Optimización

Formulario · Primavera 2021

Carlos Lezama MAT - 22211

ITAM

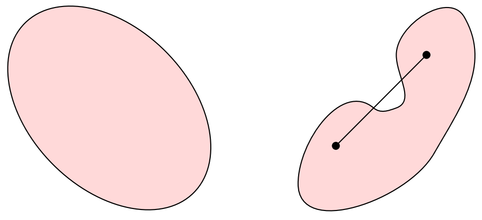
1. Optimización estática

1.1. Análisis convexo

Definición 1.1 (Conjunto convexo). *Sea $X \subseteq \mathbb{R}^n$, decimos que X es **convexo** si, para cualesquiera $\mathrm{x}, \mathrm{y} \in X$ y para toda $\lambda \in (0,1)$, se cumple:*

$$\lambda \mathrm{x} + (1-\lambda)\mathrm{y} \in X.$$

*Equivalentemente, decimos que X es **convexo** si, para todas $a \in \partial X$ y $b \in X$, existe ℓ tal que $\langle b - a, \ell \rangle \leq 0$; donde ∂X es la frontera de X y $\langle \cdot, \cdot \rangle$ denota el producto punto.*



Conjunto convexo   Conjunto no convexo

Proposición 1.1

Sean A y B dos subconjuntos convexos de $\mathbb{R}^n$, entonces:

(I) $A \cap B$ es convexo.
(II) $A + B = \{a + b : a \in A,\ b \in B\}$ es convexo.
(III) Para todo $k \in \mathbb{R}$, $kA = \{ka : a \in A\}$ es convexo.

Definición 1.2 (Función convexa). *Sea $X \subseteq R^n$ un conjunto convexo, $f : X \to \mathbb{R}$ es una **función convexa** si, para toda $\mathrm{x}_1 \neq \mathrm{x}_2 \in X$ y toda $\lambda \in (0,1)$, se tiene:*

$$f(\lambda \mathrm{x}_1 + (1-\lambda)\mathrm{x}_2) \leq \lambda f(\mathrm{x}_1) + (1-\lambda) f(\mathrm{x}_2).$$

*Si la desigualdad es estricta, se dice que la función es **estrictamente convexa**.*

Definición 1.3 (Función cóncava). *Sea $X \subseteq R^n$ un conjunto convexo, $f : X \to \mathbb{R}$ es una **función cóncava** si, para toda $\mathrm{x}_1 \neq \mathrm{x}_2 \in X$ y toda $\lambda \in (0,1)$, se tiene:*

$$f(\lambda \mathrm{x}_1 + (1-\lambda)\mathrm{x}_2) \geq \lambda f(\mathrm{x}_1) + (1-\lambda) f(\mathrm{x}_2).$$

*Si la desigualdad es estricta, se dice que la función es **estrictamente cóncava**.*

Proposición 1.2

Sean $X \subseteq \mathbb{R}^n$ un conjunto convexo, $f : X \to \mathbb{R}$ y $g : X \to \mathbb{R}$ dos funciones cóncavas, y $\alpha \in \mathbb{R}$, entonces:

(I) f es cóncava si $\alpha > 0$.
(II) f es convexa si $\alpha < 0$.
(III) $f + g$ es cóncava.

Proposición 1.3

Sean $X \subseteq \mathbb{R}^n$ un conjunto convexo, $g : X \to \mathbb{R}$ una función cóncava, y $h : Y \to \mathbb{R}$ una función cóncava y creciente tal que $g(X) \subseteq Y \subseteq \mathbb{R}$; entonces, $h \circ g$ es cóncava.

Definición 1.4 (Vector gradiente). *Sea $f \in \mathcal{C}^1(X)$, el **vector gradiente** de f está dado por:*

$$\nabla f(\mathrm{x}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}.$$

Definición 1.5 (Matriz hessiana). *Sea $f \in \mathcal{C}^2(X)$, se define la **matriz hessiana** de f como $H_f(\mathrm{x})$, donde:*

$$H_f(\mathrm{x})_{i,j} = \frac{\partial^2 f(\mathrm{x})}{\partial x_i \partial x_j}.$$

Definición 1.6 (Serie de Taylor).

$$T(\mathrm{x}) = \sum_{|\alpha| \geq 0} \frac{(\mathrm{x}-\mathrm{a})^\alpha}{\alpha!} (\partial^\alpha f)(\mathrm{a}).$$

