Tema 14. Diferenciabilidad de funciones de varias variables

Juan Medina Molina

19 de abril de 2005

Introducción

En este tema vamos a generalizar la derivabilidad de funciones reales de variable real a funciones de varias variables.

Empezaremos introduciendo los conceptos de derivada direccional y derivada parcial de funciones de varias variables. Dado que la existencia de éstas no implica en general la continuidad de la función, necesitaremos obtener una extensión de éstas que si que la implique, lo que nos llevará a presentar el concepto de diferenciabilidad de una función de varias variables. Tras estudiar las derivadas de orden superior, presentaremos el polinomio de Taylor en varias variables, que al igual que en una variable, permite aproximar funciones de varias variables por polinomios. Posteriormente abordaremos el cálculo de extremos de funciones de varias variables. La última parte del tema la dedicaremos al estudio del teorema de la función implícita y el teorema de la función inversa.

Este tema presenta los siguientes apartados:

- Derivadas direccionales y derivadas parciales.
- Diferenciabilidad de una función. Cálculo de la diferencial. Propiedades.
- Matriz jacobiana de una función. Propiedades.
- Interpretación geométrica de las derivadas parciales de una función de dos variables.
- Derivadas parciales de orden superior. Teorema de Schwarz.
- Polinomio de Taylor de una función de varias variables.

- Extremos relativos y absolutos de una función real de varias variables. Extremos condicionados, método de los multiplicadores de Lagrange.
- El teorema de la función implícita.
- El teorema de la función inversa.

Derivadas direccionales y derivadas parciales

Para empezar este tema, intentamos dar un definición de derivada que generalice el concepto de derivada de una a varias variables.

Definición 1 *Sea A un subconjunto abierto de $\mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}$ y $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.*

i) Si $a \in A$, se define la derivada direccional, en la dirección del vector $\mathbf{v}$, de f en a como

$$(\mathbf{D}_v f)(a) = \lim_{t \to 0} \frac{f(a + t\mathbf{v}) - f(a)}{t}.$$

Sea $B = \{e_1, e_2, \dots, e_n\}$ es la base canónica de $\mathbb{R}^n$.

ii) Si $i \in \{1, 2, \dots, n\}$, se define la derivada parcial i-ésima de f en el punto a y se denota por $\frac{\partial f}{\partial x_i}(a)$ o $(\mathbf{D}_i f)(a)$ como

$$\frac{\partial f}{\partial x_i}(a) = (\mathbf{D}_{e_i} f)(a) = \lim_{t \to 0} \frac{f(a + t\mathbf{e_i}) - f(a)}{t}.$$

Consideremos la función real de variable real f_i tal que

$$f_i(x_i) = f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$$

, o sea, f_i es la función que se obtiene de considerar en f todas las variables constantes excepto la i-ésima. Entonces la derivada de ésta en a sería:

$\lim_{t\to 0} \frac{f_i(a_i+t)-f_i(a_i)}{t} = \lim_{t\to 0} \frac{f(a_1,\dots,a_{i-1},a_i+t,a_{i+1},\dots,a_n)-f(a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n)}{t} = \frac{\partial f}{\partial x_i}(a)$.

Así, la derivada parcial i-ésima evaluada en a es el valor que se obtiene de sustituir a en la función que resulta de derivar f respecto de x_i considerando las otras variables constantes.

Las derivadas parciales y direccionales no son la generalización buscada de derivada ya que la existencia en un punto de éstas, no implica la continuidad en dicho punto, como muestra en $(0, 0)$ la función:

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} \mid f(x, y) = \begin{cases} \frac{x^2}{y} & \text{si } \ y \neq 0 \\ 0 & \text{si } \ y = 0 \end{cases} .$$

Definición 2 *Sea $f : A \longrightarrow \mathbb{R}^m$ donde A es un subconjunto abierto de $\mathbb{R}^n$ y sea $a \in A$. Se dice que f es diferenciable en a si existe una aplicación lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ tal que*

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\| f(a+\mathbf{h}) - f(a) - T(\mathbf{h}) \|_m}{\| \mathbf{h} \|_n} = 0.$$

En ese caso, a la aplicación lineal T se le llama diferencial de f en a y se denota por $\mathbf{d}f(a)$.

