



Facultad de Ciencias Exactas, Ingeniería y Agrimensura

Departamento de Matemática - Escuela de Ciencias Exactas y Naturales

ANÁLISIS MATEMÁTICO II

Licenciatura y Profesorado en Física, Licenciatura en Ciencias de la Computación,

Licenciatura y Profesorado en Matemática - Año 2021

Unidad 10: Nociones de diferenciabilidad en varias variables

En esta unidad extenderemos la noción de diferenciabilidad a funciones de varias variables, la cual es más compleja. Empezaremos por ver funciones que dependen sólo de una variable, y donde la diferenciabilidad se extiende por definición de una manera natural a lo conocido en los reales.

1. Curvas diferenciables

Definición 111: Una *curva* en \mathbb{R}^n es una función $\alpha : I \rightarrow \mathbb{R}^n$, donde I es un intervalo abierto en \mathbb{R} . Decimos que la curva es continua si α es una aplicación continua.

Y decimos que α es derivable en el punto $t_0 \in I$ si existe el límite

$$\alpha'(t_0) = \lim_{h \rightarrow 0} \frac{\alpha(t_0 + h) - \alpha(t_0)}{h},$$

llamado la derivada de α .

Observemos que $\alpha(t)$ es un elemento de \mathbb{R}^n para cada t , y por lo tanto tiene coordenadas

$$\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)).$$

La derivada $\alpha'(t_0)$ existe si y sólo si existen las derivadas $\alpha'_i(t_0)$ esto es,

$$\alpha'_i(t_0) = \lim_{h \rightarrow 0} \frac{\alpha_i(t_0 + h) - \alpha_i(t_0)}{h}, \quad i = 1, \dots, n$$

donde $\alpha_i = p_i \circ \alpha : I \rightarrow \mathbb{R}$ es una función de variable real y a valores reales, y su derivada se obtiene como ya conocemos. Esto resulta de un enunciado similar al del Teorema 107 pero para límites:

Lema: Sea $D \subseteq \mathbb{R}^m$ y sea $f : D \rightarrow \mathbb{R}^n$ una función con coordenadas $f_1, f_2, \dots, f_n : D \rightarrow \mathbb{R}$.

Supongamos que a es un punto de acumulación de D . Entonces

$$\lim_{x \rightarrow a} f(x) = b = (b_1, b_2, \dots, b_n) \quad \text{si y solamente si} \quad \lim_{x \rightarrow a} f_i(x) = b_i, \forall i = 1, 2, \dots, n.$$

Demostración. Si $\lim_{x \rightarrow a} f(x) = b$, entonces para cada $i = 1, 2, \dots, n$ se tiene $\lim_{x \rightarrow a} f_i(x) = b_i$ pues $|f_i(x) - b_i| \leq \|f(x) - b\|$ y esta observación permite completar la prueba con ε y δ .

Recíprocamente, si $\lim_{x \rightarrow a} f_i(x) = b_i$ para cada $i = 1, 2, \dots, n$, entonces $\lim_{x \rightarrow a} f(x) = b$ pues $|f(x) - b| \leq \sum_{i=1}^n |f_i(x) - b_i|$, lo cual permite elegir convenientemente δ dado un ε cualquiera.

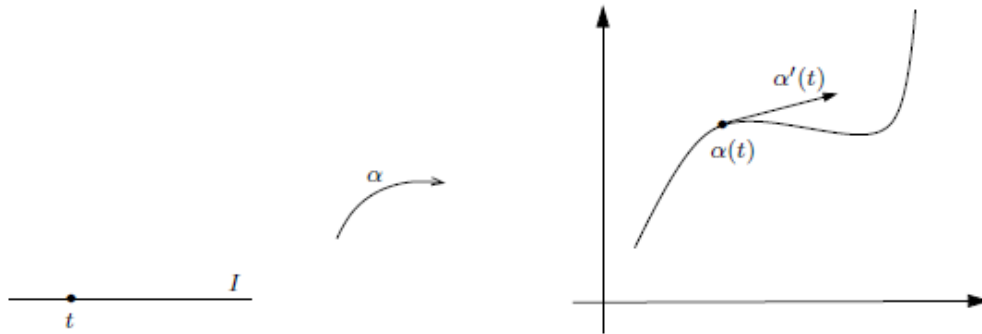
Al aplicar este resultado al cociente incremental obtenemos lo que deseamos. \square

Cuando la curva $\alpha : I \rightarrow \mathbb{R}^n$ es derivable en cada punto de I , decimos directamente que α es una curva derivable en I . En tal caso $t \rightarrow \alpha'(t)$ define una aplicación $\alpha' : I \rightarrow \mathbb{R}^n$. Si α' es continua decimos que la curva es de *clase* C^1 . Más generalmente para una curva $\alpha : I \rightarrow \mathbb{R}^n$ decimos que es de clase C^k si es derivable y α' es de clase C^{k-1} . Y para ser α de clase C^k es necesario y suficiente que cada coordenada α_i sea de clase C^k .