Teorema 1.1: de Taylor

Sea $f : \mathbb{R}^n \to \mathbb{R}$ tal que $f \in \mathcal{C}^k(\mathrm{a})$. Entonces, existe $h_\alpha : \mathbb{R}^n \to \mathbb{R}$ tal que:

$$f(\mathrm{x}) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(\mathrm{a})}{\alpha!}(\mathrm{x}-\mathrm{a})^\alpha + \sum_{|\alpha| = k} h_\alpha(\mathrm{x})(\mathrm{x}-\mathrm{a})^\alpha,$$

y

$$\lim_{\mathrm{x} \to \mathrm{a}} h_\alpha(\mathrm{x}) = 0.$$

Definición 1.7 (Matriz simétrica). *Decimos que una matriz $A \in \mathcal{M}_{n\times n}$ es **simétrica** si y solo si:*

$$A = A^T.$$

Definición 1.8 (Matriz diagonalizable). *Decimos que una matriz $A \in \mathcal{M}_{n\times n}$ es **diagonalizable** si y solo si existe una matriz $P \in \mathcal{M}_{n\times n}$ invertible tal que $P^{-1}AP$ es diagonal.*

Definición 1.9 (Matriz ortogonalmente diagonalizable). *Decimos que una matriz $A \in \mathcal{M}_{n\times n}$ es **ortogonalmente diagonalizable** si y solo si existe una matriz $T \in \mathcal{M}_{n\times n}$ invertible tal que $T^{-1}AT$ es diagonal y $T^{-1} = T^T$.*

Teorema 1.2

Si $A \in \mathcal{M}_{n\times n}$ es simétrica, sus valores propios son reales.

Teorema 1.3

Una matriz simétrica A de tamaño $n \times n$ puede determinar la forma cuadrática q_A de n variables como sigue:

$$q_A(\mathrm{x}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \mathrm{x}^T A \mathrm{x}.$$

Corolario 1.1: Clasificación de formas cuadráticas

La matriz asociada a la forma cuadrática q_A es:

(I) definida positiva si $q_A(\mathrm{x}) > 0,\ \forall \mathrm{x} \neq 0$.
(II) definida negativa si $q_A(\mathrm{x}) < 0,\ \forall \mathrm{x} \neq 0$.
(III) semidefinida positiva si $q_A(\mathrm{x}) \geq 0,\ \forall \mathrm{x} \neq 0$.
(IV) semidefinida negativa si $q_A(\mathrm{x}) \leq 0,\ \forall \mathrm{x} \neq 0$.
(V) indefinida si $q_A(\mathrm{x})$ toma tanto valores positivos como negativos.

Definición 1.10 (Menores principales). *Sea A una matriz simétrica, los **menores principales** de esta matriz son los determinantes de todas las submatrices superiores izquierdas, es decir:*

$$|A_1| = |(a_{11})|,$$

$$|A_2| = \left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right|,$$

$$|A_3| = \left| \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \right|, \ \ldots$$

Teorema 1.4: Criterio de menores principales

Sea A una matriz simétrica asociada a la forma cuadrática $q_A(\mathrm{x})$, entonces:

(I) q_A es definida positiva si y solo si los menores principales de A son todos positivos.

(II) q_A es definida negativa si y solo si los menores principales de A alternan signos de la forma:

$$|A_1| < 0, |A_2| > 0, |A_3| < 0, \ldots.$$

(III) q_A es indefinida si $|A| \neq 0$, pero no se cumplen (I) ni (II).

Teorema 1.5: Criterio de valores propios

Sea A uns matriz simétrica asociada a la forma cuadrática $q_A(\mathrm{x})$, entonces:

(I) q_A es definida positiva si y solo si todos los valores propios de A son positivos.

(II) q_A es definida negativa si y solo si todos los valores propios de A son negativos.

(III) q_A es semidefinida positiva si y solo si todos los valores propios de A son no negativos.

(IV) q_A es semidefinida negativa si y solo si todos los valores propios de A son no positivos.

(V) q_A es indefinida si y solo si la matriz A tiene valores propios positivos y negativos.

