10. Optimización de funciones de varias variables

Contenido

10.1 Extremos relativos

Los extremos relativos de funciones reales de varias variables se definen de manera análoga al caso de una variable.

Consideremos un abierto U de $\mathbb{R}^n$, una función real $f\colon U \to \mathbb{R}$ y un punto $\mathbf{a} \in U$.

La función f tiene un *máximo relativo o máximo local* en $\mathbf{a}$, si existe un entorno V de $\mathbf{a}$ tal que $f(\mathbf{x}) \leq f(\mathbf{a})$ para todo $\mathbf{x} \in V$. La función f tiene un *mínimo relativo o mínimo local* en $\mathbf{a}$, si existe un entorno V de $\mathbf{a}$ tal que $f(\mathbf{x}) \geq f(\mathbf{a})$ para todo $\mathbf{x} \in V$.

Si en estas definiciones se sustituyen las desigualdades no estrictas por desigualdades estrictas, se obtienen las definiciones de *máximo y mínimo relativos estrictos*.

Si f tiene un máximo o mínimo relativo en $\mathbf{a}$, se dice que f tiene un *extremo relativo o extremo local* en $\mathbf{a}$.

Si f es una función real diferenciable en $\mathbf{a}$ y $\nabla f(\mathbf{a}) = 0$, se dice que $\mathbf{a}$ es un punto *crítico* o *estacionario* de f. Para las funciones diferenciables, ser un punto crítico es una condición necesaria para que exista un extremo en $\mathbf{a}$:

- Si f es una función real diferenciable en un punto $\mathbf{a}$ y f tiene un extremo relativo en $\mathbf{a}$, entonces $\nabla f(\mathbf{a}) = \mathbf{0}$.

Si f tiene un punto crítico en $\mathbf{a}$ pero no tiene un extremo relativo en $\mathbf{a}$, se dice que $\mathbf{a}$ es un *punto de silla* de f.

Para funciones suficientemente regulares, es posible determinar la naturaleza de un punto crítico con la ayuda de las segundas derivadas. Supongamos que f admite todas las segundas derivadas parciales en $\mathbf{a}$. La *matriz hessiana* de f en $\mathbf{a}$ es la matriz cuadrada de orden n

$$\mathcal{H}f(\mathbf{a}) = \left(D_{ij} f(\mathbf{a})\right).$$

Notemos que, si f es de clase $\mathcal{C}^2$ en un entorno de $\mathbf{a}$, entonces $\mathcal{H}f(\mathbf{a})$ es una matriz simétrica porque, según el teorema de Schwarz, $D_{ij}f(\mathbf{a}) = D_{ji}f(\mathbf{a})$. Para $k = 1, \ldots, n$, sea $\triangle_k(f, \mathbf{a})$ el determinante de la matriz obtenida de $\mathcal{H}f(\mathbf{a})$ suprimiendo sus últimas $n-k$ filas y columnas, es decir,

$$\triangle_1(f,\mathbf{a}) = D_{11}f(\mathbf{a}), \quad \triangle_2(f,\mathbf{a}) = \begin{vmatrix} D_{11}f(\mathbf{a}) & D_{12}f(\mathbf{a}) \\ D_{21}f(\mathbf{a}) & D_{22}f(\mathbf{a}) \end{vmatrix}, \quad \ldots, \quad \triangle_n(f,\mathbf{a}) = \det \mathcal{H}f(\mathbf{a}).$$

Si $\mathbf{a}$ es un punto crítico, los signos de estos determinantes proporcionan información sobre si $\mathbf{a}$ es máximo, mínimo o punto de silla.

Condiciones suficientes de extremo y punto de silla.

Sean U un abierto de $\mathbb{R}^n$ y $f\colon U \to \mathbb{R}$ una función de la clase $\mathcal{C}^2(U)$, y supongamos que $\mathbf{a} \in U$ es un punto crítico de f.

i) Si $\triangle_k(f, \mathbf{a}) > 0$ para todo $k \in \{1, \dots, n\}$, entonces f tiene un mínimo relativo en $\mathbf{a}$.

ii) Si $(-1)^k \triangle_k(f, \mathbf{a}) > 0$ para todo $k \in \{1, \dots, n\}$, entonces f tiene un máximo relativo en $\mathbf{a}$.

iii) Si existe ℓ tal que $\triangle_\ell(f, \mathbf{a}) > 0$ y $\triangle_k(f, \mathbf{a}) \geq 0$ para todo $k \neq \ell$, entonces f tiene un mínimo relativo o un punto de silla en $\mathbf{a}$.

iv) Si existe ℓ tal que $(-1)^\ell \triangle_\ell(f, \mathbf{a}) > 0$ y $(-1)^k (f, \mathbf{a}) \triangle_k \geq 0$ para todo $k \neq \ell$, entonces f tiene un máximo relativo o un punto de silla en el punto $\mathbf{a}$.

v) Si $\triangle_k(f, \mathbf{a}) = 0$ para todo $k \in \{1, \dots, n\}$, entonces f puede tener un máximo, un mínimo o un punto de silla en el punto $\mathbf{a}$.

vi) En cualquier otro caso, f tiene un punto de silla en el punto $\mathbf{a}$.

