

Resumen FVV1

Alex Martínez Ascensión

10 de mayo de 2018

1. La geometría del espacio euclidiano

1.1. Vectores en el espacio tridimensional

Un punto/vector se puede representar en el espacio real tridimensional, \mathbb{R}^3 como una tupla (x, y, z) o de la forma $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, con \mathbf{i}, \mathbf{j} y \mathbf{k} los vectores unitarios en la base canónica en \mathbb{R}^3 : $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$.

Si queremos representar una recta que pase por dos puntos $\mathbf{a} = (x, y, z)$ y $\mathbf{b} = (x', y', z')$, entonces tomamos el vector resta $\mathbf{b} - \mathbf{a}$ y la ecuación de la recta resultante es

$$r(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a})$$

1.2. El producto interno

Si tenemos dos vectores \mathbf{u} y \mathbf{v} , ambos con la misma dimensión, entonces podemos aplicar la operación *producto interno* o *producto escalar*, definida como \cdot o $\langle \cdot, \cdot \rangle$:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \dots + u_kv_k$$

El producto interno cumple que es ≥ 0 y que es conmutativo ($\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$). Además, también es asociativo $\mathbf{a}(\mathbf{u} \cdot \mathbf{v}) = \mathbf{a}(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot \mathbf{a}\mathbf{v}$ y distributivo ($\mathbf{a} \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{a} \cdot \mathbf{u} + \mathbf{a} \cdot \mathbf{v}$)

Por otra parte, la norma de un producto, $\|\cdot\|$, queda definida como la raíz del producto escalar de un vector consigo mismo:

$$\|\mathbf{a}\| = (\mathbf{a} \cdot \mathbf{a})^{1/2}$$

Así, en relación, $\|\mathbf{a}\|^2 = \mathbf{a} \cdot \mathbf{a}$

Dados dos vectores, \mathbf{u} y \mathbf{v} , podemos definir el ángulo entre ellos con la siguiente relación:

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$$

Si el ángulo entre ambos vectores es de 90° , entonces el producto interno es cero y, en tal caso, decimos que \mathbf{u} y \mathbf{v} son ortogonales.

Esta relación puede obtenerse desde la ley de los cosenos: $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 - 2\|\mathbf{v}\|\|\mathbf{u}\|\cos \theta$

Dos desigualdades claves en cálculo, usadas frecuentemente, son la *desigualdad del triángulo* y la *desigualdad de Cauchy-Schwarz* (izquierda y derecha):

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\| \quad |\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

Por último, dados dos vectores, \mathbf{a} y \mathbf{c} , la proyección de \mathbf{c} en \mathbf{a} es:

$$\text{proy}_{\mathbf{a}}(\mathbf{c}) = \mathbf{a} \frac{\mathbf{c} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}} = \mathbf{a} \frac{\mathbf{c} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}$$

1.3. El producto en cruz

Dados dos vectores \mathbf{u} y \mathbf{v} , podemos definir la operación *producto cruz* (\times) en \mathbb{R}^3 como:

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Así, el vector resultante es ortogonal tanto a \mathbf{u} como a \mathbf{v} , siguiendo la regla de la mano derecha. Podemos extender el producto en cruz de modo que el vector resultante sea multiplicado escalarmente con otro vector, \mathbf{w} , de modo que así se crea el producto mixto:

$$\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

La noción geométrica del producto mixto es que el determinante resultante es igual al volumen generado por los tres vectores a multiplicar. Asimismo, si tenemos dos vectores en \mathbb{R}^2 , el determinante es el valor del área del paralelogramo que forman.

Una aplicación geométrica del producto escalar está en determinar ecuaciones de planos que cumplen ciertas condiciones. Un plano en \mathbb{R}^3 viene definido como el conjunto de puntos que satisfacen la ecuación $Ax + By + Cz = D$, de modo que A, B, C y D determinan la morfología del plano. Si $D = 0$ el plano estará centrado en el origen.

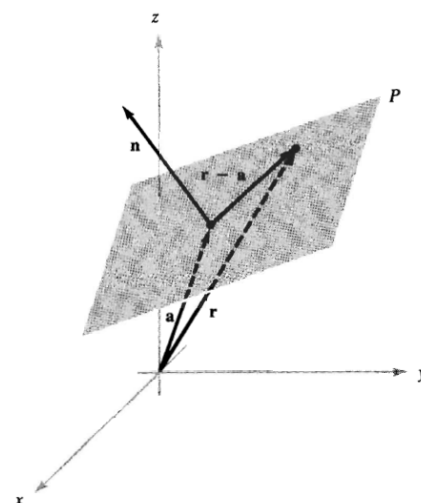
Buscamos un plano P que contenga a todo \mathbf{r} en P y que \mathbf{a} también sea un punto fijo de P . Además, como restricción imponemos que P sea perpendicular a un vector \mathbf{n} . Por tanto, \mathbf{n} también va a ser perpendicular a cualquier vector $\mathbf{r} - \mathbf{a}$. Si $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{n} = (A, B, C)$ y $\mathbf{r} = (x, y, z)$, entonces la ecuación del plano nos queda:

$$(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0 \iff A(x - a_1) + B(y - a_2) + C(z - a_3) = 0$$

Si ponemos que $D = -(Aa_1 + Ba_2 + Ca_3)$, entonces ya tenemos la ecuación del plano.