Teorema 1.6: Criterio del primer orden

Sea $X \subseteq \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$ tal que $f \in \mathcal{C}^1(X)$, entonces:

(I) f es convexa si y solo si, $\forall \mathrm{x}, \mathrm{y} \in X$, se tiene:

$$f(\mathrm{y}) \geq f(\mathrm{x}) + \nabla f(\mathrm{x})(\mathrm{y} - \mathrm{x}).$$

(II) f es cóncava si y solo si, $\forall \mathrm{x}, \mathrm{y} \in X$, se tiene:

$$f(\mathrm{y}) \leq f(\mathrm{x}) + \nabla f(\mathrm{x})(\mathrm{y} - \mathrm{x}).$$

Teorema 1.7: Criterio de segundo orden

Sea $X \subseteq \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$ tal que $f \in \mathcal{C}^2(X)$, entonces:

(I) f es convexa en X si y solo si la forma cuadrática asociada a la matriz hessiana es semidefinida positiva.

(II) f es estrictamente convexa en X si y solo si la forma cuadrática asociada a la matriz hessiana es definida positiva.

(III) f es cóncava en X si y solo si la forma cuadrática asociada a la matriz hessiana es semidefinida negativa.

(IV) f es estrictamente cóncava en X si y solo si la forma cuadrática asociada a la matriz hessiana es definida negativa.

Definición 1.11. *Sean $X \subseteq \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$, definimos:*

- *la **gráfica** de f como*

$$G_f = \{(\mathrm{x}, r) \in X \times \mathbb{R} : f(\mathrm{x}) = r\}.$$

- *el **epígrafo** de f como*

$$E_f = \{(\mathrm{x}, r) \in X \times \mathbb{R} : f(\mathrm{x}) \leq r\}.$$

- *el **hipógrafo** de f como*

$$H_f = \{(\mathrm{x}, r) \in X \times \mathbb{R} : f(\mathrm{x}) \geq r\}.$$

Teorema 1.8

Sea $X \subseteq \mathbb{R}^n$ un conjunto convexo,

(I) una función $f : X \to \mathbb{R}$ es convexa si y solo si E_f es un conjunto convexo de $\mathbb{R}^{n+1}$.

(II) una función $f : X \to \mathbb{R}$ es cóncava si y solo si H_f es un conjunto convexo de $\mathbb{R}^{n+1}$.

Definición 1.12. *Sean $X \subseteq \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$, definimos:*

- *el **contorno** de f en k como*

$$C_f(k) = \{\mathrm{x} \in X : f(\mathrm{x}) = k\}.$$

- *el **contorno superior** de f en k como*

$$CS_f(k) = \{\mathrm{x} \in X : f(\mathrm{x}) \geq k\}.$$

- *el **contorno inferior** de f en k como*

$$CI_f(k) = \{\mathrm{x} \in X : f(\mathrm{x}) \leq k\}.$$

Teorema 1.9

(I) Si $f : X \to \mathbb{R}$ es cóncava en A, $CS_f(k)$ es convexo para toda k en la imagen de f.

(II) Si $f : X \to \mathbb{R}$ es convexa en A, $CI_f(k)$ es convexo para toda k en la imagen de f.

Teorema 1.10

Sean $X \subseteq \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$, entonces:

(I) f es cuasicóncava en A si $CS_f(k)$ es convexo para toda k en la imagen de f.

(II) f es cuasiconvexa en A si $CI_f(k)$ es convexo para toda k en la imagen de f.

Teorema 1.11

Sean $X \subseteq \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$, entonces:

(I) f es cuasicóncava si f es cóncava.

(II) f es cuasiconvexa si f es convexa.

Teorema 1.12

Cualquier transformación monótona creciente de una función cuasiconvexa es cuasiconvexa. Asimismo, cualquier transformación monótona creciente de una función cuasicóncava es cuasicóncava.

Teorema 1.13: Minkowski

Sean A y B dos conjuntos de $\mathbb{R}^n$ tales que $A \cap B = \varnothing$; entonces, existe un hiperplano que los separa.

Nota: este resultado es la base para desarrollar la *Teoría de Dualidad* en programación lineal.

Teorema 1.14: Punto fijo de Brouwer

Sean $A \subset \mathbb{R}^n$ un conjunto compacto y convexo, y $f : A \to A$ una función continua; entonces, f tiene un punto fijo.

$$\text{i.e.,} \quad \exists \mathrm{x} \in A \text{ tal que } f(\mathrm{x}) = \mathrm{x}.$$

Nota: este resultado permite mostrar la existencia de equilibrios competitivos en una economía de intercambio.

