Capítulo VI

Diferenciabilidad de funciones de varias variables

La definición de diferenciabilidad para funciones $f: \mathbb{R}^n \to \mathbb{R}^m$ no tiene sentido, puesto que el cociente

$$\frac{f(a+h)-f(a)}{h}$$

no está definido, porque el cociente entre el vector $f(a+h)-f(a)$ de $\mathbb{R}^m$ y h de $\mathbb{R}^n$ carece de sentido matemático, incluso si $m=n$.

Si reescribimos la definición de derivada en una variable, podemos hallar la manera correcta de definir diferenciabilidad de funciones de varias variables. Recordemos que $f: \mathbb{R} \to \mathbb{R}$ es diferenciable en a si existe un número real que denotamos por $f'(a)$ tal que

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}.$$

Esta ecuación es equivalente a

$$\lim_{h\to 0} \left\{\frac{f(a+h)-f(a)}{h} - f'(a)\right\} = 0$$

y también a

$$\lim_{h\to 0}\left\{\frac{f(a+h)-f(a)-f'(a)h}{h}\right\}=0$$

Si observamos que la aplicación $L_a\colon \mathbb{R}\to\mathbb{R}$ tal que $L_a(h)=f'(a)h$ es lineal, podemos escribir la definición de diferenciabilidad de la siguiente manera: Decimos que $f\colon \mathbb{R}\to\mathbb{R}$ es diferenciable en a si existe una aplicación lineal $L_a\colon \mathbb{R}\to\mathbb{R}$ tal que

$$\lim_{h\to 0}\left\{\frac{f(a+h)-f(a)-L_a(h)}{h}\right\}=0$$

De manera equivalente, si existe una aplicación lineal $L_a\colon \mathbb{R}\to\mathbb{R}$ tal que

$$f(a+h)-f(a)-L_a(h)=hE_a(h)$$

para todo h suficientemente pequeño, donde

$$\lim_{h\to 0}E_a(h)=0$$

Del análisis anterior, ahora resulta natural la siguiente definición de diferenciabilidad para funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Definición: Decimos que $f\colon D\subseteq\mathbb{R}^n\to\mathbb{R}^m$ es derivable en un punto interior a de D, si existe una aplicación lineal $L_a\colon \mathbb{R}^n\to\mathbb{R}^m$ tal que

$$\lim_{h\to 0}\left\{\frac{f(a+h)-f(a)-L_a(h)}{||h||}\right\}=0 \qquad (*)$$

Observación: De la definición de diferenciabilidad se deduce que para todo h de norma suficientemente pequeña, se tiene que

$$\lim_{h\to 0}E_a(h)=0,\text{ donde }\ ||h||E_a(h)=f(a+h)-f(a)-L_a(h).$$

En consecuencia, para todo h suficientemente pequeño es válida la aproximación

$$f(a+h)\approx f(a)+L_a(h).$$

Si denotamos $a+h=x$, vemos que $h\to 0$, si y solamente si $x\to a$, razón por la cual podemos escribir la aproximación anterior de la siguiente manera:

$$f(x)\approx f(a)+L_a(x-a).$$

En tal caso vemos que la función f puede ser aproximada cerca de a por la función afín $A(x)=f(a)+L_a(x-a)$.

En el caso en que $f: \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en $\boldsymbol{a} = (\boldsymbol{x_0}, \boldsymbol{y_0})$, vemos que $z = f(x, y)$ define la gráfica de la función, mientras que la ecuación $z = f(\boldsymbol{x_0}, \boldsymbol{y_0}) + L_{(\boldsymbol{x_0}, \boldsymbol{y_0})}((x, y) - (\boldsymbol{x_0}, \boldsymbol{y_0}))$ define la grafica de una función afín, que corresponde al **plano tangente a la gráfica de la función $\boldsymbol{f}$ en el punto $(\boldsymbol{a}, \boldsymbol{f}(\boldsymbol{a}))$.** Más adelante veremos como se puede hallar explícitamente la ecuación del plano tangente.

Ejemplo 1: Si $f: \mathbb{R}^2 \to \mathbb{R}$ está definida por la ecuación $f(x, y) = x + 2y + 5$, entonces f es diferenciable en cualquier punto (a, b) de $\mathbb{R}^2$, porque existe la aplicación lineal $L_{(a,b)}: \mathbb{R}^2 \to \mathbb{R}$ tal que $L_{(a,b)}(x, y) = x + 2y$ que verifica

$$\lim_{(h,k)\to(0,0)} \left\{ \frac{f((a,b) + (h,k)) - f(a,b) - L_{(a,b)}(h,k))}{\|(h,k)\|} \right\} = 0$$

En efecto:

$$\lim_{(h,k)\to(0,0)} \left\{ \frac{f((a,b) + (h,k)) - f(a,b) - L_{(a,b)}(h,k))}{\|(h,k)\|} \right\}$$

$$= \lim_{(h,k)\to(0,0)} \left\{ \frac{a + h + 2(b + k) - a - 2b - h - 2k}{\sqrt{h^2 + k^2}} \right\}$$

$$= \lim_{(h,k)\to(0,0)} \left\{ \frac{0}{\sqrt{h^2 + k^2}} \right\} = 0$$

Ejemplo 2: Si $f: \mathbb{R}^2 \to \mathbb{R}$ está definida a través de la ecuación $f(x, y) = 2x^2 y$, entonces f es diferenciable en cualquier punto (a, b) de $\mathbb{R}^2$, porque existe la aplicación lineal $L_{(a,b)}: \mathbb{R}^2 \to \mathbb{R}$ tal que $L_{(a,b)}(x, y) = 4abx + 2a^2 y$ que verifica

$$\lim_{(h,k)\to(0,0)} \left\{ \frac{f((a,b) + (h,k)) - f(a,b) - L_{(a,b)}(h,k))}{\|(h,k)\|} \right\} = 0$$

En efecto:

$$\lim_{(h,k)\to(0,0)} \left\{ \frac{f((a,b) + (h,k)) - f(a,b) - L_{(a,b)}(h,k))}{\|(h,k)\|} \right\}$$

$$= \lim_{(h,k)\to(0,0)} \left\{ \frac{2(a + h)^2(b + k) \quad - 2a^2 b \quad - 4abh - 2a^2 k}{\sqrt{h^2 + k^2}} \right\}$$

$$= \lim_{(h,k)\to(0,0)} \left\{ \frac{2(a^2 b + 2ahb + h^2 b + a^2 k + 2ahk + h^2 k) - 2a^2 b \quad - 4abh - 2a^2 k}{\sqrt{h^2 + k^2}} \right\}$$

$$= \lim_{(h,k)\to(0,0)} \left\{ \frac{2(h^2 b + 2ahk + h^2 k)}{\sqrt{h^2 + k^2}} \right\}$$

Este último límite es 0 porque se puede separar en tres límites que valen cero, ya que $h^2 \leq h^2 + k^2$ y $|hk| \leq h^2 + k^2$.

Ejemplo 3: $f: \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal, entonces *f* es diferenciable en cualquier punto $a \in \mathbb{R}^n$, además la aplicación lineal L_a es *f*.

En efecto:

$$\lim_{h\to 0}\left\{\frac{f(a+h)-f(a)-f(h)}{\|h\|}\right\} = \lim_{h\to 0}\left\{\frac{f(a)+f(h)-f(a)-f(h)}{\|h\|}\right\} = 0.$$

Teorema: Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es derivable en un punto interior a de *D*, entonces la aplicación lineal $L_a: \mathbb{R}^n \to \mathbb{R}^m$ es única.