Por otra parte, la distancia entre un punto \mathbf{e} y un plano $P \equiv Ax + By + Cz = D$ es:

$$d(\mathbf{e}, P) = \frac{|Ae_1 + Be_2 + Ce_3 + D|}{\sqrt{A^2 + B^2 + C^2}}$$



1.4. Coordenadas esféricas y cilíndricas

Si trabajamos en \mathbb{R}^2 , podemos aplicar la transformación

$$x = r \cos \theta \quad y = r \sin \theta$$

Las coordenadas cilíndricas (r, θ, z) vienen dadas por la transformación

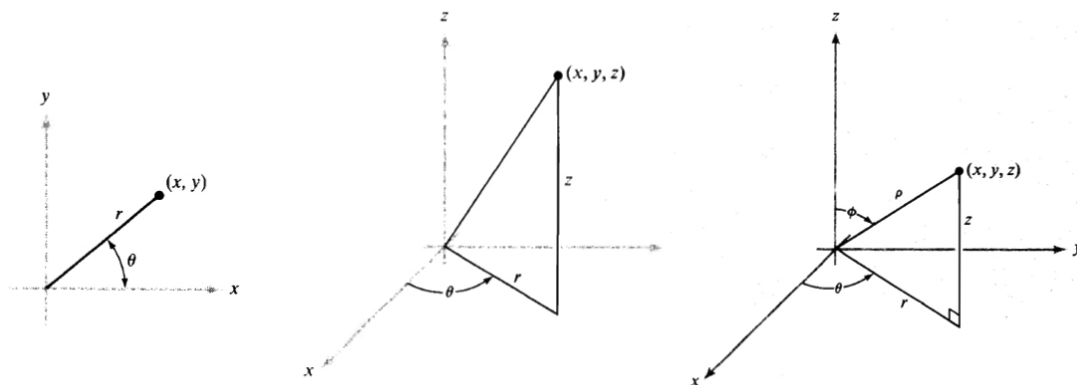
$$x = r \cos \theta \quad y = r \sin \theta \quad z = z$$

También podemos obtener la información explícitamente

$$r = \sqrt{x^2 + y^2} \quad z = z \quad \theta = \begin{cases} \tan^{-1}(y/x) & x > 0 \ y > 0 \\ \pi + \tan^{-1}(y/x) & x < 0 \\ 2\pi + \tan^{-1}(y/x) & x > 0 \ y < 0 \end{cases}$$

Las coordenadas cilíndricas (ρ, θ, ϕ) vienen dadas por la transformación

$$x = \rho \sin \phi \cos \theta \quad y = \rho \sin \phi \sin \theta \quad z = \rho \cos \phi$$



2. Diferenciación

2.1. Geometría de las funciones con valores reales

Sea $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y sea $c \in \mathbb{R}$. Entonces el conjunto de nivel de c se define como todos los puntos $\mathbf{x} \in U$ para los que $f(\mathbf{x}) = c$. Si $n = 2$, hablamos de una curva de nivel y si $n = 3$ se trata de una superficie de nivel.

Así, por ejemplo, la esfera $x^2 + y^2 + z^2 = 1$ puede describirse como un grupo de curvas de nivel de modo que $c = 1 - z^2$ y resolvemos las ecuaciones $x^2 + y^2 = c$ para varios valores de $z \in [-1, 1] \rightarrow c \in [0, 1]$.

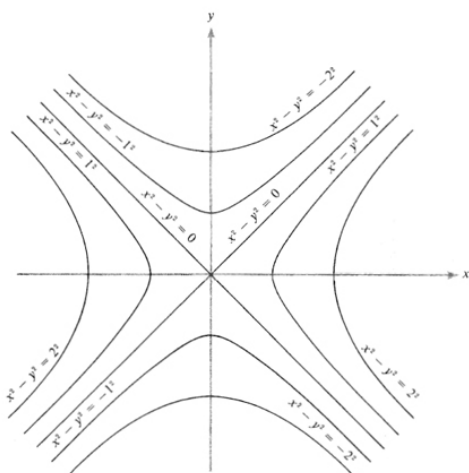


Figura 2.1.9 Curvas de nivel para la función $f(x, y) = x^2 - y^2$.

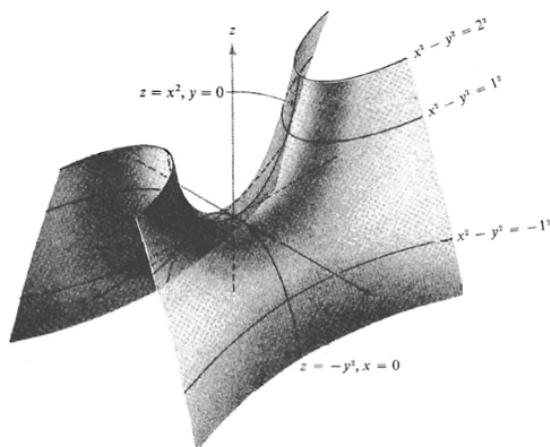


Figura 2.1.10 Algunas curvas de nivel en la gráfica de $f(x, y) = x^2 - y^2$.