En el caso de dos variables, los resultados precedentes admiten más precisión. La matriz hessiana de f en un punto (a, b) es una matriz cuadrada de orden 2. Sean $h_{ij} = D_{ij} f(\mathbf{a})$, $i, j \in \{1, 2\}$, sus cuatro términos. Los determinantes considerados se reducen ahora a dos, que son $\triangle_1(f, (a, b)) = h_{11}$ y el determinante $\triangle = \triangle_2(f, (a, b)) = h_{11}h_{22} - h_{12}h_{21}$ de la matriz hessiana.

Condiciones suficientes de extremos y puntos de silla (para dos variables).

Sean U un abierto de $\mathbb{R}^2$ y $f\colon U \to \mathbb{R}$ una función de la clase $\mathcal{C}^2(U)$. Supongamos que $(a, b) \in U$ es un punto crítico de f. Sean $\mathcal{H}f(a, b) = (h_{ij})$ la matriz hessiana de f en (a, b) y $\triangle$ su determinante. Tenemos los siguientes casos.

1. $\triangle < 0$. Entonces, f tiene un punto de silla en (a, b).

2. $\triangle = 0$.

   i) $h_{11} < 0$ o $h_{22} < 0$. Entonces, f tiene un máximo o un punto de silla en (a, b).

   ii) $\mathcal{H}f(a, b)$ es la matriz nula. Entonces, f puede tener un máximo, un mínimo o un punto de silla en (a, b).

   iii) $h_{11} > 0$ o $h_{22} > 0$. Entonces, f tiene un mínimo o un punto de silla en (a, b).

3. $\triangle > 0$.

   i) $h_{11} < 0$ o $h_{22} < 0$. Entonces, f tiene un máximo relativo en (a, b).

   ii) $h_{11} > 0$ o $h_{22} > 0$. Entonces, f tiene un mínimo relativo en (a, b).

Tres observaciones acerca del esquema anterior. Las tres se derivan de que f es de clase $\mathcal{C}^2$ en un entorno de $\mathbf{a}$, por lo que $h_{12} = h_{21}$ y $\triangle = h_{11}h_{22} - h_{12}^2 = h_{11}h_{22} - h_{21}^2$.

La primera observación hace referencia al caso 2 ii. Observemos que, si no se cumplen ni 2 i) ni 2 iii), entonces $h_{11} = h_{22} = 0$. Puesto que $\triangle = 0$, ello implica que también $h_{12} = h_{21} = 0$, es decir, que $\mathcal{H}f(a, b)$ es la matriz nula. Por tanto, en el caso $\Delta = 0$, la condición de que $\mathcal{H}f(a, b)$ sea la matriz nula puede sustituirse por la condición $h_{11} = h_{22} = 0$.

La segunda obervación concierne al apartado 3. En este caso, $h_{11}h_{22} - h_{12}^2 = \triangle > 0$ implica $h_{11}h_{22} = \triangle + h_{12}^2 > 0$, luego h_{11} y h_{22} son ambos distintos de 0 y del mismo signo. Por tanto, en cada uno de los dos apartados de 3 las dos condiciones sobre los signos h_{11} y h_{22} pueden sustituirse por sólo una de ellas.

La tercera también concierne al apartado 3. Notemos que la posibilidad $\triangle > 0$ y $h_{11} = 0$ no puede darse. En efecto, en este caso, tendríamos $0 < \triangle = h_{11}h_{22} - h_{12}^2 = -h_{12}^2$, lo que es contradictorio.

Tanto en el caso de $n > 2$ variables como en el de dos variables, el estudio de las derivadas segundas no siempre permite decidir sobre el carácter de máximo relativo, mínimo relativo o punto de silla de un punto crítico. En estos casos, consideraciones acerca de las propiedades particulares de la función concreta objeto de estudio permiten, a veces, determinar el carácter del punto crítico. Además, hay que tener en cuenta que una función puede tener un extremo relativo en un punto en el que no sea diferenciable, en cuyo caso la discusión anterior no es aplicable.

10.2 Extremos condicionados

A menudo, se desea encontrar los extremos relativos de una función de varias variables restringida a los puntos que cumplen cierta condición o condiciones. Una situación típica sucede cuando se recorre una curva sobre una superficie dada. La función cuya gráfica es la superficie puede no tener extremos relativos, pero sí puede haberlos cuando restringimos nuestra atención sólo a los puntos de la curva.