Se dice que f es diferenciable en A si lo es en cualquier punto de A.

Es sencillo demostrar que si $f : D \longrightarrow \mathbb{R}$, siendo D un subconjunto abierto de $\mathbb{R}$, f es derivable en $a \in D$ si y sólo es diferenciable en a y $\mathbf{d}f(a)(t) = f'(a)t$.

Se obtienen las siguientes tres propiedades:

Proposición 1 *Sea $f : A \longrightarrow \mathbb{R}^m$ donde A es un subconjunto abierto de $\mathbb{R}^n$ y $a \in A$. Entonces:*

i) *La diferencial $\mathbf{d}f(a)$ es única.*

ii) *f es continua en A.*

Proposición 2 *Sea $f : A \longrightarrow \mathbb{R}^m$ donde A es un subconjunto abierto de $\mathbb{R}^n$ y $a \in A$. Sean $f_1, \ldots, f_m$ las funciones coordenadas de f. Entonces f es diferenciable en a si y sólo si f_i es diferenciable en a, $1, \leq i \leq m$, y en ese caso $\mathbf{d}f(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ tiene por funciones coordenadas a $\mathbf{d}f_1(a), \ldots, \mathbf{d}f_m(a)$.*

Proposición 3 *Sea $f : A \longrightarrow \mathbb{R}$ donde A es un subconjunto abierto de $\mathbb{R}^n$ y $a \in A$ tal que f es diferenciable en a. Entonces, si $v \in \mathbb{R}^n \setminus \{0\}$, $\mathbf{d}f(a)(v) = (\mathbf{D}_v f)(a)$ y como consecuencia de esto, $\mathbf{d}f(a)(e_i) = \frac{\partial f}{\partial x_i}(a)$.*

Con respecto a las operaciones entre funciones se obtiene:

Proposición 4 *Sea $f, g : A \longrightarrow \mathbb{R}^m$ diferenciables en $a \in A$ siendo A un subconjunto abierto de $\mathbb{R}^n$ y $\alpha, \beta \in \mathbb{R}$. Entonces:*

i) *$\alpha f + \beta g$ es diferenciable en a y $\mathbf{d}(\alpha f + \beta g)(a) = \alpha(\mathbf{d}f)(a) + \beta(\mathbf{d}g)(a)$.*

ii) *Si $m = 1$ entonces fg es diferenciable en a y $\mathbf{d}(fg)(a) = (\mathbf{d}f)(a)g(a) + f(a)(\mathbf{d}g)(a)$.*

iii) *Si $m = 1$, $g(x) \neq 0$ para todo $x \in A$, $\frac{f}{g}$ es diferenciable en a y*

$$\mathbf{d}(\frac{f}{g})(a) = \frac{(\mathbf{d}f)(a)g(a) - f(a)(\mathbf{d}g)(a)}{g(a)^2}.$$

Y también se obtiene la regla de la cadena:

Teorema 1 *Sean A y $\tilde{A}$ subconjuntos abiertos de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente y sean $f : A \longrightarrow \mathbb{R}^m$ y $g : \tilde{A} \longrightarrow \mathbb{R}^l$ tales que $f(A) \subseteq \tilde{A}$. Si $a \in A$ verifica que f es diferenciable en a y g lo es en $f(a)$, entonces $g \circ f$ es diferenciable en a y $\mathbf{d}(g \circ f)(a) = \mathbf{d}g(f(a)) \circ \mathbf{d}f(a)$.*

Por definición, la diferencial de una función en un punto es una aplicación lineal luego ésta viene determinada por su matriz respecto de las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$.

