Funciones de varias variables

7 de febrero de 2008

1. Definiciones básicas

Sean a, b puntos de $\mathbb{R}^n$ (donde $n \in \mathbb{N}$) con coordenadas:

$$a = (a_1, a_2, \cdots, a_n);\ b = (b_1, b_2, \cdots, b_n)$$

Se define la **distancia euclídea** entre a y b como

$$\mathrm{d} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\mathrm{d}(a, b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

y, siendo $\mathbf{x} \in \mathbb{R}^n; \mathbf{x} = (x_1, x_2, \cdots, x_n)$, la **norma euclídea** como

$$||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Se define la **bola abierta** de centro $a \in \mathbb{R}^n$ y radio $R \in \mathbb{R}^+$ (donde $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$):

$$\mathrm{B}(a, R) = \{x \in \mathbb{R}^n : \mathrm{d}(a, x) < R\}$$

Un conjunto $A \subset \mathbb{R}^n$ es un **entorno** del punto $a \in \mathbb{R}^n$ si

$$\exists r \in \mathbb{R}^+\ \ \mathrm{B}(a, r) \subset A$$

Se llama **frontera** del conjunto $A \subset \mathbb{R}^n$ al conjunto de puntos de $\mathbb{R}^n$ tales que sus entornos contengan puntos de A y su complementario ($\bar{A}$).

Se dice que el conjunto $A \subset \mathbb{R}^n$ es **abierto** si en cada uno de sus puntos se puede trazar una bola totalmente contenida en el conjunto, es decir:

$$\text{Abierto}(A) \Leftrightarrow \forall a \in A\ \ \exists r \in \mathbb{R}^+\ \ \mathrm{B}(a, r) \subset A$$

$$\text{Cerrado}(A) \Leftrightarrow \text{Abierto}(\bar{A})$$

El punto $x \in \mathbb{R}^n$ es un **punto de acumulación** del conjunto $A \subset \mathbb{R}^n$ si x está *extremadamente cerca* del conjunto A sin – necesariamente – pertenecer a él; formalmente:

$$\forall r \in \mathbb{R}^+ \quad [\mathrm{B}(x,r) - \{x\}] \cap A \neq \emptyset$$

Un conjunto $A \subset \mathbb{R}^n$ es **acotado** si:

$$\exists r \in \mathbb{R}^+, a \in \mathbb{R}^n \;\; \mathrm{B}(a,r) \supset A$$

Si además de acotado es cerrado, entonces es **compacto**.

2. Funciones escalares

Una función escalar es de la forma:

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}; n \in \mathbb{N}$$

2.1. Conjunto y contorno de nivel

Se define el **conjunto de nivel** para k, donde $k \in \mathbb{R}$, como:

$$\mathrm{C}_k(f) = \{x \in \mathrm{Dom} f : f(x) = k\}$$

Un **contorno de nivel** para k – en el caso particular de $n = 2$ también llamado *curva de nivel* – delimita el conjunto de nivel para k de f. Los contornos de nivel permiten una visualización gráfica de la variación de f.

2.2. Cálculo del límite en $\mathbb{R}^2$

En general, se define el **límite** de una función escalar de la siguiente forma:

$$\lim_{x \to a} f(x) = L \Leftrightarrow \forall \epsilon > 0 \;\; \exists \delta > 0 : \forall x \in \mathrm{B}(a,\delta) \;\; |f(x) - L| < \epsilon$$

En el caso particular de dominios en $\mathbb{R}^2$, se pueden utilizar los siguientes métodos:

2.2.1. Límites iterados

Los límites iterados permiten descartar la existencia del límite. Sea $(x, y) \in \mathbb{R}^2; (a_1, a_2) \in \mathbb{R}^2$ y $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ una función escalar.

