Fundamentos de Optimización Continua

Alan Reyes-Figueroa
Modelación y Simulación

(Aula 20) 03.octubre.2024

Optimización Continua

Queremos resolver problemas de optimización no restricta.

Problema de Optimización:
Resolvemos el problema

$$\min_{x \in \Omega} f(x), \tag{1}$$

donde $f : \mathbb{R}^n \to \mathbb{R}$ es la **función objetivo**.

Aquí, $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ es un vector de variables independientes. Estas variables usualmente se llaman las **variables de decisión**.
EL conjunto $\Omega \subseteq \mathbb{R}^n$ se llama *conjunto factible* o **conjunto de restricciones**. En el caso de optimización no restricta, Ω es el dominio de la función f, (por lo general $\Omega = \mathbb{R}^n$).

Casi siempre, requerimos que f posea alguna propiedad de interés. Por ejemplo, f es diferenciable, f es convexa, etc. Por lo general, en la **optimización continua** se diseñan métodos y algoritmos para optimizar funciones diferenciables f diferenciables (aunque esto no es un requisito indispensable).

Ejemplos:

$$\min_{\mathbf{x}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}, \quad \text{donde } A \in \mathbb{R}^{n\times n} \text{ es simétrica.}$$

(cociente de Rayleigh)

$$\min_{\mathbf{x}\in\mathbb{R}^n} \sum_{i=1}^{n}(x_i - y_i)^2 + \lambda \sum_{i=1}^{n-1}(x_{i+1} - x_i)^2.$$

(mínimos cuadrados con regularización de Tychonoff)

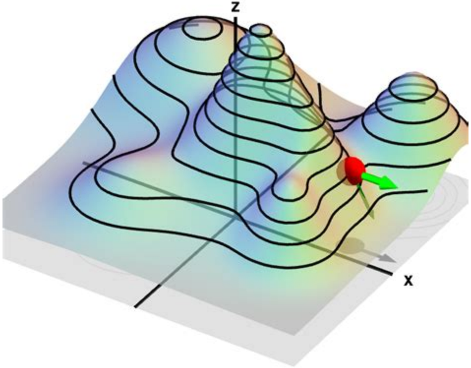
$$\min_{\mathbf{x}\in\mathbb{R}^n} \sum_{i=1}^{n-1} \left[(x_{i+1} - x_i^2)^2 + (1 - x_i)^2\right].$$

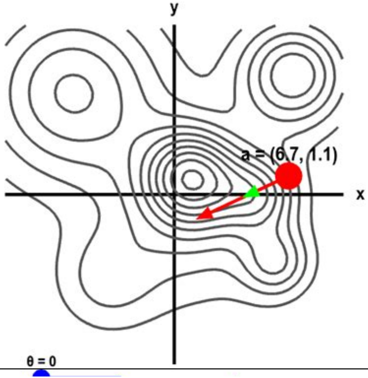
(función de ROSENBROCK).

Gradiente

Para una función diferenciable $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, recordemos que el **gradiente** de f en el punto $\mathbf{p} \in \Omega$ es el vector

$$\nabla f(\mathbf{p}) = \Big(\frac{\partial f}{\partial x_1}(\mathbf{p}), \frac{\partial f}{\partial x_2}(\mathbf{p}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{p})\Big)^T \in \mathbb{R}^n.$$





Gradiente

Propiedades del gradiente:

- El gradiente, $\nabla_{\mathbf{x}} f(\mathbf{p})$, de una función diferenciable, en el punto $\mathbf{p}$, es ortogonal al conjunto de nivel de la función f en ese punto.
- El vector de gradiente apunta en la dirección de máxima tasa de aumento de la función y el negativo del gradiente apunta en la dirección de la tasa máximo descenso de la función.
- La longitud del vector de gradiente nos dice la tasa de aumento en la dirección de aumento máximo y su negativo nos dice la tasa de disminución en la dirección de la disminución máxima.
- Similarmente, la magnitud de la derivada direccional $|\nabla_{\mathbf{x}} f(\mathbf{p})^T \mathbf{u}|$ indica la tasa de aumento/reducción de f en la dirección de $\mathbf{u}$.

Notaciones Importantes

En muchos métodos de optimización, se requiere información sobre la primera o segunda derivada de f.

