Tema 6. Funciones de varias variables

1.1. Producto escalar y norma en $\mathbb{R}^n$

Como sabes, $\mathbb{R}^n$ es un espacio vectorial en el que suele destacarse la llamada base canónica formada por los vectores $\{\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_n}\}$ donde $\mathbf{e_k}$ es el vector cuyas componentes son todas nulas excepto la que ocupa el lugar k que es igual a 1. Dados dos vectores $\mathbf{x} = (x_1, x_2, \dots, x_n)$ $\mathbf{y} = (y_1, y_2, \dots, y_n)$ se define su **producto escalar** por:

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Observa que el producto escalar de dos vectores no es un vector sino un número real. La notación $\mathbf{x.y}$ es frecuentemente usada en los libros de Física para representar el producto escalar de los vectores $\mathbf{x}$ e $\mathbf{y}$.

Las siguientes propiedades del producto escalar se deducen fácilmente de la definición:

- $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle$ para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (simetría).
- $\langle \alpha \mathbf{x} + \beta \mathbf{y} | \mathbf{z} \rangle = \alpha \langle \mathbf{x} | \mathbf{z} \rangle + \beta \langle \mathbf{y} | \mathbf{z} \rangle$ para todos $\alpha, \beta \in \mathbb{R}$ y para todos $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ (linealidad).

La **norma** (**euclídea**) de un vector $\mathbf{x}$ se define por

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle} = \sqrt{\sum_{k=1}^{n} x_k^2}$$

Observa que para $n = 1$ el producto escalar de dos números $x, y \in \mathbb{R}$ es su producto usual, y que para $x \in \mathbb{R}$ se tiene que $\|x\| = \sqrt{x^2} = |x|$.

Desigualdad de Cauchy-Schwarz.

Para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se verifica que $\left| \langle \mathbf{x} | \mathbf{y} \rangle \right| \leqslant \|\mathbf{x}\| \|\mathbf{y}\|$. Además, supuesto que $\mathbf{x}$ e $\mathbf{y}$ no son nulos, la igualdad $\left| \langle \mathbf{x} | \mathbf{y} \rangle \right| = \|\mathbf{x}\| \|\mathbf{y}\|$ equivale a que hay un número $\lambda \in \mathbb{R}$ tal que $\mathbf{x} = \lambda \mathbf{y}$ (es decir, los vectores $\mathbf{x}$ e $\mathbf{y}$ están en una misma recta que pasa por el origen).

Supuesto que $\mathbf{x}$ e $\mathbf{y}$ son vectores no nulos, la desigualdad de Cauchy-Schwarz nos dice que

$$-1 \leqslant \frac{\langle \mathbf{x} | \mathbf{y} \rangle}{\|x\| \|y\|} \leqslant 1$$

Por tanto, existe un único número $t \in [0, \pi]$ tal que

$$\cos t = \frac{\langle \mathbf{x} | \mathbf{y} \rangle}{\|x\| \|y\|}$$

Se dice que dicho número t es la *medida en radianes del ángulo que forman los vectores* $\mathbf{x}$ e $\mathbf{y}$. Naturalmente, de la definición dada se deduce que $\langle \mathbf{x} | \mathbf{y} \rangle = \|x\| \|y\| \cos t$.

Se dice que los vectores $\mathbf{x}$ e $\mathbf{y}$ son **ortogonales**, y escribimos $\mathbf{x} \perp \mathbf{y}$, cuando su producto escalar es cero.

Desigualdad triangular.

Para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se verifica que $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$. Además, supuesto que $\mathbf{x}$ e $\mathbf{y}$ no son nulos, la igualdad $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ equivale a que hay un número $\lambda > 0$ tal que $\mathbf{x} = \lambda \mathbf{y}$ (es decir, los *vectores* **x** *e* **y** *están en una misma semirrecta que pasa por el origen).*

Dados dos vectores **x** e **y**, el número $\|\mathbf{x}-\mathbf{y}\|$ se llama la **distancia** (euclídea) entre **x** e **y**.

Para $\mathbf{a}\in\mathbb{R}^n$ y $r>0$, definimos

$$B(\mathbf{a},r)=\{\mathbf{x}\in\mathbb{R}^n : \|\mathbf{x}-\mathbf{a}\|<r\}$$

Para $\mathbf{a}=(\alpha,\beta)\in\mathbb{R}^2$, tenemos que

$$B((\alpha,\beta),r)=\left\{(x,y)\in\mathbb{R}^2 : (x-\alpha)^2+(y-\beta)^2<r^2\right\}$$

es un círculo de centro (α,β) y radio r sin incluir la circunferencia que lo limita.

Para $\mathbf{a}=(\alpha,\beta,\gamma)\in\mathbb{R}^3$, tenemos que

$$B((\alpha,\beta,\gamma),r)=\left\{(x,y,z)\in\mathbb{R}^3 : (x-\alpha)^2+(y-\beta)^2+(z-\gamma)^2<r^2\right\}$$

es una bola esférica de centro (α,β,γ) y radio r sin incluir la esfera que la limita.

1.2. Conceptos topológicos

Dado un conjunto no vacío, $E\subset\mathbb{R}^n$, podemos clasificar los puntos de $\mathbb{R}^n$ con respecto al conjunto E como sigue. Un punto $\mathbf{x}\in\mathbb{R}^n$ se dice que es:

- **Interior** de E si *existe algún* $r>0$ tal que $B(\mathbf{x},r)\subset E$.
- **Exterior** a E, si es interior al complemento de E.
- **Frontera** de E, si no es interior ni exterior a E.

El conjunto de todos los puntos interiores de E se representa por $\mathrm{int}(E)$. El conjunto de todos los puntos exteriores de E se representa por $\mathrm{ext}(E)$. El conjunto de todos los puntos frontera de E se representa por $\mathrm{Fr}(E)$. Es claro que $\mathbb{R}^n=\mathrm{int}(E)\cup\mathrm{Fr}(E)\cup\mathrm{ext}(E)$ donde la unión es disjunta. Es evidente que $\mathrm{int}(E)\subset E$.

Se dice que el conjunto E es **abierto** si $\mathrm{int}(E)=E$. Se dice que el conjunto E es **cerrado** si $E=\mathrm{int}(E)\cup\mathrm{Fr}(E)$. Puesto que $\mathrm{ext}(E)=\mathrm{int}(\mathbb{R}^n\setminus E)$ se verifica que E es cerrado si, y sólo si, su complemento $\mathbb{R}^n\setminus E$ es abierto.

Para todos $\mathbf{x}\in\mathbb{R}^n$ y $r>0$, se verifica que el conjunto $B(\mathbf{x},r)$ es abierto y se llama la **bola abierta de centro a y radio** r.

Dados $\mathbf{a}\in\mathbb{R}^n$ y $r\geqslant 0$, se verifica que el conjunto

$$\overline{B}(\mathbf{a},r)=\{\mathbf{x}\in\mathbb{R}^n : \|\mathbf{x}-\mathbf{a}\|\leqslant r\}$$

es cerrado. Dicho conjunto se llama **bola cerrada de centro a y radio** r. Para el caso en que $n=2$, las bolas cerradas suelen llamarse **discos** y se usa la notación $\overline{B}((a,b),r)=D((a,b),r)$.