$$\lim_{x \to a_1} \left[\lim_{y \to a_2} f(x, y) \right] = A$$

$$\lim_{y \to a_2} \left[\lim_{x \to a_1} f(x, y) \right] = B$$

$$\exists \lim_{(x,y) \to (a_1, a_2)} f(x, y) \Longrightarrow \lim_{(x,y) \to (a_1, a_2)} f(x, y) = A = B$$

de lo que se deduce que:

$$A \neq B \Longrightarrow \nexists \lim_{(x,y) \to (a_1, a_2)} f(x, y)$$

2.2.2. Límites con coordenadas polares

Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Se puede convertir $x \in \mathbb{R}^2; x = (x_1, x_2)$ a coordenadas polares (ρ, θ) de forma que $x = \rho \cos \theta$ y $y = \rho \sin \theta$. Así:

$$\lim_{\rho \to 0} f(\rho \cos \theta, \rho \sin \theta) = \lim_{(x,y) \to (0,0)} f(x, y)$$

2.2.3. Acercamiento por rectas

Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Si se quiere calcular

$$\lim_{(x,y) \to (a_1, a_2)} f(x, y)$$

se puede utilizar un acercamiento mediante una recta $y = mx + c$ de forma que $a_2 = ma_1 + c$. Así:

$$\lim_{x \to a_1} f(x, mx + c) = A$$

$$\lim_{y \to a_2} f\left(\frac{y - c}{m}, y \right) = B$$

$$\exists \lim_{(x,y) \to (a_1, a_2)} f(x, y) \Longrightarrow \lim_{(x,y) \to (a_1, a_2)} f(x, y) = A = B$$

2.3. Continuidad

Sea $A \subset \mathbb{R}^n;\ n \in \mathbb{N}$. $f : A \longrightarrow \mathbb{R}$ es contínua en $a \in A$ si

$$\lim_{x \to a} f(x) = f(a)$$

2.4. Derivadas parciales

Sea $n \in \mathbb{N}$ y $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Se denomina **derivada parcial** de f con respecto a una componente j $(1 \leq j \leq n)$ a la función:

$$\forall x \in \mathbb{R}^n \ \frac{\partial f}{\partial x_j} : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\frac{\partial f}{\partial x_j}(a) = \lim_{x_j \to a_j} \frac{f(a_1, \cdots, x_j, \cdots, a_n) - f(a_1, \cdots, a_j, \cdots, a_n)}{x_j - a_j}$$

Como la derivada parcial es una nueva función escalar, es posible obtener la **derivada parcial segunda** – o, en general, la derivada parcial enésima – :

$$\frac{\partial f^2}{\partial x_i \partial x_j}(a) = \lim_{x_i \to a_i} \frac{\frac{\partial f}{\partial x_j}(a_1, \cdots, x_i, \cdots, a_n) - \frac{\partial f}{\partial x_j}(a_1, \cdots, a_i, \cdots, a_n)}{x_i - a_i}$$

2.4.1. Teorema de Schwarz

Sea $n \in \mathbb{N}$ y $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que $\exists \frac{\partial f^2}{\partial x_i \partial x_j}$. Si $\frac{\partial f^2}{\partial x_i \partial x_j}$ es contínua en a, entonces $\exists \frac{\partial f^2}{\partial x_j \partial x_i}$ y:

$$\frac{\partial f^2}{\partial x_i \partial x_j} = \frac{\partial f^2}{\partial x_j \partial x_i}$$

2.5. Derivadas direccionales

Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ y $\mathbf{v} \in \mathbb{R}^n$ un vector unitario ($||\mathbf{v}|| = 1$). Se define la **derivada direccional** de f según el vector $\mathbf{v}$ como

$$\mathrm{D}_{\mathbf{v}} f(a) = \lim_{h \to 0} \frac{f(a + h\mathbf{v}) - f(a)}{h}$$

La derivada parcial es un caso particular de derivada direccional:

$$\mathbf{v}_i = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \implies \frac{\partial f}{\partial x_j} = \mathrm{D}_{\mathbf{v}} f$$

2.6. Vector gradiente

Se llama **vector gradiente** de $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ en a al vector:

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

El vector gradiente indica la *dirección de máximo crecimiento* a partir del punto a.