Otro ejemplo es el de la imagen de arriba, la función $f(x, y) = x^2 - y^2$. Tomamos diferentes valores de $z = f(x, y)$ y graficamos.

2.2. Límites y continuidad

Sea $U \subset \mathbb{R}^n$. Decimos que un conjunto es abierto cuando para cualquier punto \mathbf{x}_0 en U existe algún $r > 0$ tal que $D_r(\mathbf{x}_0) = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}_0\| < r\}$ está contenido en U . Por otra parte, un punto \mathbf{x} es un punto frontera en U si todo $D_r(\mathbf{x})$ tiene por lo menos un punto perteneciente a U y otro no perteneciente.

El concepto de límite para varias variables es igual que para el de una variable, sólo que cambia el concepto de bola o vecindad para varias variables. Las propiedades de los límites para una función $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ son las siguientes:

- Si $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}$ entonces $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} cf(\mathbf{x}) = c\mathbf{b}$

- Si $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}_1$ y $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = \mathbf{b}_2$ entonces $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (f + g)(\mathbf{x}) = \mathbf{b}_1 + \mathbf{b}_2$
- Si $m=1$ y $\lim_{x \rightarrow x_0} f(x) = \mathbf{b}_1$ y $\lim_{x \rightarrow x_0} g(x) = \mathbf{b}_2$ entonces $\lim_{x \rightarrow x_0} (fg)(x) = \mathbf{b}_1 \mathbf{b}_2$
- Si $m=1$ y $\lim_{x \rightarrow x_0} f(x) = \mathbf{b} \neq 0$ y $f(x) \neq 0$ entonces $\lim_{x \rightarrow x_0} 1/f(x) = 1/\mathbf{b}$
- Si $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ entonces $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{b} = (b_1, \dots, b_m)$ sii $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_i(\mathbf{x}) = b_i$

Una función f es continua en \mathbf{x}_0 sii $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$. Los cinco puntos de propiedades de límites son también aplicables a continuidad.

La composición de funciones también es continua. Es decir, si tenemos dos funciones $f : B \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$ y $g : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, si se cumple que g es continua en $\mathbf{x} \in A$ y f es continua en $g(\mathbf{x})$, entonces $(f \circ g)(\mathbf{x}) = f(g(\mathbf{x}))$ también es continua en \mathbf{x} .

Cuando queremos resolver un límite multivariable $[\lim_{(\mathbf{x}) \rightarrow (\mathbf{x}_0)} f(\mathbf{x}) = \mathbf{b}]$ podemos emplear varias estrategias:

- Si tenemos imaginación, y a veces cuando no queda otra, recurrimos a la notación $\delta - \epsilon$ para demostrar que para todo $\delta > 0$ existe un $\epsilon > 0$ tal que si $\|\mathbf{x} - \mathbf{x}_0\| < \delta$, entonces $\|f(\mathbf{x}) - \mathbf{b}\| < \epsilon$. Sin embargo, cada caso es excepcional, y puede no ser fácilmente demostrable.
- Podemos probar a emplear límites iterados, es decir, para $\mathbf{x} = (x_1, \dots, x_m)$, podemos tomar x_1 como variable y el resto como constantes, y así ir repitiendo el límite: $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \lim_{x_1 \rightarrow a_1} (\lim_{x_2 \rightarrow a_2} (\dots (\lim_{x_n \rightarrow a_n} f(\mathbf{x}))))$ e ir repitiendo con otras combinaciones $(\lim_{x_2 \rightarrow a_2} (\lim_{x_4 \rightarrow a_4} (\dots (\lim_{x_1 \rightarrow a_1} f(\mathbf{x}))))$. Si en alguno de los casos un límite difiere, entonces tenemos que el límite no existe. Sin embargo, si en todos los casos los límites existen y son iguales, NO PODEMOS DECIR QUE EL LÍMITE EXISTA.
- Podemos aproximarnos al límite desde diferentes trayectorias ($x = 0, y = x, y = x^2 - z$) y buscar una trayectoria para la cual el límite sea diferente del resto. En este caso, probar que el límite exista en una trayectoria no demuestra que el límite exista, ya que deberíamos probar todas las trayectorias posibles.
- Podemos pasar el límite a coordenadas polares / esféricas y ver qué pasa.