Para $\alpha : I \rightarrow \mathbb{R}^n$ derivable y tal que $\alpha'(t_0) \neq 0$, tenemos definida la *recta tangente a la curva en* $\alpha(t_0)$: que es el conjunto

$$r(s) = \alpha(t_0) + s\alpha'(t_0), \quad s \in \mathbb{R}.$$

Como podemos ver esa recta pasa por $\alpha(t_0)$ y tiene dirección el vector $\alpha'(t_0)$. El vector $\alpha'(t_0)$ se denomina *vector tangente a α en t_0* .



Ejemplo 112: Dados $a \neq b$ en \mathbb{R}^n , sea $\alpha : \mathbb{R} \rightarrow \mathbb{R}^n$ el segmento de recta que pasa por a y b : $f(t) = (1 - t)a + tb$. Claramente para todo $t \in \mathbb{R}$, la curva f resulta derivable y $f'(t) = b - a$.

Ejercicio: Compute la recta tangente a f en $t = 0$ y compare con f .

De la misma manera que existe para funciones a valores reales, podemos definir las derivadas a derecha o izquierda: si t_0 no es el extremo derecho del intervalo I entonces la *derivada a derecha* de una curva $\alpha : I \rightarrow \mathbb{R}^n$ en el punto t_0 está dada por

$$\alpha'_+(t_0) = \lim_{h \rightarrow 0^+} \frac{\alpha(t_0 + h) - \alpha(t_0)}{h},$$

y de modo análogo, tenemos la *derivada a izquierda* $\alpha'_-(t_0)$ siempre que en este caso t_0 no sea el extremo izquierdo del intervalo.

Ejercicio: Pruebe que cuando t_0 es un punto interior del intervalo I , entonces la curva $\alpha : I \rightarrow \mathbb{R}^n$ es derivable en I si y solamente si existen las derivadas a derecha y a izquierda de α en t_0 , y son iguales.

Ejemplo 113: Sea $f : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por $f(t) = (t, |t|)$. Entonces para $t > 0$ se tiene $f(t) = (t, t)$ y para $t < 0$ tendremos $f(t) = (t, -t)$. Luego para todo $t \neq 0$ existe la derivada $f'(t)$ que resulta

$$f'(t) = (1, 1) \text{ para } t > 0 \text{ y } f'(t) = (1, -1) \text{ para } t < 0.$$

En $t = 0$ existen las derivadas laterales $f'_+(0) = (1, 1)$ y $f'_-(0) = (1, -1)$, que son diferentes y por lo tanto f no es derivable en $t = 0$.

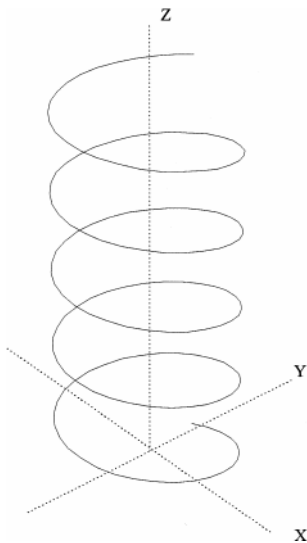
Por otro lado $g(t) = (t|t|, t^2)$ tiene la misma imagen que f pero es derivable en todos los puntos. La prueba queda como ejercicio.

Ejemplo 114: Circunferencia y hélice. Sea $f : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por

$$f(t) = (\cos(t), \sin(t))$$

y sea $g : \mathbb{R} \rightarrow \mathbb{R}^3$ dada por

$$g(t) = (\cos(t), \sin(t), t).$$



La imagen de f es una circunferencia de radio 1 centrada en el origen, denotada S^1 , y la imagen de g es una hélice (ver la figura a la izquierda), cuya proyección sobre el plano $z = 0$ es S^1 . Como ambas funciones f y g son de clase C^k para todo $k \in \mathbb{N}$, se dice que f y g son de clase C^∞ .

Ejercicio: Calcule las funciones derivadas $f'(t)$ y $g'(t)$.

Sean $\alpha, \beta : I \rightarrow \mathbb{R}^n$ curvas y $c : I \rightarrow \mathbb{R}$ una función real. Si α y β son derivables en el punto $t_0 \in I$ entonces también son derivables en t_0 las curvas $\alpha + \beta$, $c\alpha$, y también $\langle \alpha, \beta \rangle$, y $|\alpha| = \sqrt{\langle \alpha, \alpha \rangle}$, donde en este último caso debe valer $\alpha(t_0) \neq 0$.

Se cumplen además las siguientes propiedades:

1. $(\alpha + \beta)'(t_0) = \alpha'(t_0) + \beta'(t_0),$
2. $(c\alpha)'(t_0) = c'(t_0)\alpha(t_0) + c(t_0)\alpha'(t_0),$
3. $\langle \alpha, \beta \rangle'(t_0) = \langle \alpha'(t_0), \beta(t_0) \rangle + \langle \alpha(t_0), \beta'(t_0) \rangle,$

$$4. |\alpha|'(t_0) = \frac{\langle \alpha(t_0), \alpha'(t_0) \rangle}{|\alpha(t_0)|}.$$

Ejercicio: Pruebe estas propiedades, las cuales salen usando las expresiones en coordenadas de α y β .