Además, el vector gradiente es normal a un contorno de nivel en un punto de éste.

3. Funciones vectoriales y escalares

Una función vectorial es de la forma:

$$f : \mathbb{R}^m \longrightarrow \mathbb{R}^n; \;\; n, m \in \mathbb{N}$$

Una función vectorial se puede descomponer en n funciones escalares $f_i : \mathbb{R}^m \longrightarrow \mathbb{R}$, de forma que

$$f(x) = (f_1(x), f_2(x), \cdots, f_n(x))$$

3.1. Matriz jacobiana

Sea $f : \mathbb{R}^m \longrightarrow \mathbb{R}^n$. Se define la **matriz jacobiana** de f como:

$$\mathrm{J}_f(a)_{ij} = \frac{\partial f_i}{\partial x_j}(a) \begin{array}{l} 1 \leq i \leq n \\ 1 \leq j \leq m \end{array}$$

$$\mathrm{J}_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_m}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_m}(a) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \dots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix}$$

$$\mathrm{J}_f(a) \in \mathcal{M}_{n \times m}(\mathbb{R})$$

3.1.1. Regla de la cadena

Sea $f : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ y $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, de forma que $h = g \circ f;\; h : \mathbb{R}^m \longrightarrow \mathbb{R}^p$, $h(x) = (g \circ f)(x) = g(f(x))$. Si

$$f(x_1, x_2, \cdots, x_m) = (u_1, u_2, \cdots, u_n)$$

$$g(u_1, u_2, \cdots, u_n) = (y_1, y_2, \cdots, y_p)$$

$$h(x) = g(u_1, \cdots, u_n) = g(f_1(x_1, \cdots, x_m), \cdots, f_n(x_1, \cdots, x_m)) = g(f(x))$$

entonces (con $1 \leq i \leq p$ y $1 \leq j \leq m$):

$$\frac{\partial h_i}{\partial x_j}(a) = \sum_{k=1}^{n} \frac{\partial g_i}{\partial u_k}(f(a)) \cdot \frac{\partial f_k}{\partial x_j}(a)$$

por lo tanto:

$$\mathrm{J}_{(g \circ f)}(x) = \mathrm{J}_g(f(x)) \cdot \mathrm{J}_f(x)$$

3.2. Diferenciabilidad

Sea $A \subset \mathbb{R}^m$ un conjunto abierto y $f : A \longrightarrow \mathbb{R}^n$. Se dice que f es **diferenciable** en el punto $a \in A$ si admite matriz jacobiana en a y, siendo $\mathbf{h} \in \mathbb{R}^m$:

$$\lim_{||\mathbf{h}|| \to 0} \frac{||f(a+\mathbf{h}) - f(a) - \mathrm{J}_f(a) \cdot \mathbf{h}||}{||\mathbf{h}||} = 0$$

Teorema: si f admite derivadas parciales en A y todas ellas son contínuas en $a \in A$, entonces f es diferenciable en a.

Teorema: Si f es diferenciable en a, entonces es contínua en a.

3.2.1. Plano tangente en $\mathbb{R}^2$

En el caso particular de una función escalar $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ es diferenciable en $a \in \mathbb{R}^2$, su gráfica admite un **plano tangente** en $a = (a_1, a_2)$ que tiene por ecuación:

$$z - f(a) = \frac{\partial f}{\partial x} \cdot (x - a_1) + \frac{\partial f}{\partial y} \cdot (y - a_2)$$

3.3. Extremos absolutos y relativos

Siendo $A \subset \mathbb{R}^m$, se dice que la función $f : A \longrightarrow \mathbb{R}^n$ presenta:

