Optimización Numérica

Rodrigo Mendoza-Smith
rodrigo.mendoza@itam.mx

Clase 2: Conjuntos y funciones convexas I

Optimización convexa

$$\min f(x) \text{ s.t. } x \in \Omega \subset \mathbb{R}^n \tag{1}$$

- $f : \Omega \to \mathbb{R}$ es una función convexa.
- $\Omega \subset \mathbb{R}^n$ es un conjunto convexo.
- Restricción puede reescribirse como $g_i(x) \leq b_i$ para g_i función convexa.

Resolución de problemas convexos

$$\min f(x) \text{ s.t. } x \in \Omega \subset \mathbb{R}^n \tag{2}$$

- En general, (2) no admite una solución analítica.
- Los algoritmos explotan la convexidad del problema (existen muchos algoritmos).
- Muchos problemas reales se pueden formular como problemas convexos.
- Muchos problemas no convexos se pueden reformular como problemas convexos.
- Podemos pensar en:

$$\text{Probs} = \{\text{Probs convexos}\} \cup \{\text{Probs no convexos}\}$$

Conceptos básicos de convexidad

- **Conjunto afín**: $C \subset \mathbb{R}^n$ es afín si:

$$\sum_{i=1}^{m} \theta_i x_i \in C \;\; \forall\, x_i \in C,\; \theta_i \in \mathbb{R} \,\text{s.t.} \sum_i \theta_i = 1 \qquad (3)$$

- Un conjunto afín C es un subespacio $V \leq \mathbb{R}^n$ mas un sesgo $x_0 \in C$:

$$C = V + x_0 \qquad (4)$$

- El subespacio V no depende de x_0
- $\dim(C) := \dim(V)$.
- Si $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$,
  1. $C = \{x : Ax = b\}$ es afín.
  2. El subespacio asociado con C es $\text{Ker}(A)$.
  3. Todo conjunto afín puede expresarse como el espacio de solución de un sistema de ecuaciones.

Conceptos básicos de convexidad

- Si $C \subset \mathbb{R}^n$ es un conjunto de puntos, la cubierta afín se define como

$$\operatorname{aff} C = \left\{ \sum_i^m \theta_i x_i : x_i \in C, \ \sum_i \theta_i = 1 \right\}.$$

- $\operatorname{aff} C$ es el menor conjunto afín que contiene C.
- La dimensión afín de un conjunto de puntos es $\dim(\operatorname{aff} C)$.
- Si $C \subset \mathbb{R}^n$ y $\dim(\operatorname{aff} C) < n$, $\operatorname{aff} C \neq \mathbb{R}^n$.
- El interior relativo de $C \subset \mathbb{R}^n$ es

$$\operatorname{relint} C = \{x \in C : \exists r > 0 \ \text{ s.t. } \ \mathcal{B}_r(x) \cap \operatorname{aff} C \subset C\}.$$

Conjuntos convexos

- El simplejo de probabilidad $\triangle_k := \{x \in [0,1]^k : \sum x_i = 1\}$.
- $C \subset \mathbb{R}^n$ es convexo si $\forall x_1, x_2 \in C$ y $\theta \in [0,1]$

$$\theta x_1 + (1-\theta) x_2 \in C. \tag{5}$$

- $\sum_{i=1}^m \theta_i x_i$ con $\theta \in \triangle_m$ es la combinación convexa de $x_1, \ldots, x_m$.
- La cubierta convexa de C es

$$\operatorname{conv} C = \left\{ \sum_{i=1}^m \theta_i x_i : x_i \in C, \;\; \theta \in \triangle_m \right\}. \tag{6}$$

- La cubierta convexa es el _______ conjunto _______ que contiene a ______.

Combinaciones convexas

Sea $C \subset \mathbb{R}^n$ convexo.

- Si $\theta \in \triangle_\infty$ y $x_1, x_2, \cdots \in C$, entonces

$$\sum_{i=1}^{\infty} \theta_i x_i \in C \quad \text{(si la serie converge).}$$

- Si $p : \mathbb{R}^n \to \mathbb{R}$ satisface $p(x) \geq 0$ para todo $x \in C$, y $\int_C p(x) dx = 1$ entonces

$$\int_C p(x) x dx \in C \quad \text{(si la integral converge).}$$

- Si x es un vector aleatorio y $\Pr[x \in C] = 1$, entonces $\mathbb{E}[x] \in C$.

