

Tema 4

Funciones convexas y optimización convexa

José R. Berrendero

Departamento de Matemáticas
Universidad Autónoma de Madrid

Contenidos del tema 4

- ▶ Repaso de algunos resultados sobre optimización de funciones.
- ▶ Funciones convexas. Caracterizaciones.
- ▶ Operaciones que preservan la convexidad.
- ▶ Resultados generales sobre optimización convexa.

Notación

Para una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$, denotamos

- ▶ **Derivada direccional:** sea $d \in \mathbb{R}^n$ con $d \neq 0$,

$$f'(x, d) = \lim_{\lambda \rightarrow 0} \frac{f(x + \lambda d) - f(x)}{\lambda}.$$

- ▶ **Derivadas parciales:** $f'_i(x) = f'(x, e_i)$ con $e_i = (0, \dots, 1, \dots, 0)^\top$.

- ▶ **Gradiente:** $\nabla f(x) = (f'_1(x), \dots, f'_n(x))^\top$.

- ▶ **Matriz Hessiana:**

$$Hf(x) = \begin{pmatrix} f''_{11}(x) & \cdots & f''_{1n}(x) \\ \vdots & \ddots & \vdots \\ f''_{n1}(x) & \cdots & f''_{nn}(x) \end{pmatrix}$$

Propiedades básicas

- ▶ f tiene derivadas parciales continuas $\Rightarrow f$ es diferenciable $\Rightarrow f$ tiene derivadas parciales.

- ▶ f diferenciable $\Rightarrow f'(x, d) = \nabla f(x)^\top d$.

- ▶ f es diferenciable en \bar{x} si y solo si

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}) + \|x - \bar{x}\| R(\bar{x}; x - \bar{x}),$$

donde $\lim_{x \rightarrow \bar{x}} R(\bar{x}; x - \bar{x}) = 0$.

- ▶ f es diferenciable dos veces en \bar{x} si y solo si

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^\top Hf(\bar{x})(x - \bar{x}) + \|x - \bar{x}\|^2 R(\bar{x}; x - \bar{x}),$$

donde $\lim_{x \rightarrow \bar{x}} R(\bar{x}; x - \bar{x}) = 0$.

- ▶ Si f es diferenciable en \bar{x} y $\nabla f(\bar{x})^\top d < 0$ (resp. > 0), entonces existe $\delta > 0$ tal que $f(\bar{x}) > f(\bar{x} + \lambda d)$ (resp. $<$) si $0 < \lambda < \delta$.

Condiciones para óptimos locales

Condición necesaria de primer orden: Sea f diferenciable en \bar{x} . Si \bar{x} es un mínimo o máximo local de f entonces $\nabla f(\bar{x}) = 0$.

Condiciones necesarias de segundo orden: Sea f dos veces diferenciable en \bar{x} . Si \bar{x} es un mínimo local de f entonces

1. $\nabla f(\bar{x}) = 0$
2. $Hf(\bar{x})$ es semidefinida positiva

Condiciones suficientes de segundo orden: Sea f dos veces diferenciable en \bar{x} . Si

1. $\nabla f(\bar{x}) = 0$
2. $Hf(\bar{x})$ es definida positiva

entonces \bar{x} es un mínimo local (estricto) de f .

Funciones convexas

Una **función** $f : D \rightarrow \mathbb{R}$ es **convexa** si su dominio $D \subset \mathbb{R}^n$ es convexo, y para todo $x, y \in D$, para todo $\lambda \in [0, 1]$, se verifica

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

- ▶ La **convexidad estricta** requiere $<$ en lugar de \leq .
- ▶ Una función f es **(estrictamente) cóncava** si $-f$ es (estrictamente) convexa.
- ▶ f es convexa si y solo si para todo x y u , $g(t) = f(x + tu)$ es convexa (en su dominio, es decir, $\{t : x + tu \in D\}$).

Algunos ejemplos

- ▶ $f(x) = x^2$ es convexa.
- ▶ $f(x) = \max\{x_1, \dots, x_n\}$ donde $x = (x_1, \dots, x_n)^\top$ es convexa.
- ▶ Las funciones afines $f(x) = Ax + b$ son cóncavas y convexas.
- ▶ Cualquier norma $f(x) = \|x\|$ es una función convexa.
- ▶ Si S es convexo, la distancia a S ,

$$f(x) = d(x, S) = \inf_{s \in S} \|x - s\|$$

es una función convexa.