La **derivada** o **Jacobiana** de f en el punto $\mathbf{p} \in \mathbb{R}^n$, es

$$Df(\mathbf{p}) = Jf(\mathbf{p}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{p}) & \frac{\partial f}{\partial x_2}(\mathbf{p}) & \dots & \frac{\partial f}{\partial x_n}(\mathbf{p}) \end{pmatrix}.$$

El **Hessiano** de f es

$$D^2 f(\mathbf{p}) = Hf(\mathbf{p}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{p}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{p}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{p}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{p}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{p}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{p}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{p}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{p}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{p}) \end{pmatrix}.$$

Obs! Si $f \in C^2$, vale la igualdad de las segundas parciales mixtas, y $D^2 f(\mathbf{p})$ es simétrica.

Derivadas Vectoriales

En el caso general de $f : \mathbb{R}^n \to \mathbb{R}^m$, con $f = (f_1, \dots, f_m)^T \in \mathbb{R}^m$, la derivada de f es

$$Df(\mathbf{p}) = Jf(\mathbf{p}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{p}) & \frac{\partial f_1}{\partial x_2}(\mathbf{p}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{p}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{p}) & \frac{\partial f_2}{\partial x_2}(\mathbf{p}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{p}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{p}) & \frac{\partial f_m}{\partial x_2}(\mathbf{p}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{p}) \end{pmatrix}.$$

Valen las propiedades ya conocidas del cálculo. En particular, vale la pena recordar la

Regla de la Cadena:
Si $f : \mathbb{R}^m \to \mathbb{R}^k$ y $g : \mathbb{R}^n \to \mathbb{R}^m$, osn funciones diferenciables en $g(\mathbf{p}) \in \mathbb{R}^m$ y $\mathbf{p} \in \mathbb{R}^m$, respectivamente, entonces $f \circ g : \mathbb{R}^n \to \mathbb{R}^k$ es diferenciable en $\mathbf{p}$ y

$$D(f \circ g)(\mathbf{p}) = Df\big(g(\mathbf{p})\big) \cdot Dg(\mathbf{p}).$$

Tipos de Extremos

Definición

Suponga que $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ *es una función con valores reales, definida sobre* Ω*. Un punto* $\mathbf{x}^* \in \Omega$ *es un* ***mínimo local*** *o* ***minimizador local*** *de* f *si existe* $\varepsilon > 0$ *tal que*

$$f(\mathbf{x}) \geq f(\mathbf{x}^*), \qquad \textit{para todo } \mathbf{x} \in \Omega - \{\mathbf{x}^*\} \textit{ con } ||\mathbf{x} - \mathbf{x}^*|| < \varepsilon.$$

Definición

Suponga que $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ *es una función con valores reales, definida sobre* Ω*. Un punto* $\mathbf{x}^* \in \Omega$ *es un* ***mínimo global*** *o* ***minimizador global*** *de* f *sobre* Ω *si*

$$f(\mathbf{x}) \geq f(\mathbf{x}^*), \qquad \textit{para todo } \mathbf{x} \in \Omega, \ \mathbf{x} \neq \mathbf{x}^*.$$

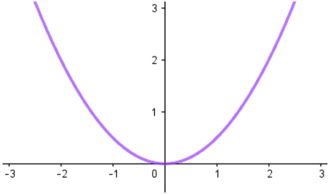
Obs! Reemplazando $\geq$ con $>$ en las definiciones anteriores obtenemos el concepto de un **mínimo local estricto** y de un **mínimo global estricto**, respectivamente.

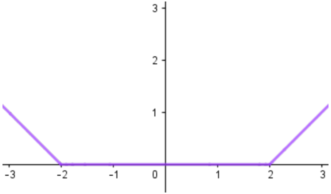
Ejemplos

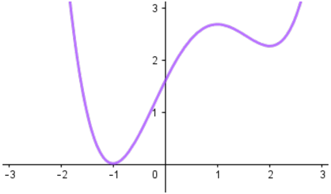
Ejemplo: La función $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ tiene un mínimo global estricto en $x = 0$. Claramente, $x = 0$ también es un mínimo local de f.

Ejemplo: La función $f : \mathbb{R} \to \mathbb{R}$, dada por $f(x) = \max\{0, |x - 2|\}$ tiene a todos los puntos de intervalo $[-2, 2]$ como mínimos globales. Estos no son estrictos.

