Diferenciación en $\mathbf{R}^n$

por Leandro Caniglia

1. Transformaciones lineales

En esta sección trazamos una rápida recorrida por nociones básicas del Algebra Lineal que suponemos conocidas por el lector. Aunque nuestro interés primordial es estudiar el espacio $\mathbf{R}^n$, sería ingenuo e incómodo disimular la existencia de espacios vectoriales más generales que $\mathbf{R}^n$. Por lo tanto, nuestro enfoque consistirá en definir con generalidad e ilustrar especialmente el caso que nos interesa. El lector debe saber que no es posible ni bueno ni desable evitar un curso dedicado íntegramente a estos temas.

Definición. Un *espacio vectorial real* es un conjunto $\mathcal{V}$ provisto de una operación de suma $(\vec{a}, \vec{b}) \mapsto \vec{a} + \vec{b}$, de $\mathcal{V} \times \mathcal{V}$ en $\mathcal{V}$, y de una acción $(\lambda, \vec{a}) \mapsto \lambda \cdot \vec{b}$, de $\mathbf{R} \times \mathcal{V}$ en $\mathcal{V}$, con las siguientes propiedades

i) [Asociatividad de la suma] $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$.
ii) [Conmutatividad de la suma] $\vec{a} + \vec{b} = \vec{b} + \vec{a}$.
iii) [Existencia de nuetro aditivo] Existe $\vec{0} \in \mathcal{V}$ tal que $\vec{a} + \vec{0} = \vec{a}$ para todo $\vec{a} \in \mathcal{V}$.
iv) [Existencia de inverso aditivo] Para cada $\vec{a} \in \mathcal{V}$ existe un $\vec{a}' \in \mathcal{V}$ tal que $\vec{a} + \vec{a}' = \vec{0}$.
v) [Acción trivial de 1] $1 \cdot \vec{a} = \vec{a}$ para todo $\vec{a} \in \mathcal{V}$.
vi) [Asociatividad de la acción] $\lambda \cdot (\mu \cdot \vec{a}) = (\lambda\mu) \cdot \vec{a}$.
vii) [Distributividad de la acción respecto a la suma] $\lambda \cdot (\vec{a} + \vec{b}) = \lambda \cdot \vec{a} + \lambda \cdot \vec{b}$.
viii) [Distributividad de la suma respecto a la acción] $(\lambda + \mu) \cdot \vec{a} = \lambda \cdot \vec{a} + \mu \cdot \vec{a}$.

Ejemplos. El más importante para nosotros es $\mathbf{R}^n$. Aquí tenemos las suma y la acción definidas en [*Nociones métricas y topológicas en* $\mathbf{R}^n$, §2]. Otro ejemplo lo constituyen el espacio $\mathbf{R}^{n\times m}$ de matrices de n filas y m columnas con la suma y la acción definidas coordenada a coordenada.

Definición. Una sucesión de *vectores* $B := \{\vec{v}_1, \dots, \vec{v}_n\}$, es decir, de elementos en un espacio vectorial $\mathcal{V}$, es una *base* si cada vector $\vec{v}$ de $\mathcal{V}$ puede expresarse de una única manera como

$$\vec{v} = \sum_{i=1}^{n} \lambda_i \vec{v}_i.$$

Si este es el caso, decimos que los escalares $\lambda_1, \dots, \lambda_n$ son las *coordenadas* de $\vec{v}$ en la base B.

Ejemlos. En $\mathbf{R}^n$ tenemos la base *canónica* $E := \{\vec{e}_1, \dots, \vec{e}_n\}$ en donde el $i$ésimo vector $\vec{e}_i$ es aquel cuyas coordenadas son todas nulas, excepto la $i$ésima la cual es 1. Así, para $n = 1$ la base canónica está formada por un único *vector*, a saber: el número real 1. En el caso $n = 2$ la base canónica es $\{(1, 0), (0, 1)\}$. En el caso $n = 3$ es $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. Para las matrices la base canónica es aquella definida por todas las matrices elementales E_{ij}, las cuales tienen todas sus entradas nulas excepto que en fila i columna j la matriz E_{ij} tiene 1. Asín en $\mathbf{R}^{2\times 3}$ tenemos

$$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad E_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Definición. Una *transformación lineal* de un espacio vectorial $\mathcal{V}$ en otro $\mathcal{W}$ es una aplicación $L\colon \mathcal{V} \to \mathcal{W}$ que verifica:

i) $L(\vec{0}_{\mathcal{V}}) = \vec{0}_{\mathcal{W}}$.

ii) $L(\vec{a} + \lambda \cdot \vec{b}) = L(\vec{a}) + \lambda \cdot L(\vec{b})$.

Ejercicio 1. Pruebe que L es una transformación lineal si, y sólo si, preserva combinaciones lineales, es decir:

$$L(\sum_{i=1}^{r} \lambda_i \cdot \vec{a}_i) = \sum_{i=1}^{r} \lambda_i \cdot L(\vec{a}_i)$$

cualesquiera sean los vectores $\vec{a}_1, \ldots, \vec{a}_r$ del dominio de L y los escalares $\lambda_1, \ldots, \lambda_r$.

Ejemplos. Las proyecciones canónicas $\pi_i\colon \mathbf{R}^n \to \mathbf{R}$ que a cada vector $\vec{a} = (a_1, \ldots, a_n)$ le asignan su $i$ésima coordanada a_i son ejemplos de transformaciones lineales. También la *rotación* $R(x, y) = (-y, x)$ es una transformación lineal de $\mathbf{R}^2$ en $\mathbf{R}^2$. En matrices cuadradas tenemos la *traza* definida como la suma de los elementos de la diagonal, es decir:

$$\begin{aligned} \mathrm{tr}\colon \mathbf{R}^{n\times n} &\longrightarrow \mathbf{R} \\ (a_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}} &\longmapsto a_{11} + \cdots + a_{nn}. \end{aligned}$$

También para las matrices tenemos la *transpuesta*

$$\begin{aligned} \mathbf{R}^{n\times m} &\longrightarrow \mathbf{R}^{m\times n} \\ (a_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}} &\longmapsto (a_{ji})_{\substack{1\le i\le n \\ 1\le j\le n}}. \end{aligned}$$

Si A es una matriz de $m \times n$, la aplicación

$$\begin{aligned} L_A\colon \mathbf{R}^n &\to \mathbf{R}^m \\ \vec{x} &\mapsto A \cdot {}^t\vec{x}, \end{aligned}$$

donde ${}^t\vec{x}$ es el vector columna que se obtiene transponiendo $\vec{x}$, es una transformación lineal. Recíprocamente, si $L\colon \mathbf{R}^n \to \mathbf{R}^m$ es una transformación lineal, podemos formar la matriz A_L cuya $i$ésima columna es el vector columna ${}^tL(\vec{e}_i)$, es decir, la imagen por L del $i$ésimo vector canónico de $\mathbf{R}^m$ en forma de columna.

En la definición de L_A estamos usando el producto de matrices, en el caso particular de una matriz de $m \times n$ por un vector columna de $n \times 1$. En general si A es una matriz de $m \times n$ y B otra de $n \times p$, se define el producto $C := AB$ como la matriz de $m \times p$ cuya coordenada en la fila i columna j es:

$$c_{ij} = \sum_{h=1}^{n} a_{ih} b_{hj}.$$

En particular el producto $A \cdot {}^t\vec{x}$ es un vector columna cuya coordenda $i$ésima es

$$\sum_{h=1}^{n} a_{ih} x_h,$$

es decir, coincide con el producto de la $i$ésima fila de A con $\vec{x}$.