Como consecuencia de este teorema, en lo que sigue hablaremos de la aplicación lineal $L_a: \mathbb{R}^n \to \mathbb{R}^m$ y la denominaremos **diferencial de *f* en *a***; también **usaremos la notación $\boldsymbol{Df(a)}$ para esta diferencial.**

Observación: Para probar que una función que $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es derivable en un punto interior a de *D*, es suficiente encontrar la aplicación lineal $L_a: \mathbb{R}^n \to \mathbb{R}^m$, y demostrar que el límite (*) es cero. Entonces, para determinar L_a es suficiente hallar sus valores en la base canónica de $\mathbb{R}^n$, o sea la matriz de L_a en la base canónica.

Para determinar esos valores de L_a comencemos por suponer que *f* es diferenciable en a y observar que si $h = tv$, para un vector no nulo cualquiera de $\mathbb{R}^n$, entonces $h \to 0$ si y solo si $t \to 0$. Nos damos cuenta que si (*) es válida entonces

$$\lim_{t\to 0}\left\{\frac{f(a+tv)-f(a)-L_a(tv)}{\|tv\|}\right\} = 0$$

De manera que

$$\lim_{t\to 0}\left\|\frac{f(a+tv)-f(a)-L_a(tv)}{tv}\right\| = \lim_{t\to 0}\frac{\|f(a+tv)-f(a)-L_a(tv)\|}{\|tv\|}$$

$$=\lim_{t\to 0}\frac{1}{\|v\|}\left\|\frac{f(a+tv)-f(a)-L_a(tv)}{t}\right\|$$

$$=\frac{1}{\|v\|}\lim_{t\to 0}\left\|\frac{f(a+tv)-f(a)-L_a(tv)}{t}\right\|=0.$$

Se deduce que

$$\lim_{t\to 0}\left\|\frac{f(a+tv)-f(a)-L_a(tv)}{t}\right\|=0.$$

En consecuencia

$$\lim_{t\to 0}\frac{f(a+tv)-f(a)-L_a(tv)}{t}=0,$$

De donde podemos deducir que

$$\lim_{t\to 0}\left\{\frac{f(a+tv)-f(a)}{t}-\frac{tL_a(v)}{t}\right\}=0,$$

Es decir

$$L_a(v)=\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t}$$

Cuando existe este límite se denota por $\frac{\partial f}{\partial v}(a)$ y se llama **derivada de *f* con respecto al vector *v* en el punto *a*.**

Casos particulares:

- Si v = u es un vector unitario, la derivada $\frac{\partial f}{\partial u}(a)$ se llama **derivada direccional** de *f* en el punto a, en la dirección *u*.
- Si v = e_i es el i-ésimo vector de la base canónica de $\mathbb{R}^n$, la derivada $\frac{\partial f}{\partial e_i}(a)$ se llama **derivada parcial de *f* en el punto a con respecto a la i-ésima coordenada** y se denota por $\frac{\partial f}{\partial x_i}(a)$.

Es claro que $\frac{\partial f}{\partial x_i}(a)=L_a(e_i)$, para todo $i\in\{1,2,\dots n\}$ y podemos escribir la representación matricial de la derivada L_a en la base canónica de $\mathbb{R}^n$.

La matriz asociada a la aplicación lineal L_a se denota por $f'(a)$ y se llama derivada de *f* en el punto a o matriz jacobiana de *f* en *a*. Entonces se tiene:

$$f'(a)=\left[\frac{\partial f}{\partial x_1}(a)\quad \frac{\partial f}{\partial x_2}(a)\,\ldots\ldots\;\; \frac{\partial f}{\partial x_n}(a)\;\right]$$

Para conocer la matriz jacobiana de una función diferenciable en un punto a del dominio de *f*, es suficiente conocer las derivadas parciales de *f*.

Si reescribimos la definición de derivada parcial de una función $f: D\subseteq\mathbb{R}^n\to\mathbb{R}^m$ en un punto $a=(a_1,a_2,\dots,a_n)$, vemos que

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

Cuando este límite existe.

Si definimos $\varphi_i(t) = f(a_1,\dots,a_{i-1},t,a_{i+1},\dots,a_n)$, se observa que

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t\to 0}\frac{\varphi(a_i+\mathrm{t})-\varphi(a_i)}{t} = \varphi_i'(a_i),$$

Por lo tanto, para hallar la derivada parcial de f que depende de las variables $x_1, x_2, \dots, x_n$, con respecto a la *i*-ésima coordenada no hay más que fijar todas las otras variables y derivar con respecto a la única variable independiente x_i, como función de una variable.

Ejemplo: Consideremos la función $f: IR^2 \to IR^3$ tal que $f(x,y) = (2x^3 - y^5, 2x, 5y)$.

La derivada parcial con respecto a la primera variable es $\frac{\partial f}{\partial x}(x,y) = (6x^2, 2{,}0)$ y con respecto a la segunda variable es $\frac{\partial f}{\partial y}(x,y) = (-5y^4, 0{,}5)$.

Si admitimos que f es diferenciable en un punto cualquiera (x,y), entonces su derivada o matriz jacobiana es:

$$f'(x,y) = \begin{bmatrix} 6x^2 & -5y^4 \\ 2 & 0 \\ 0 & 5 \end{bmatrix}$$

Teorema: Si $f: D \subseteq \mathbb{R}^n \to R^m$ es diferenciables en $a \in D$, entonces f es continua en a.

Teorema: Si $f, g: D \subseteq \mathbb{R}^n \to R^m$ son diferenciables en $a \in D$ y $\lambda \in \mathbb{R}$, entonces:

a) $f + g$ es diferenciable en a y además $\mathrm{D}(f+g)(a) = \mathrm{D}f(a) + \mathrm{D}g(a)$
b) λf es diferenciable en a y además $\mathrm{D}(\lambda f)(a) = \lambda \mathrm{D}f(a)$
c) Cuando $m = 1$, el producto fg es diferenciable en a y además se tiene
$$\mathrm{D}(fg)(a) = g(a)\mathrm{D}f(a) + f(a)\mathrm{D}g(a).$$
d) Cuando $m = 1$ y si $g(a) \neq 0$, el cociente $\frac{f}{g}$ es diferenciable en a y además se verifica
$$\mathrm{D}\left(\frac{f}{g}\right)(a) = \frac{g(a)\mathrm{D}f(a) - f(a)Dg(a)}{(g(a))^2}.$$

Observación: En términos matriciales las propiedades recién enunciadas pueden ser escritas como en cálculo de una variable:

a) $(f+g)'(a) = f'(a) + g'(a)$
b) $(\lambda f)'(a) = \lambda f(a)$
c) $(fg)'(a) = g(a)f'(a) + f(a)g'(a)$
d) $(f/g)\acute{}(a) = \frac{g(a)f'(a) - f(a)g\acute{}(a)}{(g(a))^2}$