Observación. En el ejemplo 114, se verifica que $f'(t) = (-\sin(t), \cos(t))$, vector que para cada t es perpendicular a $f(t)$. Más generalmente se puede probar que si $f : I \rightarrow \mathbb{R}^n$ es una curva tal que $|f|$ es constante, entonces $f'(t)$ es perpendicular a $f(t)$ para cada t .

En efecto, si $|f| = r$ es constante, (geométricamente esto significa que $f(t)$ está en la esfera de radio r para cada t), se tiene $\langle f(t), f(t) \rangle = r^2$ de donde al derivar respecto de t se obtiene $\langle f(t), f'(t) \rangle = 0$ (por las propiedades de arriba). Y esto prueba lo aseverado arriba.

Ejercicio Si $f : I \rightarrow \mathbb{R}^n$ es derivable y $f'(t) = 0$ (el vector 0) entonces f es constante.

Teorema 115: Regla de la cadena. Supongamos que $\varphi : I \rightarrow J$ es derivable en el punto $a \in I$ y $\alpha : J \rightarrow \mathbb{R}^n$ es una curva derivable en $b = \varphi(a)$. Entonces la curva $\alpha \circ \varphi : I \rightarrow \mathbb{R}^n$ es derivable en a y $(\alpha \circ \varphi)'(a) = \varphi'(a)\alpha'(b)$.

Demostración: Ejercicio. (Aplicamos la regla de la cadena a las funciones coordenadas $\alpha_i \circ \varphi : I \rightarrow \mathbb{R}$). □

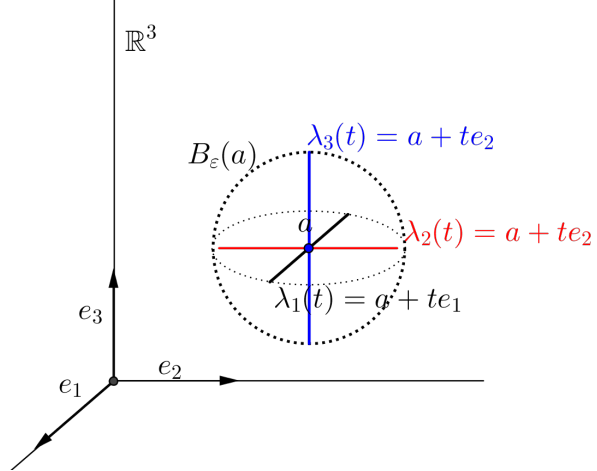
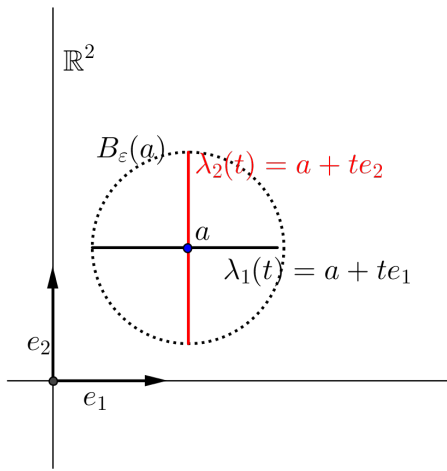
De un modo general la regla de la cadena dice que la curva $\alpha \circ \varphi$ cuya imagen está contenida en la imagen de α tiene vector velocidad un múltiplo del vector velocidad de α en $\varphi(t)$.

2. Funciones reales de n variables

En esta sección vamos a aplicar las ideas de la derivada de curvas para calcular otras derivadas y veremos algunos usos de estas nociones.

2.1. Derivadas parciales

En lo que sigue vamos a introducir las derivadas parciales. Para ello, empecemos considerando un abierto $U \subseteq \mathbb{R}^n$. Observemos que como U es abierto, dado un $a \in U$, existe una bola abierta centrada en a y completamente contenida en U . Luego existe un $\delta > 0$ tal que $a + te_i \in U$ para $t \in (-\delta, \delta)$ y donde e_i denota el vector i -ésimo de la base canónica, es decir, e_i es el vector donde todas las componentes son 0, salvo la i -ésima que es 1 (observemos que en el caso de \mathbb{R}^3 , muchas veces suele usarse la notación $e_1 = \mathbf{i}$, $e_2 = \mathbf{j}$, $e_3 = \mathbf{k}$). Por lo tanto el segmento de curva $\lambda_i(t) = a + te_i$ está bien definido para $t \in (-\delta, \delta)$.



Si ahora $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es una función cualquiera, están bien definidas las funciones $f \circ \lambda_i : \mathbb{R} \rightarrow \mathbb{R}$ en un entorno de 0, para $i = 1, \dots, n$, y por lo tanto tiene sentido analizar su derivabilidad en el sentido usual:

Definición 116: La i -ésima derivada parcial de $f : U \rightarrow \mathbb{R}$ en el punto $a \in U$ es la derivada en $t = 0$ de $f \circ \lambda_i$ y se denota por $\frac{\partial f}{\partial x_i}(a)$:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{t \rightarrow 0} \frac{f(a_1, a_2, \dots, a_i + t, \dots, a_n) - f(a)}{t},$$

si esta derivada existe.