Ejercicio 2. Dada una matriz A denotemos con $F_i(A)$ a la $i$ésima fila de A y con $C_j(A)$ a la $j$ésima columna. Pruebe que las aplicaciones $F_i\colon \mathbf{R}^{m\times n} \to \mathbf{R}^n$ y $C_j\colon \mathbf{R}^{m\times n} \to \mathbf{R}^m$ son transformaciones lineales. Demuestre también $F_i(A) = \vec{e}_i \cdot A$ y $C_j(A) = A \cdot {}^t\vec{e}_j$, donde $\vec{e}_i$ designa al $i$ésimo vector canónico de $\mathbf{R}^n$ y $\vec{e}_j$ al $j$ésimo de $\mathbf{R}^m$.

Las dos construcciones precendentes son recíprocas, es decir, (1) $A_{L_A} = A$ (la matriz asociada a la transformación lineal asociada a una matriz A es A) y (2) $L = L_{A_L}$ (la transformación lineal asociada a la matriz asociada a una transformación lineal L es L). La demostración de estos hechos queda como ejercicio para el lector.

Una transformación lineal L de $\mathbf{R}^n$ en $\mathbf{R}^m$ equivale a una matriz. Esto significa que queda determinada por nm números reales a_{ij}. Si consideramos a la matriz A como un vector de nm coordenadas, tiene sentido considerar su norma

$$\|A\| := \sqrt{\sum_{ij} a_{ij}^2}. \tag{1}$$

En particular, si como en el ejercicio anterior $F_i(A)$ designa a la $i$ésima fila de A, tenemos

$$\|A\|^2 = \sum_{i=1}^{m} \|F_i(A)\|^2.$$

Ahora dado $\vec{x} \in \mathbf{R}^n$ resulta que el vector $\vec{y} := L(\vec{x})$ verifica

$$|y_i| = \Big|\sum_{h=1}^{n} a_{ih} x_h\Big| = |\langle F_i(A), \vec{x}\rangle| \leq \|F_i(A)\| \cdot \|\vec{x}\|$$

en virtud de [*Nociones métricas y topológicas en* $\mathbf{R}^n$, Prop. 3.1]. Luego

$$\|\vec{y}\|^2 = \sum_{i=1}^{m} y_i^2 \leq \big(\sum_{i=1}^{n} \|F_i(A)\|^2\big) \|\vec{x}\|^2 \leq \|A\|^2 \|\vec{x}\|^2$$

es decir

$$\|L(\vec{x})\| \leq \|A\| \cdot \|\vec{x}\|. \tag{2}$$

En particular cuando $\|\vec{x}\| = 1$ deducimos

$$\|L(\vec{x})\| \leq \|A\|. \tag{3}$$

Estamos ahora en condiciones de dar la siguiente

Definición. La *norma* de una transformación lineal $L\colon \mathbf{R}^n \to \mathbf{R}^m$ es

$$\|L\| := \sup_{\|\vec{u}\|=1} \{\|L(\vec{u})\|\}.$$

Según (3) tal supremo existe y es menor o igual a $\|A\|$, definida en (1). Además vale la siguiente

1.1 Proposición. Dada una transformación lineal $L\colon \mathbf{R}^n \to \mathbf{R}^m$ llamemos α al ínfimo de todos los $a \in \mathbf{R}$ tales que

$$\|L(\vec{x})\| \leq a\|\vec{x}\| \quad \forall\, \vec{x} \in \mathbf{R}^n \tag{4}$$

y β al supremo

$$\beta := \sup_{\|\vec{v}\|\leq 1} \{\|L(\vec{v})\|\}.$$

Entonces

$$\alpha = \beta = \|L\|.$$

Demostración. Es claro que

$$\|L\| \leq \beta.$$

También si a verifica (4), entonces para cada $\|\vec{v}\| \leq 1$ tenemos

$$\|L(\vec{v})\| \leq a\|\vec{v}\| \leq a$$

de donde

$$\|L(\vec{v})\| \leq \alpha$$

y en consecuencia

$$\beta \leq \alpha.$$

Falta probar que $\alpha \leq \|L\|$. Dado un vector $\vec{x}$, si $\vec{x} \neq \vec{0}$, la definición de $\|L\|$ asegura que

$$\|L\left(\frac{\vec{x}}{\|\vec{x}\|}\right)\| \leq \|L\|.$$

Multiplicando por $\|\vec{x}\|$

$$\|L(\vec{x})\| \leq \|L\| \cdot \|\vec{x}\|.$$

Pero si $\vec{x} = \vec{0}$, la desigualdad anterior también es cierta; por lo tanto, $\|L\|$ es uno de los a que verifican (4). Así el ínfimo de tales a, a saber: α, es menor o igual a $\|L\|$ y la demostración queda completa ∎

La proposición anterior nos dice que el *tamaño* de una transformación lineal se define como el radio de la bola más pequeña que puede contener a la imagen de la bola de radio unitario. En este sentido, la norma $\|L\|$ se puede interpretar como un *índice de dilatación* si es mayor que 1, o de *contracción* si es menor que 1. Las rotaciones, por ejemplo, no cambian la norma de los vectores, es decir, no los dilatatan ni los contraen; esto se corresponde con el hecho de que las rotaciones tienen norma uno. Conviene aclarar que la dilatación o contracción medidas por la norma son *globales*, no *locales*. Es decir, miden lo que le ocurre a la bola unitaria en su conjunto, no a vectores individuales. Por ejemplo, la proyección $P: \mathbf{R}^2 \to \mathbf{R}^2$ que a cada (x, y) le hace corresponder el par $(x, 0)$, contrae a casi todos los vectores, excepto a los que tienen segunda coordenada nula. En efecto, $P(x, 0) = (x, 0)$ y por lo tanto $\|P(x, y)\| \leq \|(x, y)\|$, valiendo la igualdad cuando $y = 0$ y la desugualdad cuando $y \neq 0$. La imagen de la bola unitaria $B[(0, 0), 1]$ es el intervalo cerrado $[-1, 1]$. Por lo tanto, $\|P\| = 1$ a pesar que P contrae a los vectores con segunda componente no nula.

La proposición siguiente muestra que la norma definida para las transformaciones lineales tiene las propiedades que corresponden a una norma.

1.2 Proposición. Si L y L' son transformaciones lineales de $\mathbf{R}^n$ en $\mathbf{R}^m$, entonces

i) $\|L\| \geq 0$ y vale la igualdad si, y sólo si, $L = 0$ (aplicación nula).
ii) $\|\lambda L\| = |\lambda| \cdot \|L\|$.
iii) $\|L + L'\| \leq \|L\| + \|L'\|$ (desigualdad triangular).