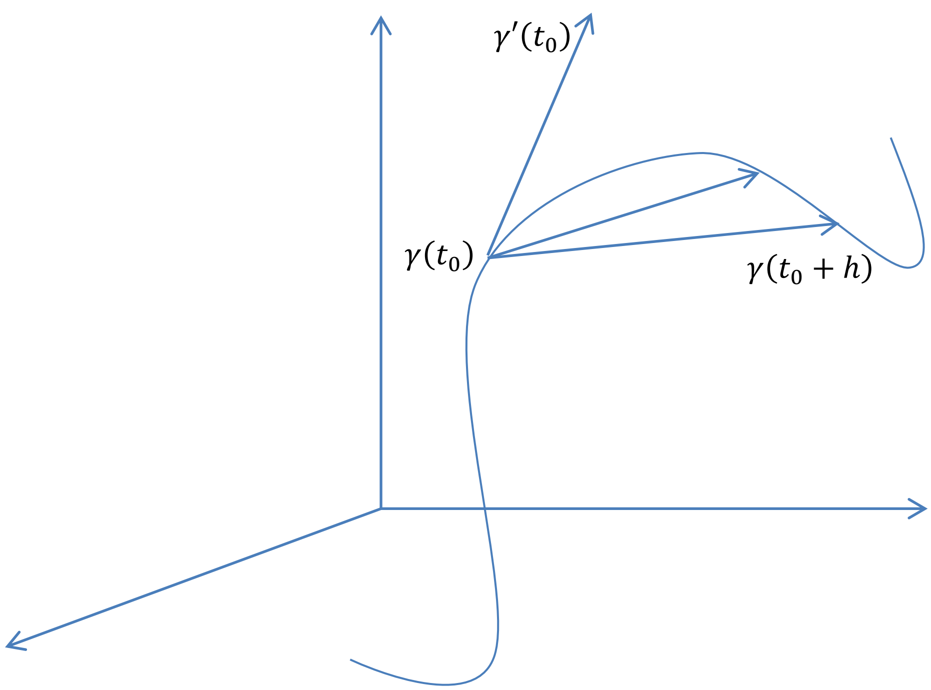
Interpretación geométrica de la derivada de una función de una variable con valores vectoriales.

Si consideramos una función $\gamma: [a,b] \subseteq \mathbb{R} \to \mathbb{R}^n$ diferenciable en un punto interior t_0 de $[a,b]$, para poder visualizar la situación, hemos escogido $n = 3$ obteniendo la figura que se puede apreciar a continuación. Entonces, entonces, para todo h suficientemente pequeño tenemos que, para todo h suficientemente pequeño (de manera que $t_0 + h \in [a,b]$),

$\gamma(t_0 + h) - \gamma(t_0) = \overrightarrow{\gamma(t_0)\gamma(t_0 + h)}$ = vector con origen $\gamma(t_0)$ y extremo $\gamma(t_0 + h)$,

por lo tanto

$\frac{\gamma(t_0+h)-\gamma(t_0)}{h}$ =vector paralelo a $\overrightarrow{\gamma(t_0)\gamma(t_0 + h)}$



Cuando h tiende a cero, vemos que $\overrightarrow{\gamma(t_0)\gamma(t_0+h)}$ tiende a la posición tangente a la curva definida por γ en el punto $\gamma(t_0)$. Por lo tanto es razonable interpretar la derivada $\gamma'(t_0)$ como el **vector tangente** a la curva en el punto $\gamma(t_0)$. También este vector se denota por $v(t_0)$ y recibe el nombre de **vector velocidad**.

Interpretación geométrica de la derivada parcial

Supongamos que $f: \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en el punto (a, b), entonces la derivada parcial con respecto a la primera variable *x* existe y además

$$\frac{\partial f}{\partial x}(a, b) = \lim_{t \to 0} \frac{f(a+t, b) - f(a, b)}{t} = \varphi_1'(a),$$

donde $\varphi_1(t) = f(t, b)$.

Si G_f representa la gráfica de la función *f*, vemos que $(a, b, c) = (a, b, f(a, b))$ es un punto de G_f. La intersección de G_f con el plano cuya ecuación es $y = b$ es una curva que pasa por el punto (a, b, c), como se muestra en la figura 1.

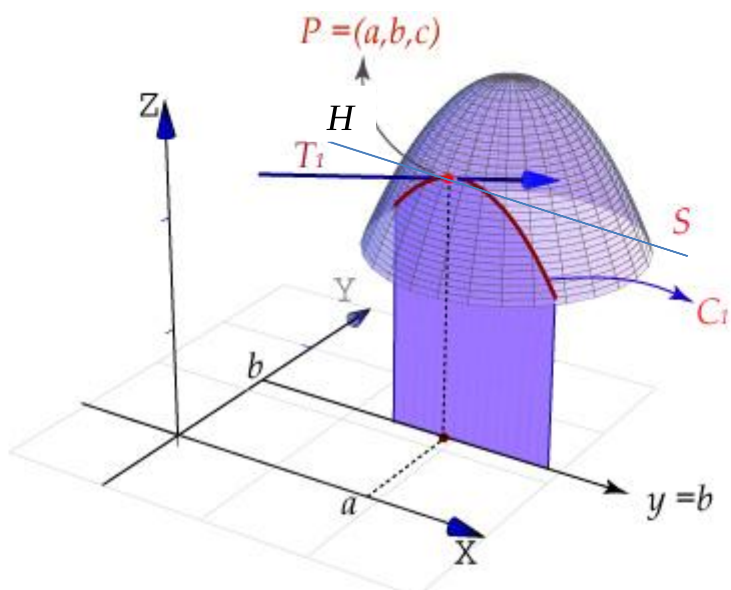


Figura 1

Del curso de cálculo de una variable sabemos que $\varphi_1'(a)$ representa la pendiente de la recta tangente a la gráfica de φ_1 en el punto $(a, \varphi_1(a))$, pero esta pendiente es igual a la tangente trigonométrica del ángulo que forma la recta tangente T_1 con la recta horizontal *H*.

De manera similar, la intersección de G_f con el plano de ecuación $x = a$ es una curva que pasa por el punto (a, b, c), como se puede apreciar en la figura 2.

Como $f: \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en el punto (a, b), entonces la derivada parcial con respecto a la segunda variable *y* existe y además

$$\frac{\partial f}{\partial y}(a, b) = \lim_{t \to 0} \frac{f(a, b+t) - f(a, b)}{t} = \varphi_2'(a),$$

donde $\varphi_2(t) = f(a, t)$.

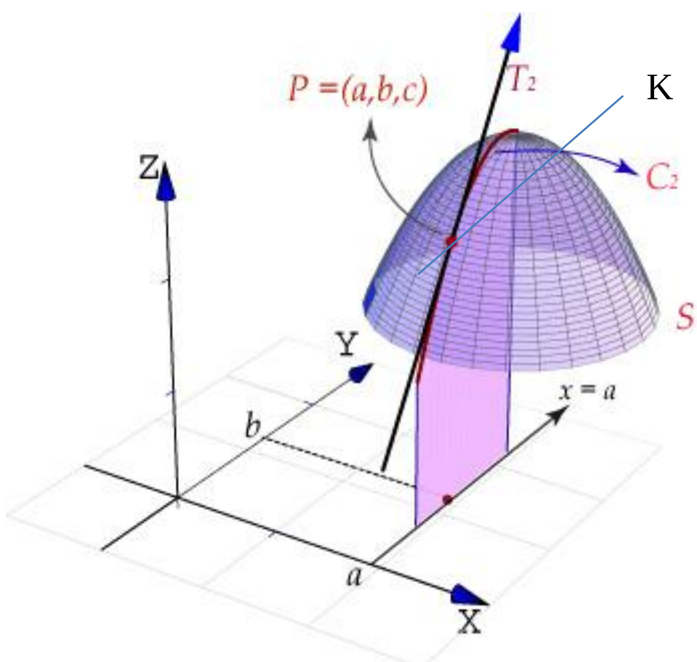


Figura 2

Por el mismo argumento usado para interpretar la derivada parcial con respecto a la primera variable, vemos que $\varphi_2'(b)$ representa la pendiente de la recta tangente a la gráfica de φ_1 en el punto $(a, \varphi_1(a))$, pero esta pendiente es igual a la tangente trigonométrica del ángulo que forma la recta tangente T_1 con la recta horizontal *K*.

