tiene*

$$D[1/f](x_0;v) = -\frac{1}{f^2(x_0)} Df(x_0;v). \tag{1.6}$$

Demostración. Las fórmulas (1.3) y (1.4) son una consecuencia directa de la Definición 1.2.1 y de las propiedades del límite de funciones. Siguiendo la Nota 1.2.1, si definimos las funciones $\phi(t) = f(x_0+tv)$ y $\xi(t) = g(x_0+tv)$, vemos que las fórmulas (1.5) y (1.6) se escriben

$$(\phi\xi)'(0) = \phi'(0)\xi(0) + \phi(0)\xi'(0) \quad \text{y} \quad \left(\frac{1}{\phi}\right)'(0) = -\frac{\phi'(0)}{\phi(0)^2}$$

respectivamente. Estas dos fórmulas fueron demostradas en el curso de cálculo de funciones de una variable real. □

Teorema 1.2.2. *Si f es una función definida en un abierto A de un e.v.n X, con valores en $\mathbb{R}^m$ y, si denotamos $f_1, ..., f_m$ las funciones componentes de f, entonces para $x_0 \in A$ y $v \in X$, la derivada parcial $Df(x_0;v)$ existe si y sólo si las derivadas parciales $Df_i(x_0;v)$ existen para todo $i = 1, ..., m$ y, en tal caso, se tiene que*

$$Df(x_0;v) = (Df_1(x_0;v), ..., Df_m(x_0;v)) \in \mathbb{R}^m. \tag{1.7}$$

Demostración. Es una consecuencia directa de la Definición 1.2.1 y de un resultado visto durante el curso que relacione el limite de una función a valores en un espacio producto con los límites de las funciones componentes respectivas. □

1.3. Diferencial

Definición 1.3.1. Una función f definida en un abierto A de un e.v.n. X, con valores en un e.v.n. Y, se dirá diferenciable en $x_0 \in A$, si existe una función $\ell \in \mathcal{L}(X,Y)$ (ℓ lineal y continua de X en Y) tal que para todo $v \in X$, con $x_0+v \in A$, se tenga

$$f(x_0+v) = f(x_0) + \ell(v) + o(v) \tag{1.8}$$

donde $o(\cdot)$ es una función de X en Y que verifica $o(0) = 0$ y

$$\lim_{\substack{v\to 0\\ v\neq 0}} \frac{o(v)}{\|v\|} = 0. \tag{1.9}$$

Haciendo el cambio de variable $x = x_0 + v \in A$, la relación (1.8) toma la forma

$$f(x) = f(x_0) + \ell(x-x_0) + o(x-x_0) \quad \text{para todo } x \in A.$$

A la función lineal continua ℓ (más adelante veremos que es única) se le llama diferencial de f en x_0 y se le denota usualmente $Df(x_0) \in \mathcal{L}(X,Y)$ o bien $df(x_0)$.

La función f se dirá diferenciable si es diferenciable en todo punto de A y llamamos diferencial de f a la función de A en $\mathcal{L}(X,Y)$ que a cada $x \in A$ le hace corresponder el diferencial $Df(x) \in \mathcal{L}(X,Y)$ de f en x. Al diferencial de f lo denotaremos Df ó df.

Nota 1.3.1. La igualdad (1.8) junto con (1.9) equivale a

$$\lim_{\substack{v\to 0\\ v\neq 0}} \frac{f(x_0+v)-f(x_0)-\ell(v)}{\|v\|} = 0 \tag{1.10}$$

lo que a su vez equivale a decir que: para todo $\varepsilon > 0$ existe $\eta > 0$ tal que

$$\|v\| \leq \eta \Rightarrow \|f(x_0+v)-f(x_0)-\ell(v)\| \leq \varepsilon\|v\|. \tag{1.11}$$

Ejemplo 1.3.1. Toda función $\ell \in \mathcal{L}(X,Y)$ es diferenciable y se tiene $D\ell(x_0) = \ell$ para todo $x_0 \in X$. Esto es una consecuencia inmediata de (1.10) y de la identidad $\ell(x_0+v)-\ell(x_0)-\ell(v) = 0$ para todo $v \in X$.

Teorema 1.3.1. *Si f es una función definida en un abierto A de un e.v.n X, con valores en un e.v.n Y, diferenciable en $x_0 \in A$, entonces existe una única función $\ell \in \mathcal{L}(X,Y)$ que verifica* (1.8).

Demostración. Si ℓ_1 y ℓ_2 verifican (1.8), usando (1.11) que es equivalente a (1.8), se obtiene fácilmente que para todo $\varepsilon > 0$ existe $\eta > 0$ tal que

$$\|v\| \leq \eta \Rightarrow \|\ell_1(v)-\ell_2(v)\| \leq \varepsilon\|v\|$$

y como $\|\frac{\eta}{\|v\|}v\| \leq \eta$ para todo $v \in X$ no nulo, se tendrá

$$\|\ell_1(\frac{\eta}{\|v\|}v)-\ell_2(\frac{\eta}{\|v\|}v)\| \leq \varepsilon\|\frac{\eta}{\|v\|}v\|.$$

Simplificando por $\frac{\eta}{\|v\|}$ obtenemos

$$\|\ell_1(v)-\ell_2(v)\| \leq \varepsilon\|v\| \quad \text{para todo } v \in X$$

y como esto se tiene para todo $\varepsilon > 0$, haciendo tender ε a 0 concluímos que

$$\|\ell_1(v)-\ell_2(v)\| = 0 \quad \text{para todo } v \in X$$

es decir, $\ell_1(v) = \ell_2(v)$ para todo $v \in X$. □

Nota 1.3.2. El diferencial de f en x_0 permite obtener la aproximación lineal afín o aproximación de primer orden de la función f en el punto x_0, que está dada por la función $h : X \to Y$ definida por

$$h(x) := f(x_0) + Df(x_0)(x-x_0)$$

es decir, la única función lineal afín que verifica $h(x_0+v)-f(x_0+v) = o(v)$.

Dos funciones f y h definidas en un abierto A de un e.v.n. X con valores en un e.v.n. Y, se dicen tangentes en $x_0 \in A$ si $f(x_0+v)-h(x_0+v) = o(v)$ donde o es una función de X en Y que verifica $o(0) = 0$ y la relación (1.9).

De lo anterior vemos que la aproximación de primer orden de la función f en $x_0 \in A$ corresponde a la función lineal afín que es tangente a la función f en x_0.

Geométricamente, este hecho se expresa diciendo que en $X \times Y$, el subespacio afín definido por el grafo de la función lineal afín h, es tangente al grafo de f en

$(x_0, f(x_0))$. Recordemos que el grafo de una función f de A en Y es el conjunto $\{(x,z) \in A \times Y \ : \ z = f(x)\}$. Cuando $Y = \mathbb{R}$ decimos que $z = h(x)$ es la ecuación del hiperplano afín tangente al grafo de f en $(x_0, f(x_0))$.

Nota 1.3.3. Si en la Definición 1.3.1 cambiamos las normas en X y Y por otras equivalentes, de (1.11) se desprende fácilmente que f sigue siendo diferenciable en x_0 y, su diferencial en ese punto es el mismo.

Teorema 1.3.2. *Si f es una función definida en un abierto A de un e.v.n X, con valores en un e.v.n Y, diferenciable en $x_0 \in A$, entonces ella es continua en x_0, parcialmente derivable en x_0 y se tiene la igualdad*

$$Df(x_0)(v) = Df(x_0; v) \quad \textit{para todo } v \in X. \tag{1.12}$$

Demostración. Usando la expresión (1.11) y haciendo el cambio de variable $x = x_0 + v$, obtenemos que para todo $\varepsilon > 0$ existe $\eta > 0$ tal que

$$\begin{aligned} \|x - x_0\| \leq \eta \quad &\Rightarrow \quad \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| \leq \varepsilon \|x - x_0\| \\ &\Rightarrow \quad \|f(x) - f(x_0)\| - \|Df(x_0)(x - x_0)\| \leq \varepsilon \|x - x_0\| \end{aligned}$$

y como $Df(x_0)$ es una función lineal continua (Lipschitz), obtenemos

$$\|f(x) - f(x_0)\| \leq (\|Df(x_0)\|_{\mathcal{L}(X,Y)} + \varepsilon)\|x - x_0\| \quad \forall x \in B(x_0, \eta)$$

lo que implica que f es continua en x_0.

Ahora, dado $v \in X$ no nulo, para $v = tv$ (abusando de la notación) la relación (1.10) implica que

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(x_0 + tv) - f(x_0) - Df(x_0)(tv)}{|t|\|v\|} = 0.$$

Multiplicando por $\|v\|$ y usando la linealidad de $Df(x_0)$ se obtiene

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(x_0 + tv) - f(x_0)}{t} = Df(x_0)(v)$$

demostrando que f es parcialmente derivable en x_0 y la igualdad (1.12).

El resultado para $v = 0$ es evidente. $\square$

Nota 1.3.4. La fórmula (1.12) nos da la relación que existe entre el diferencial de una función y su derivada parcial. Su importancia es crucial pues constituye la única forma de calcular el diferencial de una función en un punto. En el teorema anterior también demostramos que si el diferencial de una función existe en un punto de su dominio implica que ella es continua y parcialmente derivable en ese punto. Para apreciar cuanto más fuerte es la diferenciabilidad que la derivabilidad parcial, en el próximo ejemplo veremos que esta última no implica ni siquiera la continuidad.

Ejemplo 1.3.2. Considere la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x_1, x_2) := \frac{x_1 x_2^3}{x_1^2 + x_2^6}$ si $(x_1, x_2) \neq (0,0)$ y $f(0,0) := 0$. Observe que f no es continua en $(0,0)$ pero si es parcialmente derivable. En efecto, $f(k^{-1}, k^{-\frac{1}{3}}) = \frac{k^{-2}}{k^{-2} + k^{-2}} = \frac{1}{2} \neq f(0,0)$ para todo $k \in \mathbb{N}$, lo que muestra la discontinuidad de f en $(0,0)$. Por otra parte

$Df(0;v) = \lim\limits_{t\to 0} \frac{f(tv)-f(0)}{t} = \lim\limits_{t\to 0} \frac{t^4 v_1 v_2^3}{t^3 v_1^2 + t^7 v_2^6} = \lim\limits_{t\to 0} \frac{t v_1 v_2^3}{v_1^2 + t^4 v_2^6} = 0$ para todo $v \in X$, lo que muestra que f es parcialmente derivable en $(0,0)$.