Demostración. La primera propiedad es consecuencia directa de la proposición anterior puesto que con las notaciones de esa proposición, esta propieadad equivale a $\alpha \geq 0$ y $\alpha = 0$, si y sólo si $L(\vec{x}) = \vec{0}$ para todo $\vec{x} \in \mathbf{R}^n$. La segunda propiedad es evidente porque

$$\|\lambda L\| = \sup_{\|\vec{u}\|=1} \|(\lambda L)(\vec{u})\| = |\lambda| \cdot \sup_{\|\vec{u}\|=1} \|L(\vec{u})\| = |\lambda| \cdot \|L\|.$$

Para la tercera propiedad tomemos $\vec{u} \in \mathbf{R}^n$ de norma 1; entonces

$$\|(L + L')(\vec{u})\| = \|L(\vec{u}) + L'(\vec{u})\| \leq \|L(\vec{u})\| + \|L'(\vec{u})\| \leq \|L\| + \|L'\|$$

de donde, al tomar supremo sobre tales $\vec{u}$, nos queda $\|L + L'\| \leq \|L\| + \|L'\|$, como queríamos demostrar ∎

1.3 Proposición. Si $L: \mathbf{R}^n \to \mathbf{R}^m$ y $M: \mathbf{R}^m \to \mathbf{R}^p$ son dos transformaciones lineales, entonces su composición $M \circ L: \mathbf{R}^n \to \mathbf{R}^p$ también es una transformación lineal y además

$$\|M \circ L\| \leq \|M\| \cdot \|L\|.$$

Demostración. Dado $\vec{u} \in \mathbf{R}^n$ con $\|\vec{u}\| = 1$, tenemos

$$\|M \circ L(\vec{u})\| = \|M(L(\vec{u}))\| \leq \|M\| \cdot \|L(\vec{u})\| \leq \|M\| \cdot \|L\| \blacksquare$$

1.4 Proposición. Supongamos que $L: \mathbf{R}^n \to \mathbf{R}$ es una aplicación lineal con codominio en $\mathbf{R}$. Entonces su matriz tiene dimensión $1 \times n$ y por lo tanto es de la forma $(a_1, \ldots, a_n)$ para ciertos número reales a_i. Estos números verifican

$$\|L\| = \sqrt{a_1^2 + \cdots + a_n^2}.$$

En otras palabras, la norma de la transformación lineal asociada al vector $(a_1, \ldots, a_n)$ coincide con la norma de dicho vector.

Demostración. Según vimos en (3), la norma de L verifica

$$\|L\| \leq \sqrt{a_1^2 + \cdots + a_n^2}$$

y por lo tanto es suficiente demostrar la otra desigualdad. Pongamos

$$\alpha := \sqrt{a_1^2 + \cdots + a_n^2}$$

y definamos

$$\vec{v} := (a_1/\alpha, \ldots, a_n/\alpha).$$

El vector $\vec{v}$ tiene norma 1, por lo cual

$$\|L\| = \sup_{\|\vec{u}\|=1} |L(\vec{u})| \geq |L(\vec{v})| = \frac{1}{\alpha}|L(a_1, \ldots, a_n)| = \frac{1}{\alpha}(a_1^2 + \cdots + a_n^2) = \alpha \;\blacksquare$$

2. Diferenciación

Extendemos ahora la definición de diferencial que habíamos introducido en el caso de funciones de una variable. En aquella oportunidad, dijimos que la diferencial de una función f definida en un entorno del punto x_0 era una aplicación lineal de la forma $x \mapsto a(x - x_0) + f(x_0)$. El coeficiente a de tal aplicación lineal era la derivada $f'(x_0)$. Con esa definición, la diferencial no es una transformación lineal puesto que, en general, su valor en $x = 0$ no será igual a 0. Ahora que tenemos que trabajar con funciones de varias variables, vamos a introducir un cambio que asegure que la diferencial sea una transformación lineal. Eso nos permitirá aplicar lo que vimos acerca de trasnformaciones lineales a las diferenciales. La razón de no haber hecho eso desde el comienzo, es decir, cuando vimos el caso de una variable, es que ahí separar la parte lineal de la parte constante hubiera sido un tanto exagerado tratándose de funciones tan simples. Ahora, que tenemos que tratar el caso más general, conviene realizar esa separación. Por lo demás, la extensión es completamente natural.

Definición. Dadas dos funciones (de varias variables) $\vec{f}$ y $\vec{g}$ definidas en un entorno del punto $\vec{x}_0 \in \mathbf{R}^n$, decimos que $\vec{f}$ es *tangente en* $\vec{x}_0$ *a* g y escribimos $\vec{f} \sim_{\vec{x}_0} \vec{g}$, si $\vec{f}(\vec{x}_0) = \vec{g}(\vec{x}_0)$ y además

$$\lim_{\vec{x}_0} \frac{\|\vec{f}(\vec{x}) - \vec{g}(\vec{x})\|}{\|\vec{x} - \vec{x}_0\|} = 0.$$

Del mismo modo que en el caso de una variable, se puede probar que la relación '$\sim_{\vec{x}_0}$' es de equivalencia, es decir, es reflexiva, simétrica y transitiva (demuestre). También es compatible con respecto al producto por escalares y a la suma, es decir:

$$\vec{f} \sim_{\vec{x}_0} \vec{g} \quad \Rightarrow \quad \lambda\vec{f} \sim_{\vec{x}_0} \lambda\vec{g} \tag{1}$$

y

$$\vec{f}_1 \sim_{\vec{x}_0} \vec{g}_1 \quad \text{y} \quad \vec{f}_2 \sim_{\vec{x}_0} \vec{g}_2 \quad \Rightarrow \quad \vec{f}_1 + \vec{f}_2 \sim_{\vec{x}_0} \vec{g}_1 + \vec{g}_2 \tag{2}$$

como se puede comprobar fácilmente (demuestre).

2.1 Lema. Si una transformación lineal es tangente en un punto a cero, entonces es la transformación nula.

Demostración. Llamemos L a laa transformación lineal y $\vec{x}_0$ al punto en donde L es tangente a cero. Por definición tenemos que $L(\vec{x}_0) = \vec{0}$ y además

$$\lim_{\vec{x}_0} \frac{\|L(\vec{x})\|}{\|\vec{x} - \vec{x}_0\|} = 0$$

Por lo tanto, dado $\epsilon > 0$, existe $\delta > 0$ tal que $0 < \|\vec{x}-\vec{x}_0\| \leq \delta$ implica $\|L(\vec{x})\| < \|\vec{x}-\vec{x}_0\|\epsilon/2$. Por otro lado, la definición de $\|L\|$ asegura la existencia de un vector $\vec{u}$ con $\|\vec{u}\| = 1$ tal que $\|L\| - \epsilon/2 < \|L(\vec{u})\|$. Luego $\vec{x} := \vec{x}_0 + \delta\vec{u}$ verifica

$$0 < \|\vec{x} - \vec{x}_0\| = \delta$$

y en consecuencia

$$\|L\| - \epsilon/2 < \|L(\vec{u})\| = \frac{1}{\delta}\|L(\vec{x})\| < \frac{1}{\|\vec{x} - \vec{x}_0\|}\|\vec{x} - \vec{x}_0\|\epsilon/2 = \epsilon/2.$$

Así, $\|L\| < \epsilon$ y como ϵ era arbitrario, debe ser $\|L\| = 0$. Esto implica que L es la transformación nula ∎

2.2 Lema. Si dos transformaciones lineales son tangentes en un punto, entonces son iguales.

Demostración. La resta de ambas transformaciones es tangente a cero. Por el lema anterior, la resta es nula, es decir, las transformaciones son iguales ∎

Definición. Supongamos que $\vec{f}$ es una función definida en un entorno de un punto $\vec{x}_0 \in \mathbf{R}^n$, con codominio en $\mathbf{R}^m$. Por el lema anterior, puede haber *a lo sumo* una transformación lineal $L\colon \mathbf{R}^n \to \mathbf{R}^m$ tal que $\vec{f}$ es tangente en $\vec{x}_0$ a la aplicación $\vec{x} \mapsto L(\vec{x}-\vec{x}_0)+\vec{f}(\vec{x}_0)$. Si tal transformación lineal L existe, la llamaremos *diferencial de $\vec{f}$ en $\vec{x}_0$* y la denotaremos $D\vec{f}(\vec{x}_0)$. Además, si este es el caso, diremos que $\vec{f}$ es *diferenciable en $\vec{x}_0$*.

