Optimización sin restricciones

Jesús Getán y Eva Boj

Facultat d'Economia i Empresa
Universitat de Barcelona

Marzo de 2021

El problema general de optimización sin restricciones se puede enunciar así:

Dada una función

$$\begin{aligned} f : F \subseteq \mathbb{R}^n &\longrightarrow \mathbb{R} \quad , \\ \vec{x} &\mapsto z = f(\vec{x}) \end{aligned}$$

encontrar los valores $\vec{x} \in F$, tales que maximizan o minimizan el valor de la función en F.

A la función f le llamaremos función objetivo y al conjunto F conjunto factible.

Proponemos la formulación general de un programa matemático mediante:

$$\begin{array}{ll} \text{Opt} & f(\vec{x}), \\ \text{sujeta a:} & \vec{x} \in F, \end{array}$$

donde la palabra **Opt** se interpretará como máximo o mínimo según sea el problema que tratemos.

Puesto que estamos hablando de buscar máximos y mínimos de funciones, el primer paso a realizar es caracterizarlos.

Definition

Sea la función $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ donde D es el domino de la función y un punto $\vec{x}^o \in D$.
Decimos que:

El punto $\vec{x}^o$ es un máximo local de f en $D \Leftrightarrow$
$\exists\ \delta > 0$ tal que $f(\vec{x}) \leq f(\vec{x}^o)$ para todo $\vec{x} \in D \cap B_\delta(\vec{x}^o)$.

El punto $\vec{x}^o$ es un máximo global de f en $D \Leftrightarrow$
$f(\vec{x}) \leq f(\vec{x}^o)$ para todo $\vec{x} \in D$.

Definition

El punto $\vec{x}^o$ es un mínimo local de f en $D \Leftrightarrow$
$\exists\ \delta > 0$ tal que , $f(\vec{x}) \geq f(\vec{x}^o)$ para todo $\vec{x} \in D \cap B_\delta(\vec{x}^o)$.

El punto $\vec{x}^o$ es un mínimo global de f en $D \Leftrightarrow$
$f(\vec{x}) \geq f(\vec{x}^o)$ para todo $\vec{x} \in D$.

Observamos que un punto es máximo global cuando su imagen es mayor que la de cualquier otro punto del dominio; en cambio, un punto es máximo local cuando su imagen es mayor que la imagen de cualquier punto que se halle en un cierto entorno suyo, aunque, fuera de este entorno pueden existir puntos con mayor imagen que él.

a) Cuando $\vec{x} \neq \vec{x}_o$ y las desigualdades son estrictas, decimos que los óptimos son estrictos.

b) Todo óptimo global (o absoluto) es también óptimo local (o relativo).

Theorem

Sean $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ *y* $\vec{x}^o \in D$.

Si $\vec{x}^o$ *es un máximo o mínimo global de* f *en* D *entonces,* $\vec{x}^o$ *es un máximo o mínimo local de* f *en* D *respectivamente.*

Los preguntas básicas que nos podemos plantear para estudiar tras haber modelizado ciertas situaciones como programas matemáticos, son los siguientes:

- ¿Cuándo podemos afirmar que un programa matemático tiene solución?, es decir, ¿cuándo podemos asegurar la existencia de óptimos locales o globales?
- Suponiendo que existen, ¿cómo se calculan estos óptimos?

Para resolver el segundo problema, de momento, consideraremos un método geométrico consistente en dibujar el conjunto factible del problema junto con las curvas de nivel de la función y analizar los valores que toma la función sobre los puntos del conjunto factible para determinar los valores máximo y mínimo. Este método sólo se puede llevar a cabo con funciones y conjuntos factibles de una o dos variables y, aún así, el trazado de las curvas de nivel de la función sobre el conjunto factible puede ser complicado. Más adelante, y en el caso de funciones diferenciables, encontraremos condiciones para calcular los óptimos de la función.

En cuanto a la primera cuestión, la existencia y naturaleza de los óptimos depende fundamentalmente de las características de la función y del dominio.

Será objeto de estudio en otras secciones.

Dada la función $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, definimos la curva de nivel para $k \in \mathbb{R}$ como el conjunto C_k de puntos del dominio donde función f tiene valor constante k, es decir,

$$C_k = \{\vec{x} \in D \mid f(\vec{x}) = k\}.$$

Su representación sólo es posible para $n = 2$ y $n = 3$.