Teorema 1.3.3. *Dadas dos funciones f y g definidas en un abierto A de un e.v.n X, con valores en un e.v.n Y, diferenciables en $x_0 \in A$, entonces*

(i) *$f+g$ es diferenciable en x_0 y se tiene*

$$D[f+g](x_0) = Df(x_0) + Dg(x_0); \tag{1.13}$$

(ii) *Para $\lambda \in \mathbb{R}, \lambda f$ es diferenciable en x_0 y se tiene*

$$D[\lambda f](x_0) = \lambda Df(x_0); \tag{1.14}$$

(iii) *Si $Y = \mathbb{R}$, entonces $f \cdot g$ es diferenciable en x_0 y se tiene*

$$D[f \cdot g](x_0) = g(x_0)Df(x_0) + f(x_0)Dg(x_0); \tag{1.15}$$

(iv) *Si $Y = \mathbb{R}$ y $f(x_0) \neq 0$, entonces $1/f$ es diferenciable en x_0 y se tiene*

$$D\left[\frac{1}{f}\right](x_0) = -\frac{1}{f^2(x_0)} Df(x_0). \tag{1.16}$$

Demostración. Una forma directa de demostrar este teorema consiste en verificar para cada uno de los cuatro casos la igualdad (1.10). De este modo se demuestra simultáneamente la diferenciabilidad de cada una de las cuatro funciones y su respectiva fórmula.

Los casos (i) y (ii) son los más fáciles y los dejamos como ejercicio.

(iii)

$$\lim_{\substack{v\to 0\\ v\neq 0}} \frac{(f\cdot g)(x_0+v) - (f\cdot g)(x_0) - [g(x_0)Df(x_0) + f(x_0)Dg(x_0)](v)}{\|v\|} =$$

$$\lim \frac{f(x_0+v)[g(x_0+v) - g(x_0) - Dg(x_0)(v)]}{\|v\|} + \frac{g(x_0)[f(x_0+v) - f(x_0) - Df(x_0)(v)]}{\|v\|}$$

$$+ \frac{[f(x_0+v) - f(x_0)]Dg(x_0)(v)}{\|v\|} = \lim f(x_0+v) \lim \frac{g(x_0+v) - g(x_0) - Dg(x_0)(v)}{\|v\|}$$

$$+ g(x_0) \lim \frac{f(x_0+v) - f(x_0) - Df(x_0)(v)}{\|v\|} + \lim [f(x_0+v) - f(x_0)] \frac{Dg(x_0)(v)}{\|v\|} = 0.$$

El último límite de la expresión anterior es nulo debido a que la continuidad de la función lineal $Dg(x_0)$ implica que el cuociente $\frac{Dg(x_0)(v)}{\|v\|}$ es acotado y, la diferenciabilidad de f en x_0 implica su continuidad.

(iv)

$$\lim_{\substack{v\to 0\\ v\neq 0}} \frac{\frac{1}{f(x_0+v)} - \frac{1}{f(x_0)} + \frac{Df(x_0)(v)}{f(x_0)^2}}{\|v\|} = \lim \frac{f(x_0)^2 - f(x_0+v)f(x_0) + f(x_0+v)Df(x_0)(v)}{f(x_0+v)f(x_0)^2\|v\|} =$$

$$\lim \frac{1}{f(x_0+v)f(x_0)^2} \cdot \frac{Df(x_0)(v)[f(x_0+v) - f(x_0)] - f(x_0)[f(x_0+v) - f(x_0) - Df(x_0)(v)]}{\|v\|} =$$

$$\frac{1}{f(x_0)^3} \lim [f(x_0+v) - f(x_0)] \frac{Df(x_0)(v)}{\|v\|} - \frac{1}{f(x_0)^2} \lim \frac{f(x_0+v) - f(x_0) - Df(x_0)(v)}{\|v\|} = 0$$

□

TEOREMA 1.3.4. *Si f es una función definida en un abierto A de un e.v.n X, con valores en un producto (finito) de e.v.n. $Y_1 \times \ldots \times Y_m$ y, si denotamos $f_1, ..., f_m$ las funciones componentes de f, entonces f es diferenciable en $x_0 \in A$ si y solo si cada una de las m funciones f_i es diferenciable en x_0 y, en ese caso se tendrá*

$$Df(x_0) = (Df_1(x_0), ..., Df_m(x_0)). \tag{1.17}$$

DEMOSTRACIÓN. La equivalencia de la diferenciabilidad de f con la de sus funciones componentes y la fórmula (1.17) se obtienen en forma directa utilizando la igualdad (1.10) y las propiedades del límite. □

NOTA 1.3.5. Las cinco fórmulas que se dan en los dos teoremas anteriores, se pueden obtener fácilmente a partir de la fórmula (1.12) y de las respectivas fórmulas del Teorema 1.3.2 y Proposición 1.2.1. No lo hicimos así debido a que previamente había que demostrar la diferenciabilidad de cada función.

TEOREMA 1.3.5. *Si f es una función definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n Y, diferenciable en un punto $x_0 \in A$, entonces se tiene la fórmula*

$$Df(x_0)(v) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(x_0) \tag{1.18}$$

para todo $v = \sum_{i=1}^{n} v_i e_i \in \mathbb{R}^n$, donde $e_1, ..., e_n$ es la base canónica de $\mathbb{R}^n$ que usamos en la Definición 1.2.3.

DEMOSTRACIÓN. Es una consecuencia inmediata del Teorema 1.3.2 y de la definición de derivada parcial $\frac{\partial f}{\partial x_i}(x_0)$. □

DEFINICIÓN 1.3.2. Si f es una función definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, diferenciable en un punto $x_0 \in A$, se llama gradiente de f en x_0 al vector

$$\nabla f(x_0) := \left(\frac{\partial f}{\partial x_1}(x_0), ..., \frac{\partial f}{\partial x_n}(x_0) \right) \in \mathbb{R}^n. \tag{1.19}$$

NOTA 1.3.6. Con la definición anterior, cuando $Y = \mathbb{R}$, la fórmula (1.18) se escribe

$$Df(x_0)(v) = \langle \nabla f(x_0), v \rangle \tag{1.20}$$

donde $\langle \cdot, \cdot \rangle$ denota al producto interno usual en $\mathbb{R}^n$. De este modo vemos que el gradiente de una función f diferenciable en x_0 es el vector asociado a la función lineal $Df(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

Con esta notación, la aproximación de primer orden de f en x_0, definida en la Nota 1.3.2, se escribe

$$h(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle \tag{1.21}$$

y la ecuación del hiperplano afín tangente al grafo de f en $(x_0, f(x_0))$, será

$$z = \langle \nabla f(x_0), x - x_0 \rangle + f(x_0). \tag{1.22}$$

Denotando ahora $\langle \cdot, \cdot \rangle$ al producto interno en $\mathbb{R}^n \times \mathbb{R}$, esta ecuación se escribe

$$\langle (\nabla f(x_0), -1), (x, z) - (x_0, f(x_0)) \rangle = 0$$

lo que muestra que el hiperplano afín en cuestión es aquel que pasa por $(x_0, f(x_0))$ y es paralelo al hiperplano ortogonal al vector $(\nabla f(x_0), -1)$. Por esta razón, se dice que el vector $(\nabla f(x_0), -1)$, trasladado a $(x_0, f(x_0))$, es ortogonal al grafo de f en $(x_0, f(x_0))$.

NOTA 1.3.7. Observe que la fórmula (1.20) puede generalizarse a toda función f definida en un abierto A de un espacio de Hilbert X con valores en $\mathbb{R}$. En efecto, si f es diferenciable en $x_0 \in A$, como $Df(x_0)$ es una función lineal continua de X en $\mathbb{R}$, definimos el gradiente de f en x_0 como el único elemento $\nabla f(x_0) \in X$ que verifica

$$Df(x_0)(v) = \langle \nabla f(x_0), v \rangle \quad \text{para todo } v \in X \tag{1.23}$$

y con esta definición de gradiente, se deduce fácilmente del teorema anterior que, cuando $X = \mathbb{R}^n$ y $Y = \mathbb{R}$, el gradiente $\nabla f(x_0)$ está dado por la igualdad (1.19).

DEFINICIÓN 1.3.3. Si f es una función definida en un abierto A de $\mathbb{R}^n$, con valores en $\mathbb{R}^m$, si denotamos $f_1, ..., f_m$ sus m funciones componentes y suponemos f diferenciable en $x_0 \in A$, entonces llamamos Jacobiano de f en x_0 a la matriz de $m \times n$

$$J_f(x_0) = \left(\frac{\partial f_i}{\partial x_j}(x_0) \right) \tag{1.24}$$

donde $1 \leq i \leq m$ y $1 \leq j \leq n$.

NOTA 1.3.8. Con la definición anterior, cuando $Y = \mathbb{R}^m$, la fórmula (1.18) se escribe, usando notación matricial,

$$Df(x_0)(v) = J_f(x_0)v. \tag{1.25}$$

El Jacobiano de una función f en x_0 es la matriz asociada a la función lineal $Df(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, según lo estudiado en el curso de algebra lineal. Con esta notación, la aproximación de primer orden de f en x_0, definida en la Nota 1.3.2 se escribe

$$h(x) = f(x_0) + J_f(x_0)(x - x_0). \tag{1.26}$$

1.3.1. Teorema del Valor Medio.

TEOREMA 1.3.6. *Sea f una función definida en un abierto A de un e.v.n. X con valores en $\mathbb{R}$. Si f es diferenciable en todo punto de un segmento $[x_0, y_0] \subset A$, entonces existe $z \in]x_0, y_0[$ tal que*

$$f(y_0) - f(x_0) = Df(z)(y_0 - x_0). \tag{1.27}$$

DEMOSTRACIÓN. Si definimos la función $\phi(t) = f(x_0 + t(y_0 - x_0))$ para $t \in [0, 1]$ y aplicamos el teorema del valor medio para funciones de una variable real, obtenemos que existe $\eta \in]0, 1[$ tal que $\phi(1) - \phi(0) = \phi'(\eta)$ lo que corresponde exactamente a la fórmula (1.27) con $z = x_0 + \eta(y_0 - x_0) \in]x_0, y_0[$. □

NOTA 1.3.9. Si en el teorema anterior $X = \mathbb{R}^n$, de la fórmula (1.20) vemos que (1.27) puede escribirse

$$f(y_0) - f(x_0) = \langle \nabla f(z), y_0 - x_0 \rangle.$$

Ejemplo 1.3.3. Demos un ejemplo que muestre que la fórmula (1.27) no es en general válida si f toma sus valores en un e.v.n. Y. Sea $f : \mathbb{R} \to \mathbb{R}^2$ definida por $f(t) := (cos(t), sen(t))$ y sean $a := 0$ y $b := 2\pi$. Es fácil constatar que $f(b) - f(a) = (0, 0)$ y $Df(c)(b - a) = 2\pi(-sen(c), cos(c))$, lo que muestra que no existe $c \in [0, 2\pi]$ tal que se tenga la igualdad (1.27).