Ejemplo 117: Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $f(x, y) = xy/(x^2 + y^2)$, si $(x, y) \neq (0, 0)$, y $f(0, 0) = 0$. Como $f(0, y) = 0$ para todo y y $f(x, 0) = 0$ para todo x resulta que las derivadas parciales existen en $(0, 0)$ y valen $\frac{\partial f}{\partial x}(0, 0) = 0 = \frac{\partial f}{\partial y}(0, 0)$

PERO f No es CONTINUA en $(0, 0)$. En efecto, fijemos un ángulo θ cualquiera y tomemos $x = t \cos(\theta)$, $y = t \sin(\theta)$. Si calculamos $f(x, y)$ a lo largo de la curva $(t \cos(\theta), t \sin(\theta))$, obtenemos $f(x, y) = \cos(\theta) \sin(\theta)$. Por lo tanto si nos acercamos a $(0, 0)$ por estas diferentes curvas, haciendo tender t a 0, vemos que el límite depende del valor que hayamos fijado para θ . Es decir, que por distintas direcciones nos acercamos a límites diferentes.

El ejemplo arriba muestra que la existencia de las n derivadas parciales en un punto, no asegura la continuidad de la función en ese punto. Claramente la derivada parcial $\frac{\partial f}{\partial x_i}(a) = (f \circ \lambda_i)'(0)$ da información sobre la f a lo largo del segmento que une $a - \delta e_i$ con $a + \delta e_i$.

Claramente la noción de derivada parcial tiene sentido para funciones $f : U \rightarrow \mathbb{R}^n$ donde $U \subseteq \mathbb{R}^m$ es un abierto. Si $a \in U$, entonces para cada $i = 1, \dots, m$ tenemos la i -ésima derivada parcial:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t}.$$

Evidentemente $\frac{\partial f}{\partial x_i}(a)$ es un vector en \mathbb{R}^n . Si $f = (f_1, f_2, \dots, f_n)$ entonces

$$\frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x_i}(a), \dots, \frac{\partial f_n}{\partial x_i}(a) \right).$$

Sea $f : U \rightarrow \mathbb{R}$ una función que posee las n derivadas parciales en todos los puntos del abierto $U \subseteq \mathbb{R}^n$. Quedan definidas entonces las funciones

$$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \rightarrow \mathbb{R}, \quad \text{donde } \frac{\partial f}{\partial x_i} : x \rightarrow \frac{\partial f}{\partial x_i}(x).$$

Si estas funciones fueran continuas en U , diremos que f es una función de clase C^1 y escribiremos $f \in C^1$.

Una aplicación $f : U \rightarrow \mathbb{R}^n$ definida en el abierto $U \subseteq \mathbb{R}^m$ se dice de clase C^1 cuando cada una de las funciones coordenadas $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ es de clase C^1 .

Muchas propiedades importantes de las funciones de clase C^1 resultan de ser diferenciables en el siguiente sentido.

Una función $f : U \rightarrow \mathbb{R}$ definida en el abierto $U \subseteq \mathbb{R}^n$ se dice *diferenciable* en el punto $a \in U$ cuando cumple las siguientes condiciones

1. Existen las derivadas parciales $\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)$.
2. Para todo $v = (v_1, \dots, v_n)$ tal que $a + v \in U$, se tiene

$$f(a + v) - f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot v_i + r(v), \quad \text{donde } \lim_{\|v\| \rightarrow 0} \frac{r(v)}{\|v\|} = 0.$$

Observación.

- (I) Arriba, siempre que fijemos consideraciones en torno al punto a , por simplicidad escribiremos $\frac{\partial f}{\partial x_i}$ en vez de $\frac{\partial f}{\partial x_i}(a)$.
- (II) La esencia de la definición de diferenciabilidad está en la condición $\lim_{\|v\| \rightarrow 0} \frac{r(v)}{\|v\|} = 0$, pues la igualdad que define el “resto” $r(v)$ puede ser escrita para cualquier función que posea n derivadas parciales.

De $\lim_{\|v\| \rightarrow 0} \frac{r(v)}{\|v\|} = 0$ resulta que

$$\lim_{\|v\| \rightarrow 0} r(v) = 0$$

pues $r(v) = \frac{r(v)}{\|v\|} \|v\|$. Se sigue entonces que

$$\lim_{v \rightarrow 0} f(a + v) = f(a)$$

y por lo tanto *toda función diferenciable en el punto a es continua en ese punto.*

Diremos que $f : U \rightarrow \mathbb{R}$ es *diferenciable* cuando f es diferenciable en todos los puntos de U .

Cuando $n = 1$, la función $f : U \rightarrow \mathbb{R}$ es diferenciable en el punto a si y solamente si posee derivada en este punto. Y se tiene

$$f(a + v) - f(a) = f'(a)v + r(v)$$

y por lo tanto

$$\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0 \text{ si y solamente si } \lim_{v \rightarrow 0} \frac{f(a+v) - f(a)}{v} = f'(a).$$

De forma análoga puede verse que $f : I \rightarrow \mathbb{R}^n$ es derivable en t_0 si y sólo si es diferenciable en t_0 según la definición anterior.

Teorema 118. Toda función $f : U \rightarrow \mathbb{R}$ de clase C^1 es diferenciable.

Demostración: Por simplicidad vamos a suponer $U \subseteq \mathbb{R}^2$. El caso general se trata análogamente, apenas una notación más elaborada.