Conos

- C es un cono si $\theta x \in C$ para todo $x \in C$ y $\theta \geq 0$.
- C es un cono convexo si es ________ y ________.
- Esto es, si para todo $x_1, x_2 \in C$ y $\theta_1, \theta_2 \geq 0$, tenemos

$$\theta_1 x_1 + \theta_2 x_2 \in C.$$

- $\sum_{i=1}^{m} \theta_i x_i$ con $\theta_i \geq 0$ es la combinación cónica de $x_1, \ldots, x_m$.
- Si $x_i \in C$ y C es un cono convexo, toda combinación cónica de x_i está en C.
- Un conjunto C es un cono convexo si y sólo si contiene todas las combinaciones cónicas de sus elementos.
- La cubierta cónica de un conjunto C es el conjunto de todas las combinaciones cónicas de puntos en C.

Hiperplanos, bolas, elipsoides

- Si $a \in \mathbb{R}^n \setminus \{0\}$ y $b \in \mathbb{R}$, un hiperplano es el conjunto

$$\{x : a^T x = b\}. \tag{7}$$

- $\{x : a^T(x - x_0) = 0\} = x_0 + a^\perp$
- $\mathbb{R}^n = \{\{x : a^T(x - x_0) \leq 0\} \cup \{x : a^T(x - x_0) \geq 0\}$
- Un *elipsoide* es un conjunto de la forma

$$\mathcal{E} = \{x : (x - x_c)^T P^{-1}(x - x_c) \leq 1\} \tag{8}$$

con $P = P^T \succ 0$, P es s.p.d.

Bolas y conos normados

- La bola normada de radio r y centro x_c es $\{x : \|x - x_c\| \leq r\}$ (es convexa).
- El cono normado asociado con $\|\cdot\|$ es el conjunto $C = \{(x, t) : \|x\| \leq t\} \subset \mathbb{R}^{n+1}$.
- ¿Cómo se ve para $n = 2$ y $\|\cdot\| = \|\cdot\|_2$?

Poliedros

- Solución de un conjunto finito de (des)igualdades

$$\mathcal{P} = \{x : Ax \leq b, Cx = d\}.$$

- Intersección finita de semi-espacios e hiperplanos ($\leq$ desigualdad por componentes).

Símplices

- Si $v_0, v_1, \dots, v_k \in \mathbb{R}^n$ son afinmente independientes, entonces $v_1 - v_0, \dots, v_n - v_0$ son linealmente independientes.
- El simplejo está dado por

$$C = \operatorname{conv}\{v_0, \dots, v_k\} = \left\{ \sum_{i=0}^{n} \theta_0 v_0 : \theta \in \triangle_{k+1} \right\}. \quad (9)$$

- El simplejo puede ser descrito como un poliedro (tarea).
- El simplejo es de las pocas cubiertas convexas que pueden ser descritas fácilmente por poliedros (tarea)

El cono positivo semidefinido

- El conjunto de matrices simétricas stá dado por

$$\mathbf{S}^n = \left\{ X \in \mathbb{R}^{n \times n} : X = X^T \right\} \tag{10}$$

- Es un espacio vectorial de dimensión _________.
- Definimos también,

$$\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n : X \succeq 0\} \tag{11}$$
$$\mathbf{S}^n_{+} = \{X \in \mathbf{S}^n : X \succ 0\} \tag{12}$$

- $\mathbf{S}^n_{++}$ es un cono convexo.
- ¿Cómo describirían $\mathbf{S}^2_+$?

Operaciones que preservan convexidad

- (Intersecciones) Si S_α es convexo y $\alpha \in \mathcal{A}$, $\cap_{\alpha \in \mathcal{A}} S_\alpha$.
  1. *e.g.* $\mathbf{S}^n_{++} = \cap_{z \neq 0}\{X \in \mathbf{S}^n : z^T X z \geq 0\}$.
  2. Si S es convexo, $S = \cap\{\mathcal{H} : \mathcal{H} \text{ semi-espacio}, S \subset \mathcal{H}\}$.
- (Funciones afines) Si $S \subset \mathbb{R}^n$ es convexo y $f : \mathbb{R}^n \to \mathbb{R}^m$ es afín[1], $f(S)$ y $f^{-1}(S)$ son convexos.
  1. Multiplicación por escalares, translación, suma, proyeccón.
  2. Conjunto solución de desigualdades lineales de matrices $A(x) = x_1 A_1 + \cdots + x_n A_n \leq B$, el cono hiperbólico, elipsoides.
- (Fraccionales lineales y funciones de perspectiva) Tarea :)

[1] $f(x) = Ax + b$