Definición 3 *Sea $f : A \longrightarrow \mathbb{R}^m$ donde A es un subconjunto abierto de $\mathbb{R}^n$ y sea $a \in A$ tal que f es diferenciable en a. Se define la matriz jacobiana de f en a como $\mathbf{J}f(a) = \mathbf{M}_{B,B'}(\mathbf{d}f(a))$ siendo B y B' las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente.*

Como se tiene que $(\mathbf{d}f(a))(e_i) = (\mathbf{d}f_1(a))(e_i), \mathbf{d}f_2(a))(e_i), \ldots, \mathbf{d}f_m(a))(e_i)) = (\frac{\partial f_1}{\partial x_i}(a), \frac{\partial f_2}{\partial x_i}(a), \ldots, \frac{\partial f_m}{\partial x_i}(a))$, se obtiene:

$$\mathbf{J}f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \ldots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

De propiedades de aplicaciones lineales y de las propiedades de la diferencial con respecto de las aplicaciones entre funciones se obtiene:

Proposición 5 *Sea $f, g : A \longrightarrow \mathbb{R}^m$ diferenciables en $a \in A$ siendo A un subconjunto abierto de $\mathbb{R}^n$ y $\alpha, \beta \in \mathbb{R}$. Entonces:*

i) $\mathbf{J}(\alpha f + \beta g)(a) = \alpha \cdot \mathbf{J}f(a) + \beta \cdot \mathbf{J}g(a)$.

ii) *Si $m = 1$, $\mathbf{J}(fg)(a) = \mathbf{J}f(a) \cdot g(a) + f(a) \cdot \mathbf{J}g(a)$.*

iii) *Si $m = 1$, $g(x) \neq 0$ para todo $x \in A$,*

$$\mathbf{J}(\frac{f}{g})(a) = \frac{\mathbf{J}f(a) \cdot g(a) - f(a) \cdot \mathbf{J}g(a)}{g(a)^2}.$$

También se obtiene una versión de la regla de la cadena para matrices:

Proposición 6 *Con las mismas hipótesis de la regla de la cadena se obtiene que* $\mathbf{J}(g \circ f)(a) = \mathbf{J}g(f(a)) \cdot \mathbf{J}f(a)$.

La expresión matricial de la regla de la cadena permite transformar expresiones donde aparecen derivadas parciales al aplicar un cambio de coordenadas.

Definición 4 *Si* $f : A \longrightarrow \mathbb{R}$ *es diferenciable en* $a \in A$, *siendo* A *un subconjunto abierto de* $\mathbb{R}^n$, *se define el vector gradiente de* f *en* a *como* $\nabla f(a) = (\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_n}(a))$.

El siguiente resultado muestra que la continuidad de las derivadas parciales permite obtener la diferenciabilidad de una función.

Proposición 7 *Sea* $f : A \longrightarrow \mathbb{R}^m$ *siendo* A *un subconjunto abierto de* $\mathbb{R}^n$. *Supongamos que existe* $\frac{\partial f}{\partial x_i} : A \longrightarrow \mathbb{R}^n$ *y es continua en* $a \in A$, $1 \leq i \leq n$. *Entonces* f *es diferenciable en* a.

Sin embargo, la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por:

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

muestra que existen funciones diferenciables en un punto cuyas derivadas parciales no son continuas en dicho punto.

Interpretación geométrica de las derivadas parciales de una función de dos variables

Si A es un subconjunto abierto de $\mathbb{R}^2$ y $f : A \longrightarrow \mathbb{R}$ es diferenciable en $(x_0, y_0) \in A$ entonces la ecuación del plano tangente a la gráfica de f en dicho punto tiene por ecuación:

$$z - f(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0).$$

Derivadas parciales de orden superior

Sea A un subconjunto abierto de $\mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}^m$ y $i, j \in \{1, \ldots, n\}$ (Quizás $i = j$). Entonces, si existe $\frac{\partial f}{\partial x_i}$ en todo punto de A, tenemos una aplicación $\frac{\partial f}{\partial x_i} : A \longrightarrow \mathbb{R}^m$ y así, podemos plantearnos el cálculo de la derivada parcial j-ésima de ésta. Se define la derivada parcial segunda de f, primero respecto de x_i y después respecto de x_j de f en $a \in A$ como

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{d}{dx_j}\left(\frac{\partial f}{\partial x_i}\right)(a).$$

En el caso $i = j$ denotaremos por $\frac{\partial^2 f}{\partial x_i^2}(a) = \frac{\partial^2 f}{\partial x_i \partial x_i}(a)$.