Se dice que un conjunto $E\subset\mathbb{R}^n$ es **acotado** cuando hay un número $M>0$ tal que $\|\mathbf{x}\|\leqslant M$ para todo $\mathbf{x}\in E$. Se dice que un conjunto $K\subset\mathbb{R}^n$ es **compacto** cuando es cerrado y acotado.

1.3. Campos escalares

Un campo escalar de n variables es una función, $f : E \to \mathbb{R}$, definida en un subconjunto no vacío $E \subset \mathbb{R}^n$ que toma valores reales. La gráfica de dicho campo escalar es el subconjunto de R^{n+1}

$$G(f) = \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in E\} \subset \mathbb{R}^{n+1}$$

Para $n = 1$, dicha gráfica es una curva en $\mathbb{R}^2$, para $n = 2$ es una superficie en $\mathbb{R}^3$. En estos dos casos podemos visualizar la gráfica. Para campos escalares de tres o más variables su gráfica es una *hipersuperficie* en $\mathbb{R}^n$ con $n \geqslant 4$ que no se puede visualizar.

Las funciones siguientes:

a) $f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = xy + x^3y^2 + y^3.$

b) $f : \Omega_1 \to \mathbb{R}, \quad f(x,y) = \dfrac{x^2y + xy + xy^2}{x^3 + x^2y^2 + y^3}.$

c) $f : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}, \quad f(x,y) = \operatorname{sen}(\sqrt{x+y}).$

d) $f : \mathbb{R}^3 \to \mathbb{R}, \quad f(x,y,z) = xyz + z^2x^3y^2 + y^3.$

e) $f : \Omega_2 \to \mathbb{R}, \quad f(x,y,z) = \dfrac{x^2yz + z^2xy + xzy^2}{x^3z^2 + x^2y^2z + y^3}.$

f) $f : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}, \quad f(x,y,z) = \ln(x+y+z).$

son ejemplos de campos escalares. Las tres primeras son campos escalares de dos variables y las tres últimas de tres variables. Los campos escalares en a) y en d) vienen dados por funciones polinómicas en dos y tres variables respectivamente. Los campos escalares en b) y en e) vienen dados por funciones racionales de dos y tres variables respectivamente, y están definidos donde el denominador es distinto de cero, es decir, en los conjuntos $\Omega_1 = \left\{(x,y) \in \mathbb{R}^2 : x^3 + x^2y^2 + y^3 \neq 0\right\}$ y $\Omega_2 = \left\{(x,y,z) \in \mathbb{R}^3 : x^3z^2 + x^2y^2z + y^3 \neq 0\right\}$. Muchos campos escalares, como los definidos en c) y en d), se obtienen componiendo funciones elementales de una variable con funciones polinómicas o racionales de varias variables.

1.1 Definición. Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y sea $\mathbf{a} \in E$. Se dice que f es **continuo** en $\mathbf{a}$ si para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que se verifica $\| f(\mathbf{x}) - f(\mathbf{a}) \| < \varepsilon$ siempre que $\mathbf{x} \in E$ y $\| \mathbf{x} - \mathbf{a} \| < \varepsilon$.

Se dice que f es continuo en E si es continuo en todo punto de E.

1.2 Teorema (de Weierstrass). *Todo campo escalar continuo en un conjunto compacto alcanza en dicho conjunto un valor máximo absoluto y un valor mínimo absoluto.*

Dicho de otra forma, si $K \subset \mathbb{R}^n$ es un conjunto compacto y f es un campo escalar continuo en K, entonces hay puntos $\mathbf{a} \in K$, $\mathbf{b} \in K$ tales que $f(\mathbf{a}) \leqslant f(\mathbf{x}) \leqslant f(\mathbf{b})$ para todo $\mathbf{x} \in K$.

1.4. Curvas en $\mathbb{R}^n$

Una curva en $\mathbb{R}^n$ es una aplicación continua $\gamma : [a,b] \to \mathbb{R}^n$. El punto $\gamma(a)$ se llama *origen* y el punto $\gamma(b)$ *extremo* de la curva. Naturalmente, como $\gamma(t) \in \mathbb{R}^n$ podremos expresarlo por medio de sus componentes en la base canónica que serán funciones de t.

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t))$$

Las funciones $\gamma_k(t)$ se llaman funciones componentes de γ. Se dice que γ es derivable en un punto t cuando todas sus funciones componentes son derivables en dicho punto, en cuyo caso la derivada de γ en t es, por definición, el vector

$$\gamma'(t) = (\gamma_1'(t), \gamma_2'(t), \dots, \gamma_n'(t))$$

Dado un punto $\mathbf{a} = \gamma(t_0)$ tal que $\gamma'(t_0) \neq \mathbf{0}$, se define la **recta tangente** a γ en el punto $\mathbf{a}$ (aunque es más apropiado decir *en el punto* t_0) como la recta de ecuación paramétrica $\mathbf{x} = \mathbf{a} + t\,\gamma'(t_0)$, es decir, la recta que pasa por $\mathbf{a}$ con vector de dirección $\gamma'(t_0)$.

1.5. Derivadas parciales. Vector gradiente

1.3 Definición. Una **dirección** en $\mathbb{R}^n$ es un vector de norma 1.

- Dados un punto $\mathbf{a} \in \mathbb{R}^n$ y una dirección $\mathbf{u}$, la recta que pasa por $\mathbf{a}$ con dirección $\mathbf{u}$ es la imagen de la aplicación $\gamma : \mathbb{R}^n \to \mathbb{R}$ dada por $\gamma(t) = \mathbf{a} + t\mathbf{u}$, es decir, es el conjunto de puntos $\{\mathbf{a} + t\mathbf{u} : t \in \mathbb{R}\}$.
- Sea f un campo escalar definido en un conjunto abierto $E \subset \mathbb{R}^n$, sea $\mathbf{a} \in E$ y $\mathbf{u}$ una dirección. Se define la **derivada de f en a en la dirección u** como el límite

$$D_{\mathbf{u}} f(a) = \lim_{t \to 0} \frac{f(\mathbf{a} + t\,\mathbf{u}) - f(\mathbf{a})}{t} \tag{1}$$

- La derivada direccional de un campo escalar f en un punto $\mathbf{a}$ en la dirección del vector $\mathbf{e_k}$ de la base canónica, se llama **derivada parcial** de f en $\mathbf{a}$ respecto a la variable k-ésima. Está definida por

$$\begin{aligned} D_{\mathbf{e_k}} f(a) &= \lim_{t \to 0} \frac{f(\mathbf{a} + t\,\mathbf{e_k}) - f(\mathbf{a})}{t} = \lim_{t \to 0} \frac{f(a_1, \dots, a_k + t, \dots, a_n) - f(a_1, \dots, a_k, \dots, a_n)}{t} \\ &= \lim_{x_k \to a_k} \frac{f(a_1, \dots, x_k, \dots, a_n) - f(a_1, \dots, a_k, \dots, a_n)}{x_k - a_k} \end{aligned} \tag{2}$$

y se representa con los símbolos $D_k f(\mathbf{a})$ y $\dfrac{\partial f}{\partial x_k}(\mathbf{a})$.