- Un **máximo relativo** en $a \in A$ si

$$\exists r \in \mathbb{R}^+ : \forall x \in \mathrm{B}(a, r)\ f(x) \leq f(a)$$

- Un **máximo absoluto** en $a \in A$ si

$$\forall x \in A\ f(x) \leq f(a)$$

- Un **mínimo relativo** en $a \in A$ si

$$\exists r \in \mathbb{R}^+ : \forall x \in \mathrm{B}(a, r)\ f(x) \geq f(a)$$

- Un **mínimo absoluto** en $a \in A$ si

$$\forall x \in A\ f(x) \geq f(a)$$

3.3.1. Puntos críticos

Teorema: Si la función f es diferenciable en $a \in A$ y alcanza un extremo relativo en dicho punto, entonces:

$$\mathrm{J}_f(a) = \mathbf{0}$$

Es decir, esta es una *condición necesaria* (pero no suficiente) para que exista un extremo relativo en dicho punto. Los puntos donde se cumple se llaman **puntos críticos** o estacionarios.

3.3.2. Matriz hessiana en funciones escalares

Siendo $n \in \mathbb{N}$; $A \subset \mathbb{R}^n$ y la función escalar $f : A \longrightarrow \mathbb{R}$, se define la matriz hessiana de f como:

$$\mathrm{H}_f(a) \in \mathcal{M}_{n\times n}(\mathbb{R})$$

$$\mathrm{H}_f(a)_{ij} = \frac{\partial f^2}{\partial x_i \partial x_j}(a)$$

o, equivalentemente:

$$\mathrm{H}_f(a) = \begin{pmatrix} \frac{\partial f^2}{\partial x_1 \partial x_1}(a) & \frac{\partial f^2}{\partial x_1 \partial x_2}(a) & \dots & \frac{\partial f^2}{\partial x_1 \partial x_n}(a) \\ \frac{\partial f^2}{\partial x_2 \partial x_1}(a) & \frac{\partial f^2}{\partial x_2 \partial x_2}(a) & \dots & \frac{\partial f^2}{\partial x_2 \partial x_n}(a) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f^2}{\partial x_n \partial x_1}(a) & \frac{\partial f^2}{\partial x_n \partial x_2}(a) & \dots & \frac{\partial f^2}{\partial x_n \partial x_n}(a) \end{pmatrix}$$

Si la función f admite segundas derivadas y éstas son contínuas, la matriz hessiana está bien definida y, debido al teorema de Schwarz de las derivadas parciales, la matriz hessiana es simétrica:

$$\mathrm{H}_f(a) = \mathrm{H}_f(a)^{\mathrm{T}}$$

Si la matriz hessiana en a es una *matriz definida positiva*, la función presenta un **mínimo relativo** en dicho punto. Si la matriz es una *matriz definida negativa*, la función presenta un **máximo relativo** en el punto.

Es decir, si $k \in \mathbb{N}$; $1 \leq k \leq n$, definiendo H_k como el determinante

$$H_k = \begin{vmatrix} \mathrm{H}_f(a)_{11} & \dots & \mathrm{H}_f(a)_{1k} \\ \dots & \dots & \dots \\ \mathrm{H}_f(a)_{k1} & \dots & \mathrm{H}_f(a)_{kk} \end{vmatrix}$$

entonces la matriz es definida positiva si $\forall k : H_k > 0$ (y por tanto la función presenta un mínimo relativo); y es definida negativa si $\forall k : \mathrm{sgn}(H_k) = (-1)^k$ (y por tanto la función presenta un máximo relativo).

Debido a que $H_n = |\mathrm{H}_f(a)|$, en el caso particular de funciones de variable en $\mathbb{R}^2$ (es decir, $n = 2$), el procedimiento anterior se puede simplificar de la siguiente forma:

- Si $|\mathrm{H}_f(a)| > 0$ y $\frac{\partial f^2}{\partial x_1 \partial x_1}(a) > 0$, la función presenta un mínimo relativo en a;
- Si $|\mathrm{H}_f(a)| > 0$ y $\frac{\partial f^2}{\partial x_1 \partial x_1}(a) < 0$, la función presenta un máximo relativo en a.