Corrección. Para ser coherentes con la definición actual en el caso de funciones de una variable, vamos a corregir la definición de diferencial que dimos en [*Derivación* §1] de modo que la diferencial de una tal función f en un punto x_0 de su dominio sea la transformación lineal $x \mapsto f'(x_0)x$. Es decir, en lugar de llamar diferencial a la aplicación $x \mapsto f'(x_0)(x-x_0)+f(x_0)$, eliminamos ahora de $Df(x_0)$ la parte constante $f(x_0)-f'(x_0)x_0$ y nos quedamos con una lineal *pura*. Esta corrección no trae ninguna consecuencia negativa puesto que los teoremas que probamos para la aplicación $x \mapsto f'(x_0)(x-x_0)+f(x_0)$ siguen siendo válidos para la parte lineal. De hecho en esta sección vamos a volver a demostrar esos teoremas en el caso general de varias variables.

De la definición anterior surge inmediatamente que una función $\vec{f}$ es diferenciable en $\vec{x}_0$ si, y sólo si, en un entorno de $\vec{x}_0$ están definidas $\vec{f}$ y una aplicación $\vec{\alpha}$ tales que

$$\vec{f}(\vec{x}_0+\vec{s}) = \vec{f}(\vec{x}_0) + D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s})$$

de modo que para cada $\epsilon > 0$ existe $\delta > 0$ tal que

$$\|\vec{s}\| < \delta \quad \Rightarrow \quad \|\vec{\alpha}(\vec{s})\| < \epsilon\|\vec{s}\|$$

Esta caracterización será utilizada sin mayores aclaraciones en las demostraciones de esta sección.

La proposición siguiente se deduce inmediatamente de la definición anterior y de las relaciones (1) y (2).

2.3 Proposición. Supongamos que $\vec{f}$ es diferenciable en $\vec{x}_0$, entonces

i) $\vec{f}$ es continua en $\vec{x}_0$.

ii) Cualquiera sea $\lambda \in \mathbf{R}$, el producto $\lambda\vec{f}$ es diferenciable en $\vec{x}_0$ y $D(\lambda\vec{f})(\vec{x}_0) = \lambda D\vec{f}(\vec{x}_0)$.

iii) Si $\vec{g}$ es otra función diferenciable en $\vec{x}_0$, la suma $\vec{f}+\vec{g}$ también lo es y además $D(\vec{f}+\vec{g})(\vec{x}_0) = D\vec{f}(\vec{x}_0) + D\vec{g}(\vec{x}_0)$.

2.4 Proposición. Toda transformación lineal $L\colon \mathbf{R}^n \to \mathbf{R}^m$ es diferenciable en cada punto $\vec{x}_0 \in \mathbf{R}^n$ y además $DL(\vec{x}_0) = L$.

Demostración. Es trivial puesto que $L(\vec{x}) = L(\vec{x}-\vec{x}_0) + L(\vec{x}_0)$ ∎

Un tipo de aplicación lineal muy simple y que vamos a usar en esta sección y la siguiente es la *homotecia*. Una homotecia es una tranformación lineal que multiplica a cada número real ξ por un vector fijo $\vec{u}$. Se dice que $\vec{u}$ es la *razón* de la homotecia. Concretamente, si $\vec{h}$ es la homotecia de razón $\vec{u} \in \mathbf{R}^n$, entonces

$$\begin{aligned}\vec{h}\colon \mathbf{R} &\to \mathbf{R}^n\\ \xi &\mapsto \xi\cdot\vec{u}.\end{aligned}$$

El resultado anterior asegura que $\vec{h}$ es diferenciable y

$$D\vec{h}(\xi)(\lambda) = \lambda\cdot\vec{u}.$$

En algunas oportunidades vamos a utilizar una formulación equivalente a la diferenciabilidad en un punto. Tal formulación, se deduce sencillamente de las definiciones (demuestre) y es la siguiente:

Propiedad. Una función $\vec{f}$ definida en un entorno del punto $\vec{x}_0 \in \mathbf{R}^n$ con codominio en $\mathbf{R}^m$ es diferenciable en $\vec{x}_0$ si, y sólo si, existen una transformación lineal $L\colon \mathbf{R}^n \to \mathbf{R}^m$ y una función $\vec{\alpha}$ definida en un entorno V de $\vec{0} \in \mathbf{R}^n$ con codominio en $\mathbf{R}^m$ tales que

$$\vec{f}(\vec{x}_0+\vec{s}) = \vec{f}(\vec{x}_0) + L(\vec{s}) + \vec{\alpha}(\vec{s}) \quad (\forall \vec{s} \in V)$$

y

$$\lim_{\vec{0}} \frac{\|\vec{\alpha}(\vec{s})\|}{\|\vec{s}\|} = 0.$$

Además, cuando éste sea el caso, la transformación lineal L será la diferencial $D\vec{f}(\vec{x}_0)$.

2.5 Teorema. Supongamos que $\vec{f}$ es una aplicación diferenciable en un punto $\vec{x}_0 \in \mathbf{R}^n$ y que $\vec{g}$ es diferenciable en $\vec{y}_0 := \vec{f}(\vec{x}_0)$. Entonces $\vec{h} := \vec{g} \circ \vec{f}$ es diferenciable en $\vec{x}_0$ y además

$$D\vec{h}(\vec{x}_0) = D\vec{g}(\vec{y}_0) \circ D\vec{f}(\vec{x}_0)$$

Demostración. En primer lugar digamos que $\vec{h}$ está definida en un entorno de $\vec{x}_0$ en virtud de la continuidad de $\vec{f}$ en $\vec{x}_0$. En efecto, si $\vec{g}$ está definida en un entorno W de $\vec{y}_0$, entonces hay un entorno V de $\vec{x}_0$ tal que $\vec{f}$ está definida en V y además, $\vec{f}(V) \subseteq W$. Por lo tanto, $\vec{h}$ está definida en ese entorno V.