Observa que las derivadas que acabamos de definir son derivadas de funciones reales de una variable real pues, para calcular la derivada de un campo escalar f en un punto $\mathbf{a}$ en la dirección $\mathbf{u}$ lo que se hace es derivar en $t = 0$ la función $t \mapsto f(\mathbf{a} + t\,\mathbf{u})$ que es una función real de una variable real.

Observa que la segunda igualdad de (2) nos dice que, *para calcular la derivada parcial $D_k f(\mathbf{a})$, lo que se hace es derivar f respecto a la variable k-ésima considerando fijas las demás variables*. Por eso se llaman derivadas *parciales*.

Por tanto, si derivamos un campo escalar respecto de cada una de sus variables considerando que las demás variables permanecen constantes obtenemos las derivadas parciales del mismo. Las derivadas parciales respecto a las variables x, y, z, de un campo escalar f de tres variables suelen representarse con los símbolos $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}, \dfrac{\partial f}{\partial z}$ y también $D_1 f$, $D_2 f$, $D_3 f$, respectivamente.

Por ejemplo, para el campo escalar $f(x,y,z) = xyz + z^2x^3y^2 + y^3$ tenemos que

- $\dfrac{\partial f}{\partial x}(x,y,z) = D_1 f(x,y,z) = yz + 3z^2x^2y^2.$
- $\dfrac{\partial f}{\partial y}(x,y,z) = D_2 f(x,y,z) = xz + 2z^2x^3y + 3y^2.$

- $\dfrac{\partial f}{\partial z}(x,y,z) = D_3 f(x,y,z) = xy + 2zx^3y^2$.

Naturalmente, si en estas igualdades se dan valores particulares a cada variable, digamos $x = a$, $y = b$, $z = c$, se obtienen los valores de las derivadas parciales en el punto (a,b,c).

1.6. Interpretación geométrica de las derivadas parciales

Para interpretar gráficamente el significado de las derivadas parciales, consideremos un campo escalar f de dos variables definido en un conjunto $E \subset \mathbb{R}^2$. Fijemos un punto $(a,b) \in \Omega$. La derivada parcial $D_1 f(a,b)$ es la derivada de la función $x \mapsto f(x,b)$ en el punto $x = a$, y la derivada parcial $D_2 f(a,b)$ es la derivada de la función $y \mapsto f(a,y)$ en el punto $y = b$.

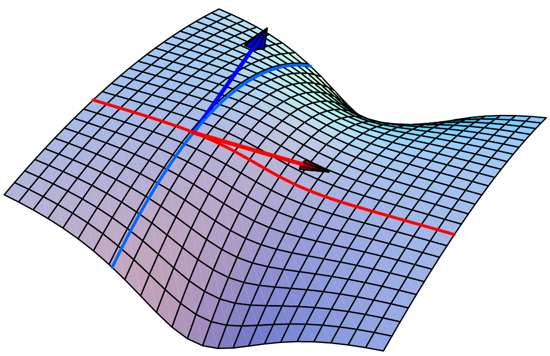
La gráfica de f, es decir, el conjunto $S = \{(x,y,f(x,y)) : (x,y) \in E\}$ es una superficie en $\mathbb{R}^3$. Las funciones

$$\gamma_1(x) = (x,b,f(x,b)), \quad \gamma_2(y) = (a,y,f(a,y))$$

son curvas contenidas en dicha superficie que pasan por el punto $(a,b,f(a,b))$. Dichas curvas se obtienen cortando la superficie S por los planos $y = b$ y $x = a$ respectivamente. Los vectores tangentes a dichas curvas en el punto $(a,b,f(a,b)) = \gamma_1(a) = \gamma_2(b)$ son, respectivamente

$$\gamma_1{}'(a) = (1,0,D_1 f(a,b)), \quad \gamma_2{}'(b) = (0,1,D_2 f(a,b))$$

En la siguiente figura se ha representado la gráfica de f y las curvas obtenidas cortándola por los planos $x = a$ e $y = b$ junto a sus vectores tangentes en el punto $(a,b,f(a,b))$.



1.4 Definición. Sea f un campo escalar. Se define el **vector gradiente** de f en un punto $\mathbf{a}$ como el vector

$$\nabla f(\mathbf{a}) = \big(D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \ldots, D_n f(\mathbf{a})\big)$$

Por ejemplo, el vector gradiente del campo escalar $f(x,y) = \operatorname{sen}(x^3y) + e^{y^3x}$ en un punto genérico (x,y) es el vector

$$\nabla f(x,y) = \big(3x^2y\cos(x^3y) + y^3 e^{y^3x},\ x^3\cos(x^3y) + 3y^2x e^{y^3x}\big).$$

1.5 Definición. Sea f un campo escalar con derivadas parciales continuas en un punto $\mathbf{a}$. El hiperplano en $\mathbb{R}^{n+1}$ de ecuación cartesiana

$$x_{n+1} = f(\mathbf{a}) + \left\langle \nabla f(\mathbf{a}) \middle| \mathbf{x} - \mathbf{a} \right\rangle \tag{3}$$

se llama hiperplano tangente a f en $\mathbf{a}$ o **hiperplano tangente** a la gráfica de f en el punto $(\mathbf{a}, f(\mathbf{a}))$.

En particular, la ecuación del plano tangente a la gráfica de un campo escalar de dos variables en un punto $(a,b,f(a,b))$ es:

$$z = f(a,b) + D_1 f(a,b)(x-a) + D_2 f(a,b)(y-b) \tag{4}$$

En lo que sigue consideraremos campos escalares que tienen derivadas parciales continuas.

1.6 Proposición. *Sean $f : E \to \mathbb{R}$ un campo escalar, $\mathbf{a} \in E$ y $\mathbf{u}$ una dirección en $\mathbb{R}^n$, se verifica que*

$$D_{\mathbf{u}} f(\mathbf{a}) = \left\langle \nabla f(\mathbf{a}) \middle| \mathbf{u} \right\rangle$$

1.7 Corolario. *Sea f un campo escalar con vector gradiente $\nabla f(\mathbf{a}) \neq \mathbf{0}$ en un punto $\mathbf{a}$.*

a) La dirección en la que la derivada direccional de f en a es máxima es la dirección dada por el gradiente, es decir, la dirección $\mathbf{u} = \dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$.

b) La dirección en la que la derivada direccional de f en a es mínima es la dirección opuesta a la dada por el gradiente, es decir, la dirección $\mathbf{v} = -\dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$.

1.7. Curvas y superficies de nivel. Rectas y planos tangentes

Recuerda que la ecuación de una recta en $\mathbb{R}^2$ es de la forma $ax + by = c$ donde a,b no son ambos nulos. Si dicha recta pasa por un punto (x_0,y_0) entonces $ax_0 + by_0 = c$ y la ecuación de la recta puede escribirse en la forma $a(x-x_0) + b(y-y_0) = 0$, es decir $\left\langle (a,b) \middle| (x-x_0, y-y_0) \right\rangle = 0$. El vector (a,b) es ortogonal a la recta.