Geométricamente podemos localizar los óptimos de un programa matemático encontrando los puntos del conjunto factible por los que pasa la curva de nivel de valor máximo o mínimo.

Introducimos un teorema de existencia de solución de un problema de optimización que depende de las características de la función objetivo y del dominio. Observamos los siguientes ejemplos:

Example

Sea la función $\begin{array}{rl} f : [0,1) \subset \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto f(x) = x^2. \end{array}$ en $[0,1)$ tiene un mínimo global en $x = 0$ y no tiene máximo global.

Example

La función $\begin{aligned} f : [0,1] \subset \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = x^2. \end{aligned}$ en $[0,1]$ tiene un mínimo global en $x = 0$ y un máximo global en $x = 1$.

Example

La función $f : [0, 2] \subset \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$f(x) = \begin{cases} x^2 & \text{si } 0 \leq x < 1, \\ \frac{1}{2} & \text{si } 1 \leq x \leq 2, \end{cases}$$

tiene un mínimo global en $x = 0$ y no tiene máximo.

Estos ejemplos nos muestran que si el dominio no es compacto (cerrado y acotado) o la función objetivo no es continua en el dominio, el programa matemático puede no tener solución, el teorema siguiente nos asegura que en caso contrario podemos afirmar la existencia de óptimos globales.

Teorema de Weierstrass

Theorem

Sea D un subconjunto compacto de $\mathbb{R}^n$ y $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función continua en D. Entonces, f posee un máximo global y un mínimo global en D, es decir,

$$\exists \quad \vec{x}^1, \vec{x}^2 \in D \text{ tal que } f(\vec{x}^1) \leq f(\vec{x}) \leq f(\vec{x}^2) \qquad (\vec{x} \in D).$$

El Teorema de Weierstrass nos asegura la existencia de óptimos globales bajo ciertas condiciones pero no afirma que el óptimo sea único (la función $y = \sin x$ en el intervalo $[0, 4\pi]$ tiene dos máximos globales y dos mínimos globales).

El Teorema de Weierstrass tampoco dice que la continuidad de la función y la compacidad del dominio sean necesarias para la existencia de óptimos globales. La función

$$f(x) = \begin{cases} \dfrac{4x}{x^2+4} & \text{si } -\infty \leq x < 0, \\ \frac{1}{2} & \text{si } x = 0, \\ \dfrac{4x}{x^2+4} & \text{si } 0 < x \leq +\infty. \end{cases}$$

tiene máximo global y mínimo global y sin embargo la función f no es continua y el dominio no es compacto.

Dada $z = f(x_1, x_2)$ tal que $f \in \mathcal{C}^2$, en primer lugar caracterizaremos los óptimos en $\vec{x}^o \in D \subseteq \mathbb{R}^n$ geométricamente, es decir mediante el plano tangente a la función en $\vec{x}^o$. La ecuación del plano tangente a f en $\vec{x}^o$ es:

$$z = f(x_1^o, x_2^o) + \left(\frac{\partial}{\partial x_1} f(x_1^o, x_2^o), \frac{\partial}{\partial x_2} f(x_1^o, x_2^o) \right) \begin{pmatrix} x_1 - x_1^o \\ x_2 - x_2^o \end{pmatrix}$$

escrito de otra forma

$$z = f(x_1^o, x_2^o) + \frac{\partial}{\partial x_1} f(x_1^o, x_2^o) (x_1 - x_1^o) + \frac{\partial}{\partial x_2} f(x_1^o, x_2^o) (x_2 - x_2^o) .$$

Si $\vec{x}^o \in D \subseteq \mathbb{R}^n$ es un óptimo de f en D, entonces el plano tangente es horizontal, es decir,

$$\frac{\partial}{\partial x_1} f(x_1^o, x_2^o) = 0 \text{ y } \frac{\partial}{\partial x_2} f(x_1^o, x_2^o) = 0,$$

luego $z = f(x_1^o, x_2^o)$.

Generalizando lo anterior podemos enunciar el siguiente teorema para la función $f : D \subset \mathbb{R}^n \to \mathbb{R}$, diferenciable en D. Vamos a dar condiciones necesarias o suficientes para que un punto del interior del dominio D sea óptimo local de la función f.

Si D es abierto estas condiciones pueden aplicarse a todos los puntos del dominio D.