2.3. Diferenciación

Recordamos el concepto de derivada en punto para funciones de una variable. Decimos que una función f es derivable en a cuando existe el límite

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

Y decimos que f es derivable en general cuando lo es para todos los puntos de \mathbb{R}

En este caso, para funciones de varias variables aplicamos el concepto de derivada parcial, de modo que la derivada queda definida de la siguiente manera:

$$\frac{\partial f}{\partial x_j}(x_1, \dots, x_n) = \lim_{h \rightarrow 0} \frac{f(x_1, x_2, \dots, x_j + h, \dots, x_n) - f(x_1, x_2, \dots, x_j, \dots, x_n)}{h}$$

Este caso es aplicable para $f : \mathbb{R}^n \rightarrow \mathbb{R}$. En el caso de $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, hablaríamos de $\frac{\partial f_k}{\partial x_j}$ a la parcial de la k -ésima función para la j -ésima variable.

Las derivadas de una función para todas las variables se resume en la matriz de la derivada de f :

$$Df(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Esta matriz también se llama **matriz jacobiana** o **diferencial jacobiana**.

Dicho esto, más precisamente, tenemos que f es diferenciable en \mathbf{x}_0 si

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{D}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$$

Cuando $\mathbf{D}f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$ entonces simplificamos la notación de matriz derivada a gradiente de f , ∇f .

Finalmente, podemos extrapolar el concepto de diferenciabilidad para todo \mathbf{x} . Sea $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Para que f sea diferenciable en \mathbf{x} todas las derivadas parciales de f han de existir y han de ser continuas en \mathbf{x} . Esto último es clave: si alguna derivada parcial existe pero no es continua en \mathbf{x} , entonces f no es diferenciable.

Si f es k -diferenciable, entonces decimos que es del tipo C^k .

2.4. Propiedades de la derivada

Las propiedades comunes de una variable sobre suma y producto de derivadas se mantienen en el caso multivariable:

- Si $h(\mathbf{x}) = cf(\mathbf{x})$, entonces $\mathbf{D}h(\mathbf{x}) = c\mathbf{D}f(\mathbf{x})$
- Si $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$, entonces $\mathbf{D}h(\mathbf{x}) = \mathbf{D}f(\mathbf{x}) + \mathbf{D}g(\mathbf{x})$
- Si $h(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$, entonces $\mathbf{D}h(\mathbf{x}) = g(\mathbf{x})\mathbf{D}f(\mathbf{x}) + f(\mathbf{x})\mathbf{D}g(\mathbf{x})$
- Si $h(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})$, entonces $\mathbf{D}h(\mathbf{x}) = \frac{g(\mathbf{x})\mathbf{D}f(\mathbf{x}) - f(\mathbf{x})\mathbf{D}g(\mathbf{x})}{[g(\mathbf{x})]^2}$

Recordamos de cálculo de una variable que si tenemos una función $f(g(x))$, entonces la derivada es, aplicando la regla de la cadena:

$$\frac{df}{dx} = \frac{df}{dg} \frac{dg}{dx} = f'(g)g'(x)$$

En el caso de funciones multivariable se extiende el concepto tal que si, por ejemplo, $h = f(g(x), k(x), j(x))$ entonces

$$\frac{dh}{dx} = \frac{\partial f}{\partial g} \frac{dg}{dx} + \frac{\partial f}{\partial k} \frac{dk}{dx} + \frac{\partial f}{\partial j} \frac{dj}{dx} = \begin{bmatrix} \frac{\partial f}{\partial g} & \frac{\partial f}{\partial j} & \frac{\partial f}{\partial k} \end{bmatrix} \begin{bmatrix} \frac{dg}{dx} \\ \frac{dj}{dx} \\ \frac{dk}{dx} \end{bmatrix} = \nabla f \cdot \mathbf{D}h$$

$h = f(g(x, y, z), k(x, y, z), j(x, y, z))$ entonces ∇f es:

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial f}{\partial g} \frac{\partial g}{\partial x} + \frac{\partial f}{\partial k} \frac{\partial k}{\partial x} + \frac{\partial f}{\partial j} \frac{\partial j}{\partial x} \\ \frac{\partial h}{\partial y} &= \frac{\partial f}{\partial g} \frac{\partial g}{\partial y} + \frac{\partial f}{\partial k} \frac{\partial k}{\partial y} + \frac{\partial f}{\partial j} \frac{\partial j}{\partial y} \\ \frac{\partial h}{\partial z} &= \frac{\partial f}{\partial g} \frac{\partial g}{\partial z} + \frac{\partial f}{\partial k} \frac{\partial k}{\partial z} + \frac{\partial f}{\partial j} \frac{\partial j}{\partial z} \end{aligned}$$

Así,

$$\nabla h = \begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial g} & \frac{\partial f}{\partial k} & \frac{\partial f}{\partial j} \end{bmatrix} \begin{bmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial k}{\partial x} & \frac{\partial k}{\partial y} & \frac{\partial k}{\partial z} \\ \frac{\partial j}{\partial x} & \frac{\partial j}{\partial y} & \frac{\partial j}{\partial z} \end{bmatrix} \rightarrow \nabla h = \nabla f \cdot \mathbf{D}h$$

2.5. Gradientes y derivadas direccionales

Si $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, la derivada direccional de f en \mathbf{x} en la dirección de un vector \mathbf{v} está dada por

$$\left. \frac{d}{dt} f(\mathbf{x} + t\mathbf{v}) \right|_{t=0} = \lim_{h \rightarrow 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}$$

Si tomamos $c(t) = \mathbf{x} + t\mathbf{v}$, entonces $\frac{d}{dt} f(c(t)) = \nabla f(c) \cdot c'(t)$. Como $c(0) = \mathbf{x}$ y $c'(t) = \mathbf{v}$, entonces la derivada direccional es $\left. \frac{d}{dt} f(\mathbf{x} + t\mathbf{v}) \right|_{t=0} = \nabla f(\mathbf{x}) \cdot \mathbf{v}$.