Fijemos $c = (a, b) \in U$ y tomemos $v = (h, k)$ tal que $c + v \in B \subseteq U$, donde B es una bola de centro c . Sea

$$r(v) = r(h, k) = f(a + h, b + k) - f(a, b) - \frac{\partial f}{\partial x} \cdot h - \frac{\partial f}{\partial y} \cdot k,$$

donde las derivadas son calculadas en el punto $c = (a, b)$. Podemos escribir

$$r(v) = r(h, k) = f(a + h, b + k) - f(a, b) + f(a, b + k) - f(a, b + k) - \frac{\partial f}{\partial x} \cdot h - \frac{\partial f}{\partial y} \cdot k.$$

Por el teorema del Valor Medio para funciones de una variable real, existen $\theta_1, \theta_2 \in (0, 1)$ tales que

$$r(v) = \frac{\partial f}{\partial x}(a + \theta_1 h, b + k) \cdot h + \frac{\partial f}{\partial y}(a, b + \theta_2 k) \cdot k - \frac{\partial f}{\partial x} \cdot h - \frac{\partial f}{\partial y} \cdot k,$$

y luego

$$\frac{r(v)}{\|v\|} = \left[\frac{\partial f}{\partial x}(a + \theta_1 h, b + k) - \frac{\partial f}{\partial x}(a, b) \right] \frac{h}{\sqrt{h^2 + k^2}} + \left[\frac{\partial f}{\partial y}(a, b + \theta_2 k) - \frac{\partial f}{\partial y}(a, b) \right] \frac{k}{\sqrt{h^2 + k^2}}.$$

Cuando hacemos $v \rightarrow 0$, los términos dentro de los corchetes tienden a 0, por la continuidad de las derivadas $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$. Además de eso, los términos fuera de los corchetes tienen valor absoluto menor o igual que 1 y por lo tanto $\lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0$ y entonces f es diferenciable. \square

Corolario 119. Toda función $f : U \rightarrow \mathbb{R}$ de clase C^1 es continua.

Teorema 120. Sean $U \subseteq \mathbb{R}^m, V \subseteq \mathbb{R}^n$ abiertos y sea $f : U \rightarrow V$ una aplicación cuyas funciones coordenadas f_1, \dots, f_n poseen derivadas parciales en el punto $a \in U$ y sea $g : V \rightarrow \mathbb{R}$ una función diferenciable en el punto $b = f(a)$. Entonces $g \circ f : U \rightarrow \mathbb{R}$ posee derivadas parciales en el punto a y vale

$$\frac{\partial(g \circ f)}{\partial x_i} = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \frac{\partial f_k}{\partial x_i}, \quad i = 1, \dots, m,$$

donde las derivadas parciales relativas a los x'_i 's son calculadas en el punto a y las relativas a y'_k 's son calculadas en el punto $b = f(a)$.

Además de eso, si f y g son de clase C^1 entonces $g \circ f \in C^1$.

Demostración: Podemos escribir

$$g(f(a + te_i)) - g(f(a)) = \sum_{k=1}^n \frac{\partial g}{\partial y_k} [f_k(a + te_i) - f_k(a)] + \rho(t) |f(a + te_i) - f(a)|$$

donde escribimos $\rho(t) = r(v)/||v||$ con $v = f(a + te_i) - f(a)$. La diferenciabilidad de g nos da $\lim_{t \rightarrow 0} \rho(t) = 0$. Entonces

$$\frac{g(f(a + te_i)) - g(f(a))}{t} = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \frac{f_k(a + te_i) - f_k(a)}{t} \pm \rho(t) \left| \frac{f(a + te_i) - f(a)}{t} \right|.$$

Luego

$$\frac{\partial (g \circ f)}{\partial x_i} = \lim_{t \rightarrow 0} \frac{g(f(a + te_i)) - g(f(a))}{t} = \sum_{k=1}^n \frac{\partial g}{\partial y_k} \frac{\partial f_k}{\partial x_i}$$

pues

$$\lim_{t \rightarrow 0} \rho(t) \quad \text{y} \quad \lim_{t \rightarrow 0} \left| \frac{f(a + te_i) - f(a)}{t} \right| = \left| \frac{\partial f}{\partial x_i}(a) \right|.$$

□

Definición 121. El *gradiente* de una función diferenciable $f : U \rightarrow \mathbb{R}$ en el punto a y donde $U \subseteq \mathbb{R}_n$ es un abierto, es el vector

$$\text{grad} f(a) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right).$$

Si v es cualquier vector en \mathbb{R}^n , la *derivada direccional* de f en el punto a en la dirección de v es por definición

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}. \quad (1)$$

Estas definiciones permiten enunciar los siguientes corolarios de la regla de la cadena. Si f es diferenciable en el punto a , la derivada direccional $\frac{\partial f}{\partial v}(a)$ existe para cualquier dirección v . Más aún $\frac{\partial f}{\partial v}(a)$ da una expresión en términos de las derivadas parciales de f y las coordenadas del vector v .