Theorem

Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *difenciable en* D, *con* D *abierto. Si* $\vec{x}^o \in D$ *es un óptimo local de* f *en* D, *entonces* $\nabla f(\vec{x}^o) = \vec{0}$.

DEM: Sea $\vec{x}^o \in D$ es un máximo local, sus derivadas parciales son

$\frac{\partial}{\partial x_i} f(x_1^o, x_2^o) = \lim_{h\to 0} \frac{f(\vec{x}^o + h\vec{e}_i) - f(\vec{x}^o)}{h}$ y $i = 1...n$. Entonces, para cada i tenemos

$$\lim_{h\to 0^+} \frac{f(\vec{x}^o + h\vec{e}_i) - f(\vec{x}^o)}{h} \qquad \left(\frac{\text{numerador } < 0}{\text{denominador } > 0}\right) = < 0,$$

$$\lim_{h\to 0^-} \frac{f(\vec{x}^o + h\vec{e}_i) - f(\vec{x}^o)}{h} \qquad \left(\frac{\text{numerador } > 0}{\text{denominador } > 0}\right) = > 0,$$

como el límite existe, la única posibilidad es que sea cero. Por inducción se acaba la demostración del teorema. $\square$

El recíproco no es cierto, como demuestra el ejemplo,

$f(x_1, x_2) = x_1 x_2$ en $\vec{x}^o = (0, 0)$,

en el cual se anula el gradiente pero no es óptimo es punto de silla.

Llamamos **puntos críticos** a los puntos en los que se anula el gradiente de la función objetivo.

Theorem

Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$, *con* $f \in \mathcal{C}^2$ *y* D *abierto. Sea* $\vec{x}^o \in D$ *tal que* $\nabla f(\vec{x}^o) = \vec{0}$. *Entonces,*

el punto $\vec{x}^o \in D$ *es mínimo local de* f *en* $D \Rightarrow Hf(\vec{x}^o)$ *semidefinida positiva.*

el punto $\vec{x}^o \in D$ *es máximo local de* f *en* $D \Rightarrow Hf(\vec{x}^o)$ *semidefinida negativa.*

El recíproco no es cierto, como demuestra el ejemplo,

$f(x_1, x_2) = x_1^3 + x_2^2$ en $\vec{x}^o = (0, 0)$,

en el cual se anula el gradiente, la hessiana es semidefinida positiva pero no es mínimo, es punto de silla.

Definition

Dado $\vec{x}^o \in D$ con D abierto. El punto $\vec{x}^o$ es un punto de silla de f en D $\Leftrightarrow$ se cumplen las dos condiciones siguientes

$$\begin{cases} (i)\, \vec{x}^o \text{ es punto crítico,} \\ (ii)\, \forall r > 0\,,\ \exists \vec{x}^1\,,\ \vec{x}^2 \in B_r(\vec{x}^o) \cap D \text{ tales que } f(\vec{x}^1) < f(\vec{x}^o) < f(\vec{x}^2). \end{cases}$$

Theorem

Sea $f \in \mathcal{C}^2$. *Sea* $\vec{x}^o \in D$ *con* D *conjunto abierto tal que* $\nabla f(\vec{x}^o) = \vec{0}$.

Que $Hf(\vec{x}^o)$ *sea definida positiva* $\Rightarrow \vec{x}^o$ *es mínimo local.*

Que $Hf(\vec{x}^o)$ *sea definida negativa* $\Rightarrow \vec{x}^o$ *es máximo local.*

Que $Hf(\vec{x}^o)$ *sea semidefinida positiva en todo un entorno de* $\vec{x}^o \Rightarrow \vec{x}^o$ *es mínimo local.*

Que $Hf(\vec{x}^o)$ *sea semidefinida negativa en todo un entorno de* $\vec{x}^o$ $\Rightarrow \vec{x}^o$ *es máximo local.*

Que $Hf(\vec{x}^o)$ *sea indefinida* $\Rightarrow \vec{x}^o$ *es punto de silla.*

El teorema recíproco no es cierto, como muestran los siguientes ejemplos.

(i) $f(x_1, x_2) = x_1^4 + x_2^4$ en $\vec{x}^o = (0, 0)$,

(ii) $f(x_1, x_2) = x_1^3 + x_2^2$ en $\vec{x}^o = (0, 0)$.

En (i) la hessiana es semidefinida positiva y en el punto la función alcanza su valor mínimo.

En (ii) la hessiana es semidefinida positiva y en el punto la función posee un punto de silla.