Como visión geométrica, tenemos que el gradiente apunta en la dirección normal al plano tangente de la superficie de la gráfica en el punto. Asimismo, el gradiente también apunta en la dirección en la que f crece más rápido. Por ende, el negativo del gradiente indica la dirección en la que f decrece más rápido.

Por tanto, si \mathbf{v} es el vector tangente de una trayectoria $c(t)$ de una superficie S en $t = 0$, entonces el gradiente de la función f en ese será ortogonal a \mathbf{v} , es decir, $\nabla f(c(0)) \cdot \mathbf{v} = 0$.

Así, extendiendo lo que vimos en 2.3, el plano tangente a una superficie S en un punto (x_0, y_0, z_0) es

$$P \equiv \nabla f \cdot (x - x_0, y - y_0, z - z_0) = 0$$

2.6. Derivadas parciales iteradas

Hasta la fecha hemos realizado sólo una diferenciación a las funciones, y hemos tratado con funciones C^1 . Sin embargo, podemos diferenciar varias veces, de modo que podemos extender la notación de derivada parcial a una derivada parcial iterada:

$$\frac{\partial^k f}{\partial x_1 \partial x_2 \cdots \partial x_k} = \frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_2} \left(\cdots \left(\frac{\partial f}{\partial x_k} \right) \right) \right) \quad \frac{\partial^k f}{\partial x^k} = \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \left(\cdots \left(\frac{\partial f}{\partial x} \right) \right) \right)$$

Si la función es C^2 , el teorema de la derivada mixta afirma que

$$f_{yx} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = f_{xy}$$

Este concepto puede extenderse a funciones de cualquier tipo C^k .

3. Funciones con valores vectoriales

3.1. Trayectorias y velocidad

Una trayectoria es una función $\sigma : [a, b] \rightarrow \mathbb{R}^n$. Los puntos $\sigma(a)$ y $\sigma(b)$ son los extremos de la trayectoria. La imagen de la trayectoria en \mathbb{R}^n se llama curva.

Así, una trayectoria $\sigma(t)$ puede definirse en \mathbb{R}^3 como una función $\sigma(t) = (x(t), y(t), z(t))$. La derivada de la trayectoria viene definida como:

$$\sigma'(t) = \lim_{h \rightarrow 0} \frac{\sigma(t_0 + h) - \sigma(t_0)}{h} = (x'(t), y'(t), z'(t))$$

La recta tangente a una trayectoria es $\mathbf{l}(\lambda) = \sigma(t_0) + \lambda \sigma'(t_0)$

A nivel físico, si tenemos una trayectoria $\mathbf{r}(t)$, a velocidad es $\mathbf{v}(t) = \mathbf{r}'(t)$, la rapidez es $S(t) = \|\mathbf{v}(t)\|$ y la aceleración es $\mathbf{a}(t) = \mathbf{r}''(t)$.

3.2. Longitud de arco

Si tenemos una trayectoria $\sigma : [a, b] \rightarrow \mathbb{R}^n$, la longitud de arco entre a y b está definida como:

$$l(\sigma) = \int_a^b \|\sigma'(t)\| dt = \int_a^b \sqrt{[x'_1(t)]^2 + [x'_2(t)]^2 + \cdots + [x'_n(t)]^2} dt$$

Por lo general no nos interesa obtener la longitud de arco entre dos puntos, sino dado un a fijo, obtener una función de longitud de arco para un tiempo t :

$$s(t) = \int_a^t \|\sigma'(\tau)\| d\tau$$

Así, por el Primer Teorema del Cálculo, tenemos que $s'(t) = \|\sigma'(t)\|$

3.3. Campos vectoriales

Un campo vectorial en \mathbb{R}^n es una función $\mathbf{F} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ que asigna a cada vector \mathbf{x} en A su vector $\mathbf{F}(\mathbf{x})$ en \mathbb{R}^m . Así, por ejemplo, para una partícula posicionada en $\mathbf{r} = (x, y, z)$, su fuerza de atracción localizada en el origen es un campo vectorial $\mathbf{F} = -\frac{mMG}{\|\mathbf{r}\|^3} \mathbf{r}$.