Corolario 122. Sea $f : U \rightarrow \mathbb{R}$ diferenciable en el abierto $U \subseteq \mathbb{R}^n$ con $a \in U$. Dado el vector $v = (v_1, \dots, v_n)$, si $\lambda : (-\delta, \delta) \rightarrow U$ es cualquier curva diferenciable tal que $\lambda(0) = a$ y $\lambda'(0) = v$, se tiene

$$(f \circ \lambda)'(0) = \langle \text{grad} f(a), v \rangle = \frac{\partial f}{\partial v}(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) v_i. \quad (2)$$

Demostración: Basta aplicar directamente la fórmula

$$(f \circ \lambda)' = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{d\lambda_i}{dt},$$

observando que $\lambda(t) = (\lambda_1(t), \dots, \lambda_n(t))$ y se tiene $v_i = \frac{d\lambda_i}{dt}(0)$. □

Observación: Notemos que $\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$ con $\lambda(t) = a + tv$, pues $\lambda'(0) = v$.

Corolario 123. Teorema del Valor Medio. Dada $f : U \rightarrow \mathbb{R}$ diferenciable en el abierto $U \subseteq \mathbb{R}^n$, si el segmento de recta que une a y $a + v$ está contenido en U , entonces existe $\theta \in (0, 1)$ tal que

$$f(a + v) - f(a) = \frac{\partial f}{\partial v}(a + \theta v) = \langle \text{grad} f(a + \theta v), v \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + \theta v) v_i,$$

donde $v = (v_1, \dots, v_n)$.

Demostración:

En efecto considerando el segmento de curva $\lambda : [0, 1] \rightarrow U$, dado por $\lambda(t) = a + tv$, vemos que $f(a + v) - f(a) = (f \circ \lambda)(1) - (f \circ \lambda)(0)$. Por el Teorema del Valor Medio para funciones de una variable real, existe $\theta \in (0, 1)$ tal que $(f \circ \lambda)(1) - (f \circ \lambda)(0) = (f \circ \lambda)'(\theta)$. Por la regla de la cadena

$$(f \circ \lambda)'(\theta) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + \theta v) v_i = \frac{\partial f}{\partial v}(a + \theta v) = \langle \text{grad} f(a + \theta v), v \rangle. \quad \square$$

Recordemos que dada $f : U \rightarrow \mathbb{R}$, para cada $c \in \mathbb{R}$ el conjunto $f^{-1}(c) = \{x \in U \mid f(x) = c\}$ es llamado el conjunto de nivel de c de la función f . Cuando $n = 2$ ese conjunto puede ser llamado línea de nivel y cuando $n = 3$, se denomina superficie de nivel.

Volvamos al gradiente de una función f .

Proposición 124. Sea $f : U \rightarrow \mathbb{R}$ una función diferenciable de clase C^1 . Fijemos $a \in U$ y supongamos $\text{grad} f(a) \neq 0$. Entonces

- (I) El gradiente apunta en la dirección en la cual la función es creciente.
- (II) De entre todas las direcciones a lo largo de las cuales f crece, la dirección del gradiente es la de crecimiento más rápido.
- (III) El gradiente de f en el punto a es ortogonal al conjunto de nivel que pasa por a .

Demostración:

(I) Recordemos que la derivada direccional en la dirección de cualquier vector v satisface

$$\frac{\partial f}{\partial v}(a) = \langle \text{grad} f(a), v \rangle, \quad (3)$$

y por lo tanto tomando $v = \text{grad} f(a)$ tendremos $\langle \text{grad} f(a), \text{grad} f(a) \rangle > 0$. Esto quiere decir que la derivada direccional en la dirección de $v = \text{grad} f(a)$ es positiva, y por lo tanto si $\lambda : (-\varepsilon, \varepsilon) \rightarrow U$ es una curva C^1 tal que $\lambda(0) = a$ y $\lambda'(0) = \text{grad} f(a)$, entonces la función $t \rightarrow f \circ \lambda(t)$ tiene derivada positiva en el punto $t = 0$. Luego $(f \circ \lambda)'$ es continua, y por lo tanto $(f \circ \lambda)'(t)$ será positiva en un entorno de 0. Achicando ε si fuese necesario, vemos que $f \circ \lambda : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ será una función creciente.

(II) Volvamos a (3). Observemos que los vectores v para los cuales $\frac{\partial f}{\partial v}(a) > 0$ son direcciones a lo largo de las cuales f crece (con el mismo razonamiento que hicimos para el gradiente). Claramente se tiene $\langle \text{grad} f(a), v \rangle > 0$ con lo cual, v forma un ángulo agudo con $\text{grad} f(a)$. Recordemos para esto que el ángulo entre dos vectores no nulos, v y w viene dado por la fórmula

$$\cos(\theta) = \frac{\langle v, w \rangle}{\|v\| \|w\|}.$$

Y además sabemos que la igualdad se da si y sólo si w está en la misma dirección que v , $w = cv$ con $c \in \mathbb{R}$.

Aplicando la desigualdad de Cauchy-Schwartz a (3) al gradiente y a cualquier otra dirección v en la cual f crece, tenemos

$$\frac{\partial f}{\partial v}(a) = \langle \text{grad} f(a), v \rangle = \cos(\theta) \|\text{grad} f(a)\| \|v\| \leq \|\text{grad} f(a)\|^2 = \frac{\partial f}{\partial(\text{grad} f(a))}(a).$$

(III) Supongamos que $f(a) = c$. La ortogonalidad de $w \in \mathbb{R}^n$ con el conjunto de nivel $f^{-1}(c)$ quiere decir que dada cualquier curva diferenciable $\lambda : (-\varepsilon, \varepsilon) \rightarrow f^{-1}(c)$ en $t = 0$ con $\lambda(0) = a$, se tiene $\langle w, \lambda'(0) \rangle = 0$.