Sea $f : A \to \mathbb{R}$ un campo escalar de dos variables. Dado un número $c \in f(A)$, el conjunto $\Gamma_c = \{(x,y) \in A : f(x,y) = c\}$ es una curva en el plano que se llama **curva de nivel** de f. Dicha curva es la proyección en el plano XY de la curva que se obtiene cortando la gráfica de f por el plano $z = c$. Se dice que dicha curva está **implícitamente definida** por la ecuación $f(x,y) - c = 0$. Observa que las curvas de nivel no se cortan. Las curvas de nivel son las que se representan en los mapas topográficos.

Se verifica que el vector gradiente de un campo escalar de dos variables, f, con derivadas parciales continuas es ortogonal en todo punto en el que no se anula a la tangente a la curva de nivel que pasa por dicho punto. En consecuencia, supuesto que $f(u,v) = c$ y que $\nabla f(u,v) \neq (0,0)$, la ecuación de la tangente a la curva de nivel Γ_c en (u,v) es $\left\langle \nabla f(u,v) \middle| (x-u, y-v) \right\rangle = 0$.

1.8 Ejemplos. Esta forma de calcular tangentes es muy sencilla y generaliza lo que ya sabes para el caso particular de la tangente a la gráfica $y = g(x)$ de una función derivable $g : I \to \mathbb{R}$. En efecto, definiendo $f(x,y) = y - g(x)$ para $(x,y) \in I \times \mathbb{R}$, tenemos que la curva de nivel

$$\Gamma_0 = \{(x,y) \in I \times \mathbb{R} : f(x,y) = 0\} = \{(x, g(x)) : x \in I\}$$

es la gráfica de g. Si $u \in I$ y $v = g(u)$, entonces $(u,v) \in \Gamma_0$ y la tangente a la gráfica de g en (u,v) viene dada por

$$0 = \left\langle \nabla f(u,v) \middle| (x-u, y-v) \right\rangle = \left\langle (-g'(u),1) \middle| (x-u, y-g(u)) \right\rangle = -g'(u)(x-u) + y - g(u) \Longleftrightarrow$$
$$y = g(u) + g'(u)(x-u)$$

que es la conocida ecuación de la recta tangente a la gráfica de g en el punto $(u,g(u))$.

Consideremos una elipse dada por

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

Dicha elipse es una curva de nivel del campo $f(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$; concretamente, es la curva de nivel dada por $f(x,y) = 1$. Sea (u,v) un punto de la elipse, es decir $f(u,v) = 1$. Tenemos que $\nabla f(u,v) = \left(\frac{2u}{a^2}, \frac{2v}{b^2}\right)$. Claramente $\nabla f(u,v) \neq (0,0)$. Por tanto el vector $\nabla f(u,v)$ es ortogonal a la tangente a la elipse en (u,v), por lo que la ecuación de dicha tangente es

$$0 = \left\langle \left(\frac{u}{a^2}, \frac{v}{b^2}\right) \middle| (x-u, y-v) \right\rangle = \frac{ux}{a^2} + \frac{vy}{b^2} - \frac{u^2}{a^2} - \frac{v^2}{b^2} = \frac{ux}{a^2} + \frac{vy}{b^2} - 1$$

Esto es, la ecuación de la tangente a la elipse en el punto (u,v) es

$$\frac{ux}{a^2} + \frac{vy}{b^2} = 1$$

La ecuación de un plano en $\mathbb{R}^3$ es de la forma $ax+by+cz = d$ donde a,b,c no son todos nulos. Si dicho plano pasa por (x_0,y_0,z_0) entonces $ax_0+by_0+cz_0 = d$ y la ecuación del plano puede escribirse en la forma $a(x-x_0)+b(y-y_0)+c(z-z_0) = 0$, es decir $\left\langle (a,b,c) \middle| (x-x_0, y-y_0, z-z_0) \right\rangle = 0$. El vector (a,b,c) es ortogonal al plano.

Sea $f : A \to \mathbb{R}$ un campo escalar de tres variables. Dado un número $c \in f(A)$, el conjunto

$$S_c = \{(x,y,z) \in A : f(x,y,z) = c\}$$

es una superficie en el espacio que se llama **superficie de nivel** de f. Se dice que dicha superficie está **implícitamente definida** por la ecuación $f(x,y,z) - c = 0$. Observa que las superficies de nivel no se cortan.

Se verifica que el vector gradiente de un campo escalar de tres variables, f, con derivadas parciales continuas es ortogonal en todo punto en el que no se anula al plano tangente a la superficie de nivel que pasa por dicho punto. En consecuencia, supuesto que $f(u,v,w) = c$ y que $\nabla f(u,v,w) \neq (0,0,0)$, la ecuación del plano tangente a la superficie de nivel S_c en (u,v,w) es $\left\langle \nabla f(u,v,w) \middle| (x-u, y-v, z-w) \right\rangle = 0$.

1.9 Ejemplos. Esta forma de calcular planos tangentes es muy sencilla y generaliza lo que ya sabemos para el caso particular del plano tangente a la gráfica $z = g(x,y)$ de un campo escalar de dos variables $g : A \to \mathbb{R}$. En efecto, definiendo $f(x,y,z) = z - g(x,y)$ para $(x,y,z) \in A \times \mathbb{R}$, tenemos que la superficie de nivel $S_0 = \{(x,y,z) \in A \times \mathbb{R} : f(x,y,z) = 0\} = \{(x,y,g(x)) : (x,y) \in A\}$ es la gráfica de g. Si $(u,v) \in A$ y $w = g(u,v)$, entonces $(u,v,w) \in S_0$ y la tangente a la gráfica de g en (u,v,w) viene dada por

$$0 = \left\langle \nabla f(u,v,w) \middle| (x-u, y-v, z-w) \right\rangle = \left\langle (-D_1g(u,v), -D_2g(u,v), 1) \middle| (x-u, y-v, z-g(u,v)) \right\rangle \Longleftrightarrow$$
$$0 = -D_1g(u,v)(x-u) - D_2g(u,v)(y-v) + z - g(u,v) \Longleftrightarrow$$
$$z = g(u,v) + D_1g(u,v)(x-u) + D_2g(u,v)(y-v)$$

que es la misma ecuación (4) del plano tangente a la gráfica de g en el punto $(u,v,g(u,v))$.