Desigualdad de Jensen

Teorema: Si $f : D \rightarrow \mathbb{R}$ es convexa, entonces para todo x_1, \dots, x_k de su dominio y $\lambda_1, \dots, \lambda_k \geq 0$ con $\lambda_1 + \dots + \lambda_k = 1$ se cumple

$$f(\lambda_1 x_1 + \dots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \dots + \lambda_k f(x_k).$$

Demostración: Por inducción sobre k , observando que si $x = \lambda_1 x_1 + \dots + \lambda_{k+1} x_{k+1}$, entonces $x = \lambda y + (1 - \lambda) x_{k+1}$, donde $y = \sum_{i=1}^k (\lambda_i / \lambda) x_i$, con $\lambda = 1 - \lambda_{k+1}$.

Si X es una v.a. que toma los valores x_i con probabilidad λ_i , entonces $f[\mathbb{E}(X)] \leq \mathbb{E}[f(X)]$.

Epigrafo

El **epigrafo** de una función $f : D \rightarrow \mathbb{R}$ se define como:

$$\text{epi}(f) = \{(x, t) : x \in D, f(x) \leq t\} \subset \mathbb{R}^{n+1}.$$

Teorema: Una función es convexa si y solo si su epigrafo es un conjunto convexo.

Teorema: Sea $\{f_i : i \in I\}$ una familia de funciones convexas definidas sobre un conjunto convexo no vacío S tal que para todo $x \in S$, el conjunto $\{f_i(x) : i \in I\}$ está acotado superiormente. Entonces la función $f(x) = \sup\{f_i(x) : i \in I\}$, $x \in S$, es convexa.

Demostración: $\text{epi}(f) = \bigcap_{i \in I} \text{epi}(f_i)$.

Funciones convexas diferenciables

Teorema: Sea $f : D \rightarrow \mathbb{R}$ diferenciable sobre un dominio convexo y abierto D . Entonces, f es convexa si y solo si

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}),$$

para todo $x, \bar{x} \in D$.

Demostración:

(\Rightarrow) Si $x, \bar{x} \in D$ y f es convexa en D ,

$$f(\lambda x + (1 - \lambda)\bar{x}) \leq \lambda f(x) + (1 - \lambda)f(\bar{x}) \Leftrightarrow [f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x})]/\lambda \leq f(x) - f(\bar{x}).$$

Se toman límites cuando $\lambda \rightarrow 0$.

(\Leftarrow) Si $x \neq y$ y $\lambda \in [0, 1]$, se aplica la condición a x e y con $\bar{x} = \lambda x + (1 - \lambda)y$:

$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x})$ y $f(y) \geq f(\bar{x}) + \nabla f(\bar{x})(y - \bar{x})$. Se multiplica la primera desigualdad por λ , la segunda por $1 - \lambda$ y se suman.

Funciones convexas diferenciables dos veces

Teorema: Sea $f : D \rightarrow \mathbb{R}$ con dos derivadas continuas sobre un dominio convexo y abierto D . Entonces, f es convexa si y solo si $Hf(x)$ es semidefinida positiva para todo $x \in D$.

Demostración: Para $n = 1$ se demuestra en los ejercicios. Para $n > 1$ se considera $g(\lambda) = f(x + \lambda d)$ y se tiene en cuenta que $g''(\lambda) = d^\top Hf(x + \lambda d)d$.

- ▶ Una función dos veces diferenciable es cóncava si D es convexo y $Hf(x)$ es semidefinida negativa para todo $x \in D$.
- ▶ Si $Hf(x)$ es definida positiva para todo $x \in D$, la función es estrictamente convexa, pero el recíproco no es cierto (por ejemplo, $f(x) = x^4$).

Operaciones que preservan la convexidad

- ▶ El **supremo de funciones convexas** es una función convexa.
- ▶ **Sumas ponderadas no negativas:** Si f_i convexa y $w_i \geq 0$, para $i = 1, \dots, n$, entonces $f = w_1 f_1 + \dots + w_n f_n$ es convexa.
- ▶ **Composición con una aplicación afín:** $g(x) = f(Ax + b)$ es convexa si f es convexa.
- ▶ **Composición:** Sea $D \subset \mathbb{R}^n$ y $f : D \rightarrow \mathbb{R}$ convexa. Sea $I \subset f(D)$ un intervalo tal que $g : I \rightarrow \mathbb{R}$ es creciente y convexa. Entonces la composición $g \circ f$ es convexa.
- ▶ **Minimización parcial:** Si f es convexa en (x, y) y C es un conjunto convexo no vacío, la función $g(x) = \inf_{y \in C} f(x, y)$ es convexa en x .