Ahora si $\lambda(t) \subset f^{-1}(c)$ para todo $t \in (-\varepsilon, \varepsilon)$, entonces $f(\lambda(t)) = c$. Derivando esta expresión en $t = 0$ obtenemos

$$0 = (f \circ \lambda)'(0) = \langle \text{grad} f(a), \lambda'(0) \rangle.$$

Así el vector $\text{grad} f(a)$ es ortogonal al vector velocidad en el punto $a = \lambda(0)$ de cualquier curva diferenciable λ contenida en el conjunto de nivel $f^{-1}(c)$. \square

Generalizando lo hecho en funciones reales, definimos un *punto crítico* de $f : U \rightarrow \mathbb{R}$ a un $a \in U$ tal que $\text{grad} f(a) = 0$. La aplicación de este concepto se verá más adelante.

Una aplicación $f : U \rightarrow \mathbb{R}^n$ definida en un abierto $U \subseteq \mathbb{R}^m$ se dice *diferenciable* en el punto $a \in U$ cuando cada una de sus funciones coordenadas $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ son diferenciables en ese punto a .

Si este es el caso entonces para todo $v = (v_1, v_2, \dots, v_n)$ tal que $a + v \in U$ y para cada $i = 1, \dots, n$ se tiene

$$f_i(a + v) - f_i(a) = \sum_{j=1}^m \frac{\partial f_i}{\partial x_j}(a) v_j + r_i(v) \quad \text{con} \quad \lim_{v \rightarrow 0} \frac{r_i(v)}{\|v\|} = 0.$$

La matrix $n \times m$ cuya fila i es el vector gradiente de f_i y que se denota $Jf(a) = \left[\frac{\partial f_i}{\partial x_j}(a) \right]$ se denomina *matriz jacobiana* de f en el punto a .

La transformación lineal $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ cuya matriz en las bases canónicas de \mathbb{R}^m y \mathbb{R}^n es $Jf(a)$, se llama *la diferencial de f* en el punto a . A veces $f'(a)$ también se escribe df_a .

De acuerdo con la definición de matriz de una transformación real, para todo $v = (v_1, \dots, v_m) \in \mathbb{R}^m$, tenemos

$$f'(a)v = (w_1, \dots, w_n) \quad \text{donde} \quad w_i = \sum_{j=1}^m \frac{\partial f_i}{\partial x_j}(a) v_j = \frac{\partial f_i}{\partial v}(a).$$

Esto significa que el resultado $f'(a)v$ es un vector en \mathbb{R}^n , cuya componente i resulta de hacer el producto interno entre la fila i de la matriz y el vector columna v . Con la otra notación $f'(a)v = df_a(v)$.

Como es natural, definimos la *derivada direccional* de f en el punto a y en la dirección del vector v como

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t},$$

de donde obtenemos inmediatamente que

$$\frac{\partial f}{\partial v}(a) = \left(\frac{\partial f_1}{\partial v}(a), \dots, \frac{\partial f_n}{\partial v}(a) \right) = f'(a)v.$$

De la regla de la cadena y de la definición de arriba, surge que

$$\frac{\partial f}{\partial v}(a) = (f \circ \lambda)'(0)$$

donde $\lambda(t) = a + tv$ o donde λ puede ser cualquier curva diferenciable que pasa por a y tiene vector velocidad v en $t = 0$: $\lambda(0) = a, \lambda'(0) = v$.

Puesto que cada función coordenada $f_i : U \rightarrow \mathbb{R}$ es diferenciable, cada vez que $a + v \in U$ tendremos

$$f_i(a + v) - f_i(a) = \langle \text{grad } f_i(a), v \rangle + r_i(v), \quad \text{donde} \quad \lim_{v \rightarrow 0} \frac{r_i(v)}{\|v\|} = 0.$$

Si reunimos estas expresiones para $i = 1, \dots, n$, obtenemos en coordenadas la expresión siguiente

$$f(a + v) - f(a) = f'(a)v + r(v), \quad \text{donde} \quad \lim_{v \rightarrow 0} \frac{r(v)}{\|v\|} = 0.$$

Claramente $r(v) = (r_1(v), \dots, r_n(v))$.

El siguiente teorema muestra una caracterización equivalente de diferenciabilidad que algunos textos toman como definición. Esta noción aunque más abstracta ahora, sirve para comprender el papel de la diferencial o derivada de f en a .

Teorema 125. Sea $U \subseteq \mathbb{R}^m$ un abierto y sea $f : U \rightarrow \mathbb{R}^n$ una función. Son equivalentes:

(I) f es diferenciable.