Teorema 1.3.7. *Sea f una función definida en un abierto A de un e.v.n. X con valores en un e.v.n. Y. Si f es diferenciable en todo punto de un segmento $[x_0, y_0] \subset A$ y si L es una constante que verifica $L \geq \|Df(x)\|$ para todo $x \in [x_0, y_0]$, entonces se tiene la desigualdad*

$$\|f(y_0) - f(x_0)\| \leq L\|y_0 - x_0\|. \tag{1.28}$$

Demostración. Supongamos lo contrario, es decir, que existe $\delta > 0$ tal que$\|f(y_0) - f(x_0)\| - L\|y_0 - x_0\| = \delta$. Si definimos $m = \frac{x_0+y_0}{2}$, puesto que $\|f(y_0) - f(m)\| + \|f(m) - f(x_0)\| \geq \|f(y_0) - f(x_0)\|$ y $\|y_0 - x_0\| = \|y_0 - m\| + \|m - x_0\|$, se tendrá

$$\|f(y_0) - f(m)\| - L\|y_0 - m\| \geq \frac{\delta}{2}$$

o bien

$$\|f(m) - f(x_0)\| - L\|m - x_0\| \geq \frac{\delta}{2}.$$

Si se tiene la primera de estas desigualdades definimos $y_1 := b$ y $x_1 := m$, en caso contrario definimos $y_1 := m$ y $x_1 := x_0$. Aplicando sucesivamente el mismo procedimiento obtenemos una sucesión de intervalos encajonados $[x_n, y_n]$ con $\|y_n - x_n\| = \frac{\|y_0 - x_0\|}{2^n}$ y tales que

$$\|f(y_n) - f(x_n)\| - L\|y_n - x_n\| \geq \frac{\delta}{2^n}.$$

Puesto que $\|x_n - y_n\| \to 0$ sabemos que las sucesiones $\{x_n\}$ y $\{y_n\}$ deben converger a un mismo $w \in [x_0, y_0]$. Por otra parte por ser f diferenciable en w, de la desigualdad anterior podemos escribir

$$\begin{aligned}
\frac{\delta}{2^n} &\leq \|f(y_n) - f(w) - (f(x_n) - f(w))\| - L\|y_n - x_n\| \\
&= \|Df(w)(y_n - w) + o(y_n - w) - Df(w)(x_n - w) - o(x_n - w)\| - L\|y_n - x_n\| \\
&\leq \|Df(w)\|\|y_n - x_n\| + \|o(y_n - w)\| + \|o(x_n - w)\| - L\|y_n - x_n\|
\end{aligned}$$

y como $\|y_n - w\| < \|y_n - x_n\|$ y $\|x_n - w\| < \|y_n - x_n\|$, dividiendo la desigualdad anterior por $\|y_n - x_n\|$ podemos escribir

$$\frac{\delta}{\|y_0 - x_0\|} \leq \|Df(w)\| + \frac{\|o(w - y_n)\|}{\|w - y_n\|} + \frac{\|o(w - x_n\|}{\|w - x_n\|} - L$$

y tomando límite sobre n obtenemos una contradicción con la hipótesis $\|Df(x)\| \leq L$ para todo $x \in [x_0, y_0]$. $\square$

Nota 1.3.10. Del teorema anterior se deduce que si f es diferenciable en todo punto de un conjunto convexo $C \subset A$ y si $L \geq \|Df(x)\|$ para todo $x \in C$, entonces $\|f(x) - f(y)\| \leq L\|x - y\|$ para todo $x, y \in C$.

Definición 1.3.4. Un conjunto C en un e.v.n. X se dirá conexo si no existen conjuntos abiertos no vacíos C_1 y C_2 en X que intersecten C, tales que $C_1 \cap C_2 = \phi$ y $C \subset C_1 \cup C_2$.

Teorema 1.3.8. *Sea f una función diferenciable definida en un abierto A de un e.v.n. X con valores en un e.v.n. Y. Si el diferencial de f es nulo en todo punto de un conjunto abierto conexo $C \subset A$, entonces la función f será constante en C.*

Demostración. Demostremos primero que f es constante en toda bola $B(x_0, \delta) \subset C$. Sea $z \in B(x_0, \delta)$, puesto que $[x_0, z] \subset C$ y $Df(x) = 0$ para todo $x \in [x_0, z]$, del teorema anterior concluímos que $\|f(z) - f(x_0)\| \leq 0$, lo que equivale a decir que $f(z) = f(x_0)$.

Sea ahora $x_0 \in C, C_1 = \{x \in C : f(x) = f(x_0)\}$ y $C_2 = \{x \in C : f(x) \neq f(x_0)\}$. Puesto que la función f es continua (ver Teorema 1.3.2) y C es un conjunto abierto, es fácil demostrar que C_2 es un conjunto abierto.

Mostremos finalmente que C_1 es también un conjunto abierto. Sea $x_0 \in C_1$ y $B(x_0, \delta) \subset C$, de la primera parte de esta demostración concluímos que f es constante en $B(x_0, \delta)$, y se tendrá entonces $f(x) = f(x_0)$ para todo $x \in B(x_0, \delta)$, lo que implica que $B(x_0, \delta) \subset C_1$. Lo anterior muestra que C_1 es abierto. Como $C_1 \cap C_2 = \phi$, como $C \subset C_1 \cup C_2$ y como $C_1 \neq \phi$ (en efecto, $x_0 \in C_1$), del hecho que C es conexo concluímos que $C_2 = \phi$, es decir, $C = C_1$. □

1.4. Funciones de clase $\mathcal{C}^1$

Definición 1.4.1. Una función f definida en un abierto A de $\mathbb{R}^n$, con valores en un e.v.n Y se dirá de clase $\mathcal{C}^1$ si las n derivadas parciales de f con respecto a $x_1, ..., x_n$ (ver Nota 1.2.3) existen y son continuas.

Teorema 1.4.1. *Si f es una función de clase $\mathcal{C}^1$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n Y, entonces ella es diferenciable.*

Demostración. Con el único objeto de simplificar la notación, haremos la demostración para el caso en que $n = 2$ y usaremos la norma del máximo.

Dado un elemento cualquiera $x_0 := a_1 e_1 + a_2 e_2 \in A$, vamos a demostrar que f es diferenciable en x_0. De acuerdo al Teorema 1.3.5, debemos probar entonces que la función lineal $\ell(v) := v_1 \frac{\partial f}{\partial x_1}(x_0) + v_2 \frac{\partial f}{\partial x_2}(x_0)$ (donde $v = (v_1, v_2)$) es el diferencial de f en x_0.

Dado $\eta > 0$ tal que $B(x_0, \eta) \subset A$, escribamos para $v \in \mathbb{R}^2$ con $\|v\| \leq \eta$ la desigualdad

$$\frac{\|f(x_0 + v) - f(x_0) - \ell(v)\|}{\|v\|} \leq \frac{\|f(a_1 + v_1, a_2 + v_2) - f(a_1, a_2 + v_2) - v_1 \frac{\partial f}{\partial x_1}(x_0)\|}{\|v\|}$$
$$+ \frac{\|f(a_1, a_2 + v_2) - f(a_1, a_2) - v_2 \frac{\partial f}{\partial x_2}(x_0)\|}{\|v\|}.$$

Para concluir debemos probar que el límite cuando $v \to 0$ de cada uno de los dos sumandos de la derecha de esta desigualdad es cero.

La desigualdad

$$\frac{\|f(a_1, a_2 + v_2) - f(a_1, a_2) - v_2 \frac{\partial f}{\partial x_2}(x_0)\|}{\|v\|} \leq \left\| \frac{f(x_0 + v_2 e_2) - f(x_0)}{v_2} - \frac{\partial f}{\partial x_2}(x_0) \right\|$$

y la definición de $\frac{\partial f}{\partial x_2}(x_0)$, nos muestra que el segundo de estos límites, es cero.

Definamos ahora la función $\phi(t) := f(a_1+t, a_2+v_2) - f(a_1, a_2+v_2) - t\frac{\partial f}{\partial x_1}(x_0)$. Como f es de clase C^1, es claro que ϕ es diferenciable en todo punto del intervalo J (donde $J = [0, v_1]$ si $v_1 > 0$, $= [v_1, 0]$ si $v_1 < 0$) y $D\phi(t)(v) = [\frac{\partial f}{\partial x_1}(a_1+t, a_2+v_2) - \frac{\partial f}{\partial x_1}(x_0)]v$. Escribiendo entonces para la función ϕ la desigualdad $\|\phi(v_1) - \phi(0)\| \leq L|v_1|$
con $L = \max_{t\in J} \|D\phi(t)\|$, dada por el Teorema 1.3.7, obtenemos

$$\|f(a_1+v_1, a_2+v_2) - f(a_1, a_2+v_2) - v_1\frac{\partial f}{\partial x_1}(x_0)\| \leq |v_1| \max_{t\in J} \|\frac{\partial f}{\partial x_1}(a_1+t, a_2+v_2) - \frac{\partial f}{\partial x_1}(x_0)\|.$$

Puesto que $\frac{\partial f}{\partial x_1}$ es una función continua en x_0 y que $\frac{|v_1|}{\|v\|} \leq 1$, dividiendo por $\|v\|$ esta desigualdad, vemos que el lado derecho tiende a cero cuando $v \to 0$, lo que nos permite concluir nuestra demostración. □

Teorema 1.4.2. *Si f es una función de clase $\mathcal{C}^1$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. Y, entonces f es continua.*

Demostración. Es una consecuencia inmediata del teorema anterior y del Teorema 1.3.2. □

Definición 1.4.2. Una función definida en un abierto A de un e.v.n X, con valores en un e.v.n Y se dirá continuamente diferenciable si ella es diferenciable en todo punto de A y la función $Df : A \subseteq X \to \mathcal{L}(X, Y)$ es continua.