Del mismo modo, se definen las derivadas parciales terceras, cuartas, etcétera.

Se dice que f es una función de clase $\mathcal{C}^k$ si tiene todas las derivadas k-ésimas continuas en A. Entonces escribiremos $f \in \mathcal{C}^k(A, \mathbb{R}^m)$.

En este punto nos planteamos la siguiente cuestión:

Dados $i, j \in \{1, \ldots, n\}$, ¿Bajo que condiciones $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$?

Se tiene el siguiente resultado de Schwarz:

Teorema 2 *Sea $f : A \longrightarrow \mathbb{R}^m$, siendo A un subconjunto abierto de $\mathbb{R}^n$ y supongamos que $\frac{\partial f}{\partial x_i}$ y $\frac{\partial f}{\partial x_j}$ son continuas en A siendo $i, j \in \{1, \ldots, n\}$ distintos. Si existe $\frac{\partial^2 f}{\partial x_i \partial x_j} : A \longrightarrow \mathbb{R}^m$ y es continua en $a \in A$ entonces existe $\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ y $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$.*

Polinomio de Taylor de una función de varias variables

La idea del polinomio de Taylor en varias variables es la misma que una variable, dada $f : A \longrightarrow \mathbb{R}$ siendo $A \subseteq \mathbb{R}^n$ y $a \in A$, encontrar un polinomio de n variables que aproxime a f en un entorno de a.

Para enunciar el siguiente teorema, introducimos el concepto de intervalo en $\mathbb{R}^n$:

Definición 5 *Si $a, b \in \mathbb{R}^n$, se define el segmeto de extremos a y b como*

$$[a, b] = \{a + t(b - a) \mid t \in [0, 1]\}.$$

Claramente, si $a \in A$ siendo A un subconjunto abierto de $\mathbb{R}^n$ entonces existe $x \in A$, $x \neq a$ tal que $[a, x] \subseteq A$.

Entonces se tiene el teorema de Taylor:

Teorema 3 *Supongamos que $f \in \mathcal{C}^{k+1}(A, \mathbb{R})$ siendo A un subconjunto abierto de $\mathbb{R}^n$ y sea $a = (a_, \ldots, a_n) \in A$. Entonces, para todo $x = (x_1, \ldots, x_n) \in A$ tal que $[a, x] \subseteq A$, existe $\theta \in]0, 1[$ tal que se tiene $f(x) = f(a) + \frac{1}{1!} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} f(a)(x_i - a_i)(x_j - a_j) + \ldots + \sum_{i_1,\ldots,i_k=1}^{n} \frac{d^k f}{dx_1 \ldots x_k} f(a)(x_{i_1} - a_{i_1}) \ldots (x_{i_k} - a_{i_k}) + \frac{1}{(k+1)!} \sum_{i_1,\ldots,i_{k+1}=1}^{k} \frac{d^k f}{dx_1 \ldots x_{k+1}} f(a + \theta(x-a))(x_{i_1} - a_{i_1}) \ldots (x_{i_{k+1}} - a_{i_{k+1}})$.*

A $p_k(x_1, \ldots x_n) = f(a) + \frac{1}{1!} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} f(a)(x_i - a_i)(x_j - a_j) + \ldots + \sum_{i_1,\ldots,i_k=1}^{k} \frac{d^k f}{dx_1 \ldots x_k} f(a)(x_{i_1} - a_{i_1}) \ldots (x_{i_k} - a_{i_k})$ se le llama polinomio de Taylor de grado k de f en a y a la expresión $R_k(x_1, \ldots x_n) = \frac{1}{(k+1)!} \sum_{i_1,\ldots,i_{k+1}=1}^{k} \frac{d^k f}{dx_1 \ldots x_{k+1}} f(a + \theta(x-a))(x_{i_1} - a_{i_1}) \ldots (x_{i_{k+1}} - a_{i_{k+1}})$ se le llama resto de Lagrange de dicho polinomio.