Si tenemos en cuenta que la derivada de una función $g: \mathbb{R} \to \mathbb{R}$ es $g'(a)$ y que la ecuación paramétrica de la recta tangente es $(x, y) = (a, b) + t(1, g'(a))$, vemos que las ecuaciones de las rectas tangentes T_1 y T_2 son respectivamente:

T_1: $\quad (x, y, z) = (a, b, c) + t(1,0, \frac{\partial f}{\partial x}(a, b))$

T_2: $\quad (x, y, z) = (a, b, c) + t(0,1, \frac{\partial f}{\partial y}(a, b))$

Desde luego que estas dos rectas tangentes determinan un plano, del cual esperamos que sea tangente a la gráfica de *f*.

Puesto que (a, b, c) pertenece a este plano, para hallar su ecuación cartesiana es suficiente encontrar un vector normal a dicho plano. Salta a la vista que el producto cruz de los vectores directores de las rectas tangentes nos proporciona un tal vector normal.

Como los vectores directores son $v_1 = (1,0, \frac{\partial f}{\partial x}(a, b))$ y $v_2 = (0,1, \frac{\partial f}{\partial y}(a, b))$, entonces un vector es

$$n = v_1 \times v_2 = \begin{array}{c} (\; 1, \quad 0, \quad \frac{\partial f}{\partial x}(a, b) \;) \\ \times \\ (\; 0, \quad 1, \quad \frac{\partial f}{\partial y}(a, b) \;) \end{array} = (-\frac{\partial f}{\partial x}(a, b), -\frac{\partial f}{\partial y}(a, b), 1)$$

Por lo tanto, la ecuación vectorial de dicho plano es $\overrightarrow{AP} \cdot \vec{n} = 0$, o sea

$$(x - a, y - b, z - c) \cdot = \left(-\frac{\partial f}{\partial x}(a, b), -\frac{\partial f}{\partial y}(a, b), 1\right) = 0,$$

o mejor

$$\boldsymbol{\frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) - (z - f(a, b)) = 0.}$$

Por ejemplo, si consideramos la función $f: \mathbb{R}^2 \to \mathbb{R}$ tal que $f(x, y) = x^2 + y^2$, entonces

$\frac{\partial f}{\partial x}(x, y) = 2x$ y $\frac{\partial f}{\partial y}(x, y) = 2y$, por lo tanto $\frac{\partial f}{\partial x}(1,2) = 2$ y $\frac{\partial f}{\partial y}(1,2) = 4$, por consiguiente la ecuación del plano en el punto $(1,2,5)$ de la gráfica de f es

$$2(x - 1) + 4(y - 2) - (z - 5) = 0,$$

que expresado en su forma general es

$$2x + 4y - z - 5 = 0.$$

Cabe observar que la ecuación de este plano también puede ser expresada en la forma

$$\left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b), -1\right) \cdot (x-a, y-b, z-c) = 0.$$

Es importante recordar que la ecuación del plano tangente a la gráfica de $z = f(x,y)$ en el punto $(\boldsymbol{x_0}, \boldsymbol{y_0}, (f(\boldsymbol{x_0}, \boldsymbol{y_0}))$ **es** $z = f(\boldsymbol{x_0}, \boldsymbol{y_0}) + L_{(\boldsymbol{x_0}, \boldsymbol{y_0})}((x,y) - (\boldsymbol{x_0}, \boldsymbol{y_0}))$, o sea

$$z = f(\boldsymbol{x_0}, \boldsymbol{y_0}) + \begin{pmatrix} \frac{\partial f}{\partial x}(\boldsymbol{x_0}, \boldsymbol{y_0}) & \frac{\partial f}{\partial y}(\boldsymbol{x_0}, \boldsymbol{y_0}) \end{pmatrix} \begin{pmatrix} x - \boldsymbol{x_0} \\ y - \boldsymbol{y_0} \end{pmatrix}$$

O bien,

$$z = f(\boldsymbol{x_0}, \boldsymbol{y_0}) + \left(\frac{\partial f}{\partial x}(\boldsymbol{x_0}, \boldsymbol{y_0}), \quad \frac{\partial f}{\partial y}(\boldsymbol{x_0}, \boldsymbol{y_0})\right) \cdot (x - \boldsymbol{x_0}, y - \boldsymbol{y_0})$$

En el caso de nuestro dibujo, la ecuación del plano tangente se escribe así:

$$z - f(a,b) = \left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b)\right) \cdot (x-a, y-b)$$

Como podemos ver, en la última ecuación aparece el vector cuyas componentes son las derivadas parciales de f en el punto (a,b) y que es conocido como **vector gradiente**.

Definición: El vector gradiente de una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto $a \epsilon D$ es

$$\nabla f(a) = gradf(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right),$$

siempre que las derivadas parciales existan.

Ejemplo: Si $f: \mathbb{R}^2 \to \mathbb{R}$ es la función tal que $f(x,y) = x^2y + 5y^2 + x + y$, entonces

$\nabla f(x,y) = (2xy + 1, x^2 + 10y + 1)$ y $\nabla f(2,3) = (13,35)$.

Propiedades: Si $f, g: D \subseteq \mathbb{R}^n \to \mathbb{R}$ tienen derivadas parciales en $a \in D$ y $\lambda \in \mathbb{R}$, entonces:

e) $\nabla(f+g)(a) = \nabla f(a) + \nabla g(a)$
f) $\nabla(\lambda f)(a) = \lambda \nabla f(a)$
g) $\nabla(fg)(a) = g(a)\nabla f(a) + f(a)\nabla g(a)$
h) $\nabla\left(\frac{f}{g}\right)(a) = \frac{g(a)\nabla f(a) - f(a)\nabla g(a)}{(g(a))^2}$, si $g(a) \neq 0$.

Regla de la cadena

Recordemos que en cálculo de una variable la regla de la cadena permite derivar funciones compuestas, siendo posible la derivación de muchas funciones. En cálculo de varias variables también juega un papel importante en la derivación de funciones compuestas.

Teorema: Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $a \in D$ y $g: E \subseteq \mathbb{R}^m \to \mathbb{R}^p$ es diferenciable en $b = f(a)$, con $f(D) \subseteq E$, entonces $g \circ f: D \subseteq \mathbb{R}^n \to \mathbb{R}^p$ es diferenciable en a. Además,

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

En términos matriciales esta última ecuación se escribe de la misma manera que la regla de la cadena en cálculo de una variable, es decir

$$(g \circ f)'(a) = g'\big(f(a)\big)f'(a),$$

donde las derivadas son las matrices jacobianas respectivas en los puntos indicados.

Observación: Un caso importante se presenta cuando $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ y $g = \gamma: [a, b] \subseteq \mathbb{R} \to D \subseteq \mathbb{R}^n$ es una curva diferenciable en un punto $t_0 \in]a, b[$, en cuyo caso $F = f \circ \gamma: [a, b] \subseteq \mathbb{R} \to \mathbb{R}$ es una función diferenciable en t_0 , y además:

$$F'(t_0) = f'\big(\gamma(t_0)\big)\gamma'(t_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\gamma(t_0)) & \frac{\partial f}{\partial x_2}(\gamma(t_0)) & . \ . \ . & \frac{\partial f}{\partial x_n}(\gamma(t_0)) \end{pmatrix} \begin{pmatrix} \gamma_1'(t_0) \\ \vdots \\ \gamma_n'(t_0) \end{pmatrix}$$

$$= \left(\frac{\partial f}{\partial x_1}(\gamma(t_0)),\ \ \frac{\partial f}{\partial x_2}(\gamma(t_0)), . \ . \ . \ , \frac{\partial f}{\partial x_n}(\gamma(t_0))\right)(\gamma_1'(t_0), \dots, \gamma_n'(t_0)),$$

Por lo tanto,

$$\boldsymbol{F'(t_0) = \nabla f\big(\gamma(t_{0})\big) \cdot \gamma'(t_0)}$$

Conjuntos de nivel.