Teorema 1.4.3. *Si f es una función de clase $\mathcal{C}^1$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n Y, entonces ella es continuamente diferenciable.*

Demostración. Del Teorema 1.4.1 sabemos que f es diferenciable en x_0. Verifiquemos ahora la continuidad de Df en $x_0 \in A$. De (1.18) se tiene

$$\begin{aligned}
\|Df(x_0) - Df(x)\| &= \sup_{v\neq 0} \frac{\|[Df(x_0) - Df(x)](v)\|}{\|v\|} = \\
&= \sup_{v\neq 0} \frac{\|[\frac{\partial f}{\partial x_1}(x_0) - \frac{\partial f}{\partial x_1}(x)]v_1 + ... + [\frac{\partial f}{\partial x_n}(x_0) - \frac{\partial f}{\partial x_n}(x)]v_n\|}{\|v\|} \\
&\leq \sup_{v\neq 0} \frac{\|(\frac{\partial f}{\partial x_1}(x_0) - \frac{\partial f}{\partial x_1}(x))v_1\|}{\|v\|} + ... + \sup_{v\neq 0} \frac{\|(\frac{\partial f}{\partial x_n}(x_0) - \frac{\partial f}{\partial x_n}(x))v_n\|}{\|v\|} \\
&\leq \left\|\frac{\partial f}{\partial x_1}(x_0) - \frac{\partial f}{\partial x_1}(x)\right\| + ... + \left\|\frac{\partial f}{\partial x_n}(x_0) - \frac{\partial f}{\partial x_n}(x)\right\|
\end{aligned}$$

por lo tanto, haciendo tender x a x_0 y puesto que las funciones $\frac{\partial f}{\partial x_i}(\cdot)$ son continuas en x_0, obtenemos

$$\lim_{x\to x_0} \|Df(x_0) - Df(x)\| = 0$$

lo que es equivalente a $\lim_{x\to x_0} Df(x) = Df(x_0)$ mostrando así la continuidad de Df en x_0. □

Teorema 1.4.4. *Si f es una función continuamente diferenciable, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. Y, entonces f es de clase $\mathcal{C}^1$.*

Demostración. Vamos a demostrar la continuidad de la función $\frac{\partial f}{\partial x_i} : A \to Y$ en un punto $x_0 \in A$. Del Teorema 1.3.2 deducimos que

$$\begin{aligned} \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(x_0) &= Df(x)(e_i) - Df(x_0)(e_i) \\ &= [Df(x) - Df(x_0)](e_i) \end{aligned}$$

y de la desigualdad $\|\ell(x)\|_Y \leq \|\ell\|_{\mathcal{L}(X,Y)} \; \|x\|_X$ para lineales continuas, aplicada a la función lineal continua $[Df(x) - Df(x_0)]$, vemos que

$$(*) \qquad \left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(x_0) \right\| \leq \|Df(x) - Df(x_0)\|_{\mathcal{L}(X,Y)} \|e_i\|.$$

Como por hipótesis la función $Df : A \to \mathcal{L}(\mathbb{R}^n, F)$ es continua en $x_0 \in A$, obtenemos $\lim\limits_{x \to x_0} Df(x) = Df(x_0)$, lo que implica a partir de (*) que

$$\lim_{x \to x_0} \frac{\partial f}{\partial x_i}(x) = \frac{\partial f}{\partial x_i}(x_0)$$

concluyendo la continuidad de $\frac{\partial f}{\partial x_i}$ en x_0. $\square$

Nota 1.4.1. Si f es una función definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. Y, entonces los teoremas 1.3.2, 1.4.1, 1.4.3 y 1.4.4 se resumen en el siguiente diagrama

$$\begin{aligned} f \text{ es de clase } \mathcal{C}^1 &\Leftrightarrow f \text{ es continuamente diferenciable} \\ &\Rightarrow f \text{ es diferenciable} \\ &\Rightarrow f \text{ es parc. derivable y } f \text{ es continua.} \end{aligned}$$

Teorema 1.4.5. *Si f y g son dos funciones continuamente diferenciables definidas en un abierto A de un e.v.n. X con valores en un e.v.n. Y, entonces*

(i) *$f + g$ es continuamente diferenciable.*
(ii) *Si $\lambda \in \mathbb{R}, \lambda f$ es continuamente diferenciable.*
(iii) *Si $Y = \mathbb{R}, f \cdot g$ es continuamente diferenciable.*
(iv) *Si $Y = \mathbb{R}$ y $f(x) \neq 0$ para todo $x \in A, 1/f$ es continuamente diferenciable.*

Demostración. Del Teorema 1.3.3 vemos que las cuatro funciones del enunciado serán diferenciables. Por otra parte las fórmulas (1.13) a (1.16) muestran que para todo $x \in A$

(i) $D[f + g](x) = Df(x) + Dg(x)$
(ii) $D[\lambda f](x) = \lambda Df(x)$
(iii) $D[f \cdot g](x) = g(x)Df(x) + f(x)Dg(x)$.
(iv) $D[1/f](x) = -\frac{1}{f(x)^2} Df(x)$.

Dado que por hipótesis las funciones f, Df, g y Dg son continuas, de las fórmulas anteriores deducimos que $D[f+g], D[\lambda f], D[f \cdot g]$ y $D[\frac{1}{f}]$ son continuas. Lo anterior nos permite concluir que $f+g, \lambda f, f \cdot g$ y $1/f$ son continuamente diferenciables. $\square$

Teorema 1.4.6. *Si f es una función definida en un abierto A de un e.v.n X con valores en un producto de e.v.n. $Y_1 \times ... \times Y_m$ y, si denotamos $f_1, ..., f_m$ las funciones componentes de f, entonces ella será continuamente diferenciable si y solo si sus m funciones componentes son continuamente diferenciables.*

Demostración. Del Teorema 1.3.4 vemos que f es diferenciable si y solo si sus m funciones componentes lo son. Por otra parte la fórmula (1.17) muestra que

$$Df(x) = (Df_1(x), ..., Df_m(x))$$

lo que significa que Df es continua si y solo si las funciones $Df_1, ..., Df_m$ son continuas. Lo anterior nos permite concluir que f es continuamente diferenciable si y solo si las funciones $f_1, ..., f_m$ también lo son. □

Teorema 1.4.7. *Si f es una función continuamente diferenciable, definida en un abierto A de un e.v.n. X con valores en un e.v.n. Y, entonces para todo $x_0 \in A$ y todo $\varepsilon > 0$ existe $\delta > 0$ tal que*

$$x, y \in B(x_0, \delta) \Rightarrow \|f(x) - f(y)\| \leq (\|Df(x_0)\| + \varepsilon)\|x - y\|. \tag{1.29}$$

La función f se dice entonces localmente Lipschitziana en A.

Demostración. Sea $x_0 \in A$ y $\varepsilon > 0$. Como Df es una función continua en x_0, existirá $\delta > 0$ tal que

$$\|Df(z)\| \leq \|Df(x_0)\| + \varepsilon \text{ para todo } z \in B(x_0, \delta).$$

Como para todo $x, y \in B(x_0, \delta)$ se tiene que $[x, y] \subset B(x_0, \delta)$, del Teorema 1.3.7 obtenemos la desigualdad (1.29). □

Teorema 1.4.8. *Sea f una función continuamente diferenciable, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. Y. Entonces ella es Lipschitziana en toda bola $B(x_0,r) \subset A$, con constante de Lipschitz $L := \max\limits_{z \in B(x_0,r)} \|Df(z)\|$.*

Demostración. Dado que una función continua definida sobre un conjunto compacto a valores en $\mathbb{R}$ alcanza su máximo, aplicando a Df en $B(x_0, r)$, concluímos que L está bien definido. El Teorema 1.3.7 nos permite entonces concluir. □

1.5. Composición de funciones diferenciables

Teorema 1.5.1. *Sea f una función de un abierto A de un e.v.n. X en un e.v.n. Y y sea g una función de un abierto B del e.v.n. Y, en un e.v.n. G (suponemos que $f(A) \subset B$). Entonces si f y g son diferenciables en $x_0 \in A$ y $f(x_0) \in B$ respectivamente, la función $h := g$ o f es diferenciable en x_0 y se tiene la fórmula*

$$Dh(x_0) = Dg(f(x_0)) \circ Df(x_0). \tag{1.30}$$

Demostración. De acuerdo a la relación (1.11) debemos probar que dado $\varepsilon > 0$ existe $\eta > 0$ tal que

$$(*) \quad \|v\| \leq \eta \Rightarrow \|(g \circ f)(x_0 + v) - (g \circ f)(x_0) - [Dg(f(x_0)) \circ Df(x_0)](v)\| \leq \varepsilon \|v\|.$$

Para simplificar la notación escribamos $h := g \circ f$, $u := f(x_0)$ y $w := f(x_0 + v) - f(x_0)$. Entonces

$$\|h(x_0 + v) - h(x_0) - [Dg(u) \circ Df(x_0)](v)\| \leq$$
$$\|g(u + w) - g(u) - Dg(u)(w)\| + \|Dg(u)(w) - [Dg(u) \circ Df(x_0)](v)\|$$

y además

$$\begin{aligned}
\|Dg(u)(w) - [Dg(u) \circ Df(x_0)](v)\| &= \|Dg(u)(w - Df(x_0)(v))\| \\
&\leq \|Dg(u)\| \ \|f(x_0 + v) - f(x_0) - Df(x_0)(v)\|
\end{aligned}$$

obteniendo la desigualdad

(1.31)
$$\|h(x_0+v)-h(x_0)-[Dg(u)\circ Df(x_0)](v)\| \leq$$
$$\|g(u+w)-g(u)-Dg(u)(w)\| + \|Dg(u)\|\ \|f(x_0+v)-f(x_0)-Df(x_0)(v)\|.$$

Sea $\varepsilon > 0$. Como f es diferenciable en x_0, de acuerdo al Teorema 1.3.2 (ver su demostración), ella será también Lipschitziana (localmente) en x_0. Existirá entonces $\eta_1 > 0$ tal que

$$\|v\| \leq \eta_1 \Rightarrow \|f(x_0+v)-f(x_0)-Df(x_0)(v)\| \leq \frac{\varepsilon}{2\|Dg(u)\|}\|v\| \tag{1.32}$$

y

$$\|v\| \leq \eta_1 \Rightarrow \|f(x_0+v)-f(x_0)\| \leq L\|v\|. \tag{1.33}$$

Por otra parte, como g es diferenciable en u, existirá $\eta_2 > 0$ tal que

$$\|w\| \leq \eta_2 \Rightarrow \|g(u+w)-g(u)-Dg(u)(w)\| \leq \frac{\varepsilon}{2L}\|w\|. \tag{1.34}$$

De las relaciones (1.31), (1.32), (1.33) y (1.34) definiendo $\eta = \text{mín}\{\eta_1, \frac{\eta_2}{L}\}$ obtenemos directamente la desigualdad (*). □

NOTA 1.5.1. De acuerdo a la Definición 1.3.2 y a la Nota 1.3.8 vemos que si en el teorema anterior $X = \mathbb{R}^n, Y = \mathbb{R}^m$ y $G = \mathbb{R}^p$, entonces el jacobiano de la función $h = g \circ f$ en x_0, será igual al producto de los jacobianos de las funciones g y f, esto es

$$J_h(x_0) = J_g(f(x_0)) \cdot J_f(x_0) \tag{1.35}$$

que corresponde a la fórmula (1.30) escrita matricialmente. Se verifica fácilmente que el elemento (i, j) de la matriz $J_h(x_0)$ (de p filas y n columnas) es $\sum\limits_{k=1}^{m} \partial_k g_i(f(x_0))\partial_j f_k(x_0)$ donde $\partial_j f_k(x_0)$ representa la derivada parcial en x_0 de la función componente f_k con respecto a la j-ésima variable y, $\partial_k g_i(f(x_0))$ representa la derivada parcial en $f(x_0)$ de la función componente g_i con respecto a la k-ésima variable.