Consideremos un elipsoide dado por

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

Dicho elipsoide es una superficie de nivel del campo $f(x,y,z)=\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}$; concretamente, es la superficie de nivel dada por $f(x,y,z)=1$. Sea (u,v,w) un punto del elipsoide, es decir $f(u,v,w)=1$. Tenemos que $\nabla f(u,v,w)=\left(\frac{2u}{a^2},\frac{2v}{b^2},\frac{2w}{c^2}\right)$. Claramente $\nabla f(u,v,w)\neq(0,0,0)$. Por tanto el vector $\nabla f(u,v,w)$ es ortogonal al plano tangente al elipsoide en (u,v,w), por lo que la ecuación de dicho plano tangente es

$$0=\left\langle\left(\frac{u}{a^2},\frac{v}{b^2},\frac{w}{c^2}\right)\Big|(x-u,y-v,z-w)\right\rangle=\frac{ux}{a^2}+\frac{vy}{b^2}+\frac{wz}{c^2}-\frac{u^2}{a^2}-\frac{v^2}{b^2}-\frac{w^2}{c^2}=\frac{ux}{a^2}+\frac{vy}{b^2}+\frac{wz}{c^2}-1$$

Esto es, la ecuación del plano tangente al elipsoide en el punto (u,v,w) es

$$\frac{ux}{a^2}+\frac{vy}{b^2}+\frac{wz}{c^2}=1$$

Cuando una curva Γ en $\mathbb{R}^3$ viene dada como intersección de dos superficies S_1 y S_2, la tangente en un punto $(a,b,c)\in\Gamma$ a la curva Γ es la recta intersección de los planos tangentes a las superficies en dicho punto. Por ejemplo, si las superficies vienen dadas por sus ecuaciones implícitas.

$$\begin{cases} S_1=\left\{(x,y,z)\in\mathbb{R}^3 : f(x,y,z)=0\right\} \\ S_2=\left\{(x,y,z)\in\mathbb{R}^3 : g(x,y,z)=0\right\}\end{cases} \qquad \Gamma=S_1\cap S_2=\left\{(x,y,z)\in\mathbb{R}^3 : g(x,y,z)=f(x,y,z)=0\right\}$$

Entonces, las ecuaciones implícitas de la recta tangente a Γ en un punto $(a,b,c)\in\Gamma$ son

$$\begin{cases} \left\langle\nabla f(a,b,c)\big|(x-a,y-b,z-c)\right\rangle=0 \\ \left\langle\nabla g(a,b,c)\big|(x-a,y-b,z-c)\right\rangle=0\end{cases}$$

Donde se supone que los vectores gradiente $\nabla f(a,b,c)$, $\nabla g(a,b,c)$ son linealmente independientes pues, en otro caso, la recta tangente a la curva Γ en (a,b,c) no está definida.

1.8. Derivadas parciales de orden superior

Supongamos un campo escalar f que tiene derivadas parciales $D_k f$ en un conjunto $E\subset\mathbb{R}^n$. Las funciones $D_k f$ son también campos escalares que podemos, cuando se dejen, volver a derivar parcialmente en puntos de E. Obtenemos de esta forma las *derivadas parciales de segundo orden* de f, es decir las funciones $D_j(D_k f)$, que se representan simbólicamente de las formas

$$D_{jk}f(\mathbf{x}),\qquad \frac{\partial^2 f}{\partial x_j\,\partial x_k}(\mathbf{x}),\qquad \frac{\partial^2 f}{\partial x_k^2}(\mathbf{x})$$

De forma análoga se definen las derivadas parciales de tercer orden de f como las derivadas parciales de las derivadas parciales de segundo orden de f y se representan por

$$D_{jkm}f(\mathbf{x}),\qquad \frac{\partial^3 f}{\partial x_j\,\partial x_k\,\partial x_m}(\mathbf{x});\qquad \frac{\partial^3 f}{\partial x_k^3}(\mathbf{x});\qquad \frac{\partial^3 f}{\partial x_k^2\partial x_j}(\mathbf{x})$$

Es natural preguntarse si el orden en que se realizan las derivadas debe ser o no tenido en cuenta. Afortunadamente, en la mayoría de los casos podemos olvidarlo porque se verifica el siguiente resultado.

1.10 Definición. Se dice que un campo escalar f es de clase C^k en un abierto $E \subset \mathbb{R}^n$ si f tiene derivadas parciales de orden k continuas en E.

1.11 Teorema. *Las derivadas parciales de orden menor o igual que k de un campo escalar de clase C^k solamente dependen del número de veces que se deriva parcialmente respecto de cada variable, pero el orden en que se realicen dichas derivaciones no afecta para nada al resultado final.*

1.9. Extremos relativos

1.12 Definición. Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$. Se dice que f tiene un **máximo relativo** (resp. **mínimo relativo**) en un punto $\mathbf{a} \in E$, si $\mathbf{a}$ es un punto interior de E y existe un número $r > 0$ tal que $B(\mathbf{a}, r) \subset E$ y $f(\mathbf{x}) \leqslant f(\mathbf{a})$ (resp. $f(\mathbf{a}) \leqslant f(\mathbf{x})$) para todo $\mathbf{x} \in B(\mathbf{a}, r)$. Cuando estas desigualdades se verifican de forma estricta se dice que el máximo o el mínimo relativo es estricto.

Los puntos en los que f tiene un máximo o un mínimo relativos se llaman **extremos relativos** de f.

1.13 Proposición (Condición necesaria de extremo relativo). *Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y supongamos que f tiene un extremo relativo en un punto $\mathbf{a} \in E$ y además que el vector gradiente de f en $\mathbf{a}$ está definido. Entonces se verifica que $\nabla f(\mathbf{a}) = \mathbf{0}$. Es decir, las derivadas parciales de primer orden de f en $\mathbf{a}$ son todas nulas.*

1.14 Definición. Los puntos donde se anula el gradiente de un campo escalar f se llaman **puntos críticos** de f. Los puntos críticos de un campo escalar que no son extremos relativos se llaman **puntos de silla**.

Si f es un campo escalar diferenciable, en los puntos críticos el hiperplano tangente es "horizontal".

La condición necesaria de extremo relativo no es suficiente. Por ejemplo, el campo escalar $f(x,y) = x^2 - y^2$ tiene un punto crítico en $(0,0)$, pero no tiene extremo relativo en dicho punto pues en toda bola centrada en $(0,0)$ toma valores positivos y negativos.

Al igual que para funciones de una variable, la derivada segunda proporciona una condición suficiente de extremo relativo, para campos escalares de varias variables las derivadas parciales de segundo orden nos van a permitir dar una condición suficiente de extremo relativo.

1.15 Definición. Sea f un campo escalar de n variables que tiene derivadas parciales de segundo orden continuas en un punto $\mathbf{a}$. La matriz $n \times n$

$$H(f, \mathbf{a}) = \big(D_{ij} f(\mathbf{a})\big)_{1 \leqslant i,j \leqslant n}$$

se llama **matriz hessiana** de f en $\mathbf{a}$.

Observa que la matriz hessiana es simétrica porque $D_{ij} f(\mathbf{a}) = D_{ji} f(\mathbf{a})$. Dicha matriz define una forma cuadrática, que representaremos por $Q(f, \mathbf{a})$, que viene dada para todo $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ por

$$Q(f, \mathbf{a})(\mathbf{x}) = \mathbf{x}.H(f, \mathbf{a}).\mathbf{x}^t = \sum_{j=1}^{n} \sum_{k=1}^{n} D_{jk} f(\mathbf{a}) x_k x_j$$

donde el punto "**.**" indica producto matricial y $\mathbf{x}^t$ es el vector columna $\mathbf{x}$.