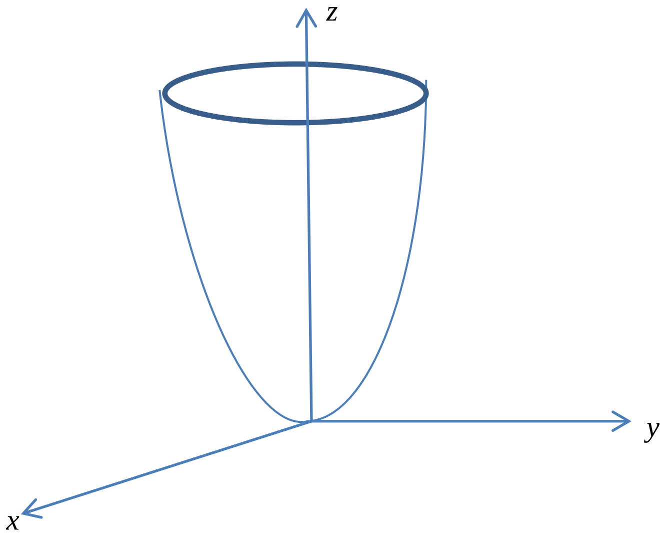
Antes de dar una interpretación geométrica del vector gradiente, es necesario introducir la noción de conjunto de nivel; en particular las nociones de superficie de nivel y de líneas de nivel.

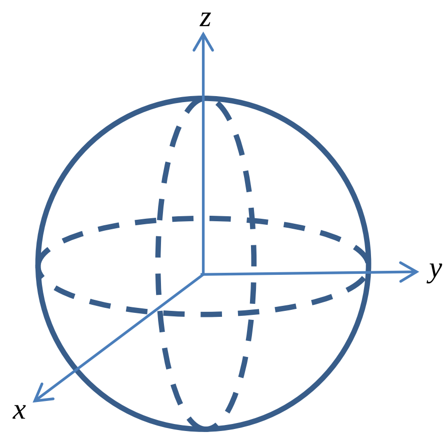
El conjunto de nivel $N_c(f)$ de una función $f: \mathbb{R}^n \to \mathbb{R}$, donde $c \in \mathbb{R}$, es el conjunto de todos los puntos de $\mathbb{R}^n$ cuya imagen bajo *f* es *c*, es decir,

$$N_c(f) = \{x \in \mathbb{R}^n / f(x) = c\}$$

Ejemplos:

1) Si $f: \mathbb{R}^3 \to \mathbb{R}$ está definida por la ecuación $f(x, y, z) = x^2 + y^2 - z$, entonces $N_0(f)$ es el conjunto de todos los puntos $(x, y, z) \in \mathbb{R}^3$ tales que $f(x, y, z) = 0$, es decir, aquellos puntos $(x, y, z) \in \mathbb{R}^3$ tales que $z = x^2 + y^2$, o sea un paraboloide como el que se muestra en la siguiente figura (lado izquierdo).

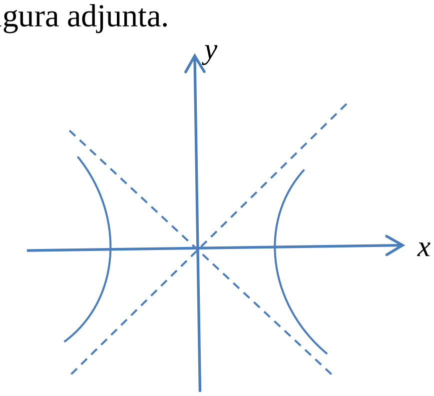




2) Si $f: \mathbb{R}^3 \to \mathbb{R}$ está definida por la ecuación $f(x, y, z) = x^2 + y^2 + z^2$, entonces $N_4(f)$ es el conjunto de todos los puntos $(x, y, z) \in \mathbb{R}^3$ tales que $f(x, y, z) = 4$, es decir, aquellos puntos $(x, y, z) \in \mathbb{R}^3$ tales que $x^2 + y^2 + z^2 = 4$, o sea una esfera centrada en el origen (0,0,0), de radio 2, como se muestra en la figura anterior (lado derecho)

En ejemplos como los dos anteriores, en vez de hablar de conjuntos de nivel usaremos el término superficie de nivel.

3) Si $f: \mathbb{R}^2 \to \mathbb{R}$ está definida por la ecuación $f(x, y) = x^2 - y^2$, entonces $N_1(f)$ es el conjunto de todos los puntos $(x, y) \in \mathbb{R}^2$ tales que $f(x, y) = 1$, es decir, aquellos puntos $(x, y) \in \mathbb{R}^3$ tales que $x^2 - y^2 = 1$, o sea una hipérbola equilátera, como la que se muestra en la figura adjunta.



Interpretación geométrica del vector gradiente

La interpretación geométrica del vector gradiente se obtiene como consecuencia de la regla de la cadena para funciones vectoriales de una variable real y de la interpretación geométrica de la derivada de este tipo de funciones en un punto.

Supongamos que la ecuación $f(x, y, z) = 0$ define una superficie de nivel $S = N_0(f)$ y que $a = (a_1, a_2, a_3) \in S$, donde f es una función de clase C^1. Consideremos entonces una curva C contenida en S y que pasa por a, con la siguiente parametrización $C = Im\gamma$, donde $\gamma: [a, b] \to \mathbb{R}$ es una función de clase C^1 en $]a, b[$ y $\gamma(t_0) = a$, para cierto $t_0 \in]a, b[$.

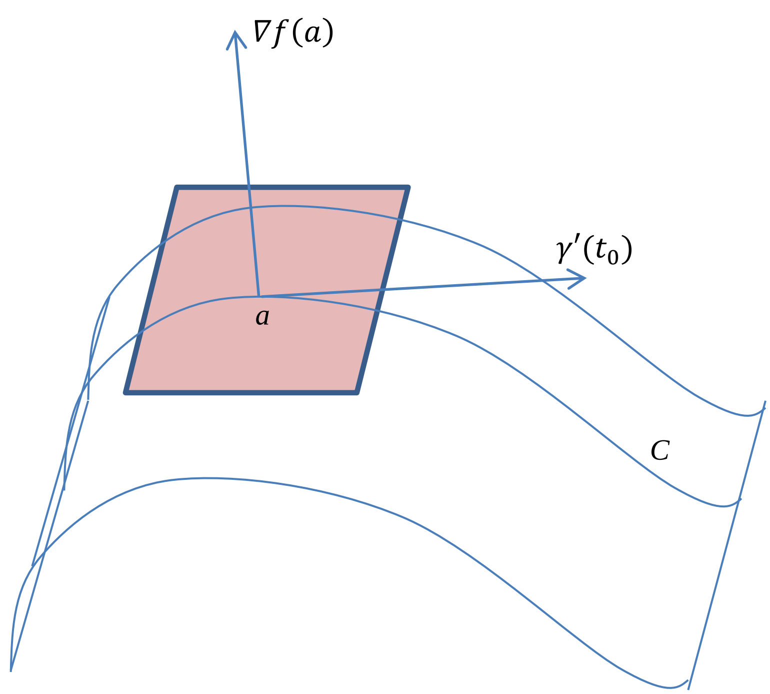
Como $\gamma(t) \in S$, para todo $t \in [a, b]$ entonces $F(t) = f\big(\gamma(t)\big) = 0$, para todo $t \in [a, b]$, entonces se tiene:

- $F'(t_0) = 0$, porque F es constante en $[a, b]$.
- $F'(t_0) = \nabla f\big(\gamma(t_0)\big) \cdot \gamma'(t_0)$, por regla de la cadena.