Una línea de flujo para \mathbf{F} es una trayectoria $\sigma(t)$ tal que

$$\sigma'(t) = \mathbf{F}(\sigma(t))$$

Es decir, el argumento del campo vectorial es la propia trayectoria, y su imagen ha de ser igual a la derivada de la trayectoria. En condiciones iniciales, tomamos $\sigma(0) = (x_0, y_0, z_0)$.

Para agilizar la notación, definimos $\phi(\mathbf{x}, t)$ como la posición del punto en la línea de flujo tras haber transcurrido un tiempo t . Así, la ecuación anterior se reescribe como (además de una segunda condición inicial):

$$\phi(\mathbf{x}, 0) = \mathbf{x} \quad \frac{\partial}{\partial t} \phi(\mathbf{x}, t)$$

3.4. Divergencia y rotacional de un campo vectorial

Usualmente tomamos el gradiente como una función asociada a una función. Sin embargo, el gradiente como tal es un operador, es decir, un elemento matemático que combinado a una función modifica sus valores. Así, el gradiente como tal está definido como:

$$\nabla = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z}$$

Y cuando hacemos ∇f en realidad hacemos el producto escalar entre ∇ y f .

La *operación rotacional* asocia a cada campo vectorial C^1 \mathbf{F} el siguiente campo vectorial:

$$\text{rot } \mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix}$$

Si $\nabla \times \mathbf{F} = 0$, entonces decimos que \mathbf{F} es irrotacional.

Otra función es la *divergencia* de un campo vectorial, definida como:

$$\text{div } \mathbf{F} = \nabla \cdot \mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$$

En este campo, la divergencia es un campo escalar, a diferencia del rotacional, que es un campo vectorial.

Como curiosidad, es importante recordar que la divergencia de un rotacional siempre es 0:

$$\nabla \cdot (\nabla \times \mathbf{F}) = 0$$

Además del operador gradiente existen otros muchos. Uno de los más importantes es el *operador laplaciano*, definido como

$$\Delta f = (\nabla \cdot \nabla)f = \nabla^2 f = \sum_i^n \frac{\partial^2 f}{\partial x_i^2}$$

Cuando una función cumple que su laplaciano es nulo, es decir, $\nabla^2 f = 0$ decimos que es una *función armónica*.

4. Derivadas de orden superior; Máximos y mínimos

4.1. Teorema de Taylor

Si recordamos el teorema para una variable, dada una función f derivable k veces en un punto a , entonces podemos escribir una aproximación de f como:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + R_k(x, a)$$

con R_k :

$$R_k(x, a) = \int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t) dt$$

Así, conforme mayor es k , menor es R_k , y obtenemos una mejor aproximación de f .

Si desarrollamos el teorema para varias variables con derivadas hasta de tercer orden tenemos la expresión:

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_i^n h_i \frac{\partial f}{\partial x_i}(\mathbf{x}_0) + \frac{1}{2!} \sum_{i,j} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) + R_2(\mathbf{h}, \mathbf{x}_0)$$

Aquí cambiamos un poco la notación por conveniencia. Si definimos $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$, entonces $f(\mathbf{x}_0 + \mathbf{h})$, donde \mathbf{x}_0 sería el equivalente multidimensional de a , entonces $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$, que sería el equivalente multidimensional de $(x - a)$. Por tanto, la expresión final nos va a quedar en función de \mathbf{h} .

La expresión anterior se puede expandir hasta el orden que deseemos, k :

$$\begin{aligned} f(\mathbf{x}_0 + \mathbf{h}) = & f(\mathbf{x}_0) + \sum_{m_1}^n h_{m_1} \frac{\partial f}{\partial x_{m_1}}(\mathbf{x}_0) + \frac{1}{2!} \sum_{m_1, m_2}^n h_{m_1} h_{m_2} \frac{\partial^2 f}{\partial x_{m_1} \partial x_{m_2}}(\mathbf{x}_0) + \cdots \\ & + \frac{1}{k!} \sum_{m_1, m_2, \dots, m_k}^n h_{m_1} h_{m_2} \cdots h_{m_k} \frac{\partial^k f}{\partial x_{m_1} \partial x_{m_2} \cdots \partial x_{m_k}}(\mathbf{x}_0) + R_k(\mathbf{h}, \mathbf{x}_0) \end{aligned}$$

4.2. Extremos de funciones con valores reales

Si $U \subset \mathbb{R}^n$ es abierto, $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable y $\mathbf{x}_0 \in U$, entonces $\mathbf{D}f(\mathbf{x}_0)$ es un punto crítico de f . Por tanto, para hallar extremos críticos, realizamos la derivada de f e igualamos a cero su matriz derivada. La condición que haga que todos los elementos de la matriz derivada sean 0 es un punto crítico.