De (1.35) y (1.25) vemos que para todo $v \in \mathbb{R}^n$ se tendrá

$$Dh(x_0)(v) = J_g(f(x_0))J_f(x_0)^t v^t \tag{1.36}$$

que también se escribe

$$Dh(x_0)(v) = (\sum_{j=1}^{n}\sum_{k=1}^{m} v_j \partial_j f_k(x_0)\partial_k g_1(f(x_0)), \quad ... \quad , \sum_{j=1}^{n}\sum_{k=1}^{m} v_j \partial_j f_k(x_0)\partial_k g_p(f(x_0))).$$

En particular, para $v = e_j \in \mathbb{R}^n$, se tendrá para todo $j = 1, ..., n$ e $i = 1, ..., p$ la fórmula

$$\partial_j h_i(x_0) = \sum_{k=1}^{m} \partial_j f_k(x_0)\partial_k g_i(f(x_0)). \tag{1.37}$$

Esta fórmula se llama usualmente regla de la cadena para el cálculo de las derivadas parciales de la función $h = g \circ f$.

TEOREMA 1.5.2. *Si suponemos que las funciones f y g del teorema anterior son continuamente diferenciables, entonces $h := f$ o g también será continuamente diferenciable.*

DEMOSTRACIÓN. Es una consecuencia inmediata de la fórmula (1.30) y del hecho que composición de funciones continua es una función continua. $\square$

1.6. Diferencial Parcial

En esta sección vamos a introducir la noción de diferencial parcial de una función, que en cierto sentido generaliza la noción de derivada parcial y, daremos después un resultado que generaliza el Teorema 1.3.5.

DEFINICIÓN 1.6.1. Dados $n+1$ e.v.n. $X_1, ..., X_n, Y$, una función f definida en un abierto A del e.v.n. $X_1 \times ... \times X_n$ con valores en Y y $x_0 \in A$, denotaremos por $D_j f(x_0)$ (para $j = 1, ..., n$) al diferencial en $x_j \in X_j$ ($x_0 = (x_1, ..., x_n) \in X_1 \times ... \times X_n$) de la función $f(x_1, ..., x_{j-1}, \cdot, x_{j+1}, ..., x_n)$, cuando existe. Se tiene entonces que $D_j f(x_0) \in \mathcal{L}(X_j, Y)$.

Al diferencial $D_j f(x_0)$ lo llamamos diferencial parcial de f en x_0 respecto a la variable j.

NOTA 1.6.1. Con los datos de la definición anterior definimos las tres funciones $p_j : X_1 \times ... \times X_n \to X_j$, $i_j : X_j \to X_1 \times ... \times X_n$ e $I_j : X_j \to X_1 \times ... \times X_n$ por

$$p_j(x_1, ..., x_n) := x_j \tag{1.38}$$

$$i_j(x_j) := (0, ..., 0, x_j, 0, ..., 0) \tag{1.39}$$

$$I_j(x_j) := x_0 + i_j(x_j - x_j). \tag{1.40}$$

Se deduce entonces fácilmente que

$$D_j f(x_0) = D[f \circ I_j](p_j(x_0)). \tag{1.41}$$

TEOREMA 1.6.1. *Dados $n+1$ e.v.n. $X_1, ..., X_n, Y$, un abierto A en el e.v.n. $X_1 \times ... \times X_n$, un elemento $x_0 \in A$ y una función $f : A \to Y$ diferenciable en $x_0 \in A$, entonces para todo $v \in X_1 \times ... \times X_n$ se tiene la fórmula*

$$Df(x_0)(v) = \sum_{j=1}^{n} D_j f(x_0)(v_j) \tag{1.42}$$

donde $v_j := p_j(v)$ (ver (1.38)).

DEMOSTRACIÓN. De la igualdad (1.41), usando el Teorema 1.5.1 y el Ejemplo 1.3.1 que nos muestra que $DI_j(p_j(x_0)) = i_j$, vemos que

$$\begin{aligned} D_j f(x_0) &= D[f \circ I_j](p_j(x_0)) \\ &= Df(I_j(p_j(x_0))) \circ DI_j(p_j(x_0)) \qquad (*) \\ &= Df(x_0) \circ i_j. \end{aligned}$$

Si componemos la igualdad (*) con p_j y sumamos sobre j, obtenemos

$$\sum_{j=1}^{n} D_j f(x_0) \circ p_j = \sum_{j=1}^{n} Df(x_0) \circ i_j \circ p_j$$

y como $Df(x_0)$ es lineal y $\sum_{j=1}^{n} i_j \circ p_j$ es la identidad en $X_1 \times ... \times X_n$, concluimos que

$$\sum_{j=1}^{n} D_j f(x_0) \circ p_j = Df(x_0)$$

que corresponde exactamente a la igualdad (1.42). $\square$

Nota 1.6.2. Si en la Definición 1.6.1 se tiene para todo $j \in \{1, ..., n\}$ que $X_j := \mathbb{R}^{k_j}$, $Y = \mathbb{R}^m$ y si $f_1, ..., f_m$ son las funciones componentes de f, entonces la fórmula (1.42) se escribe, usando notación matricial

$$Df(x_0)(v) = \sum_{j=1}^{n} J_j(x_0) {v_j}^t \tag{1.43}$$

donde $J_j(x_0)$ es el jacobiano asociado al diferencial $D_j f(x_0)$, esto es la matriz de coeficientes $\partial_\ell f_i(x_0)$ donde el indice $i = 1, ..., m$ indica la fila y $\ell : n_j, ..., n_{j+1} - 1$ indica la columna ($n_j = \sum_{p=1}^{j} k_p$).

1.7. Teoremas de la función inversa y de la función implícita

Nota 1.7.1. Sea $f : \mathbb{R} \to \mathbb{R}$ una función continuamente derivable y tal que $f'(a) \neq 0$. Del curso de cálculo de funciones de una variable sabemos que existe un intervalo $I :=]a - \varepsilon, a + \varepsilon[$, tal que la restricción $\tilde{f} : I \to f(I)$ de la función f es biyectiva y su función inversa $\tilde{f}^{-1}$ es continuamente derivable. Se tiene además que $(\tilde{f}^{-1})'(f(x_0)) = 1/f'(x_0)$. En el teorema que sigue vamos a generalizar este resultado al caso en que f es una función continuamente diferenciable, definida en un abierto A de un espacio de Banach X.

Antes de enunciar el teorema veamos un ejemplo donde lo señalado anteriormente no se tiene pues falla la continuidad de la derivada.

Ejemplo 1.7.1. En este ejemplo mostraremos una función $f : \mathbb{R} \to \mathbb{R}$ que es derivable, con derivada no nula en 0 y, que sin embargo no existe un intervalo $I =]-\varepsilon, \varepsilon[$ tal que la restricción de f a I sea inyectiva. Esta función está definida por la fórmula

$$f(x) = \begin{cases} \frac{x}{2} + x^2 sen(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

Que no exista $\varepsilon > 0$ tal que f sea inyectiva en $]-\varepsilon, \varepsilon[$ es equivalente a decir que en todo intervalo de la forma $]-\varepsilon, \varepsilon[$ la función f no es estrictamente creciente (o decreciente) lo que equivale a decir que en todo intervalo $]-\varepsilon, \varepsilon[$ la derivada f' cambia de signo. Este hecho es evidente puesto que

$$f'(x) = \begin{cases} \frac{1}{2} + 2x sen(\frac{1}{x}) - cos(\frac{1}{x}) & \text{si } x \neq 0 \\ \frac{1}{2} & \text{si } x = 0. \end{cases}$$

Teorema 1.7.1. *Sea f una función continuamente diferenciable, definida en un abierto A de un espacio de Banach X con valores en X. Si $Df(x_0) \in \mathcal{L}(X, X)$ es biyectiva, entonces existe un abierto V en X que contiene a x_0 y, tal que la* *restricción $\tilde{f} : V \to f(V)$ de la función f es biyectiva y su función inversa $\tilde{f}^{-1}$ es continuamente diferenciable. Se tendrá además para $b := f(x_0)$ que*

$$D\tilde{f}^{-1}(b) = [Df(x_0)]^{-1}. \tag{1.44}$$

Demostración. Haremos la demostración en cuatro etapas. En la primera veremos que sin perder generalidad podemos suponer que $x_0 = 0, f(x_0) = 0$ y $Df(x_0) = I$ (función identidad en X), con lo cual la demostración se simplifica notablemente. Luego, en la segunda parte, encontraremos $\delta > 0$ tal que la restricción $\tilde{f}$, de la función f al conjunto abierto $V := f^{-1}[B(0, \frac{\delta}{2})] \cap B(0, \delta)$, es una biyección de ese conjunto en $f(V)$. En la tercera etapa mostraremos que $\tilde{f}^{-1}$ es Lipschitziana con constante de Lipschitz 2. Finalmente probaremos que $\tilde{f}^{-1}$ es continuamente diferenciable.

1a. Parte. Por razones pedagógicas la desarrollaremos al final.

2a. Parte. De acuerdo a la primera parte, podemos suponer $x_0 = 0, f(0) = 0$ y $Df(0) = I$ (función identidad en X). Definamos entonces la función g de A en X por $g(x) = x - f(x)$. Como g es continuamente diferenciable y $Dg(0) = 0$ (función constante en X igual a 0), del Teorema 1.4.7 (tomando $\varepsilon = 1/3$) vemos que existe $\delta > 0$ tal que

$$x, y \in B(0, \delta) \Rightarrow \|g(x) - g(y)\| \leq \frac{1}{3}\|x - y\| \tag{*}$$

y como $g(0) = 0$, si hacemos $y = 0$ en (*) vemos que $g(B(0, \delta)) \subseteq B(0, \frac{\delta}{3}) \subseteq B(0, \frac{\delta}{2})$. Vamos ahora a demostrar que para todo $u \in \overline{B}(0, \frac{\delta}{2})$ (bola cerrada) existe un único $x \in \overline{B}(0, \delta)$ tal que $u = f(x)$, lo que implica que la restricción $\tilde{f}$ de la función f al conjunto $V := f^{-1}(B(0, \frac{\delta}{2})) \cap B(0, \delta)$ es biyectiva de V en $f(V)$. Para esto definamos la función ℓ de $B(0, \delta)$ en X por $\ell(y) := g(y) + u$. Como ℓ también verifica la desigualdad (*), vemos por una parte que $\|y\| \leq \delta \Rightarrow \|\ell(y)\| \leq \frac{1}{3}\|y\| + \|u\| \leq \delta$ lo que significa que ℓ es una función de $\overline{B}(0, \delta)$ en $\overline{B}(0, \delta)$ y por otra parte (*) nos muestra que ℓ es contractante. Dado que X es un espacio de Banach y $\overline{B}(0, \delta)$ un conjunto cerrado, podemos deducir, de acuerdo al teorema del punto fijo, la existencia de un único $x \in \overline{B}(0, \delta)$ tal que $\ell(x) = x$, lo que equivale a decir $u = f(x)$. La continuidad de f garantiza que el conjunto V es abierto.