Se concluye que $\nabla f\big(\gamma(t_0)\big) \cdot \gamma'(t_0) = 0$, y entonces $\nabla f\big(\gamma(t_0)\big) \perp \gamma'(t_0)$, o sea

$$\nabla f(a) \perp \gamma'(t_0)$$

Entonces, el vector gradiente es perpendicular al vector tangente a la curva C en el punto a; pero como esta curva es arbitraria, entonces ocurre lo mismo para cualquier otra curva que satisfaga las mismas condiciones que γ. Resulta entonces que el vector gradiente $\nabla f(a)$ es perpendicular al plano tangente a la superficie de nivel en el punto a: ***Decimos que*** $\boldsymbol{\nabla f(a)}$ ***es perpendicular a la superficie de nivel en el punto a .***



Un punto p pertenece al plano tangente si y solo si $\overrightarrow{ap}$ es perpendicular con $\nabla f(a)$, por lo tanto p pertenece al plano tangente si y solo si

$$\overrightarrow{ap} \cdot \nabla f(a) = 0,$$

Esta última ecuación es conocida como ecuación del plano tangente a la superficie de nivel en el punto $\big(a, f(a)\big)$.

Aplicación al cálculo de derivadas direccionales

Sabemos que si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $a \in D$, entonces existe la derivada direccional y además

$$\frac{\partial f}{\partial \hat{u}}(a) = Df(a)(\hat{u}).$$

Por lo tanto es cierta la igualdad

$$\frac{\partial f}{\partial \hat{u}}(a) = \nabla f(a) \cdot \hat{u}$$

De la desigualdad de Cauchy Schwarz tenemos que

$$\left|\frac{\partial f}{\partial \hat{u}}(a)\right| = |\nabla f(a) \cdot \hat{u}| \leq \|\nabla f(a)\| \ \|\hat{u}\| = \|\nabla f(a)\|\,,$$

entonces

$$-\|\nabla f(a)\| \ \leq \frac{\partial f}{\partial \hat{u}}(a) \leq \|\nabla f(a)\|$$

Si $\nabla f(a) \neq 0$ entonces el vector $\hat{u} = \frac{\nabla f(a)}{\|\nabla f(a)\|}$ define una dirección y se tiene

$$\frac{\partial f}{\partial \hat{u}}(a) = \nabla f(a) \cdot \hat{u} = \nabla f(a) \cdot \frac{\nabla f(a)}{\|\nabla f(a)\|} = \frac{\|\nabla f(a)\|^2}{\|\nabla f(a)\|} = \|\nabla f(a)\|.$$

Por lo tanto, el valor máximo de la derivada direccional en el punto a se obtiene en la dirección del vector gradiente en a y este valor máximo es exactamente $\|\nabla f(a)\|$.

Por la misma razón, el mínimo de la derivada direccional en a se obtiene en la dirección definida por el opuesto del vector gradiente en a y ese valor mínimo es – $\|\nabla f(a)\|$.

Ejemplo: El valor máximo de la derivada direccional de $f: \mathbb{R}^2 \to \mathbb{R}$, donde f está definida por la ecuación $f(x, y) = x^2 + xy^2$ en el punto $a = (1,2)$ es $\|\nabla f(a)\| = \|(6,4)\| = \sqrt{52} = 2\sqrt{13}$, puesto que $\nabla f(x, y) = (2x + y^2, 2xy)$. Además este valor se alcanza en la dirección $\hat{u} = \frac{(6,4)}{2\sqrt{13}} = \left(\frac{3}{\sqrt{13}}, \frac{2}{\sqrt{13}}\right)$.

Máximos y mínimos

Definición:

- Decimos que $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ tiene un máximo local o relativo en $a \in D$, si existe $r > 0$, tal que $B[a, r] \subseteq D$ y $f(x) \leq f(a)$, para todo $x \in B[a, r]$.
- Decimos que $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ tiene un mínimo local o relativo en $a \in D$, si existe $r > 0$, tal que $B[a, r] \subseteq D$ y $f(x) \geq f(a)$, para todo $x \in B[a, r]$.

Ejemplos:

1) $f: \mathbb{R}^2 \to \mathbb{R}$, tal que $f(x, y) = x^2 + y^2$ tiene un mínimo local en $a = (0,0)$. El valor mínimo es $f(0,0) = 0$.
2) $f: \mathbb{R}^2 \to \mathbb{R}$, tal que $f(x, y) = (1 - x^2 - y^2)^{1/2}$ tiene un máximo local en $a = (0,0)$. El valor máximo es $f(0,0) = 1$
3) $f: [0,1] \times [0,1] \to \mathbb{R}$, tal que $f(x, y) = x^2 + y^2$ tiene un mínimo absoluto en $a = (0,0)$ y un máximo absoluto en $b = (1,1)$. Sin embargo, *f* no tiene ni máximo ni mínimo local en estos puntos, puesto que no existe $r > 0$ tal que $B[a, r] \subseteq D$, ni existe $r > 0$ tal que $B[b, r] \subseteq D$.

En cálculo de una variable se prueba que si $f: D \subseteq \mathbb{R} \to \mathbb{R}$ es diferenciable en $a \in D$ y *f* tiene un máximo o un mínimo local en *a*, entonces $f'(a) = 0$. En varias variables un teorema análogo es el siguiente:

Teorema: Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $a \in D$ y *f* tiene un máximo o un mínimo local en *a*, entonces $f'(a) = 0$.

Demostración: Es suficiente observar que la función $\varphi_i(t) = f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$ tiene un máximo o un mínimo local en a_i, para cada $i = 1,2, \dots, n$.

Observación: La igualdad $f'(a) = 0$ es equivalente a $\nabla f(a) = 0$ y también es equivalente a las *n* igualdades $D_i f(a) = 0$, para $i = 1,2, \dots, n$.

Los puntos $a \in D$ tales que $\nabla f(a) = 0$ se llaman puntos estacionarios de *f*.

Consecuencia de este teorema es que los puntos en que *f* presenta máximo o mínimo local se encuentran entre las soluciones de la ecuación $\nabla f(x) = 0$.

En lo que sigue enunciaremos un teorema análogo al criterio de la segunda derivada para máximos y mínimos. Previo a hacerlo, para una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ cuyas derivadas parciales hasta el orden dos en un punto $a \in D$ existen, definimos la matriz hessiana de *f* en

a como la matriz cuyos coeficiente son las derivada parciales de segundo orden de f en a, es decir

$$H_f(a) = \left(D_{ij}f(a)\right)$$

Observación: Si f es de clase C^2 en una bola abierta centrada en a, entonces la matriz hessiana es simétrica, porque $D_{ij}f(a) = D_{ji}f(a)$.

Criterio de la segunda derivada: Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es de clase C^2 en una bola abierta que contiene al punto $a \in D$ y a es un punto estacionario de *f*, entonces se tiene:

a) Si todos los valores propios de $H_f(a)$ son negativos, entonces f tiene un máximo local en a.
b) Si todos los valores propios de $H_f(a)$ son positivos, entonces f tiene un mínimo local en a.
c) Si $H_f(a)$ tiene valores propios positivos y otros negativos, pero no tiene valores propios nulos, entonces f tiene un punto de ensilladura en a.
d) Si $H_f(a)$ tiene al menos un valor propio nulo, no se obtiene información.