1.16 Definición. Una forma cuadrática $Q(\mathbf{x}) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j$ se llama:

- **Definida positiva** si $Q(\mathbf{x}) > 0$ para todo $\mathbf{x} \in \mathbb{R}^n$ con $\mathbf{x} \neq \mathbf{0}$.
- **Definida negativa** si $Q(\mathbf{x}) < 0$ para todo $\mathbf{x} \in \mathbb{R}^n$ con $\mathbf{x} \neq \mathbf{0}$.
- **No definida o indefinida** si hay vectores $\mathbf{x}$ para los que $Q(\mathbf{x}) > 0$ y hay vectores $\mathbf{x}$ para los que $Q(\mathbf{x}) < 0$.

1.17 Teorema. *Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y supongamos que f tiene derivadas parciales de segundo orden continuas en un punto $\mathbf{a}$ **interior** de E que además **es un punto crítico** de f. Sea $Q(f,\mathbf{a})$ la forma cuadrática asociada a la matriz hessiana de f en $\mathbf{a}$.*

$$Q(f,\mathbf{a})(\mathbf{x}) = \mathbf{x}.H(f,\mathbf{a}).\mathbf{x}^t = \sum_{j=1}^{n}\sum_{k=1}^{n} D_{jk} f(\mathbf{a}) x_k x_j$$

a) Si la forma cuadrática $Q(f,\mathbf{a})$ es definida positiva entonces f tiene en $\mathbf{a}$ un mínimo relativo estricto.

b) Si la forma cuadrática $Q(f,\mathbf{a})$ es definida negativa entonces f tiene en $\mathbf{a}$ un máximo relativo estricto.

c) Si la forma cuadrática $Q(f,\mathbf{a})$ es no definida entonces f tiene un punto de silla en $\mathbf{a}$.

Para poder usar el resultado anterior hay que saber clasificar una forma cuadrática. Hay varios procedimientos sencillos para ello. Los dos que siguen a continuación son los que me parecen más cómodos.

Clasificación de formas cuadráticas

Sean $\mathcal{A} = \left(a_{ij}\right)_{1 \leqslant i,j \leqslant n}$ una matriz simétrica de números reales y

$$Q_{\mathcal{A}}(\mathbf{x}) = \mathbf{x}.\mathcal{A}.\mathbf{x}^t = \sum_{i,j=1}^{n} a_{ij} x^i x^j \tag{5}$$

la forma cuadrática definida por $\mathcal{A}$. Los *valores propios* de $\mathcal{A}$ son las raíces del polinomio característico $p(\lambda)$, que se define como el determinante de la matriz $\mathcal{A} - \lambda I$:

$$p(\lambda) = \left|\mathcal{A} - \lambda I\right|$$

Se verifica que, en la situación que estamos considerando, las raíces de dicho polinomio son todas reales.

- La forma cuadrática $Q_{\mathcal{A}}$ es definida positiva si, y sólo si, todos los valores propios de $\mathcal{A}$ son positivos.
- La forma cuadrática $Q_{\mathcal{A}}$ es definida negativa si, y sólo si, todos los valores propios de $\mathcal{A}$ son negativos.
- La cuadrática $Q_{\mathcal{A}}$ es no definida si, y sólo si, $\mathcal{A}$ tiene valores propios positivos y negativos.

Otro criterio para estudiar el carácter de la forma cuadrática (5) se basa en la sucesión de signos de los *menores principales* de la matriz $\mathcal{A}$. El menor principal de orden k de la matriz $\mathcal{A}$ es el determinante $\Delta_k = \left|a_{ij}\right|_{1 \leqslant i,j \leqslant k}$ de la matriz formada por las primeras k filas y k columnas de la matriz $\mathcal{A}$. Se verifica que:

- *La forma cuadrática es definida positiva si, y sólo si, todos los menores principales son positivos* .

- *La forma cuadrática es definida negativa si, y sólo si, los menores principales de orden par son positivos y los menores principales de orden impar son negativos.*

- *Si los menores principales son nulos a partir de uno de ellos en adelante y los no nulos son positivos o van alternando signo siendo el primero de ellos negativo, no puede afirmarse nada.*

- *En los demás casos la forma cuadrática es no definida.*

Observa que cuando la dimensión n es par, si el determinante de la matriz $\mathcal{A}$ es negativo entonces la forma cuadrática es no definida.

Podemos particularizar este criterio para el caso de dos dimensiones.

Sea $A \subset \mathbb{R}^2$ un conjunto abierto y sea f un campo escalar definido en A que tiene derivadas parciales de segundo orden continuas. Supongamos que $(a,b) \in A$ es un punto crítico de f y sea

$$H(f,(a,b)) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a,b) & \dfrac{\partial^2 f}{\partial x \partial y}(a,b) \\ \dfrac{\partial^2 f}{\partial x \partial y}(a,b) & \dfrac{\partial^2 f}{\partial y^2}(a,b) \end{pmatrix}$$

la matriz hessiana de f en (a,b) y notemos $\det H(f,(a,b)$ su determinante.

- Si $\det H(f,(a,b)) > 0$ y $\dfrac{\partial^2 f}{\partial x^2}(a,b) > 0$ entonces f tiene en (a,b) un mínimo relativo estricto.
- Si $\det H(f,(a,b)) > 0$ y $\dfrac{\partial^2 f}{\partial x^2}(a,b) < 0$ entonces f tiene en (a,b) un máximo relativo estricto.
- Si $\det H(f,(a,b)) < 0$ entonces f no tiene extremo relativo en (a,b). Se dice que (a,b) es un punto de silla de f.
- Cuando $\det H(f,(a,b)) = 0$ el conocimiento de la matriz hessiana no permite decidir si hay o no hay extremo relativo en (a,b). Cuando esto sucede puede ser interesante estudiar el comportamiento de las curvas $f(a,t+b)$ y $f(a+t,b)$. Si alguna de dichas curvas no tiene extremo relativo o tienen extremos relativos de distinta naturaleza en $t=0$, podemos concluir que en (a,b) no hay extremo relativo de f.

1.10. Cálculo de extremos absolutos en conjuntos compactos

Como consecuencia del teorema de Weierstrass, y de que todo campo escalar con derivadas parciales continuas es continuo, se verifica que todo campo escalar con derivadas parciales continuas en un conjunto compacto $K \subset \mathbb{R}^2$ alcanza en dicho conjunto un valor máximo absoluto y un valor mínimo absoluto. Dichos valores o bien se alcanzan en el interior de K, en cuyo caso deben ser puntos críticos de f, o bien se alcanzan en la frontera. Cuando la frontera de K está formada por curvas conocidas es fácil calcular los puntos de la frontera en los que el campo puede alcanzar sus extremos absolutos. Una vez calculados todos estos puntos, se evalúa en ellos el campo para saber en cuales se alcanzan los extremos absolutos. El siguiente ejemplo indica la forma de proceder.