Extremos relativos y absolutos de funciones reales de varias variables

Para empezar esta sección, introducimos los conceptos sobre extremos de una función real de varias variables:

Definición 6 *Sea $f : A \longrightarrow \mathbb{R}$, siendo $A \subseteq \mathbb{R}^n$ y sea $a \in A$.*

i) *Diremos que a es un máximo absoluto de f si $f(a) \geq f(x)$ para todo $x \in A$.*

ii) *Diremos que a es un mínimo absoluto de f si $f(a) \leq f(x)$ para todo $x \in A$.*

iii) *Diremos que a es un máximo relativo de f si existe $\epsilon > 0$ tal que $f(a) \geq f(x)$ para todo $x \in \mathrm{B}(a, \epsilon) \cap A$.*

iv) *Diremos que a es un mínimo relativo de f si existe $\epsilon > 0$ tal que $f(a) \leq f(x)$ para todo $x \in \mathrm{B}(a, \epsilon) \cap A$.*

v) *Diremos que a es un extremo relativo de f si es un máximo relativo o un mínimo relativo de f.*

Bajo ciertas condiciones, el siguiente resultado de Weierstrass afirma la existencia de extremos absolutos:

Teorema 4 *Sea $f : A \longrightarrow \mathbb{R}$ una función continua, siendo A un subconjunto cerrado y acotado de $\mathbb{R}^n$. Entonces existen $x_m, x_M \in A$ tales que $f(x_m) \leq f(x) \leq f(x_M)$ para todo $x \in A$.*

Evidentemente, todo extremo absoluto es extremo relativo.

El cálculo de los extremos absolutos de una función continua en un conjunto cerrado y acotado consta de dos etapas:

1. Cálculo de los extremos relativos en el interior del conjunto.

2. Cálculo de los restantes posibles extremos absolutos pertenecientes a la frontera del conjunto.

De todos los candidatos obtenidos anteriormente, aquellos cuya imagen sea máxima serán los máximos absolutos y aquellos cuya imagen sea mínima, los mínimos relativos.

Para la primera etapa se tiene:

Proposición 8 *Sea A un subconjunto abierto de $\mathbb{R}^n$ y supongamos que $f \in \mathcal{C}^1(A, \mathbb{R})$. Si $a \in A$ es un extremo relativo de f, entonces $\frac{\partial f}{\partial x_i}(a) = 0$, $1 \leq i \leq n$.*

Los puntos críticos son aquellos que anulan las derivadas parciales primeras. Las funciones $f : \mathbb{R}^2 \longrightarrow \mathbb{R} \mid f(x,y) = x^2 - y^2$ o $f(x,y) = (y - x^2)(y - 2x^2)$ poseen a $(0,0)$ como punto crítico que no es extremo relativo, por lo que el recíproco del resultado anterior no es cierto.

Introducimos la matriz Hessiana:

Definición 7 *Sea $f \in \mathcal{C}^2(A, \mathbb{R}^n)$ siendo A un subconjunto abierto de $\mathbb{R}^n$ y sea $a \in A$. Se define la matriz Hessiana de f en A como:*

$$\mathrm{H}f(a) = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \dots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{vmatrix}.$$

Si $i \in \{1, \dots n\}$ se define:

$$\Delta_i f(a) = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_i}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_i}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_i \partial x_1}(a) & \frac{\partial^2 f}{\partial x_i \partial x_2}(a) & \dots & \frac{\partial^2 f}{\partial x_i^2}(a) \end{vmatrix}.$$

Se verifica que si $a \in A$ es un punto crítico de f, entonces:

i) Si $\Delta_i f(a) > 0$ $1 \leq i \leq n$, entonces a es un mínimo relativo de f.

ii) Si $\Delta_1 f(a) < 0$, $\Delta_2 f(a) > 0$ y se siguen alternando los signos, entonces a es un máximo relativo.