3a. Parte. Mostremos que

$$u, v \in f(V) \Rightarrow \|\tilde{f}^{-1}(u) - \tilde{f}^{-1}(v)\| \leq 2\|u - v\|.$$

Dados $u, v \in f(V)$, sean $x, y \in V$ tales que $x = \tilde{f}^{-1}(u)$ y $y = \tilde{f}^{-1}(v)$, puesto que de la segunda parte sabemos que $f + g = I$ y que g verifica (*), se deduce

$$\begin{aligned} \|x - y\| &= \|(f + g)(x) - (f + g)(y)\| \\ &\leq \|g(x) - g(y)\| + \|f(x) - f(y)\| \\ &\leq \frac{1}{2}\|x - y\| + \|f(x) - f(y)\| \end{aligned}$$

lo que muestra $\|x - y\| \leq 2\|f(x) - f(y)\|$ que equivale a escribir

$$\|\tilde{f}^{-1}(u) - \tilde{f}^{-1}(v)\| \leq 2\|u - v\|.$$

4a. Parte. Demostremos ahora que $\tilde{f}^{-1}$ es diferenciable. Dado $b \in B(0, \frac{\delta}{2})$ probemos que $D\tilde{f}^{-1}(b) = [Df(x_0)]^{-1}$, donde $x_0 = \tilde{f}^{-1}(b)$. Dados b y $b + k \in B(0, \frac{\delta}{2})$, definiendo $h := \tilde{f}^{-1}(b+k) - \tilde{f}^{-1}(b)$ vemos que, de acuerdo a la tercera parte de la demostración

$$\|h\| = \|\tilde{f}^{-1}(b+k) - \tilde{f}^{-1}(b)\| \leq 2\|k\|$$

por lo tanto

$$\begin{aligned}\frac{\|\tilde{f}^{-1}(b+k) - \tilde{f}^{-1}(b) - Df(x_0)^{-1}(k)\|}{\|k\|} &\leq \frac{2\|h - Df(x_0)^{-1}(f(x_0+h) - f(x_0))\|}{\|h\|} \\ &\leq 2\frac{\|Df(x_0)^{-1}[Df(x_0)(h) - (f(x_0+h) - f(x_0))]\|}{\|h\|} \\ &\leq 2\|Df(x_0)^{-1}\|\frac{\|f(x_0+h) - f(x_0) - Df(x_0)(h)\|}{\|h\|}.\end{aligned}$$

Como f es diferenciable y como cuando $k \to 0$ se tiene que $h \to 0$, obtenemos

$$\lim_{\substack{k\to 0\\ k\neq 0}} \frac{\tilde{f}^{-1}(b+k) - \tilde{f}^{-1}(b) - Df(x_0)^{-1}(k)}{\|k\|} = 0$$

que muestra que $\tilde{f}^{-1}$ es diferenciable y que $D\tilde{f}^{-1}(b) = Df(x_0)^{-1}$.
1a. Parte. Si definimos la función $h : A - \{x_0\} \longrightarrow X$ por

$$h(x) := Df(x_0)^{-1}[f(x_0+x) - f(x_0)]$$

vemos que $h(0) = 0$ y que h es continuamente diferenciable si y solo si f es continuamente diferenciable, en efecto $Dh(x) = Df(x_0)^{-1} \circ Df(x_0+x)$ y en particular $Dh(0) = I$. Vemos también que si existen abiertos $B_1, B_2 \subset X$ que contienen a 0 tales que la restricción $\tilde{h}$, de h a B_1 sea una biyección de B_1 en B_2, entonces la restricción $\tilde{f}$, de f a $A_1 := B_1 + \{x_0\}$, será una biyección de A_1 en $A_2 := Df(x_0)(B_2) + \{f(x_0)\}$, en efecto $\tilde{f}^{-1}(u) = \tilde{h}^{-1}(Df(x_0)^{-1}(u - f(x_0))) + x_0$. Finalmente, de la última igualdad queda claro que si $\tilde{h}^{-1}$ es continuamente diferenciable, entonces $\tilde{f}^{-1}$ también lo será y se tiene

$$D\tilde{f}(u) = D\tilde{h}^{-1}(Df(x_0)^{-1}(u - f(x_0))) \circ Df(x_0)^{-1}.$$

La demostración de que $\tilde{f}$ es continuamente diferenciable es dejada como ejercicio. □

Nota 1.7.2. Dada una función $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ nos preguntamos si es posible despejar la variable y (en función de la variable x) en la ecuación

$$f(x,y) = 0 \tag{1.45}$$

es decir, ¿existe una función $\phi : \mathbb{R} \to \mathbb{R}$ tal que "$f(x,\phi(x)) = 0$ para todo $x \in \mathbb{R}$", o lo que es equivalente, tal que "$f(x,y) = 0 \Leftrightarrow y = \phi(x)$"?
Es fácil ver que "globalmente" esto no es posible, en general no existe ϕ que verifique $f(x,\phi(x)) = 0$ para todo $x \in \mathbb{R}$.
Sin embargo, cuando f es continuamente diferenciable, dado $(a,b) \in \mathbb{R} \times \mathbb{R}$ que verifica la ecuación (1.45), es decir $f(a,b) = 0$, si $\partial_2 f(a,b) \neq 0$, entonces existe $\varepsilon > 0$ y una función continuamente derivable $\phi : [a-\varepsilon, a+\varepsilon] \to \mathbb{R}$ tal que

$$b = \phi(a) \quad \text{y} \quad f(x,\phi(x)) = 0 \quad \text{para todo} \quad x \in [a-\varepsilon, a+\varepsilon]. \tag{1.46}$$

Aplicando la regla de la cadena para derivar la función nula $x \in [a-\varepsilon, a+\varepsilon] \to f(x,\phi(x))$, se tendrá

$$\phi'(a) = -\frac{\partial_1 f(a,b)}{\partial_2 f(a,b)}. \tag{1.47}$$

Desarrollemos esta idea para la ecuación $f(x,y)=0$ definida por la función de clase $\mathcal{C}^1$ $f(x,y):=(x-1)^2+(y-1)^2-1$. Consideremos entonces $(a,b):=(\frac{1}{2},1-\frac{\sqrt{3}}{2})$, que evidentemente verifica $f(a,b)=0$ y además $\partial_2 f(a,b)=-\sqrt{3}\neq 0$. De acuerdo a lo que decíamos, debe entonces existir una función $\phi:[a-\varepsilon,a+\varepsilon]\to\mathbb{R}$ continuamente derivable tal que $\phi(a)=b$ y $f(x,\phi(x))=0$ para todo $x\in[a-\varepsilon,a+\varepsilon]$. No es difícil deducir que la función $\phi:[\frac{1}{4},\frac{3}{4}]\to\mathbb{R}$ definida por $\phi(x)=1-\sqrt{1-(x-1)^2}$ cumple con este requerimiento. Se tiene además $\phi'(\frac{1}{2})=\frac{1}{\sqrt{3}}=-\frac{\partial_1 f(\frac{1}{2},1-\sqrt{\frac{3}{2}})}{\partial_2 f(\frac{1}{2},1-\frac{\sqrt{3}}{2})}$.

Este mismo razonamiento lo podemos hacer para cualquier punto $(a,b)\in\mathbb{R}\times\mathbb{R}$ de la forma $(a,1\pm\sqrt{1-(a-1)^2})$ con $a\in]0,2[$, pues todos esos puntos verifican la ecuación $f(x,y)=0$ y además la derivada parcial de f con respecto a la segunda variable, en esos puntos, no se anula. Esto no sería posible sin embargo, en $(0,1)$ y $(2,1)$ donde $\partial_2 f$ se anula, hecho que en este ejemplo es fácil de evaluar directamente. El teorema de la función implícita que daremos a continuación generaliza este problema a la ecuación $f(x,y)=0$, definida por una función continuamente diferenciable $f:A\to X_2$, donde A es un abierto en el e.v.n. $X_1\times X_2$, siendo X_1 y X_2 dos espacios de Banach.

Teorema 1.7.2. *Dados dos espacios de Banach X_1 y X_2, un abierto $A\subset X_1\times X_2$ y una función continuamente diferenciable $f:A\to X_2$, definimos la ecuación*

$$f(x_1,x_2)=0. \tag{1.48}$$

Entonces, si $x_0=(x_1,x_2)\in A$ verifica la ecuación (1.48) *y si $D_2f(x_0)$ (ver Definición 1.6.1) es biyectiva en X_2, se tiene que existe una bola $B(x_1,\varepsilon)\subset X_1$ y una función continuamente diferenciable $\phi:B(x_1,\varepsilon)\to X_2$ tales que*

$$x_2=\phi(x_1)\ \ \text{y}\ \ f(x,\phi(x))=0\quad \textit{para todo}\quad x\in B(x_1,\varepsilon). \tag{1.49}$$

Se tiene además la fórmula

$$D\phi(x_1)=D_2f(x_0)^{-1}\circ D_1f(x_0). \tag{1.50}$$

Demostración. Definamos la función $\varphi:A\subseteq X_1\times X_2\to X_1\times X_2$ por $\varphi(x):=(x_1,f(x))$, es decir, $\varphi:=(p_1,f)$ (ver (1.38) para la definición de p_1). Entonces, del Teorema 1.4.6, concluimos que φ es continuamente diferenciable y $D\varphi(x_0)=(p_1,Df(x_0))$ (pues la proyección p_1 es una función lineal continua). Por otra parte, del Teorema 1.6.1 sabemos que $Df(x_0)(v)=D_1f(x_0)(v_1)+D_2f(x_0)(v_2)$ para todo $v\in X_1\times X_2$, lo que muestra que $D\varphi(x_0)$ es biyectiva y

$$D\varphi(x_0)^{-1}(u_1,u_2)=(u_1,[-D_2f(x_0)^{-1}\circ D_1f(x_0)](u_1)+D_2f(x_0)^{-1}(u_2)).$$

Concluimos entonces, del Teorema 1.7.1 que existe un abierto $V\subset X_1\times X_2$ que contiene a x_0 tal que la restricción $\tilde{\varphi}:V\to\varphi(V)$ de la función φ, es biyectiva y $\tilde{\varphi}^{-1}:\varphi(V)\to V$ es continuamente diferenciable. Además se tiene que

$$D\tilde{\varphi}^{-1}(b)=D\varphi(x_0)^{-1}$$

donde $b:=\varphi(x_0)$. Es fácil entonces verificar que la función $\phi:B(x_1,\varepsilon)\to X_2$ definida por $\phi(u_1):=\tilde{\varphi}_2^{-1}(u_1,0)$, donde $\tilde{\varphi}_2^{-1}$ es la función componente de $\tilde{\varphi}^{-1}$ en X_2 (es decir $\tilde{\varphi}^{-1}=(\tilde{\varphi}_1^{-1},\tilde{\varphi}_2^{-1})$) y donde $\varepsilon>0$ es tal que $B(x_1,\varepsilon)\times\{0\}\subset\varphi(V)$, verifica (1.49) y (1.50). $\square$

1.8. Derivadas parciales de orden superior

DEFINICIÓN 1.8.1. Una función f definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. Y, se dirá de clase $\mathcal{C}^2$ si ella y sus derivadas parciales $\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}$ son de clase $\mathcal{C}^1$. La derivada parcial de la función $\frac{\partial f}{\partial x_i}$ con respecto a la variable x_j la denotaremos $\frac{\partial^2 f}{\partial x_j \partial x_i}$ o bien $\partial^2_{j,i} f$. Estas funciones se llaman derivadas parciales de segundo orden de la función f. Por recurrencia sobre k, diremos que f es de clase $\mathcal{C}^k$ si ella es de clase $\mathcal{C}^{k-1}$ y sus derivadas parciales de orden $k-1$, que denotaremos por

$$\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} ... \partial x_{i_1}} \quad \text{para todo } \; i_1, ..., i_{k-1} \in \{1, ..., n\}, \tag{1.51}$$

son de clase $\mathcal{C}^1$. Las derivadas parciales de orden k de f las denotaremos $\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} ... \partial x_{i_1}}$ o bien $\partial^k_{i_k, i_{k-1}, ..., i_1} f$.