1.18 Ejemplo. Calcula los extremos absolutos de la función $f(x,y) = 4x^2 + y^2 - 4x - 3y$ en el conjunto

$$M = \left\{(x,y) \in \mathbb{R}^2 : y \geqslant 0,\ 4x^2 + y^2 \leqslant 4\right\}$$

Solución. El conjunto $M = \left\{(x,y) \in \mathbb{R}^2 : y \geqslant 0,\ x^2 + \dfrac{y^2}{4} \leqslant 1\right\}$ es la parte de la elipse centrada en el origen de semiejes 1 y 2 que queda en el semiplano superior. Se trata de un conjunto compacto (cerrado porque incluye a su frontera y acotado) y, como la función f es continua, el teorema de Weierstrass asegura que f alcanza un máximo y un mínimo absolutos en M. Dichos extremos deben alcanzarse o bien en el interior de M o en la frontera de M. Los puntos del interior de M que sean extremos relativos de f tienen que ser puntos críticos de f, es decir, deben ser puntos de M donde se anule el gradiente de f.

$$\frac{\partial f}{\partial x}(x,y)=8x-4=0, \qquad \frac{\partial f}{\partial y}(x,y)=2y-3=0$$

Por tanto, f tiene un único punto crítico que es $(1/2,3/2)$ el cual, efectivamente, está en M.

Los extremos absolutos de f en M pueden alcanzarse en la frontera de M. La frontera de M está formada por la parte superior de la elipse $4x^2+y^2=4$ y por el segmento $\{(x,0),-1\leqslant x\leqslant 1\}$.

Los extremos de f en el segmento son fáciles de calcular pues son los extremos de la función $h(x)=f(x,0)=4x^2-4x$ donde $x\in[-1,1]$. Como $h'(x)=8x-4$, los únicos posibles valores extremos de f en el segmento son $h(1/2)=f(1/2,0)$, $h(-1)=f(-1,0)$ y $h(1)=f(1,0)$.

Finalmente, calculemos los valores extremos de f en la parte superior de la elipse $4x^2+y^2=4$. Nuestro problema, pues, es calcular los extremos de f cuando f toma valores en el conjunto $E=\left\{(x,y)\in\mathbb{R}^2 : 4x^2+y^2=4, y\geqslant 0\right\}$.

Para $(x,y)\in E$ se tiene que

$$f(x,y)=4x^2+y^2-4x-3y=4-4x-3y=4-4x-3\sqrt{4-4x^2}$$

Por tanto, los valores de la restricción de f a E están dados por la función $g(x)=4-4x-3\sqrt{4-4x^2}$ donde $-1\leqslant x\leqslant 1$. La derivada de g se anula en un único punto $x_0=2/\sqrt{13}$ que está en $]-1,1[$. Por tanto, los extremos de f en E han de alcanzarse en alguno de los puntos $(-1,0)$, $\left(2/\sqrt{13},6/\sqrt{13}\right)$, $(1,0)$. Los extremos absolutos de f en M han de alcanzarse en alguno de dichos puntos o en los puntos $(1/2,0)$, $(1/2,3/2)$. Tenemos que

$$f(1,0)=0,\quad f(-1,0)=8,\quad f\left(2/\sqrt{13},6/\sqrt{13}\right)=4-2\sqrt{13},\quad f(1/2,3/2)=-\frac{13}{4},\quad f(1/2,0)=-1$$

El valor máximo absoluto de f en M es igual a 8 y se alcanza en el punto $(-1,0)$. El valor mínimo absoluto de f en M es $=-\dfrac{13}{4}$ y se alcanza en el punto $(1/2,3/2)$.

1.11. Integrales dobles

Sea $f:A\to\mathbb{R}$ un campo escalar de dos variables definido en un rectángulo $A=[a,b]\times[c,d]$. Sean

$$P=\{a=x_0,x_1,x_2,\dots,x_{p-1},x_p=b\}, \qquad Q=\left\{c=y_0,y_1,y_2,\dots,y_{q-1},y_q=d\right\}$$

particiones de los intervalos $[a,b]$ y $[c,d]$ respectivamente. Dichas particiones determinan una partición, que notamos $P\times Q$, del rectángulo $A=[a,b]\times[c,d]$ en subrectángulos $[x_{i-1},x_i]\times[y_{j-1},y_j]$, donde $1\leqslant i\leqslant p$, $1\leqslant j\leqslant q$.

Una **suma de Riemann** de f para la partición $P\times Q$ es un número que se obtiene eligiendo puntos $(s_i,t_j)\in[x_{i-1},x_i]\times[y_{j-1},y_j]$ y calculando la suma

$$\sum_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}} f(s_i,t_j)(x_i-x_{i-1})(y_j-y_{j-1}) \tag{6}$$

Se verifica que cuando la mayor de las longitudes de los intervalos de las particiones P y Q se hace arbitrariamente pequeña (o sea, tiende a 0), las sumas de Riemann de f se aproximan tanto como se quiera a un número real que es, por definición, la integral de Riemann de f en el rectángulo $[a,b]\times[c,d]$, que se representa por

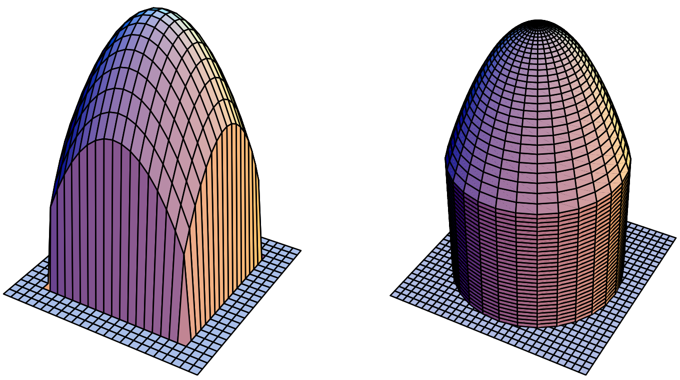
$$\iint\limits_{[a,b]\times[c,d]} f(x,y)\,\mathrm{d}(x,y)$$

Interpretaciones de las integrales dobles

Sea $f : A \to \mathbb{R}$ un campo escalar de dos variables definido en un conjunto $A \subset \mathbb{R}^2$. Supongamos que $f(x,y) \geqslant 0$ para todo $(x,y)\in A$. Consideremos el "cilindro" en $\mathbb{R}^3$ que tiene como base el conjunto A y como tapadera la gráfica de f, es decir el conjunto

$$C(f,A) = \left\{(x,y,z)\in\mathbb{R}^3 : (x,y)\in A,\ 0 \leqslant z \leqslant f(x,y)\right\}.$$

Las siguientes figuras muestran este conjunto para la función $f(x,y) = 4-x^2-y^2$ y los conjuntos $A = [-1,1]\times[-1,1]$ y $A = \left\{(x,y) : x^2+y^2 \leqslant 2\right\}$. En esta situación, una suma de Riemann del tipo