Ejemplos:

1) Clasificar los puntos estacionarios de la función $f: \mathbb{R}^2 \to \mathbb{R}$, definida por la igualdad $f(x,y) = x^2y + 2x^2 - 2xy - 6x + 4y + 2$.
Solución: Para determinar los puntos estacionarios de la función *f*, resolvemos la ecuación $\nabla f(x,y) = (0{,}0)$, o sea $(2xy + 4x - 2y - 6, x^2 - 2x + 4) = (0{,}0)$, que podemos escribir como sistema de ecuaciones:

$$\begin{cases} 2xy + 4x - 2y - 6 = 0 \\ x^2 - 2x + 4 = 0 \end{cases}$$

De la segunda ecuación obtenemos $x = 2$, que al ser sustituido en la primera ecuación se obtiene $4y + 4 - 2y - 6 = 0$, de donde $y = -1$. Entonces *f* posee solo un punto estacionario $(x,y) = (2,-1)$.
La matriz Hessiana de f en cualquier punto (x,y) es

$$H_f(x,y) = \begin{bmatrix} 2y + 4 & 2x - 2 \\ 2x - 2 & 0 \end{bmatrix}$$

Por lo tanto

$$H_f(2,-1) = \begin{bmatrix} 2 & 2 \\ 2 & 0 \end{bmatrix}$$

Entonces la ecuación característica es $(2 - \lambda)(0 - \lambda) - 4 = 0$ o $\lambda^2 - 2\lambda - 4 = 0$, por lo tanto los valores propios son $\lambda_1 = 1 - \sqrt{5} < 0$ y $\lambda_2 = 1 + \sqrt{5} > 0$. En consecuencia, *f* tiene un punto de inflexión en el punto estacionario$(x,y) = (2,-1)$.

2) Verifiquemos que $f: \mathbb{R}^2 \to \mathbb{R}$, tal que $f(x,y) = (1 - x^2 - y^2)^{1/2}$ tiene un máximo local en $a = (0,0)$.
Solución: Primero observamos que $\nabla f(x,y) = (0,0)$ si y sólo si

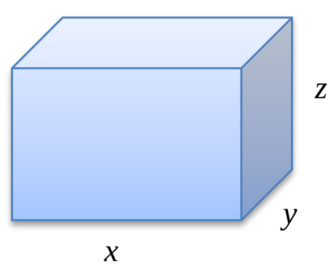
$$\left(\frac{-x}{(1 - x^2 - y^2)^{1/2}}, \frac{-y}{(1 - x^2 - y^2)^{1/2}}\right) = (0,0)$$

De donde se deduce que $(x,y) = (0,0)$ es el único pinto estacionario de *f*. Solo falta verificar que los valores propios de $H_f(0,0)$ son negativos.
Como

$$H_f(0,0) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Se ve que los valores propios de la matriz Hessiana $H_f(0,0)$ son $\lambda_1 = \lambda_2 = -1 < 0$, por consiguiente f tiene un máximo local en el origen de coordenadas.

3) Se desea construir una caja recta rectangular sin tapa en hojalata, de volumen $V = 8.000$ cm^3, empleando el mínimo de material. Hallar las dimensiones que debe tener dicha caja para que en su construcción se emplee un mínimo de material.
Solución: En primer lugar bosquejamos la situación planteada como se muestra en la figura adjunta.



Si denotamos por x, y y z las longitudes de los lados de la caja, vemos que $V = xyz$, de manera que $z = \frac{V}{xy}$.

El área de la superficie de la caja es $A(x,y,z) = xy + yz + xz$

Sustituyendo la expresión de z en esta última ecuación hallamos la función que nos piden minimizar: $A(x,y) = xy + \frac{V}{x} + \frac{V}{y}$.

Para determinar los puntos estacionarios de A, resolvemos la ecuación $\nabla A(x,y) = (0,0)$, que es equivalente al sistema de ecuaciones

$$\begin{cases} y - \dfrac{V}{x^2} = 0 \\ x - \dfrac{V}{y^2} = 0 \end{cases}$$

Despejando y en la segunda ecuación y sustituyendo en la primera ecuación se obtiene

$$x = \frac{x^4}{V}$$

Simplificando por x que es no nulo, se obtiene $x^3 = V$, de donde $x = \sqrt[3]{V}$. Entonces A tiene un único punto estacionario, $(x, y) = (\sqrt[3]{V}, \sqrt[3]{V}) = (20{,}20)$.
La matriz Hessiana de A en un punto (x, y) cualquiera es

$$H_A(x, y) = \begin{bmatrix} \frac{2V}{x^3} & 1 \\ 1 & \frac{2V}{y^3} \end{bmatrix}$$

Y en el punto estacionario este punto estacionario $(x, y) = (20{,}20)$ es

$$H_A(20{,}20) = \begin{bmatrix} \frac{16000}{8000} & 1 \\ 1 & \frac{16000}{8000} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Los valore propios de esta matriz son $\lambda_1 = 1 > 0$ y $\lambda_2 = 3 > 0$, por lo tanto A tiene un mínimo local en (20,20). Como no hay puntos en la frontera del dominio de f, donde la función esté definida, no puede alcanzar un máximo absoluto en la frontera, razón por la cual este mínimo local es también absoluto. Como $xyz = 8000$, vemos que $400z = 8000$, de donde $z = 20$.
En consecuencia, la caja de área mínima es un cubo de arista 20 cm.

Método de multiplicadores de Lagrange

Supongamos que tenemos dos superficies de nivel S_1 y S_2 definidas respectivamente por las ecuaciones $F_1(x, y, z) = 0$ y $F_2(x, y, z) = 0$, donde F_1 y F_2 son funciones que toman valores reales y son de clase C^1 en $\mathbb{R}^3$. Si además $f: D \subseteq \mathbb{R}^3 \to \mathbb{R}$ es una función de clase C^1 en el conjunto abierto D, que contiene a $S_1 \cap S_2$, entonces nos proponemos determinar máximos y mínimos de f sobre $C = S_1 \cap S_2$; es decir, hallar máximos y mínimos de f con las condiciones $F_1(x, y, z) = 0$ y $F_2(x, y, z) = 0$.

Para nuestro razonamiento supondremos que C es una curva suave en $\mathbb{R}^3$. Si f tiene un máximo o un mínimo en un punto $A = (x_0, y_0, z_0) \in C$, entonces podemos hallar una parametrización $\gamma:]a, b[\to C$ de clase C^1, de una parte de la curva C que contiene a A, es decir, que existe $t_0 \in]a, b[$ tal que $\gamma(t_0) = A$.