Para funciones de dos variables, además se obtiene:

iii) Si $\Delta_2 f(a) < 0$, entonces a no es extremo relativo de f.

Para los casos no contemplados anteriormente es necesario un análisis en las proximidades del punto crítico para determinar si éste es o no un extremo relativo de la función. Para ello utilizaremos el método de las regiones.

Para acabar esta sección, presentamos el método de los multiplicadores de Lagrange en $\mathbb{R}^2$ y $\mathbb{R}^3$ para determinar los extremos relativos de una función de n variables en "$(n-1)$-variedades".

En $\mathbb{R}^2$:

Supongamos que $f \in \mathcal{C}^2(A, \mathbb{R})$, siendo A un subconjunto abierto de $\mathbb{R}^2$ y $B = \{(x, y) \in A \mid g(x, y) = 0\}$ siendo $g \in \mathcal{C}^1(A, \mathbb{R})$ con gradiente no nulo en A. Veamos la forma de calcular los extremos relativos de f en B.

En primer lugar consideramos la función $F : A \times \mathbb{R} \longrightarrow \mathbb{R} \mid F(x, y, \lambda) = f(x, y) + \lambda g(x, y)$. Entonces planteamos el sistema:

$$\left\{ \begin{array}{c} \frac{\partial F}{\partial x}(x, y, \lambda) = 0 \\ \frac{\partial F}{\partial y}(x, y, \lambda) = 0 \\ g(x, y) = 0 \end{array} \right. .$$

Entonces, los puntos cuyas componentes son las dos primeras componentes de las soluciones de dicho sistema son los candidatos a extremos relativos de f en B.

Para ver si uno de tales puntos (x_0, y_0) es un extremo relativo, siendo (x_0, y_0, λ_0) la solución asociada del sistema, planteamos la ecuación

$$\frac{\partial g}{\partial x}(x_0, y_0)a + \frac{\partial g}{\partial y}(x_0, y_0)b = 0,$$

y de ésta despejamos una de las dos incógnitas a o b en función de la otra.

Entonces sustituimos dicha incognita en la expresión:

$\Delta(x_0, y_0, \lambda) = \frac{\partial^2 F}{\partial x^2}(x_0, y_0, \lambda_0)a^2 + 2\frac{\partial^2 F}{\partial x \partial y}(x_0, y_0, \lambda_0)ab + \frac{\partial^2 F}{\partial y^2}(x_0, y_0, \lambda_0)b^2$. Entonces:

i) Si $\Delta(x_0, y_0, \lambda_0) > 0$ entonces (x_0, y_0) es un mínimo relativo de f en B.

ii) Si $\Delta(x_0, y_0, \lambda_0) < 0$ entonces (x_0, y_0) es un máximo relativo de f en B.

iii) Si $\Delta(x_0, y_0, \lambda_0) = 0$ tendremos que analizar en las proximidades de (x_0, y_0) si éste es o no un extremo relativo.

En $\mathbb{R}^3$:

Supongamos que $f \in \mathcal{C}^2(A, \mathbb{R})$, siendo A un subconjunto abierto de $\mathbb{R}^3$.

Para el caso de que $B = \{(x, y, z) \in A \mid g(x, y, z) = 0\}$, procederemos de la misma forma que en $\mathbb{R}^2$.

Supongamos ahora que $B = \{(x, y, z) \in A \mid g_1(x, y, z) = 0, g_2(x, y, z) = 0\}$ siendo $g_1, g_2 \in \mathcal{C}^1(A, \mathbb{R})$. Veamos la forma de calcular los extremos relativos de f en B.