Sea U un abierto de $\mathbb{R}^n$, y consideremos el conjunto C de los puntos de U que verifican las ecuaciones $g_i(x_1, \ldots, x_n) = 0$ para $i = 1, \ldots m$, con $m < n$, denominadas *condiciones*, *restricciones* o *ligaduras*. Con la notación $\mathbf{x} = (x_1, \ldots, x_n)$ y $G = (g_1, \ldots, g_m)$, las condiciones pueden escribirse $G(\mathbf{x}) = \mathbf{0}$. Sea f una función real de dominio U y $\mathbf{a} \in C \cap U$. Se dice que en $\mathbf{a}$ la función f tiene un *máximo* (respectivamente, *mínimo*) *condicionado* por las mencionadas ligaduras si existe una bola $\mathcal{B}_r(\mathbf{a})$ tal que $f(\mathbf{x}) \leq f(\mathbf{a})$ (respectivamente, $f(\mathbf{x}) \geq f(\mathbf{a})$ para todo $\mathbf{x} \in \mathcal{B}_r(\mathbf{a}) \cap C$.

La técnica principal para hallar extremos relativos sujetos a ligaduras es la de los multiplicadores de Lagrange, que se basa en el siguiente teorema.

Teorema de Lagrange. Sean U un abierto de $\mathbb{R}^n$ y $f, g_1, \ldots, g_m$, con $m < n$, funciones reales de la clase $\mathcal{C}^1(U)$. Sean $G = (g_1, \ldots, g_m)$ y $C = \{\mathbf{x} \in U \ : \ G(\mathbf{x}) = \mathbf{0}\}$. Supongamos que $\mathbf{a} \in C$ es un extremo relativo de f condicionado por las ligaduras $G(\mathbf{x}) = \mathbf{0}$ y tal que el rango de la matriz jacobiana $\mathcal{J}G(\mathbf{a})$ es m. Entonces, existen números reales $\lambda_1, \ldots, \lambda_m$

tales que si

$$F(x_1,\ldots,x_n) = f(x_1,\ldots,x_n) + \sum_{j=1}^{m} \lambda_j g_j(x_1,\ldots,x_n),$$

el punto **a** es una solución del sistema de $n+m$ ecuaciones

$$\frac{\partial F}{\partial x_i}(x_1,\ldots,x_n) = 0, \quad 1 \leq i \leq n, \qquad g_j(x_1,\ldots,x_n) = 0, \quad 1 \leq j \leq m. \tag{1}$$

A efectos prácticos, se considera la función de $n+m$ variables

$$F^*(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) = f(x_1,\ldots,x_n) + \sum_{j=1}^{m} \lambda_j g_j(x_1,\ldots,x_n).$$

En condiciones de regularidad adecuadas,

$$\frac{\partial F^*}{\partial \lambda_j}(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) = g_j(x_1,\ldots,x_n),$$

de forma que las condiciones (1) pueden ponerse en la forma $\nabla F^* = \mathbf{0}$. Entonces, los puntos críticos $(x_1^*,\ldots,x_n^*,\lambda_1^*,\ldots,\lambda_m^*)$ de F^* proporcionan los candidatos a valores $\lambda_1,\ldots\lambda_m$ y a extremos $(x_1,\ldots,x_n)$. De este modo, se restringe el conjunto en el que se encuentran los extremos.

En general, no hay un criterio sencillo que permita decidir qué puntos de este conjunto son extremos y de qué clase. En el caso de que se puedan despejar m variables en el sistema de ligaduras en función de las $n-m$ restantes, se sustituyen en la expresión de la función f y queda una función de $n-m$ variables cuyos extremos relativos son los extremos condicionados buscados. A veces son las propias características geométricas o físicas del problema concreto en consideración las que permiten decidir sobre el carácter del punto crítico.

10.3 Extremos absolutos

Recordemos que, según el teorema de Weierstrass, si una función real es continua en un compacto de $\mathbb{R}^n$, entonces admite máximo y mínimo absolutos en él. La teoría garantiza la existencia de estos extremos absolutos, pero hallarlos de forma efectiva no siempre es fácil.

Las técnicas usuales consisten en hallar un conjunto tan restringido como sea posible de puntos en los que puedan alcanzarse el máximo y el mínimo absolutos, y decidir, mediante las consideraciones adecuadas, en cuáles se alcanza efectivamente el extremo absoluto.

Consideremos el caso particular, pero no infrecuente, de una función continua $f\colon K \to \mathbb{R}$, definida sobre un compacto K de $\mathbb{R}^2$ cuya frontera está definida por una curva $g(x,y) = 0$. En este caso, los extremos absolutos de f en K sólo pueden alcanzarse

(i) en los puntos críticos de f en el interior de K;
(ii) en los puntos del interior de K en los que f no es diferenciable;
(iii) en los extremos de f condicionados por $g(x,y) = 0$.