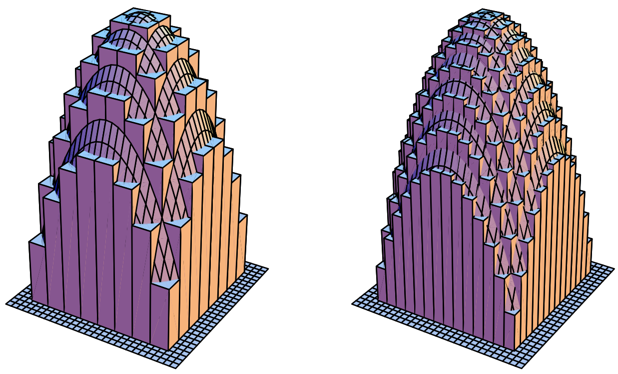


(6) representa una aproximación del volumen del conjunto $C(f,A)$. Pues lo que hacemos en (6) es sumar los volúmenes de pequeños ortoedros de base los rectángulos $R_{ij} = [x_{i-1},x_i]\times[y_{j-1},y_j]$ y altura $f(s_i,t_j)$. Es claro que la suma de todos estos volúmenes es una aproximación del volumen del conjunto $C(f,A)$. La aproximación es tanto mejor cuanto más pequeños sean los lados de los rectángulos R_{ij} y, en el límite, el volumen del conjunto $C(f,A)$ viene dado por la integral doble de f en A.

$$\iint\limits_{A} f(x,y)\,\mathrm{d}(x,y) = \mathrm{volumen}(C(f,A)) \tag{7}$$

Las siguientes figuras muestran aproximaciones al volumen del primero de los dos conjuntos representados en la figura anterior.

Naturalmente, pueden darse otras muchas interpretaciones. Por ejemplo, la función f puede representar una densidad superficial de masa o de carga eléctrica en una lámina plana A. En tal caso la integral doble $\iint\limits_{A} f(x,y)\,\mathrm{d}(x,y)$ proporciona, respectivamente, la masa o la carga total de la lámina A.



Cálculo de integrales dobles

El cálculo de una integral doble se reduce al cálculo de dos integrales simples gracias al siguiente resultado:

$$\iint\limits_{[a,b]\times[c,d]} f(x,y)\,\mathrm{d}(x,y) = \int_a^b \left[\int_c^d f(x,y)\,\mathrm{d}y\right]\mathrm{d}x = \int_c^d \left[\int_a^b f(x,y)\,\mathrm{d}x\right]\mathrm{d}y \tag{8}$$

Las integrales que figuran a la derecha en esta igualdad se llaman *integrales iteradas*. Observa que las integrales iteradas son dos integrales simples. Para calcular $\int_c^d f(x,y)\,\mathrm{d}y$ lo que se hace es integrar respecto a la variable y considerando x fija. Para ello lo que se hace es obtener una primitiva de la función $y \mapsto f(x,y)$ y usar la regla de Barrow. Análogo procedimiento se sigue para calcular la otra integral iterada.

Con frecuencia el campo escalar está definido en un conjunto A de tipo I o de tipo II. Esto es

$$\begin{aligned} A &= \{(x,y) : a \leqslant x \leqslant b,\ g(x) \leqslant y \leqslant h(x)\} \qquad \text{(tipo I)} \\ A &= \{(x,y) : c \leqslant y \leqslant d,\ \varphi(y) \leqslant x \leqslant \psi(y)\} \qquad \text{(tipo II)} \end{aligned}$$

En tales casos tenemos que

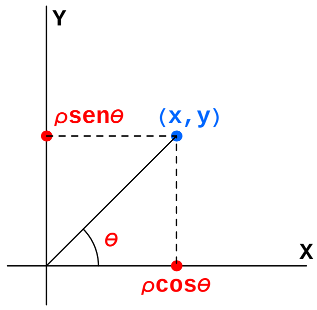
$$\iint\limits_A f(x,y)\,\mathrm{d}(x,y) = \int_a^b \left[\int_{g(x)}^{h(x)} f(x,y)\,\mathrm{d}y\right]\mathrm{d}x \tag{9}$$

$$\iint\limits_A f(x,y)\,\mathrm{d}(x,y) = \int_c^d \left[\int_{\varphi(y)}^{\psi(y)} f(x,y)\,\mathrm{d}x\right]\mathrm{d}y \tag{10}$$

Cambio a coordenadas polares

Coordenadas polares

El par de números (ρ,ϑ) dados por $x=\rho\cos\vartheta$, $y=\rho\operatorname{sen}\vartheta$ donde $\rho>0$ y $-\pi<\vartheta\leqslant\pi$ se llaman coordenadas polares del punto de coordenadas cartesianas (x,y).



Se verifica que

$$\iint\limits_A f(x,y)\,\mathrm{d}(x,y)=\iint\limits_B f(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\rho\,\mathrm{d}(\rho,\vartheta) \tag{11}$$

Para aplicar esta fórmula hay que determinar el conjunto B. Dicho conjunto viene dado por

$$B=\left\{(\rho,\vartheta)\in\mathbb{R}^+\times]-\pi,\pi]:(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\in A\right\}$$

Si, por ejemplo, el conjunto A es de tipo I, $A=\{(x,y):a\leqslant x\leqslant b,\ g(x)\leqslant y\leqslant h(x)\}$, entonces $B=\{(\rho,\vartheta)\in\mathbb{R}^+\times]-\pi,\pi]:a\leqslant\rho\cos\vartheta\leqslant b,\ g(\rho\cos\vartheta)\leqslant\rho\operatorname{sen}\vartheta\leqslant h(\rho\cos\vartheta)\}$. Es importante describir bien el conjunto B porque para calcular la integral $\iint\limits_B f(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\rho\,\mathrm{d}(\rho,\vartheta)$ tienes que calcular las integrales iteradas. Si, por ejemplo, $B=\{(\rho,\vartheta):\alpha\leqslant\vartheta\leqslant\beta,\ g(\vartheta)\leqslant\rho\leqslant h(\vartheta)\}$, entonces

$$\iint\limits_A f(x,y)\,\mathrm{d}(x,y)=\iint\limits_B f(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\rho\,\mathrm{d}(\rho,\vartheta)=\int_\alpha^\beta\left[\int_{g(\vartheta)}^{h(\vartheta)} f(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\rho\,\mathrm{d}\vartheta\right]\mathrm{d}\rho \tag{12}$$

Las coordenadas polares son especialmente útiles cuando el conjunto A es un círculo, o un sector circular o una corona circular, pues en estos casos el conjunto B es muy sencillo. Si, por ejemplo, A es el disco $D((0,0),R)$ de centro el origen y radio R, $D((0,0),R)=\left\{(x,y):x^2+y^2\leqslant R^2\right\}$, entonces

$$B=\left\{(\rho,\vartheta)\in\mathbb{R}^+\times]-\pi,\pi]:\rho\leqslant R\right\}=]0,R]\times]-\pi,\pi]$$

Por tanto

$$\iint\limits_{D((0,0),R)} f(x,y)\,\mathrm{d}(x,y)=\int_0^R\left[\int_{-\pi}^{\pi} f(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\rho\,\mathrm{d}\vartheta\right]\mathrm{d}\rho=\int_{-\pi}^{\pi}\left[\int_0^R f(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)\rho\,\mathrm{d}\rho\right]\mathrm{d}\vartheta \tag{13}$$