NOTA 1.8.1. Es una consecuencia casi inmediata de la definición anterior que si f y g son dos funciones de clase $\mathcal{C}^k$, definidas en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. Y, entonces :

(i) las funciones $f+g$ y λf también son de clase $\mathcal{C}^k$ y se tiene

$$\partial^k_{i_k,...,i_1}(f+g)(x_0) = \partial^k_{i_k,...,i_1} f(x_0) + \partial^k_{i_k,...,i_1} g(x_0) \quad \text{para todo } x_0 \in A \tag{1.52}$$

$$\partial^k_{i_k,...,i_1}(\lambda f)(x_0) = \lambda \partial^k_{i_k,....,i_1} f(x_0) \quad \text{para todo } x_0 \in A. \tag{1.53}$$

(ii) Cuando $Y = \mathbb{R}$ la función $f \cdot g$ también es de clase $\mathcal{C}^k$.
Cuando $Y = \mathbb{R}^m$ se tendrá que f es de clase $\mathcal{C}^k$ si y solo si cada una de sus m funciones componentes es de clase $\mathcal{C}^k$, y en ese caso

$$\partial^k_{i_k,...,i_1} f(x_0) = (\partial^k_{i_k,...,i_1} f_1(x_0), ..., \partial^k_{i_k,...,i_1} f_m(x_0)) \tag{1.54}$$

TEOREMA 1.8.1. *Sea f una función de clase $\mathcal{C}^2$, definida en un abierto A de $\mathbb{R}^2$ con valores en un e.v.n. Y. Entonces se tendrá para todo $x_0 \in A$, la fórmula*

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0). \tag{1.55}$$

DEMOSTRACIÓN. Sean $x_0 \in A$ y $\varepsilon > 0$. Por continuidad de las derivadas parciales de orden dos, existirá $\eta > 0$ tal que

$$\|x\|_\infty \leq \eta \Rightarrow \|\partial^2_{2,1} f(x_0+x) - \partial^2_{2,1} f(x_0)\| \leq \varepsilon \;\text{ y }\; \|\partial^2_{1,2} f(x_0+x) - \partial^2_{1,2} f(x_0)\| \leq \varepsilon. \tag{1.56}$$

Si e_1 y e_2 son los vectores de la base canónica en $\mathbb{R}^2$ y $x = x_1 e_1 + x_2 e_2$, denotemos

$$\Delta^2 f(x) := f(x_0 + x_1 e_1 + x_2 e_2) - f(x_0 + x_1 e_1) - f(x_0 + x_2 e_2) + f(x_0)$$

y demostremos que

$$\|x\|_\infty \leq \eta \Rightarrow \|\Delta^2 f(x) - x_1 x_2 \partial^2_{2,1} f(x_0)\| \leq \varepsilon |x_1||x_2|. \tag{1.57}$$

Si definimos $\varphi(t) := f(x_0 + te_1 + x_2e_2) - f(x_0 + te_1) - tx_2\partial^2_{2,1}f(x_0)$, aplicando el teorema del valor medio (t.v.m.) a esta función, vemos que existe $\bar{t} \in [0, x_1]$ tal que

$$\begin{aligned}\Delta^2 f(x) - x_1x_2\partial^2_{2,1}f(x_0) &= \varphi(x_1) - \varphi(0) = \varphi'(\bar{t})x_1 \\ &= [\partial_1 f(x_0 + \bar{t}e_1 + x_2e_2) - \partial_1 f(x_0 + \bar{t}e_1) - \partial^2_{2,1}f(x_0)x_2]x_1\end{aligned}$$

y si definimos $\theta(\tau) := \partial_1 f(x_0 + \bar{t}e_1 + \tau e_2) - \tau\partial^2_{2,1}f(x_0)$ aplicando el t.v.m. a esta función, vemos que existe $\bar{\tau} \in [0, x_2]$ tal que

$$\begin{aligned}\Delta^2 f(x) - x_1x_2\partial^2_{2,1}f(x_0) &= [\theta(x_2) - \theta(0)]x_1 = \theta'(\bar{\tau})x_2x_1 \\ &= [\partial^2_{2,1}f(x_0 + \bar{t}e_1 + \bar{\tau}e_2) - \partial^2_{2,1}f(x_0)]x_2x_1.\end{aligned}$$

De esta última igualdad, usando la primera desigualdad en (1.56), obtenemos fácilmente la implicación (1.57).
Del mismo modo se puede obtener la implicación

$$\|x\|_\infty \leq \eta \Rightarrow \|\Delta^2 f(x) - x_1x_2\partial^2_{1,2}f(x_0)\| \leq \varepsilon|x_1||x_2|. \tag{1.58}$$

De (1.57) y (1.58), sumando se obtiene que

$$\|\partial^2_{2,1}f(x_0) - \partial^2_{1,2}f(x_0)\| \leq 2\varepsilon$$

y como ε es cualquiera, concluimos que $\partial^2_{2,1}f(x_0) = \partial^2_{1,2}f(x_0)$. □

TEOREMA 1.8.2. *Sea f una función de clase $\mathcal{C}^k$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. Y. Entonces para toda k-tupla $(i_1, ..., i_k) \in \{1, ..., n\}^k$ y toda biyección σ en $\{1, ..., k\}$ se tendrá para todo $x_0 \in A$ la fórmula*

$$\frac{\partial^k f}{\partial x_{i_1} ... \partial x_{i_k}}(x_0) = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} ... \partial x_{i_{\sigma(k)}}}(x_0).$$

DEMOSTRACIÓN. Si $k = 2$ es una consecuencia inmediata del teorema anterior. Si $k \geq 3$, se deduce del teorema anterior por inducción sobre k. □

1.9. Desarrollos limitados

DEFINICIÓN 1.9.1. Se llama polinomio de grado $N \in \mathbb{N}_*$ (conjunto de los naturales incluyendo el 0) en $\mathbb{R}^n$ con coeficientes en $\mathbb{R}$ a toda función $P : \mathbb{R}^n \to \mathbb{R}$ de la forma

$$P(v_1, ..., v_n) = \sum_{|i| \leq N} \alpha_i \; v_1^{i_1} \; v_2^{i_2} ... \; v_n^{i_n} \tag{1.59}$$

donde $i := (i_1, ..., i_n) \in \mathbb{N}^n_*$ es una n-tupla de componentes naturales, $|i| = \sum_{j=1}^{n} i_j$ y $\alpha_i \in \mathbb{R}$.

DEFINICIÓN 1.9.2. Sea A un conjunto abierto en $\mathbb{R}^n$ y f una función de A en $\mathbb{R}$. Se llama desarrollo limitado de orden N de la función f en un punto $x_0 \in A$ a todo polinomio $P : \mathbb{R}^n \to \mathbb{R}$ de grado N tal que para todo $v \in \mathbb{R}^n$ con $x_0 + v \in A$ se tiene

$$f(x_0 + v) - P(v) = o^N(v) \tag{1.60}$$

donde o^N es una función de $\mathbb{R}^n$ en $\mathbb{R}$ que verifica $o^N(0) = 0$ y

$$\lim_{\substack{v \to 0 \\ v \neq 0}} \frac{o^N(v)}{\|v\|^N} = 0. \tag{1.61}$$

NOTA 1.9.1. De la definición anterior es evidente que en (1.61) se puede usar cualquier norma en $\mathbb{R}^n$ y por lo tanto el desarrollo limitado de una función f no depende de la norma. Además es fácil ver que P es un desarrollo limitado de orden N de f en x_0, entonces $P(0) = f(x_0)$ deduciendo así que f es continua en x_0.

NOTA 1.9.2. De la Definición 1.3.1 se desprende de inmediato que f tiene un desarrollo limitado de orden 1 en un punto $x_0 \in A$ si y solo si es diferenciable en x_0. En este caso se tendrá

$$P(v) = f(x_0) + Df(x_0)(v) \tag{1.62}$$

que de acuerdo a la fórmula (1.20), podemos escribir

$$P(v) = f(x_0) + \langle \nabla f(x_0), v \rangle \tag{1.63}$$

que efectivamente es un polinomio de grado 1 en $\mathbb{R}^n$.

NOTA 1.9.3. De la Nota 1.3.2 vemos que el desarrollo limitado de orden 1 en x_0 de una función f, dado por (1.63), define la aproximación de primer orden de f en x_0, mediante la fórmula $h(x) = P(x - x_0)$. Del Teorema 1.3.1 concluimos que el desarrollo limitado de orden 1 de una función en un punto (que está dado por (1.62)) es siempre único.

NOTA 1.9.4. Dada una función f definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, su desarrollo limitado de orden N en $x_0 \in A$ define la llamada aproximación de orden N de f en x_0 que está dada por la función $h(x) := P_N(x - x_0)$.

Dos funciones f y h definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, se dicen tangentes de orden N en un punto $x_0 \in A$, si $f(x_0 + v) - h(x_0 + v) = o^N(v)$ donde o^N es una función de $\mathbb{R}^n$ en $\mathbb{R}$ que verifica $o^N(0) = 0$ y la relación (1.61).
De lo anterior vemos que la aproximación de orden N de la función f en $x_0 \in A$, corresponde al polinomio de grado N que es tangente de orden N a la función f en x_0.