Por hipótesis, dado $0 < \epsilon < 1$, existe $\delta > 0$ tal que si $\|\vec{s}\| < \delta$ y $\|\vec{t}\| < \delta$, entonces

$$\vec{f}(\vec{x}_0 + \vec{s}) = \vec{f}(\vec{x}_0) + D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s}) \tag{3}$$

$$\vec{g}(\vec{y}_0 + \vec{t}) = \vec{g}(\vec{y}_0) + D\vec{g}(\vec{y}_0)(\vec{t}) + \vec{\beta}(\vec{t}) \tag{4}$$

con

$$\|\vec{\alpha}(\vec{s})\| \leq \epsilon\|\vec{s}\| \tag{5}$$

$$\|\vec{\beta}(\vec{t})\| \leq \epsilon\|\vec{t}\|. \tag{6}$$

Por otro lado, si llamamos $a := \|D\vec{f}(\vec{x}_0)\|$ y $b := \|D\vec{g}(\vec{y}_0)\|$, tenemos

$$\|D\vec{f}(\vec{x}_0)(\vec{s})\| \leq a\|\vec{s}\| \quad \text{y} \quad \|D\vec{g}(\vec{y}_0)(\vec{t})\| \leq b\|\vec{t}\|$$

por lo tanto

$$\|D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s})\| \leq \|D\vec{f}(\vec{x}_0)(\vec{s})\| + \epsilon\|\vec{s}\| \leq (a+1)\|\vec{s}\| \tag{7}$$

para $\|\vec{s}\| \leq \delta$. Luego, para $\|\vec{s}\| \leq \delta/(a+1)$ de (6) y (7) nos queda

$$\|\vec{\beta}(D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s}))\| \leq \epsilon\|D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s})\| \leq \epsilon(a+1)\|\vec{s}\|$$

y también

$$\|D\vec{g}(\vec{y}_0)(\vec{\alpha}(\vec{s}))\| \leq b\|\vec{\alpha}(\vec{s})\| \leq b\epsilon\|\vec{s}\|.$$

Luego

$$\begin{aligned}
\vec{h}(\vec{x}_0 + \vec{s}) &= \vec{g}(\vec{f}(\vec{x}_0 + \vec{s})) \\
&= \vec{g}(\vec{f}(\vec{x}_0) + D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s})) \\
&= \vec{g}(\vec{f}(\vec{x}_0)) + D\vec{g}(\vec{y}_0)(D\vec{f}(\vec{x}_0)(\vec{s})) + D\vec{g}(\vec{y}_0)(\vec{\alpha}(\vec{s})) + \vec{\beta}(D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s})) \\
&= \vec{h}(\vec{x}_0) + D\vec{g}(\vec{y}_0) \circ D\vec{f}(\vec{x}_0)(\vec{s}) + \gamma(\vec{s})
\end{aligned}$$

donde

$$\gamma(\vec{s}) := D\vec{g}(\vec{y}_0)(\vec{\alpha}(\vec{s})) + \vec{\beta}(D\vec{f}(\vec{x}_0)(\vec{s}) + \vec{\alpha}(\vec{s})).$$

Así, para $\|\vec{s}\| < \delta/(a+1)$ resulta

$$\|\vec{\gamma}(\vec{s})\| \leq b\epsilon\|\vec{s}\| + \epsilon(a+1)\|\vec{s}\| \leq (b+a+1)\epsilon\|\vec{s}\|$$

y como ϵ era arbitrario, esto termina la demostración ∎

Si $\vec{f}: D \to \mathbf{R}^m$ es una función definida en un subconjunto D de $\mathbf{R}^n$, habíamos visto en [*Nociones métricas y topológicas en* $\mathbf{R}^n$, §2] que $\vec{f}$ era de la forma $(f_1, \ldots, f_m)$, la función $f_i: D \to \mathbf{R}$ es la $i$ésima coordenada de $\vec{f}$, es decir, $f_i = \pi_i \circ \vec{f}$, donde $\pi_i: \mathbf{R}^m \to \mathbf{R}$ es la $i$ésima proyección canónica. Con estas notaciones podemos enunciar la siguiente

2.6 Proposición. Para que una función $\vec{f}$ con codominio en $\mathbf{R}^m$ sea diferenciable en un punto $\vec{x}_0 \in \mathbf{R}^n$ es necesario y suficiente que cada una de sus coordenadas f_i sea diferenciable en el mismo punto. Además, si este es el caso, es

$$D\vec{f}(\vec{x}_0) = (Df_1(\vec{x}_0), \ldots, Df_m(\vec{x}_0))$$

o sea

$$\pi_i \circ D\vec{f}(\vec{x}_0) = Df_i(\vec{x}_0).$$

Demostración. La condición es necesaria porque si $\vec{f}$ es diferenciable en $\vec{x}_0$, como π_i también lo es por ser una transformación lineal y además, por la misma razón, es $D\pi_i(\vec{x}_0) = \pi_i$, del teorema anterior resulta que f_i es diferenciable en $\vec{x}_0$ y

$$Df_i(\vec{x}_0) = D(\pi_i \circ \vec{f})(\vec{x}_0) = D\pi_i(\vec{f}(\vec{x}_0)) \circ D\vec{f}(\vec{x}_0) = \pi_i \circ D\vec{f}(\vec{x}_0).$$

Es suficiente porque si llamamos $L := (Df_1(\vec{x}_0), \ldots, Df_m(\vec{x}_0))$ resulta que la $i$ésima coordenada del vector

$$\vec{f}(\vec{x}_0 + \vec{s}) - \vec{f}(\vec{x}_0) - L(\vec{s})$$

es

$$f_i(\vec{x}_0 + \vec{s}) - f_i(\vec{x}_0) - Df_i(\vec{s})$$

con lo cual si todos los límites

$$\lim_{\vec{s} \to \vec{0}} \frac{|f_i(\vec{x}_0 + \vec{s}) - f_i(\vec{x}_0) - Df_i(\vec{x}_0)(\vec{s})|}{\|\vec{s}\|}$$

son nulos, también lo será

$$\lim_{\vec{s} \to \vec{0}} \frac{\|\vec{f}(\vec{x}_0 + \vec{s}) - \vec{f}(\vec{x}_0) - L(\vec{s})\|}{\|\vec{s}\|}$$ ∎

Al considerar funciones de varias variables es normal que formemos nuevas funciones haciendo la siguiente

Construcción. Dadas $\vec{f}: S \to \mathbf{R}^n$ y $\vec{g}: T \to \mathbf{R}^m$, donde S y T son subconjuntos de $\mathbf{R}^s$ y $\mathbf{R}^t$ respectivamente, formamos la función *producto* $\vec{f} \times \vec{g}$ del siguiente modo

$$\begin{aligned} U \times V: &\longrightarrow \mathbf{R}^{n+m} \\ (\vec{x}, \vec{y}) &\mapsto (\vec{f}(\vec{x}), \vec{g}(\vec{y})). \end{aligned}$$

Es decir, la función $\vec{f} \times \vec{g}$ toma $(s+t)$uplas en $U \times V$ y las aplica en $(n+m)$uplas en $\mathbf{R}^{n+m}$ dejando que $\vec{f}$ actúe sobre las s primeras coordenadas y $\vec{g}$ sobre las últimas t coordenadas.