Entonces tenemos:

a) $F = f \circ \gamma$ tiene un máximo o un mínimo local en t_0, por lo tanto $F'(t_0) = 0$, pero como $F'(t_0) = \nabla f\big(\gamma(t_0)\big) \cdot \gamma'(t_0) = \nabla f(A) \cdot \gamma'(t_0)$, entonces

$$\nabla f(A) \perp \gamma'(t_0)$$

b) Por otra parte, como $C \subseteq S_1$ y $\gamma'(t_0)$ es un vector tangente a γ en $\gamma(t_0)$, entonces

$$\nabla F_1(A) \perp \gamma'(t_0)$$

c) Análogamente, como $C \subseteq S_2$ y $\gamma'(t_0)$ es un vector tangente a γ en $\gamma(t_0)$, entonces

$$\nabla F_2(A) \perp \gamma'(t_0)$$

Observando a), b) y c), vemos que $\nabla f(A)$, $\nabla F_1(A)$ y $\nabla F_2(A)$ son perpendiculares a $\gamma'(t_0)$, por lo tanto los tres gradientes se encuentran en un mismo plano, razón por la cual, si $\nabla F_1(A)$ y $\nabla F_2(A)$ son *linealmente independientes*, entonces existen escalares λ_1 y λ_2 tales que

$$\nabla f(A) = \lambda_1 \nabla F_1(A) + \lambda_2 \nabla F_2(A)$$

De donde podemos concluir que *Los puntos donde se presentan los máximos y mínimos de f bajo las condiciones $F_1(x, y, z) = 0$ y $F_2(x, y, z) = 0$, se encuentran entre las soluciones del sistema de ecuaciones*

$$\begin{cases} F_1(x, y, z) = 0 \\ F_2(x, y, z) = 0 \\ \nabla f(x, y, z) = \lambda_1 \nabla F_1(x, y, z) + \lambda_2 \nabla F_2(x, y, z) \end{cases}$$

Los números reales λ_1 y λ_2 reciben el nombre de multiplicadores de Lagrange, y este método para hallar máximos y mínimos condicionados es conocido como método de los multiplicadores de Lagrange.

Ejemplo: Hallar los máximos y mínimos de la función $f: \mathbb{R}^3 \to \mathbb{R}$ definida por $f(x,y,z) = x + y - z$ sobre la elipse intersección del cilindro $x^2 + y^2 = 1$ con el plano cuya ecuación es $x - z = 0$.

Solución: En primer lugar observemos que la función *f* es continua en todo $\mathbb{R}^3$, por lo tanto en la elipse cuyos puntos satisfacen $x^2 + y^2 = 1$ y $x - z = 0$. Como esta elipse es un conjunto compacto, entonces *f* alcanza un máximo y un mínimo absolutos sobre ésta.
Para hallar los puntos donde se presentan el máximo y el mínimo usaremos el método de los multiplicadores de Lagrange.

En este caso $F_1(x,y,z) = x^2 + y^2 - 1$ y $F_2(x,y,z) = x - z$ que claramente son funciones de clase C^1 en $\mathbb{R}^3$. Además, *f* es de clase C^1 en $\mathbb{R}^3$.
Por lo tanto, para determinar dichos puntos resolvemos el sistema de ecuaciones

$$\begin{cases} F_1(x,y,z) = 0 \\ F_2(x,y,z) = 0 \\ \nabla f(x,y,z) = \lambda_1 \nabla F_1(x,y,z) + \lambda_2 \nabla F_2(x,y,z) \end{cases} \text{ o } \begin{cases} x^2 + y^2 - 1 = 0 \\ x - z = 0 \\ (1,1,-1) = \lambda_1(2x, 2y, 0) + \lambda_2(1,0,-1), \end{cases}$$

de donde:

$$\begin{cases} x^2 + y^2 - 1 = 0 \\ x - z = 0 \\ 2\lambda_1 x + \lambda_2 = 1 \\ 2\lambda_1 y = 1 \\ -\lambda_2 = -1 \end{cases}$$

Por lo tanto, $\lambda_2 = 1$, $2\lambda_1 x = 0$, $2\lambda_1 y = 1$ $(por\ lo\ tanto\ \lambda_1 \neq 0)$, o sea $x = 0$ e $y = \frac{1}{2\lambda_1}$. Sustituyendo en la primera ecuación del sistema se obtiene la igualdad $0 + \frac{1}{4\lambda_1^2} = 1$, que equivale a $4\lambda_1^2 = 1$ y entonces $\lambda_1 = \pm\frac{1}{2}$. Por lo tanto $x = 0$, $y = \pm 1$ y $z = 0$.

Se obtienen las soluciones $(0,-1,0)$ y $(0,1,0)$ para el sistema de ecuaciones. Además $\nabla F_1(0,-1,0) = (0,-2,0)$ y $\nabla F_2(0,-1,0) = (1,0,-1)$ son linealmente independientes, y también lo son $\nabla F_1(0,1,0) = (0,2,0)$ y $\nabla F_2(0,1,0) = (1,0,-1)$. Como

$f(0,-1,0) = -1$ y $f(0,1,0) = 1$, vemos que los valores máximo y mínimo de *f* en la elipse son 1 y -1 respectivamente.

Este método se generaliza y permite realizar el estudio de máximos y mínimos de funciones de varias variables con varias condiciones. El teorema sobre el cual se basa este método es el siguiente:

Teorema (de los multiplicadores de Lagrange): Supongamos que $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ y $F = (F_1, \ldots, F_m): D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ son funciones de clase C^1 en el conjunto abierto D, donde $m < n$, y definamos $S = \{x \in D \,/ F(x) = 0\}$.
Si $a \in D$ verifica las siguientes condiciones:

a) Existe $r > 0$ tal que $f(x) \leq f(a)$, para todo $x \in B(a, r) \cap D$ o bien, existe $r > 0$ tal que $f(x) \geq f(a)$, para todo $x \in B(a, r) \cap D$.
b) $\nabla F_1(a), \ldots, \nabla F_m(a)$ son linealmente independiente.

Entonces existen números reales λ_1, $\lambda_2, \ldots$, λ_m, llamados multiplicadores de Lagrange tales que

$$\nabla f(a) = \sum_{i=1}^{m} \lambda_i \nabla F_i(a)$$

Desde luego que el método es aplicable a funciones con solamente una condición, como veremos en el siguiente ejemplo.

Ejemplo: Hallar los máximos y mínimos de la función $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por $f(x, y, z) = x + y - z$ sobre la esfera unitaria $x^2 + y^2 + z^2 = 1$.

Solución: En primer lugar observemos que la función *f* es continua en todo $\mathbb{R}^3$, por lo tanto también es continua en la esfera unitaria S^2. Como esta esfera es un conjunto compacto, entonces *f* alcanza un máximo y un mínimo absolutos sobre ésta.
Para hallar los puntos donde se presentan el máximo y el mínimo usaremos el método de los multiplicadores de Lagrange.
En este caso tanto *f* como $F_1(x, y, z) = x^2 + y^2 + z^2 - 1$ son funciones de clase C^1 en $\mathbb{R}^3$.
Por lo tanto, para determinar dichos puntos resolvemos el sistema de ecuaciones

$$\begin{cases} F_1(x, y, z) = 0 \\ \nabla f(x, y, z) = \lambda_1 \nabla F_1(x, y, z) \end{cases} \quad \text{o} \quad \begin{cases} x^2 + y^2 + z^2 - 1 = 0 \\ (1,1,-1) = \lambda_1 (2x, 2y, 2z), \end{cases}$$

de donde:

$$\begin{cases} x^2 + y^2 + z^2 - 1 = 0 \\ \\ 2\lambda_1 x = 1 \\ 2\lambda_1 y = 1 \\ 2\lambda_1 z = 1 \end{cases}$$

Observemos que $2\lambda_1 x = 1$, por lo tanto $\lambda_1 \neq 0,$ y entonces $x = y = z = \frac{1}{2\lambda_1}$. Sustituyendo en la primera ecuación del sistema se obtiene $3x^2 = 1$, de donde $x == y = z = \pm\frac{\sqrt{3}}{3}$.

Se obtienen las soluciones $(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3})$ y $(-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3})$ para el sistema de ecuaciones. Además $\nabla F_1(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3})$ es linealmente independientes porque es no nulo. Como $f\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right) = \frac{\sqrt{3}}{3}$ y $f\left(-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}\right) = -\frac{\sqrt{3}}{3}$, vemos que los valores máximo y mínimo de *f* en la elipse son $\frac{\sqrt{3}}{3}$ y $-\frac{\sqrt{3}}{3}$ respectivamente.