Ejemplo: La función $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3x^4 - 8x^3 - 6x^2 + 12x$ tiene mínimos locales estricto en $x = -1$ y $x = 2$. De éstos, sólo $x = -1$ es un mínimo global.





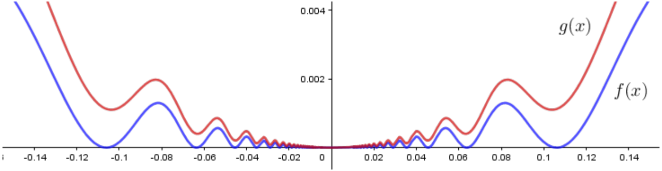


Definición

Un punto $\mathbf{x}^* \in \Omega$ *es un* ***mínimo local aislado*** *de* $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, *si* $\mathbf{x}^*$ *es mínimo local de* f *y existe una vecindad* $U \subset \mathbb{R}^n$ *de* $\mathbf{x}^*$ *tal que* $\mathbf{x}^*$ *es el único mínimo local de* f *en* U.
Un punto $\mathbf{x}^* \in \Omega$ *es un* ***mínimo local no aislado*** *de* f, *si para toda vecindad* U *de* $\mathbf{x}^*$, *existe* $\mathbf{x} \in U$, $\mathbf{x} \neq \mathbf{x}^*$, *tal que* $\mathbf{x}$ *también es mínimo local de* f.

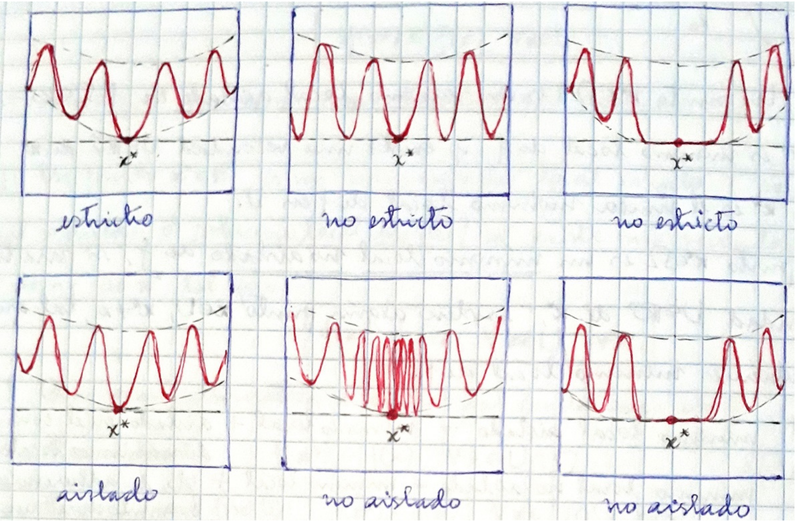
Ejemplo: La función $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 \cos \frac{1}{x} + x^2$, $f(0) = 0$, posee un mínimo local no estricto y no aislado en $x = 0$.

Ejemplo: La función $g : \mathbb{R} \to \mathbb{R}$, $g(x) = x^2 \cos \frac{1}{x} + 2x^2$, $g(0) = 0$, posee un mínimo local estricto y no aislado en $x = 0$.



UNIVERSIDAD DEL VALLE DE GUATEMALA

Ejemplos

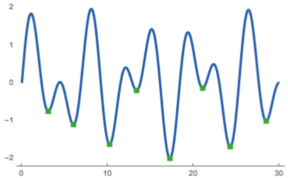


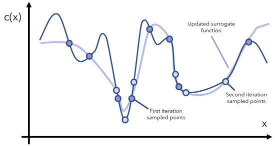
x^*
estricto
x^*
no estricto
x^*
no estricto
x^*
aislado
x^*
no aislado
x^*
no aislado

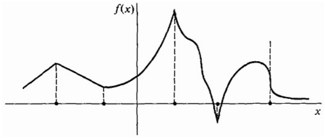
Dificultades

Dificultades de la optimización global:

- funciones con muchos mínimos;
- los algoritmos tienden a quedarse atrapados en mínimos locales,
- en el caso de métodos de búsqueda, puede que el óptimo global esté en una región no explorada;
- cuando el mínimo se encuentra dentro de una región donde la función es muy plana (curvatura cercana a 0), los métodos de optimización suelen ser muy lentos;
- no-diferenciabilidad en un punto mínimo.







UVG UNIVERSIDAD DEL VALLE DE GUATEMALA