Introducción

El objetivo del curso es estudiar el problema

$$\min_{x \in X} f(x)$$ desde un punto de vista teórico y algorítmico.

$X \subset \mathbb{R}^n$ en general

$f : \mathbb{R}^n \to \mathbb{R}$

Def

- $x^* \in \mathbb{R}^n$ es un mínimo local de f si $\exists\, \varepsilon > 0 \,/\, f(x^*) \leq f(x) \quad \forall x \in B(x^*, \varepsilon)$
- $x^* \in \mathbb{R}^n$ es un mínimo global de f si $f(x^*) \leq f(x) \quad \forall x \in \mathbb{R}^n$

Los mínimos son estrictos cuando las desigualdades son estrictas para $x \neq x^*$

Ejemplo:

- Regresión

- $$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$$

- $$\min \|Ax - b\|_2$$
  $$\text{s.t. } \|x\|_2 \leq R$$

En el curso veremos:

- existencia de mínimos, condiciones de optimalidad
- Algoritmos para hallar mínimos. Análisis de convergencia.
- Teoría dualidad
- Otros métodos (funciones no diferenciables, ADMM)
- Stochastic gradient descent.

Convexidad

Def

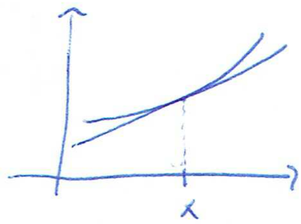
$f: \mathbb{R}^n \to \mathbb{R}$ es convexa si $f(tx + (1-t)y) \leq t f(x) + (1-t) f(y) \quad \forall x, y \quad t \in [0,1]$

Decimos que es estrictamente convexa si la desigualdad es estricta para $x \neq y$ $\quad 0 < t < 1$.

Prop

Si $f: \mathbb{R}^n \to \mathbb{R}$ es convexa y diferenciable $\Rightarrow f(z) \geq f(x) + \nabla f(x)^T (z - x) \quad \forall x, z$

Dem



Sean x, z, y consideremos $x + \alpha(z - x) \quad \alpha \in [0,1]$

$(\alpha z + (1-\alpha) x)$

Por convexidad: $f(x + \alpha(z-x)) \leq \alpha f(z) + (1-\alpha) f(x) \quad \forall \alpha \in [0,1]$

$$\Rightarrow \frac{f(x + \alpha(z-x))}{\alpha} \leq f(z) + \frac{f(x)}{\alpha} - f(x) \Rightarrow \frac{f(x+\alpha(z-x)) - f(x)}{\alpha} \leq f(z) - f(x) \quad \forall \alpha \in [0,1]$$

$$\Rightarrow \lim_{\alpha \to 0} \frac{f(x + \alpha(z-x)) - f(x)}{\alpha} \leq f(z) - f(x)$$

$$\parallel$$

$$\nabla f(x)^T (z - x) \leq f(z) - f(x) \Rightarrow f(z) \geq f(x) + \nabla f(x)^T (z - x)$$

Prop

Si $f: \mathbb{R}^n \to \mathbb{R}$ tiene derivadas segundas en todo punto. Entonces:

- Si $\nabla^2 f(x) \succeq 0 \quad \forall x \Rightarrow f$ es convexa.
- Si $\nabla^2 f(x) \succ 0 \quad \forall x \Rightarrow f$ es estrictamente convexa.

Ejemplos:

- $f(x) = e^{ax}$ es convexa $\forall a$
- $f(x) = x^a$ es convexa si $a \geq 1$ o $a \leq 0$
- $f(x) = |x|^p$ es convexa si $p \geq 1$
- $f(x) = \log(x)$ es cóncava en $\mathbb{R}^+$
- $f(x) = x \log(x)$ es convexa en $\mathbb{R}^+$

- Las normas en $\mathbb{R}^n$ son convexas
- $f(x) = \max\{x_1, x_2, \dots x_n\}$ es convexa
- $f(X) = \log\det(X)$ es cóncava en $S^n_{++} = \{X \,/\, X = X^T, X > 0\}$

Operaciones que preservan convexidad

- Combinación lineal de funciones convexas, con pesos no-negativos
- Composición con mapa lineal: $f(Ax+b)$ f convexa
- Máximo y supremo de funciones convexas $f(x) = \max\{f_1(x), f_2(x)\}$

> Máximo valor propio de una matriz simétrica ($\lambda_{max}(X)$)
>
> $f(X) = \sup\{y^T X y \,/\, \|y\| = 1\}$

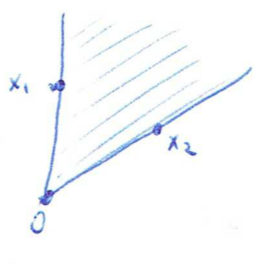
Conjuntos Convexos

Def

Un conjunto C es convexo sii $\forall x, y \in C$, $tx + (1-t)y \in C$.