3.3.3. Extremos absolutos en un conjunto

Teorema (de Weierstrass): Toda función contínua definida en un conjunto compacto (cerrado y acotado) de $\mathbb{R}^n$ alcanza máximos y mínimos absolutos en el conjunto.

Los puntos candidatos a ser extremos absolutos de $f : A \longrightarrow \mathbb{R}; A \subset \mathbb{R}^n$ en un conjunto son:

- **Puntos críticos** de f en el interior del conjunto;
- Puntos del interior del conjunto donde **no exista alguna derivada parcial**;
- Puntos de la **frontera del conjunto**.

3.4. Extremos condicionados en funciones escalares

3.4.1. Introducción

Sea $n \in \mathbb{N};\ A \subset \mathbb{R}^n$ y la función escalar diferenciable $f : A \longrightarrow \mathbb{R}$. Se desea obtener los máximos o mínimos de la función f sujeta a una condición S definida a partir de la función diferenciable $g : A \longrightarrow \mathbb{R}$, de forma que:

$$S = \{x \in A : g(x) = 0\}$$

Ha de observarse que S es el conjunto de nivel para 0 de g ($S = \mathrm{C}_0(g)$). Se considerará el contorno de nivel delimitado por S (es decir, puntos $x \in S$ tales que $\nabla g(x) \neq \mathbf{0}$). Al recorrerlo es posible que se atraviesen contornos de nivel de f (es decir, f está variando en esos puntos).

Si en el recorrido se está en un entorno de un punto de un contorno de nivel de f y recorriendo dicho contorno de forma tangencial, la componente tangencial del vector gradiente de f en dicho punto es nula; es decir, considerando la función f restringida al contorno de g que es está recorriendo, el punto es un punto crítico. Al ser el contorno de g tangente al contorno de f

en el entorno de tal punto los vectores gradiente de f y g en ese punto han de tener la misma dirección:

$$\exists\lambda \in \mathbb{R} \;\; \nabla g(x) = \lambda \nabla f(x)$$

Formalmente se puede expresar mediante el siguiente **teorema**: Sea $a \in A$; si $a \in S$, $\nabla g(a) \neq \mathbf{0}$ y la función f restringida al contorno S presenta un extremo relativo, entonces

$$\exists\lambda \in \mathbb{R} \;\; \nabla g(x) = \lambda \nabla f(x)$$

3.4.2. Multiplicadores de Lagrange

El método de los multiplicadores de Lagrange permite localizar los puntos candidatos a extremos condicionados por S. Para ello, se define la función auxiliar (*función lagrangiana*) $F : A \times \mathbb{R} \longrightarrow \mathbb{R}$, teniendo en cuenta $x \in A$, como:

$$F(x, \lambda) = f(x) + \lambda g(x)$$

cuyo gradiente es

$$\nabla F(x, \lambda) = \left(\frac{\partial f}{\partial x_1}(x) + \lambda \frac{\partial g}{\partial x_1}(x), \cdots, \frac{\partial f}{\partial x_n}(x) + \lambda \frac{\partial g}{\partial x_n}(x), g(x) \right)$$

y se calculan los puntos donde el gradiente es el vector $\mathbf{0}$ (con lo cual cumplirán la condición del teorema anterior y que g sea 0).

El método se puede generalizar para obtener puntos que cumplan $r \in \mathbb{N}$ condiciones:

$$\lambda = (\lambda_1, \cdots, \lambda_r)$$

$$F(x, \lambda) = f(x) + \sum_{i=1}^{r} \lambda_i g_i(x)$$

$$\nabla F(x, \lambda) = \mathbf{0}$$