Ejemplo. Supongamos que queremos expresar a la función $f: \mathbf{R}^2 \to \mathbf{R}$ que a cada par (x, y) le hace corresponder la suma $x^2 + y^3$ como composición de funciones más simples. Llamemos $p_2: \mathbf{R} \to \mathbf{R}$ a la función $x \mapsto x^2$, $p_3: \mathbf{R} \to \mathbf{R}$ a $x \mapsto x^3$ y $s: \mathbf{R}^2 \to \mathbf{R}$ a la suma $(x, y) \mapsto x + y$. Entonces la función origianl f se puede escribir como la siguiente composición:

$$f = s \circ p_2 \times p_3.$$

Si quisiéramos calcuar ahora la diferencial de f podríamos aplicar el teorema anterior y obtener:

$$Df(a, b) = Ds(a^2 + b^3) \circ D(p_2 \times p_3)(a, b).$$

Como s es una transformación lineal, su diferencial en cualquier punto coincide con s por lo tanto sólo nos falta averiguar al diferencial de la función producto. Ahora, dado que la diferencial en un punto es la aproximación lineal de la función en ese punto, es lógico pensar que la diferencial del producto es el producto de las diferenciales, ya que cada diferencial es una aproximación en las coordenadas que le correspondes. Es decir, el producto $Dp_2(a) \times Dp_3(b)$ es un buen candidado para una aproximación lineal de $p_2 \times p_3$; de hecho es una función lineal que aproxima en cada coordenada. Aceptando por un momento este resultado, llegamos a

$$Df(a, b) = s \circ (Dp_2(a) \times Dp_3(b))$$

Es decir,

$$Df(a, b)(x, y) = Dp_2(a)(x) + Dp_3(b)(y) = 2ax + 3b^2 y.$$

Queda por ver la propiedad del producto. Ese el proósito de la proposición siguiente. Pero primero veamos un

Ejercicio 1. Demuestre que si $\vec{f_1} \sim_{\vec{x}_0} \vec{g}_1$ y $\vec{f_2} \sim_{\vec{y}_0} \vec{g}_2$, entonces $\vec{f_1} \times \vec{f_2} \sim_{(\vec{x}_0, \vec{y}_0)} \vec{g}_1 \times \vec{g}_2$. Es decir, la tangencia se preversva por productos.

2.7 Proposición. Si dos funciones $\vec{f}$ y $\vec{g}$ son diferenciales respectivamente en $\vec{x}_0$ y $\vec{y}_0$, entonces su producto $\vec{f} \times \vec{g}$ resulta diferenciable en $(\vec{x}_0, \vec{y}_0)$ y además

$$D(\vec{f} \times \vec{g})(\vec{x}_0, \vec{y}_0) = D\vec{f}(\vec{x}_0) \times D\vec{g}(\vec{y}_0).$$

Es decir, la diferencial del producto es el producto de las diferenciales.

Demostración. Llamemos L a la diferencial de $\vec{f}$ en $\vec{x}_0$ y M a la de $\vec{g}$ en $\vec{y}_0$. El producto $L \times M$ es una transformación lineal (demuestre) tangente a $\vec{f} \times \vec{g}$ en $(\vec{x}_0, \vec{y}_0)$ (ejercicio anterior) ∎

3. Teorema del Valor Medio

El Teorema de Valor Medio que vimos en [*Derivación* §2] también cobra sentido cuando la función considerada toma valores en $\mathbf{R}^m$. El enunciado y la demostración de este caso más general son totalmente análogos al caso de una variable. Lo mismo ocurre con varios corolarios. Aquí repetimos las demostraciones cambiando los detalles que tienen que ver con el hecho de considerar $\mathbf{R}^m$ en lugar de $\mathbf{R}$.

3.1 Teorema. Supongamos que $\vec{f}\colon [a,b] \to \mathbf{R}^m$ y $\varphi\colon [a,b] \to \mathbf{R}$ son funciones continuas en $[a,b]$ y diferenciables en (a,b) tales que para todo $\xi \in (a,b)$ vale que

$$\|D\vec{f}(\xi)\| \leq \varphi'(\xi).$$

Entonces

$$\|\vec{f}(b) - \vec{f}(a)\| \leq \varphi(b) - \varphi(a).$$

Demostración. La demostración es prácticamente idéntica al caso de una variable. Por lo tanto, vamos a repetir aquí lo que hicimos en [*Derivación*, Teorema 2.1], cambiándo sólo lo que sea necesario para el caso de varias variables.

Vamos a probar que dado $\epsilon > 0$ es $\|\vec{f}(b) - \vec{f}(a)\| \leq \varphi(b) - \varphi(a) + \epsilon(b-a)$. Como el lado izquierdo de la desigualdad no depende de ϵ, esto será suficiente. Fijemos entonces un $\epsilon > 0$ y definamos

$$A := \{\xi \in [a,b] \mid a \leq x < \xi \Rightarrow \|\vec{f}(x) - \vec{f}(a)\| \leq \varphi(x) - \varphi(a) + \epsilon(x-a)\}.$$

El conjunto A no es vacío puesto que $a \in A$ (el antecedente de la implicación que determina la pertenencia a A es falso cuando $\xi = a$ y por lo tanto la implicación es verdadera en ese caso). Además el conjunto A tiene la siguiente propiedad: si contiene a un elemento ξ, contiene a todos los elementos del intervalo $[a,\xi]$ (demuestre). Por lo tanto, si llamamos $s := \sup A$, debe ser $A = [a,s)$ o $[a,s]$ (el supremo existe porque A no es vacío y está acotado superiormente por b). Veamos primero que, en realidad, $s \in A$. Si no fuera así, habría un $t \in [a,s)$ con $\|\vec{f}(t) - \vec{f}(a)\| > \varphi(t) - \varphi(a) + \epsilon(t-a)$. Por definición de supremo habría un $\xi \in A$ tal que $t < \xi$. Pero por ser ξ un elemento de A y t un elemento entre a y ξ debería verificarse la desigualdad contraria $\|\vec{f}(t) - \vec{f}(a)| \leq \varphi(t) - \varphi(a) + \epsilon(t-a)$. Esta contradicción muestra que debe ser $s \in A$ y en consecuencia $A = [a,s]$. Falta probar que $s = b$. Supongamos que $s < b$. Por ser $\vec{f}$ y φ diferenciables en s, existe $\delta > 0$ tal que

$$\|\vec{f}(x) - \vec{f}(s) - D\vec{f}(s)(x-s)\| \leq \frac{\epsilon}{2}(x-s) \quad \text{y} \quad |\varphi(x) - \varphi(s) - \varphi'(s)(x-s)| \leq \frac{\epsilon}{2}(x-s)$$

para todo $x \in [s, s+\delta)$. Luego, para tales x tenemos

$$\begin{aligned}
\|\vec{f}(x) - \vec{f}(s)\| &\leq \|\vec{f}(x) - \vec{f}(s) - D\vec{f}(s)(x-s)\| + \|D\vec{f}(s)(x-s)\| \\
&\leq \frac{\epsilon}{2}(x-s) + \|D\vec{f}(s)\| \cdot |x-s| \\
&\leq \frac{\epsilon}{2}(x-s) + \varphi'(s)(x-s) \\
&= \frac{\epsilon}{2}(x-s) + (\varphi'(s)(x-s) - \varphi(x) + \varphi(s)) + (\varphi(x) - \varphi(s)) \\
&\leq \frac{\epsilon}{2}(x-s) + \frac{\epsilon}{2}(x-s) + \varphi(x) - \varphi(s) \\
&\leq \varphi(x) - \varphi(s) + \epsilon(x-s)
\end{aligned}$$

de donde

$$\begin{aligned}\|\vec{f}(x)-\vec{f}(a)\| &\le \|\vec{f}(x)-\vec{f}(s)\| + \|\vec{f}(s)-\vec{f}(a)\| \\ &\le \varphi(x)-\varphi(s)+\epsilon(x-s)+\varphi(s)-\varphi(a)+\epsilon(s-a) \\ &= \varphi(x)-\varphi(a)+\epsilon(x-a)\end{aligned}$$

por lo que $s+\delta \in A$. Esto contradice el hecho de que s era el supremo de A. Luego no puede ser $s<b$ y en consecuencia es $s=b$ como debíamos demostrar ∎

Ejercicio 1. Explique por qué razón el Teorema del Valor Medio para varias variables es un corolario directo del mismo teorema para una variable.