TEOREMA 1.9.1. *Sea f una función definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Si el polinomio*

$$P_N(v) := \sum_{|i| \leq N} \alpha_i \; v_1^{i_1} \; v_2^{i_2} \ldots \; v_n^{i_n}$$

es un desarrollo limitado de orden N de la función f en $x_0 \in A$, entonces para todo $K < N$, el polinomio

$$P_K(v) = \sum_{|i| \leq K} \alpha_i \; v_1^{i_1} \; v_2^{i_2} \ldots \; v_n^{i_n}$$

es un desarrollo limitado de orden K de la función f en x_0.

DEMOSTRACIÓN. Puesto que

$$\lim_{v\to 0}\frac{\dfrac{f(x_0+v)-\sum\limits_{|i|\le K}\alpha_i\ v_1^{i_1}\dots\ v_n^{i_n}}{\|v\|^K}-\dfrac{\sum\limits_{K+1\le|i|\le N}\alpha_i\ v_1^{i_1}\dots\ v_n^{i_n}}{\|v\|^K}}{\|v\|^{N-K}}=0$$

y también (usando la norma infinito)

$$\lim_{v\to 0}\frac{\sum\limits_{K+1\le|i|\le N}\alpha_i\ v_1^{i_1}\dots\ v_n^{i_n}}{\|v\|^K}=0$$

deducimos que

$$\lim_{v\to 0}\frac{f(x_0+v)-\sum\limits_{|i|\le K}\alpha_i\ v_1^{i_1}\dots\ v_n^{i_n}}{\|v\|^K}=0$$

lo que equivale a decir que P_K es un desarrollo limitado de orden K de la función f en x_0. $\square$

TEOREMA 1.9.2. *Si una función f definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, admite un desarrollo limitado P_N de orden N en $x_0 \in A$, entonces éste es único.*

DEMOSTRACIÓN. Hagamos la demostración por inducción. En la Nota 1.9.3 vimos que la unicidad se tiene cuando $N = 1$. Supongamos ahora que f admite un desarrollo limitado P_N de orden N y que su desarrollo limitado P_{N-1} de orden $N-1$ es único. Demostremos entonces la unicidad de P_N.
Si denotamos P_N y P'_N dos desarrollos limitados de orden N de f en x_0, del Teorema 1.9.1 y del hecho que el desarrollo limitado de orden $N-1$ es único, podemos escribir

$$\begin{aligned} P_N(v) &= P_{N-1}(v)+\sum_{|i|=N}\alpha_i\ v_1^{i_1}\dots\ v_n^{i_n}\\ P'_N(v) &= P_{N-1}(v)+\sum_{|i|=N}\alpha'_i\ v_1^{i_1}\dots\ v_n^{i_n}\end{aligned}$$

y aplicando a P_N y P'_N la igualdad (1.60) y luego restando, se obtiene

$$\phi(v):=\sum_{|i|=N}(\alpha_i-\alpha'_i)\ v_1^{i_1}\dots\ v_n^{i_n}=o^N(v). \tag{*}$$

De este modo para todo $v \in \mathbb{R}^n$ y todo $\lambda \in \mathbb{R}$ se tiene $\phi(\lambda v) = \lambda^N\phi(v)$, es decir $\phi(v) = \frac{\phi(\lambda v)}{\lambda^N}$. De la igualdad (*) se obtiene entonces

$$\lim_{\lambda\to 0}\frac{\phi(\lambda v)}{\lambda^N}=\|v\|^N\lim_{\lambda\to 0}\frac{\phi(\lambda v)}{\|\lambda v\|^N}=0.$$

Esto implica que $\phi(v) = 0$ para todo $v \in \mathbb{R}^n$ y por lo tanto todos los coeficientes del polinomio ϕ deben ser nulos, esto es $\alpha_i - \alpha'_i = 0$ para todo $i \in \mathbb{N}^n_*$ con $|i| = N$. Todo esto nos permite concluir que $P_N = P'_N$. $\square$

TEOREMA 1.9.3. *Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Entonces para todo $x_0 \in A, f$ tiene un desarrollo limitado P_2 de orden dos en x_0. Se tiene además que*

$$P_2(v)=f(x_0)+\langle\nabla f(x_0),v\rangle+\frac{1}{2}\langle v,H(x_0)v\rangle \tag{1.64}$$

donde $H(x_0)$ es la matriz de $n \times n$, llamada matriz Hessiana de f en x_0, cuyos coeficientes son las derivadas parciales de orden dos de f en x_0, esto es

$$H(x_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) \quad \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x_0) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) \dots & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{pmatrix} \tag{1.65}$$

Demostración. Sea $r > 0$ tal que $B(x_0; r) \subset A$. Sea $v \in \mathbb{R}^n$ tal que $\|v\| \leq r$. Definamos entonces la función $\varphi : [0,1] \to \mathbb{R}$ por $\varphi(t) := f(x_0 + tv)$. Esta función es de clase $\mathcal{C}^1$ y de acuerdo al teorema fundamental del calculo se tiene

$$\varphi(1) = \varphi(0) + \int_0^1 \varphi'(t)dt \tag{*}$$

donde $\varphi'(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0 + tv)v_i$. Como las funciones $\frac{\partial f}{\partial x_i}$ son de clase $\mathcal{C}^1$, ellas son diferenciables y por lo tanto, para cada $i \in \{1, ..., n\}$ se tiene

$$\frac{\partial f}{\partial x_i}(x_0 + tv) = \frac{\partial f}{\partial x_i}(x_0) + \sum_{j=1}^n \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x_0)tv_j + o_i(tv)$$

y reemplazando en (*) obtenemos

$$\begin{aligned} f(x_0 + v) = f(x_0) \quad &+ \quad \int_0^1 \sum_{i=1}^n \left[\frac{\partial f}{\partial x_i}(x_0) + \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0)v_j t + o_i(tv)\right] v_i dt \\ &= \quad f(x_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0)v_i + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0)v_j v_i + \int_0^1 \sum_{i=1}^n o_i(tv)v_i dt \\ &= \quad f(x_0) + \langle \nabla f(x_0), v\rangle + \frac{1}{2}\langle v, H(x_0)v\rangle + \int_0^1 \sum_{i=1}^n o_i(tv)v_i dt. \end{aligned}$$

De acuerdo a la Definición 1.9.2. Para terminar la demostración debemos probar que la función

$$\sigma(v) := \int_0^1 \sum_{i=1}^n o_i(tv)v_i dt$$

verifica $\lim\limits_{\substack{v \to 0 \\ v \neq 0}} \frac{\sigma(v)}{\|v\|^2} = 0$.

Por definición de $o_i(\cdot)$, para todo $\varepsilon > 0$ existe $\eta > 0$ tal que

$$\|v\| < \eta \Rightarrow |o_i(v)| \leq \varepsilon\|v\| \quad \text{para todo} \quad i \in \{1, ..., \ n\}$$

entonces, con $\tilde{\eta} := \min\{\eta, r\}$ obtenemos, usando como ha sido habitual la norma ∞ en $\mathbb{R}^n$,

$$\|v\| \leq \tilde{\eta} \Rightarrow |\sigma(v)| \leq \int_0^1 \sum_{i=1}^n |v_i| \ \varepsilon \ \|v\| \ t \ dt \leq \varepsilon \ n\|v\|^2 \frac{1}{2}$$

lo que demuestra que $\lim\limits_{\substack{v\to 0\\ v\neq 0}} \frac{\sigma(v)}{\|v\|^2} \leq \varepsilon\frac{n}{2}$ y, como ε es cualquiera, concluimos que el límite es 0. □

Nota 1.9.5. Del teorema anterior vemos que si f es una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, su aproximación de orden 2 o aproximación cuadrática en $x_0 \in A$, definida en la Nota 1.9.4, está dada por la fórmula

$$h(x) = f(x_0) + \langle \nabla f(x_0), x - x_0\rangle + \frac{1}{2}\langle (x-x_0), H(x_0)(x-x_0)\rangle. \tag{1.66}$$

Nota 1.9.6. Es fácil verificar que la aproximación cuadrática de una función cuadrática definida en $\mathbb{R}^n$ con valores $\mathbb{R}$ es, en todo punto de $\mathbb{R}^n$, ella misma.

Teorema 1.9.4. *Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Entonces la función ∇f definida en A con valores en $\mathbb{R}^n$ es de clase $\mathcal{C}^1$ y*

$$D[\nabla f](x)(v) = H(x)v. \tag{1.67}$$

Demostración. Si f es de clase $\mathcal{C}^2$ se tendrá que las funciones $\partial_1 f, ..., \partial_n f$ son de clase $\mathcal{C}^1$, lo que es equivalente de acuerdo a los teoremas 1.4.1 y 1.4.3, a decir que ellas son continuamente diferenciables. Aplicando entonces el Teorema 1.4.6 a la función $\nabla f := (\partial_1 f, ..., \partial_n f)$ concluimos que ella es continuamente diferenciable y de la fórmula (1.17) vemos que

$$D[\nabla f](x) = (D[\partial_1 f](x), ..., D[\partial_n f](x)).$$

Aplicando ahora el Teorema 1.3.5 a las funciones $\partial_i f$, obtenemos para cada $i = 1, ..., n$

$$\begin{aligned} D[\partial_i f](x)(v) &= \sum_{j=1}^{n} \partial_{j,i} f(x) v_j \\ &= (\partial_{1,i} f(x), ..., \partial_{n,i} f(x))v \end{aligned}$$

que no es otra que la fórmula (1.67) escrita componente a componente. □

Nota 1.9.7. Terminaremos esta sección enunciando el teorema que generaliza la fórmula (1.64) al caso en que la función f es de clase $\mathcal{C}^N$.

Teorema 1.9.5. *Sea f una función de clase $\mathcal{C}^N$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Entonces para todo $x_0 \in A, f$ tiene un desarrollo limitado de orden N en x_0, dado por el polinomio*

$$P_N(v) = \sum_{|i|\leq N} \alpha_i \quad v_1^{i_1} v_2^{i_2} .... v_n^{i_n} \tag{1.68}$$

donde

$$\alpha_i = \frac{1}{i_1! ... i_n!} \quad \frac{\partial^{|i|} f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(x_0), \quad \alpha_0 = f(x_0) \tag{1.69}$$

con la convención $0! = 1$.

Nota 1.9.8. Del Teorema anterior vemos que si f es una función de clase $\mathcal{C}^N$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, su aproximación de orden N en $x_0 \in A$, definida en la Nota 1.9.4, está dada por la fórmula

$$h(x) = \sum_{|i| \leq N} \alpha_i (x_1 - a_1)^{i_1} (x_2 - a_2)^{i_2} ... (x_n - a_n)^{i_n} \tag{1.70}$$

donde los coeficientes α_i están definidos por (1.69).
El polinomio (1.70) con los coeficientes definidos por (1.69) se llama también desarrollo de Taylor de orden N de la función f en x_0.

Índice alfabético