(II) Existe una transformación lineal $T_a : \mathbb{R}^m \rightarrow \mathbb{R}^n$ tal que cada vez que $a + v \in U$, el siguiente límite existe y vale

$$\lim_{v \rightarrow 0} \frac{\|f(a+v) - f(a) - T_a(v)\|}{\|v\|} = 0. \quad (4)$$

Demostración: (\Rightarrow) Tomamos $T_a = f'(a)$ y el resultado resulta pues $\frac{\|f(a+v) - f(a) - T_a(v)\|}{\|v\|} = \frac{\|r(v)\|}{\|v\|}$.

Veamos la vuelta. Supongamos que tenemos T_a lineal como en (ii) y sea $r(v) = f(a+v) - f(a) - T_a(v)$.

Reemplacemos v arriba por tv , donde ahora suponemos que $\|v\| = 1$. Entonces tendremos

$$\frac{f(a+tv) - f(a)}{t} = T_a(v) \pm \frac{r(tv)}{\|tv\|}.$$

Tomemos límite cuando $t \rightarrow 0$ y apliquemos lo que ya sabemos:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t} &= T_a(v) \\ &= \frac{\partial f}{\partial v}(a) = f'(a)v \end{aligned}$$

Observemos que probamos que si se satisface (4) entonces la transformación lineal que da el límite debe satisfacer $T_a = f'(a)$. \square

Definición 126. Una función $f : U \rightarrow \mathbb{R}^n$ con $U \subseteq \mathbb{R}^m$ es *diferenciable* en $a \in U$ si satisface cualquiera de las condiciones equivalentes del Teorema 125.

Diremos que f es diferenciable en U cuando es diferenciable en cualquiera de los puntos de U .

Corolario 127. Toda función $f : U \rightarrow \mathbb{R}^n$ con $U \subseteq \mathbb{R}^m$ diferenciable es continua.

En efecto, tomemos $v \in \mathbb{R}^m$ con $\|v\| = 1$ y escribamos

$$\|f(a+tv) - f(a)\| = \frac{\|f(a+tv) - f(a) - f'(a)(tv) + f'(a)(tv)\|}{\|tv\|} \|tv\| \leq \frac{\|f(a+tv) - f(a) - f'(a)(tv)\|}{\|tv\|} + \|tf'(a)\|$$

de donde tomando límite cuando $t \rightarrow 0$ a la derecha, vemos que

$$\lim_{t \rightarrow 0} \|f(a+tv) - f(a)\| = 0,$$

para cualquier $v \in \mathbb{R}^m$.

Ejemplos.

1. Sea $f : I \rightarrow \mathbb{R}^n$ una curva diferenciable. Entonces f es una aplicación diferenciable. En efecto, su derivada en el punto $a \in I$ es el vector velocidad

$$f'(a) = \left(\frac{df_1}{dt}(a), \dots, \frac{df_n}{dt}(a) \right).$$

2. Si $f : U \rightarrow \mathbb{R}$ es una función definida en el abierto $U \subseteq \mathbb{R}^m$, diferenciable en el punto $a \in U$, su derivada en a es el vector gradiente en a

$$f'(a) = \text{grad}f(a),$$

de modo que $f'(a)v = \langle \text{grad}f(a), v \rangle$.

3. Sea $f : U \rightarrow \mathbb{R}^m$ una función constante, entonces $f'(x) = 0$ para todo $x \in U$. Recíprocamente si $f'(x) = 0$ en una bola $B_r(a)$ entonces f es constante en esa bola.
4. Si consideramos una transformación lineal $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$, entonces la derivada de T es $T'(a) = T$, para todo $a \in \mathbb{R}^m$.

La verificación de estas afirmaciones queda como ejercicio.

Sea $f : U \rightarrow \mathbb{R}$, con $U \subset \mathbb{R}^3$ una función diferenciable. Supongamos que $S = f^{-1}(c)$ es el conjunto de nivel de $c \in \mathbb{R}$ de modo que para todo $x \in S$ el gradiente $\text{grad}f(x) \neq 0$. Entonces el subespacio de \mathbb{R}^3 que es ortogonal a $\text{grad}f(x)$ se llama el *espacio tangente* a S en x y se denota $T_x S$:

$$T_x S = \{u \in \mathbb{R}^3 : \langle u, \text{grad}f(x) \rangle = 0\},$$

el cual es un subespacio de dimensión dos.

Ejemplo. Tomemos la función $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por $f(x, y, z) = x^2 + y^2 + z^2$. Entonces la esfera unitaria en \mathbb{R}^3 es el conjunto de nivel $f^{-1}(1)$;

$$S^2 = \{x \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}.$$

El gradiente de f en un punto p de coordenadas $p = (x_0, y_0, z_0)$ está dado por el vector

$$\text{grad}f(x_0, y_0, z_0) = (2x_0, 2y_0, 2z_0),$$

el cual es no nulo si $p \neq (0, 0, 0)$ y esto ocurre siempre en S^2 . Luego el plano tangente a S^2 en p es el conjunto ortogonal a $\text{grad}f(p)$:

$$T_p S^2 = \{(x, y, z) \in \mathbb{R}^3 : xx_0 + yy_0 + zz_0 = 0\},$$

esto es, el espacio tangente en p consiste de todos los vectores de \mathbb{R}^3 que son ortogonales a p . Esto tiene una interpretación geométrica, que se puede ver con ayuda de un programa para graficar.