Siguen ahora los mismos colorarios que en el caso de una variable.

3.2 Corolario. Si $\vec{f}\colon [a,b] \to \mathbf{R}^m$ es una función continua en un intervalo compacto $[a,b]$, y para cada $\xi \in (a,b)$ existe la diferencial $D\vec{f}(\xi)$ y es $\|D\vec{f}(\xi)\| \le M$ para cierta constante M, entonces $\|\vec{f}(b)-\vec{f}(a)\| \le M(b-a)$.

Demostración. Es suficiente tomar en el teorema φ como la función $x \mapsto M(x-a)$ ∎

3.3 Corolario. Si una función $\vec{f}\colon [a,b] \to \mathbf{R}^m$ es continua tiene diferencial nula en todo punto del intervalo abierto (a,b), entonces es una función constante.

Demostración. Dado un punto cualquiera $\xi \in [a,b]$, el corolario anterior aplicado en el intervalo $[a,\xi]$ con $M=0$ nos dice que $\vec{f}(a)=\vec{f}(\xi)$. Como ξ es arbitrario esto significa que $\vec{f}$ es la función constante $\vec{f}(a)$ ∎

3.4 Corolario. Supongamos que $\vec{f}\colon A \to \mathbf{R}^m$ es una función continua definida en un abierto A de $\mathbf{R}^n$ y que el segmento que une a dos puntos $\vec{x}_0$ y $\vec{x}_0+\vec{t}$ de A está incluido en A (es decir, para todo $\lambda \in [0,1]$ el punto $\vec{x}_0+\lambda\cdot\vec{t}$ pertenece a A). Entonces si $\vec{f}$ es diferenciable en todo punto del segmento y M es una cota superior para la norma $\|D\vec{f}(\vec{x}_0+\xi\cdot\vec{t})\|$ de la diferencial de $\vec{f}$ a lo largo del segmento, vale que

$$\|\vec{f}(\vec{x}_0+\vec{t})-\vec{f}(\vec{x}_0)\| \le M\|\vec{t}\|.$$

Demostración. Definamos primero $\vec{h}$, la homotecia de razón $\vec{t}$ como

$$\vec{h}(\xi)=\xi\cdot\vec{t}.$$

Como $\vec{h}$ es lineal, resulta diferenciable e igual a su diferencial en cualquier punto. En particular

$$D(\vec{x}_0+\vec{h})(\xi)=\vec{h}$$

cualquiera sea $\xi \in (0,1)$. Definamos ahora $\vec{g}\colon [0,1] \to \mathbf{R}^m$ como la composición

$$\vec{g}(\xi)=\vec{f}\circ(\vec{x}_0+\vec{h}).$$

Por ser $\vec{g}$ una composición de dos funciones diferenciables, a saber: $\vec{f}$ y $\vec{x}_0+\vec{h}$, resulta que $\vec{g}$ también es diferenciable y además

$$D\vec{g}(\xi)=D\vec{f}(\vec{x}_0+\xi\cdot\vec{t})\circ\vec{h}.$$

Por lo tanto

$$\|D\vec{g}(\xi)\| \le M\cdot\|\vec{h}\| = M\cdot\|\vec{t}\|$$

y el resultado se sigue del coro 2.8 aplicado al intervalo $[0,1]$ y a la función $\vec{g}$ ∎

4. Derivadas parciales y direccionales

La diferencial de una función $\vec{f}$ de n variables da una aproximación lineal en un punto $\vec{x}_0$ del dominio de $\vec{f}$. Esta aproximación lineal toma en cuenta el comportamiento de $\vec{f}$ en las proximidades de $\vec{x}_0$. En el caso de dimensión superior a 1, uno puede aproximarse a un punto por caminos totalmente diferentes. Por ejemplo, en el caso de dos variables x e y uno puede aproximarse al punto (a, b) por rectas paralelas a los ejes coordenados, es decir, por puntos de la forma $(a+\xi, b)$ con $\xi \to 0$ o de la forma $(a, b+\xi)$ con $\xi \to 0$. Pero también puede acercarse siguiendo la dirección de otro vector (u, v) por puntos de la forma $(a+\xi u, b+\xi v)$ donde $\xi \to 0$. Incluso estas formas de acercarse a (a, b) no agotan todas las posiblidades porque es posible seguir la trayectoria de una curva cualquiera que termine en (a, b). En esta sección vamos a estudiar qué ocurre cuando el acercamiento a $\vec{x}_0$ sigue la dirección de algún vector que puede ser paralelo a los ejes, o más generalmente, un vector cualquiera. Ese estudio nos va a conducir a las derivadas *parciales* y *direccionales* y su relación con la diferencial.

Definición. Supongamos que $\vec{f}$ es una función definida en un entorno de $\vec{x}_0 \in \mathbf{R}^n$ con codominio en $\mathbf{R}^m$. Dado un vector $\vec{u}$ de norma 1, decimos que $\vec{f}$ es *diferenciable en $\vec{x}_0$ en la dirección de $\vec{u}$* si la función $\vec{g}$ definida en un entorno de $0 \in \mathbf{R}$ como

$$\xi \mapsto \vec{f}(\vec{x}_0 + \xi \cdot \vec{u})$$

es diferenciable en 0. En tal caso escribimos

$$D_{\vec{u}}\vec{f}(\vec{x}_0) := D\vec{g}(0).$$

4.1 Proposición. Si $\vec{f}$ es diferenciable en $\vec{x}_0$, entonces resulta diferenciable en cuaquier dirección $\vec{u}$ y además

$$D_{\vec{u}}\vec{f}(\vec{x}_0)(\xi) = D\vec{f}(\vec{x}_0)(\xi \cdot \vec{u}) = \xi \cdot D\vec{f}(\vec{x}_0)(\vec{u}),$$

es decir, la diferencial en la dirección de $\vec{u}$ coincide con la homotecia de razón $D\vec{f}(\vec{x}_0)(\vec{u})$.