En primer lugar consideramos la función $F : A \times \mathbb{R}^2 \longrightarrow \mathbb{R} \mid F(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda g_1(x, y, z) + \mu g_2(x, y, z)$. Entonces planteamos el sistema:

$$\left\{\begin{array}{c} \frac{\partial F}{\partial x} F(x, y, z, \lambda, \mu) = 0 \\ \frac{\partial F}{\partial y} F(x, y, z, \lambda, \mu) = 0 \\ \frac{\partial F}{\partial z} F(x, y, z, \lambda, \mu) = 0 \\ g_1(x, y, z) = 0 \\ g_2(x, y, z) = 0 \end{array}\right. .$$

Entonces, los puntos cuyas componentes son las tres primeras componentes de las soluciones de dicho sistema son los candidatos a extremos relativos de f en B.

Para ver si uno de tales puntos (x_0, y_0, z_0) es un extremo relativo, siendo $(x_0, y_0, z_0, \lambda_0, \mu_0)$ la solución asociada del sistema, planteamos el sistema formado por las dos ecuaciones siguientes:

$$\frac{\partial g_i}{\partial x}(x_0, y_0, z_0)a + \frac{\partial g_i}{\partial y}(x_0, y_0, z_0)b + \frac{\partial g_i}{\partial z}(x_0, y_0, z_0)c = 0,\ i = 1, 2,$$

y en éste despejamos dos de las tres incógnitas a, b y c en función de la otra.

Entonces sustituimos dichas incognitas en la expresión:

$\Delta(x_0, y_0, z_0, \lambda, \mu) = \frac{\partial^2 F}{\partial x^2}(x_0, y_0, z_0, \lambda_0, \mu_0)a^2 + 2\frac{\partial^2 F}{\partial x \partial y}(x_0, y_0, z_0, \lambda_0, \mu_0)ab + 2\frac{\partial^2 F}{\partial x \partial z}(x_0, y_0, z_0, \lambda_0, \mu_0)ac + \frac{\partial^2 F}{\partial y^2}(x_0, y_0, z_0, \lambda_0, \mu_0)b^2 + 2\frac{\partial^2 F}{\partial y \partial z}(x_0, y_0, z_0, \lambda_0, \mu_0)bc + \frac{\partial^2 F}{\partial z^2}(x_0, y_0, z_0, \lambda_0, \mu_0)c^2$. Entonces:

i) Si $\Delta(x_0, y_0, z_0 \lambda_0, \mu_0) > 0$ entonces (x_0, y_0, z_0) es un mínimo relativo de f en B.

ii) Si $\Delta(x_0, y_0, z_0 \lambda_0, \mu_0) < 0$ entonces (x_0, y_0, z_0) es un máximo relativo de f en B.

iii) Si $\Delta(x_0, y_0, z_0 \lambda_0, \mu_0) = 0$, tendremos que analizar en las proximidades de (x_0, y_0, z_0) si éste es o no un extremo relativo.

El teorema de la función implícita

Supongamos que tenemos una ecuación o sistema de ecuaciones con una serie de variables y queremos expresar una o varias de estas variables en función de las otras. Este problema aparece algunas veces en la resolución de una ecuación diferencial cuando la función solución viene dada de forma implícita por una ecuación.

El teorema de la función implícita resuelve en algunas ocasiones este problema:

Teorema 5 *Sea A un subconjunto abierto de $\mathbb{R}^n \times \mathbb{R}^m$, supongamos que $f_i = f_i(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathcal{C}^k(A, \mathbb{R})$ $1 \leq i \leq m$ y consideramos un punto $(a, b) = (a_1, \ldots, a_n, b_1, \ldots, b_m) \in A$ tal que $f_i(a, b) = 0$, $1 \leq i \leq m$. Supongamos que*

$$\begin{vmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & \frac{\partial f_1}{\partial y_2}(a,b) & \ldots & \frac{\partial f_1}{\partial y_m}(a,b) \\ \frac{\partial f_2}{\partial y_1}(a,b) & \frac{\partial f_2}{\partial y_2}(a,b) & \ldots & \frac{\partial f_2}{\partial y_m}(a,b) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1}(a,b) & \frac{\partial f_m}{\partial y_2}(a,b) & \ldots & \frac{\partial f_m}{\partial y_m}(a,b) \end{vmatrix} \neq 0.$$