Si suponemos que f tiene derivadas de segundo orden, entonces el hessiano de f en \mathbf{x}_0 es:

$$Hf(\mathbf{x}_0)(\mathbf{h}) = \frac{1}{2} \sum_{i,j} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)$$

Esta expresión puede reescribirse en forma matricial como:

$$H(\mathbf{h}) = \frac{1}{2} [h_1 \ h_2 \ \cdots \ h_n] B \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \quad \mathbf{H} = B = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Para saber si un punto crítico es un máximo, mínimo o punto silla, nos fijaremos en la matriz hessiana, \mathbf{H} . El punto crítico será máximo cuando \mathbf{H} sea definida negativa, y mínimo cuando sea definida positiva. Si recordamos de álgebra, una matriz en base \mathcal{B} puede transformarse a una base \mathcal{B}^* aplicando transformaciones en filas y columnas, de modo que la matriz resultante es una diagonal. Así, si la matriz es definida positiva, será de la forma de la izquierda, y si es definida negativa es de la forma de la derecha.

$$\begin{bmatrix} + & 0 & 0 & \cdots & 0 \\ 0 & + & 0 & \cdots & 0 \\ 0 & 0 & + & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & + \end{bmatrix} \quad \begin{bmatrix} - & 0 & 0 & \cdots & 0 \\ 0 & - & 0 & \cdots & 0 \\ 0 & 0 & - & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & - \end{bmatrix}$$

Visto esto, el criterio de clasificación es el siguiente:

- Si el punto crítico es un mínimo, entonces $\frac{\partial^2 f}{\partial x_1^2} > 0$ y el determinante de todas las submatrices es positivo.
- Si el punto crítico es máximo, entonces $\frac{\partial^2 f}{\partial x_1^2} < 0$ y el determinante de las submatrices alterna el signo ($D_1 < 0, D_2 > 0, D_3 < 0, \dots$).
- Si los determinantes no siguen que todos son positivos o alternantes, pero son mayores que cero, entonces tenemos un punto silla (la matriz no es ni definida positiva ni negativa).
- Si el determinante es cero, entonces no podemos aplicar el criterio y tenemos que analizar la función por otros métodos.

En algunas situaciones nos piden analizar los máximos y mínimos de una función f restringida a un conjunto U . Entonces los pasos a seguir son:

1. Localizar los puntos críticos en U . No hace falta saber si son máximos o mínimos.
2. Localizar los puntos críticos en la región frontera de U .
3. Calcular el valor de todos los puntos críticos, y seleccionar el mayor y el menor.

4.3. Extremos restringidos y multiplicadores de Lagrange

En muchas ocasiones queremos maximizar/minimizar una función sujeta a ciertas condiciones. Si f es la función que queremos optimizar, y g es una función de restricción, definida en un conjunto de nivel c : $g(\mathbf{x}) = c$, entonces si f restringida en g tiene un máximo/mínimo en \mathbf{x}_0 , entonces existe un número real λ tal que $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$.

Si consideramos la función $h(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda[g(\mathbf{x}) - c]$, entonces para hallar los puntos críticos tenemos que resolver el siguiente sistema de ecuaciones:

$$\begin{cases} \frac{\partial h}{\partial x_1} = 0 \iff \frac{\partial f}{\partial x_1} = \lambda \frac{\partial g}{\partial x_1} \\ \frac{\partial h}{\partial x_2} = 0 \iff \frac{\partial f}{\partial x_2} = \lambda \frac{\partial g}{\partial x_2} \\ \dots \\ \frac{\partial h}{\partial x_n} = 0 \iff \frac{\partial f}{\partial x_n} = \lambda \frac{\partial g}{\partial x_n} \\ \frac{\partial h}{\partial \lambda} = 0 \iff g(\mathbf{x}) = c \end{cases}$$

Como tenemos $n + 1$ incógnitas (las n componentes de \mathbf{x} y λ), entonces tenemos $n + 1$ ecuaciones, siendo la última la propia restricción del enunciado.

En muchas ocasiones tenemos más de una restricción. En este caso podemos generalizar el problema a k restricciones añadiendo las diferentes restricciones al sistema. La ecuación a igualar sería en este caso:

$$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla g_1(\mathbf{x}_0) + \lambda_2 \nabla g_2(\mathbf{x}_0) + \dots + \lambda_k \nabla g_k(\mathbf{x}_0)$$

El sistema de ecuaciones a resolver es, por tanto:

$$\begin{cases} \frac{\partial h}{\partial x_1} = 0 \iff \frac{\partial f}{\partial x_1} = \lambda \frac{\partial g}{\partial x_1} \\ \frac{\partial h}{\partial x_2} = 0 \iff \frac{\partial f}{\partial x_2} = \lambda \frac{\partial g}{\partial x_2} \\ \dots \\ \frac{\partial h}{\partial x_n} = 0 \iff \frac{\partial f}{\partial x_n} = \lambda \frac{\partial g}{\partial x_n} \\ \frac{\partial h}{\partial \lambda_1} = 0 \iff g_1(\mathbf{x}) = c_1 \\ \frac{\partial h}{\partial \lambda_2} = 0 \iff g_2(\mathbf{x}) = c_2 \\ \dots \\ \frac{\partial h}{\partial \lambda_k} = 0 \iff g_k(\mathbf{x}) = c_k \end{cases}$$

Que en este caso es un sistema de $n + k$ ecuaciones con $n + k$ incógnitas.