Más ejemplos de funciones convexas

- ▶ $f(x) = -\log x$ es convexa.
- ▶ $f(x) = e^{ax}$ es convexa en \mathbb{R} para todo $a \in \mathbb{R}$.
- ▶ $f(x) = x^a$ ($x > 0$) es convexa si $a \leq 0$ o $a \geq 1$, y cóncava si $0 \leq a \leq 1$.
- ▶ $f(x) = x^\top Ax + a^\top x + c$, donde A es simétrica, es convexa si y solo si A es semidefinida positiva.
- ▶ Sea X una matriz $n \times p$ y sean $\beta \in \mathbb{R}^p$ e $y \in \mathbb{R}^n$, entonces $f(\beta) = \|y - X\beta\|$ es convexa, donde $\|\cdot\|$ es cualquier norma.

Problema general de optimización convexa

La *optimización convexa* trata el problema general de minimizar una función convexa, sobre un conjunto factible también convexo:

$$\text{minimizar } f(x) \text{ s.a. } x \in S, \quad (1)$$

donde $f : D \rightarrow \mathbb{R}$ es convexa y $S \subset D \subset \mathbb{R}^n$ es convexo.

Casos particulares:

- ▶ Optimización lineal.
- ▶ Optimización cuadrática.
- ▶ Mínimos cuadrados.

Teorema local-global

En un problema convexo no hay distinción entre mínimos globales y locales.

Teorema: Todo mínimo local de (1) es también un mínimo global de (1).

Demostración:

- ▶ Sea \bar{x} un mínimo local. Existe $R > 0$ tal que $f(\bar{x}) \leq f(x)$, para todo $x \in S \cap B(\bar{x}, R)$.
- ▶ Spg. que existe $y \in S$ tal que $f(y) < f(\bar{x})$.
- ▶ Consideramos $z = (1 - \theta)\bar{x} + \theta y$, para $\theta > 0$ suficientemente pequeño.
- ▶ $z \in S \cap B(\bar{x}, R)$ pero $f(z) < f(\bar{x})$.

Mínimos globales bajo convexidad y diferenciabilidad

Teorema: Consideremos el problema (1) en el que se supone además que f es diferenciable. Un punto $\bar{x} \in S$ es mínimo global de (1) si y solo si

$$\nabla f(\bar{x})^\top (x - \bar{x}) \geq 0, \text{ para todo } x \in S.$$

Demostración:

(\Leftarrow) Por la convexidad de f , para todo $x \in S$,

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^\top (x - \bar{x}) \geq f(\bar{x}).$$

(\Rightarrow)

1. Spg. existe $x \in S$ con $\nabla f(\bar{x})^\top (x - \bar{x}) < 0$. Entonces, $x - \bar{x}$ es una dirección de descenso local.
2. Existe $0 < \lambda < 1$ tal que $f(\bar{x}) > f(\bar{x} + \lambda(x - \bar{x})) = f(\lambda x + (1 - \lambda)\bar{x})$.
3. $\lambda x + (1 - \lambda)\bar{x} \in S$ porque S es convexo.
4. Por lo tanto, \bar{x} no es un mínimo global.

Observaciones

1. Si S es abierto (por ejemplo, $S = \mathbb{R}^n$) entonces la condición del teorema anterior se reduce a $\nabla f(\bar{x}) = 0$.
(Considera $x = \bar{x} - \lambda \nabla f(\bar{x})$, para $\lambda > 0$ suf. pequeño.)
2. Sin embargo, si S no es abierto la condición sirve para detectar óptimos en la frontera. ($\min x^2$ s.a. $1 \leq x \leq 2$.)
3. Si $\nabla f(\bar{x}) \neq 0$, entonces $-\nabla f(\bar{x})$ define un hiperplano soporte a S en \bar{x} .
4. Aplicación a un problema de optimización lineal.
5. Aplicación a un problema convexo con restricciones de igualdad:

$$\min f(x) \text{ s.a. } Ax = b,$$

donde f es convexa y diferenciable.

Ejemplo

Resuelve gráficamente el problema:

$$\begin{array}{ll} \text{minimizar} & (x_1 - 4)^2 + (x_2 - 6)^2 \\ \text{s.a.} & x_2 \geq x_1^2 \\ & x_2 \leq 4 \end{array}$$

Demuestra analíticamente que la solución obtenida gráficamente es el mínimo global del problema.

Una versión más explícita del problema convexo

En la práctica consideramos problemas convexos para los que el conjunto factible se conoce más explícitamente:

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{s.a.} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & a_i^\top x = b_i, \quad i = 1, \dots, p, \end{array}$$

donde las funciones f, f_1, \dots, f_n son convexas.

Equivalentemente,

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{s.a.} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b, \end{array}$$

donde las filas de A son los vectores a_i^\top .

Denotamos por D la intersección de los dominios de todas las funciones del problema.