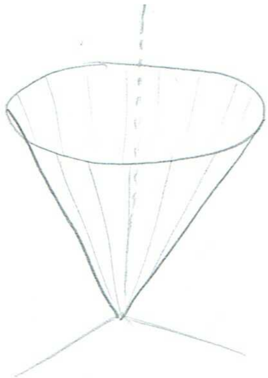
Ejemplos:

- Conjuntos afines (Soluciones de $Ax = b$)
- Bolas
- Cono convexo: C es un cono convexo sii $\forall x_1, x_2 \in C$, $\theta_1, \theta_2 \geq 0$
  $\theta_1 x_1 + \theta_2 x_2 \in C$



- Norm cone: $C = \{(x,t) \,/\, \|x\| \leq t\} \subset \mathbb{R}^{n+1}$

Con la norma $\|\cdot\|_2$ se denomina "Second order cone"



- Positive Semidefinite Cone:

$$S^n = \{X \in \mathcal{M}_{n\times n}(\mathbb{R}) \,/\, X = X^T\}$$

$$S^n_+ = \{X \in S^n \,/\, X \succcurlyeq 0\}$$

(Recordamos: X es semidefinida positiva, $X \succcurlyeq 0$, si $z^T X z \geq 0 \quad \forall z \in \mathbb{R}^n$)

Problema de Optimización Convexa

Un problema de optimización convexa es:

$$\boxed{\begin{array}{lll} \min\limits_{x} & f(x) & \\ \text{s.t.} & f_i(x) \leq 0 & i = 1 \ldots m \\ & a_i^T x = b_i & i = 1 \ldots p \end{array} \quad \text{con } f, f_i \text{ convexas.}}$$

o

$$\boxed{\min_{x \in C} f(x) \qquad \text{con } \begin{array}{l} f \text{ función convexa} \\ C \text{ conjunto convexo} \end{array}}$$

Prop

En un problema convexo, un mínimo local es también mínimo global.

Ejemplos

- Mínimos cuadrados
- Problemas lineales (LP) :

$$\begin{aligned} \min_x \quad & c^T x + d \\ \text{s.t.} \quad & Gx \leq h \\ & Ax = b \end{aligned}$$

- Problemas cuadráticos (QCQP) :

$$\begin{aligned} \min_x \quad & x^T P x + q^T x + r \\ \text{s.t.} \quad & x^T P_i x + q_i^T x + r_i \leq 0 \\ & Ax = b \end{aligned}$$

con $P, P_i \in S^n_+$

- Second-order cone programming (SOCP) :

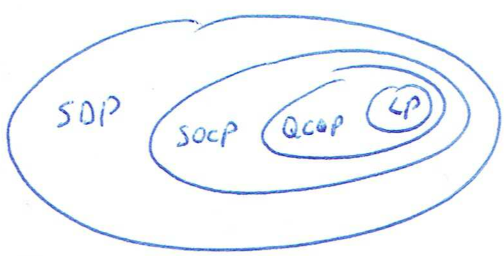
$$\begin{aligned} \min_x \quad & q^T x \\ \text{s.t.} \quad & \|A_i x + b_i\|_2 \leq c_i^T x + d_i \quad i=1 \dots m \\ & Fx = g \end{aligned}$$

- Semidefinite programming (SDP) :

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & F_0 + x_1 F_1 + \dots + x_n F_n \preccurlyeq 0 \\ & Ax = b \end{aligned}$$

con $F_i \in S^k$.

Obs:



Ejemplo

$$\min_x \quad \lambda_{max}(A(x))$$

$A(x) = A_0 + x_1 A_1 + \dots + x_n A_n$

$A_i \in S^k$ dados

$$\downarrow$$

$$\begin{aligned} \min_x \quad & t \\ \text{s.t.} \quad & A(x) \preccurlyeq tI \end{aligned}$$

Condiciones de Optimalidad

Ahora f no es necesariamente convexa, a no ser que se especifique.

Condiciones para problemas sin restricciones

Prop (Cond. necesaria de optimalidad)

Sea $f: \mathbb{R}^n \to \mathbb{R}$ diferenciable y x^* un mínimo local. Entonces $\nabla f(x^*) = 0$.

Si además $f \in C^2$, entonces $\nabla^2 f(x^*)$ es semidefinida positiva.

Dem

Sea $v \in \mathbb{R}^n$.

$$0 \leq \lim_{h \to 0^+} \frac{f(x^* + hv) - f(x^*)}{h} = \nabla f(x^*)^T v$$

Hacemos lo mismo para $-v$ $\Rightarrow \nabla f(x^*) = 0$ $\left(\text{pues } \nabla f(x^*)^T v = 0 \ \forall v \in \mathbb{R}^n \right)$

Si además $f \in C^2$: $\quad f(x^* + hv) - f(x^*) = \frac{h^2}{2} v^T \nabla^2 f(x^*) v + o(h^2)$

$$\Rightarrow \quad 0 \leq \frac{f(x^* + hv) - f(x^*)}{h^2} = \frac{1}{2} v^T \nabla^2 f(x^*) v + \frac{o(h^2)}{h^2}$$

Tomando límite cuando $h \to 0$, resulta $v^T \nabla^2 f(x^*) v \geq 0$

Prop

Sea $f: \mathbb{R}^n \to \mathbb{R}$ convexa, entonces $\nabla f(x^*) = 0$ es condición necesaria y suficiente para que x^* sea mínimo global.

Dem

Sabemos que $\quad f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*) \quad \forall x$

Si $\nabla f(x^*) = 0$, resulta $f(x) \geq f(x^*)$.