Demostración. Llamemos $\vec{\eta}$ a la aplicación $\xi \mapsto \vec{x}_0 + \xi \cdot \vec{u}$. como $\vec{\eta}$ es la homotecia $\vec{h}$ de razón $\vec{u}$ más la constante $\vec{x}_0$, resulta que

$$D\vec{\eta}(0) = \vec{h}.$$

Segúnn las notaciones de la definición precendente tenemos:

$$\vec{g} = \vec{f} \circ \vec{\eta}$$

y en consecuencia

$$D_{\vec{u}}\vec{f}(\vec{x}_0) = D\vec{g}(0) = D\vec{f}(\vec{x}_0) \circ D\vec{\eta}(0) = D\vec{f}(\vec{x}_0) \circ \vec{h}$$

es decir

$$D_{\vec{u}}\vec{f}(\vec{x}_0)(\xi) = D\vec{f}(\vec{x}_0)(\vec{h}(\xi)) = D\vec{f}(\vec{x}_0)(\xi \cdot \vec{u})) = \xi \cdot D\vec{f}(\vec{x}_0)(\vec{u}) \blacksquare$$

Definición. Si $\vec{f}\colon A \to \mathbf{R}^m$ es una función definida en un entorno de un punto $\vec{x}_0 \in \mathbf{R}^n$, decimos que $\vec{f}$ es *diferenciable en $\vec{x}_0$ con respecto a la iésima variable* si es diferenciable en $\vec{x}_0$ en la dirección de el $i$ésimo vector canónico $\vec{e}_i$. En ese caso escribimos

$$D_i \vec{f}(\vec{x}_0)$$

para denotar a $D_{\vec{e}_i}\vec{f}(\vec{x}_0)$. En el caso $m = 1$, es decir, cuando $\vec{f}$ es una función $f\colon A \to \mathbf{R}$ con dominio en $\mathbf{R}$, la $i$ésima derivada parcial es la multiplicación por un número real, a saber: la derivada en 0 de la función $\xi \to f(\vec{x}_0 + \xi \cdot \vec{e}_i)$. En ese caso, a tal número real se lo denota

$$\left.\frac{\partial f}{\partial x_i}\right|_{\vec{x}_0}$$

y se lo llama *iésima derivada parcial de f en $\vec{x}_0$*.

4.2 Proposición. Si f es diferenciable en $\vec{x}_0$, entonces

$$\left.\frac{\partial f}{\partial x_i}\right|_{\vec{x}_0} = Df(\vec{x}_0)(\vec{e}_i)$$

o sea

$$Df(\vec{x}_0) = \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x_i}\right|_{\vec{x}_0} \pi_i,$$

donde π_i es la $i$ésima proyección canónica. Es decir, dado $\vec{a} = (a_1, \ldots, a_n) \in \mathbf{R}^n$ se tiene

$$Df(\vec{x}_0)(\vec{a}) = \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x_i}\right|_{\vec{x}_0} a_i.$$

Demostración. La primera igualdad es un caso particular de la proposición anterior. Las dos siguientes son consecuencia directa de la primera ∎

Si una función $f\colon A \to \mathbf{R}$, con dominio en un subconjunto abierto A de $\mathbf{R}^n$ y codominio en $\mathbf{R}$, tiene derivada parcial $i$ésima en cada punto de A, podemos pensar en la función

$$\frac{\partial f}{\partial x_i}\colon A \longrightarrow \mathbf{R}$$
$$\vec{x} \longmapsto \left.\frac{\partial f}{\partial x_i}\right|_{\vec{x}}$$

que a cada punto $\vec{x} \in A$ le asigna la $i$ésima derivada parcial en $\vec{x}$. La proposición anterior dice cómo están relacionadas las derivadas parciales con la diferencial *cuando la función es diferenciable*. El siguiente resultado permite determinar la diferenciabilidad de una función a partir de la existencia y continuidad de las derivadas parciales.

4.3 Teorema. Supongamos que $f\colon A \to \mathbf{R}$ es una función definida en un conjunto abierto $A \subseteq \mathbf{R}^n$. Una condición suficiente para que f sea diferenciable en cada punto de

A es todas las derivadas parciales $\partial f/\partial x_i$ estén definidas en A y que al menos $n-1$ de ellas sean continuas en A.

Demostración. Vamos a hacer la demostración para el caso $n = 2$ y dejar para el lector en el caso general. Sin pérdida de generalidad podemos suponer que la segunda derivada parcial es continua en A. Tomemos un punto $\vec{a} = (a_1, a_2) \in A$ y escribamos

$$\begin{aligned} f(a_1 + t_1, a_2 + t_2) - f(a_1, a_2) = f(a_1 + t_1, a_2 + t_2) - f(a_1 + t_1, a_2) \\ + f(a_1 + t_1, a_2) - f(a_1, a_2). \end{aligned} \tag{1}$$

Para simplifacar la notación introducimos

$$\partial_1 := \frac{\partial f}{\partial x_1} \quad \text{y} \quad \partial_2 := \frac{\partial f}{\partial x_2}.$$

Fijemos ahora un $\epsilon > 0$ arbitrario. Por la definición de derivada parcial existe $\delta > 0$ tal que

$$|t_1| < \delta \quad \Rightarrow \quad |f(a_1 + t_1, a_2) - f(a_1, a_2) - \partial_1(a_1, a_2)t_1| \leq \epsilon|t_1|. \tag{2}$$

Consideremos ahora un $\delta' > 0$ tal que todos los puntos de la forma $(a_1 + t_1, a_2 + t_2)$ con $|t_1|, |t_2| < \delta'$ pertenezcan a A. Ese δ' existe porque A es abierto. Para tales puntos está definida la función

$$\begin{aligned} \eta_{t_1} \colon (-r, r) &\longrightarrow \mathbf{R} \\ t &\mapsto f(a_1 + t_1, a_2 + t) - \partial_2(a_1 + t_1, a_2)t. \end{aligned}$$

Como esta función es diferenciable, podemos aplicarle el coro. 2.10 y deducir

$$|\eta_{t_1}(t) - \eta_{t_1}(0)| \leq M|t_2|$$

donde

$$M := \sup_{|t|<|t_2|} |\eta'_{t_1}(t)|.$$

Escribiendo estas relaciones en términos de f nos queda

$$|f(a_1 + t_1, a_2 + t_2) - \partial_2(a_1 + t_1, a_2)t_2 - f(a_1 + t_1, a_2)| \leq M|t_2| \tag{3}$$

donde

$$M = \sup_{|t|<|t_2|} |\partial_2(a_1 + t_1, a_2 + t) - \partial_2(a_1 + t_1, a_2)|.$$

Pero como ∂_2 es continua, es posible tomar δ' suficientemente chico como para que sea

$$M < \epsilon. \tag{4}$$

Apelando una vez más a la continuidad de ∂_2, existe $\delta'' > 0$ tal que

$$|\partial_2(a_1 + t_1, a_2) - \partial_2(a_1, a_2)| < \epsilon \tag{5}$$

si $|t_1| < \delta''$. Luego, si $r := \min(\delta, \delta', \delta'')$ usando (2), (3), (4) y (5) en (1) nos queda que para $|t_1|, |t_2| < r$ es

$$\begin{aligned}|f(a_1+t_1, a_2+t_2) - f(a_1, a_2) - \partial_1(a_1, a_2)t_1 - \partial_2(a_1, a_2)t_2|& \\ \leq |f(a_1+t_1, a_2+t_2) - f(a_1+t_1, a_2) - \partial_2(a_1+t_1, a_2)t_2|& \\ + |f(a_1+t_1, a_2) - f(a_1, a_2) - \partial_1(a_1, a_2)t_1|& \\ + |(\partial_2(a_1+t_1, a_2) - \partial_2(a_1, a_2))t_2|& \\ < 3\epsilon \max(|t_1|, |t_2|)& \\ \leq 3\epsilon \|(t_1, t_2)\|&\end{aligned}$$

lo cual demuestra que f es diferenciable en (a_1, a_2) con diferencial

$$(t_1, t_2) \longmapsto \partial_1(a_1, a_2)t_1 + \partial_2(a_1, a_2)t_2$$

como queríamos demostrar ∎