Una vez obtengamos los puntos críticos del problema de optimización, tenemos que saber si son máximos o mínimos. Para ello empleamos el criterio del hessiano limitado. Tomando el determinante:

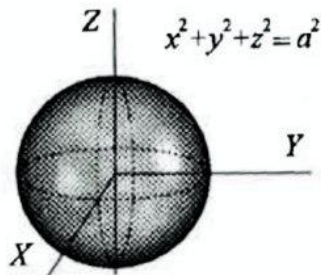
$$|\overline{H}| = \begin{vmatrix} 0 & -\frac{\partial g}{\partial x_1} & -\frac{\partial g}{\partial x_2} & \dots & -\frac{\partial g}{\partial x_n} \\ -\frac{\partial g}{\partial x_1} & \frac{\partial^2 h}{\partial x_1^2} & \frac{\partial^2 h}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 h}{\partial x_1 \partial x_n} \\ -\frac{\partial g}{\partial x_2} & \frac{\partial^2 h}{\partial x_2 \partial x_1} & \frac{\partial^2 h}{\partial x_2^2} & \dots & \frac{\partial^2 h}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial g}{\partial x_n} & \frac{\partial^2 h}{\partial x_n \partial x_1} & \frac{\partial^2 h}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 h}{\partial x_n^2} \end{vmatrix}$$

Realizamos los subdeterminantes $k \times k$ para $k \geq 3$ y aplicamos el criterio del hessiano al revés: si todos los subdeterminantes son positivos, entonces el punto crítico es un máximo local, mientras que si los signos de los subdeterminantes se alternan, el punto crítico es un mínimo local. Si $|\overline{H}| = 0$ entonces el criterio no es concluyente.

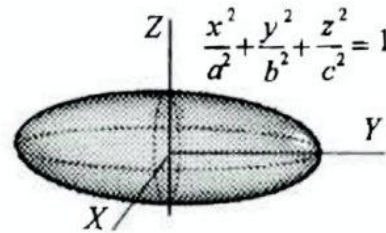
Para $k = 3$ el criterio se simplifica a que si $|\overline{H}_2| = \begin{vmatrix} 0 & -\frac{\partial g}{\partial x_1} & -\frac{\partial g}{\partial x_2} \\ -\frac{\partial g}{\partial x_1} & \frac{\partial^2 h}{\partial x_1^2} & \frac{\partial^2 h}{\partial x_1 \partial x_2} \\ -\frac{\partial g}{\partial x_2} & \frac{\partial^2 h}{\partial x_2 \partial x_1} & \frac{\partial^2 h}{\partial x_2^2} \end{vmatrix} > 0$ entonces es máximo local y si $|\overline{H}_2| < 0$ entonces es un mínimo local.

Superficies Cuadráticas

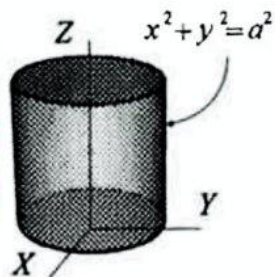
Esfera



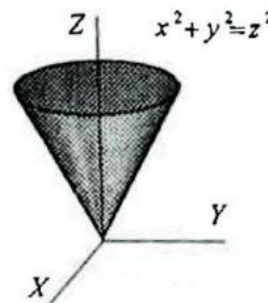
Elipsoide



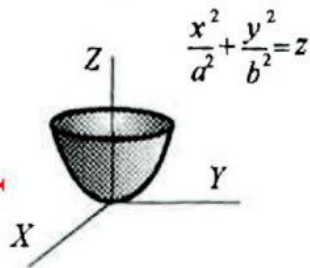
Cilindro circular



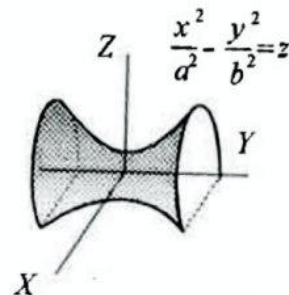
Cono



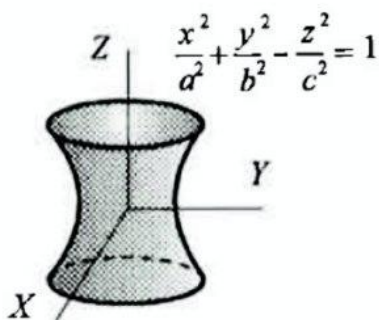
Paraboloide elíptico



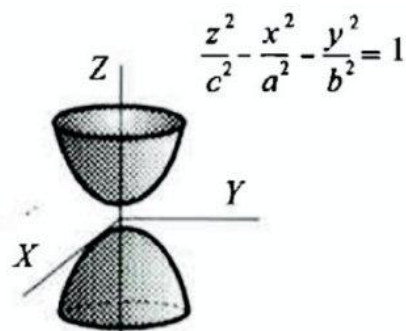
Paraboloide hiperbólico



Hiperboloide una hoja



Hiperboloide dos hojas



$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz = 0$$