Teorema 1.15

Sean $A \subseteq \mathbb{R}^n$ un conjunto convexo, $f : A \to \mathbb{R}$ una función cóncava y x_0 un punto interior de A entonces, existe un vector $\bar{p} \in \mathbb{R}^n$ tal que:

$$f(\mathrm{x}) - f(\mathrm{x}_0) \leq \bar{p}\,(\mathrm{x} - \mathrm{x}_0), \quad \forall \mathrm{x} \in A.$$

Por definición, llamaremos a dicho vector $\bar{p}$ **súpergradiente**.
Asimismo, sea f una función convexa, existe un vector $\bar{q} \in \mathbb{R}^n$ tal que:

$$f(\mathrm{x}) - f(\mathrm{x}_0) \geq \bar{q}\,(\mathrm{x} - \mathrm{x}_0), \quad \forall \mathrm{x} \in A.$$

Por definición, llamaremos a dicho vector $\bar{q}$ **subgradiente**.

1.2. Optimización estática

Definición 1.13 (Máximo global). *Sea $X \subseteq \mathbb{R}^n$ un conjunto, decimos que $x_0 \in X$ es un **máximo global** de la función $f : X \to \mathbb{R}$ si, para toda $x \in X$, $f(x_0) \geq f(x)$. Si dicha desigualdad es estricta, se le conoce como **máximo global estricto**.*

Definición 1.14 (Mínimo global). *Sea $X \subseteq \mathbb{R}^n$ un conjunto, decimos que $x_0 \in X$ es un **mínimo global** de la función $f : X \to \mathbb{R}$ si, para toda $x \in X$, $f(x_0) \leq f(x)$. Si dicha desigualdad es estricta, se le conoce como **mínimo global estricto**.*

Definición 1.15 (Máximo local). *Sean (X, d_X) un espacio métrico y $f : X \to \mathbb{R}$ una función. Decimos que $x_0 \in X$ es un **máximo local** de la función f si existe $\varepsilon > 0$ tal que, para toda $x \in X$, $d_X(x, x_0) < \varepsilon$ y $f(x_0) \geq f(x)$. Si dicha desigualdad es estricta, se le conoce como **máximo local estricto**.*

Definición 1.16 (Mínimo local). *Sean (X, d_X) un espacio métrico y $f : X \to \mathbb{R}$ una función. Decimos que $x_0 \in X$ es un **mínimo local** de la función f si existe $\varepsilon > 0$ tal que, para toda $x \in X$, $d_X(x, x_0) < \varepsilon$ y $f(x_0) \leq f(x)$. Si dicha desigualdad es estricta, se le conoce como **mínimo local estricto**.*

Teorema 1.16: Condiciones necesarias de primer orden

Si $f(\mathrm{x})$ es derivable en x_0 y f alcanza un extremo local en x_0, entonces, todas sus derivadas parciales se anulan en x_0.

Definición 1.17 (Punto crítico). *Dados un conjunto $A \subseteq \mathbb{R}^n$ y una función $f : A \to \mathbb{R}$, decimos que x_0 es un **punto crítico** de f si sus derivadas parciales se anulan en x_0, o bien, alguna de ellas no existe en x_0.*

$$\exists\, \mathrm{x}_0 \quad \text{tal que} \quad \nabla f(\mathrm{x}_0) = 0.$$

Definición 1.18 (Punto de silla). *Un punto crítico x_0 es un **punto de silla** si existen direcciones por las que la función crece y otras por las que la función decrece a partir del punto x_0.*

Teorema 1.17: Condiciones suficientes de segundo orden

Sea x_0 un punto crítico de una función diferenciable f, entonces:

(i) f alcanza un máximo local si, en x_0, la forma cuadrática asociada a la matriz hessiana H_f es definida negativa.

(ii) f alcanza un mínimo local si, en x_0, la forma cuadrática asociada a la matriz hessiana H_f es definida positiva.

(iii) x_0 es un punto de silla si la forma cuadrática asociada a la matriz hessiana H_f es indefinida.

1.3. Optimización restringida

2. Cálculo de variaciones

3. Teoría de control óptimo

4. Elementos de programación dinámica