Entonces existe un subconjunto abierto U de $\mathbb{R}^n$ tal que $(a_1, \ldots, a_n) \in U$, un subconjunto abierto V de $\mathbb{R}^m$ tal que $(b_1, \ldots, b_m) \in V$ y una única función $\varphi \in \mathcal{C}^k(U, V)$ con funciones coordenadas $\varphi_1, \ldots, \varphi_m$ tales que:

i) $U \times V \subseteq A$.

ii) $f_i(x_1, \ldots, x_n, \varphi_1(x_1, \ldots, x_n), \ldots, \varphi_m(x_1, \ldots, x_n)) = 0$ $1 \leq i \leq m$ *si* $(x_1, \ldots, x_n) \in U$ *y si* $(x_1, \ldots, x_n, y_1, \ldots, y_m) \in U \times V$ *verificando que* $f(x_1, \ldots x_n, y_1, \ldots, y_m) = 0$ *entonces* $\varphi(x_1, \ldots, x_n) = (y_1, \ldots, y_m)$.

iii) $\varphi(a_1, \ldots, a_n) = (b_1, \ldots, b_m)$.

Aunque este resultado no nos da una expresión de φ, de las ecuaciones $f_i(x_1, \ldots, x_n, \varphi(x_1, \ldots, x_n)) = 0$, $1 \leq i \leq m$, se pueden obtener las derivadas parciales de φ en $(a_1, \ldots, a_n)$ y así, se puede obtener una aproximación de φ en un entorno de $(a_1, \ldots, a_n)$ mediante un polinomio de Taylor.

El teorema de la función inversa

Ya conocemos un cambio de coordenadas en $\mathbb{R}^2$ que es el cambio de coordenadas cartesianas a coordenadas polares. Estos permiten en algunas ocasiones simplificar expresiones, ya que una expresión en coordenadas cartesianas complicada puede ser sencilla en coordenadas polares. Como se ve en el capítulo de integrales múltiples, éstos son utilizados en ocasiones para resolver integrales múltiples. Para que una ecuación sea un cambio de coordenadas, debe de ser biyectiva además de cumplir ciertas condiciones de diferenciabilidad. En algunas ocasiones, el siguiente resultado, conocido como el teorema de la función inversa, es útil para demostrar que una aplicación es un cambio de coordenadas en un entorno de un punto.

Teorema 6 *Sea A un subconjunto abierto de $\mathbb{R}^n$, $f \in \mathcal{C}^k(A, \mathbb{R}^n)$ y $a \in A$ tal que $|\mathbf{J}f(a)| \neq 0$. Entonces existen U, V subconjuntos abiertos de $\mathbb{R}^n$ con $a \in U$ y $f(a) \in V$ tales que $f_{|U} : U \longrightarrow V$ es biyectiva y $f^{-1} : V \longrightarrow U$ cumple que $f^{-1} \in \mathcal{C}^k(V, U)$ y $\mathbf{J}f^{-1}(f(y)) = (\mathbf{J}f(x))^{-1}$ para todo $y \in V$ y $x \in U$ tal que $f(x) = y$.*

Bibliografía

1. T. Apostol, *Calculus Vol. 2.* Ed. Reverte.

2. G. Bradley, K. Smith, *Cálculo de varias variables.* Ed. Prentice Hall.

3. J. Burgos, *Cálculo infinitesimal de varias variable.* Ed. McGraw-Hill.

4. J. S. Canovas, J. A. Murillo, *Fundamentos Matemáticos de la Ingeniería.* Ed. Diego Marín.

5. A. de la Villa, A. García, A. López, G. Rodríguez, S. Romero, *Teoría y problemas de funciones de varias variables.* CLAGSA.

6. J. Fernández, M. Sánchez, *Ejercicios y complementos de análisis matemático II.* Ed. Tecnos.

7. M. Franco, F. Martínez, R. Molina, *Lecciones de cálculo infinitesimal II.* S. P. Universidad de Murcia.

8. G. Thomas, R. Finney, *Cálculo de varias variable.* Ed. Addison Wesley.