Análisis Matemático para Inteligencia Artificial

Martín Errázquin (merrazquin@fi.uba.ar)

Especialización en Inteligencia Artificial

Optimización convexa

Región Convexa

Un subconjunto $S \subset \mathbb{R}^n$ es convexo si para cada par de puntos $x, y \in S, \alpha \in [0,1]$ se verifica que $z = \alpha x + (1-\alpha)y \in S$.

Obs: X_1, X_2 son dos conjuntos convexos, entonces,

- $X_1 \cap X_2$ es convexo.
- si L es una transformación lineal, $L(X_1)$ es convexa.

Función Convexa

Sea $M \subset \mathbb{R}^n, M \neq \emptyset$ convexo, $f : M \to \mathbb{R}$. Entonces se dice que:

- f es convexa en M sii $\forall x, y \in M, \forall \alpha \in [0,1]$ se verifica que:
$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)f(y)$$
- f es estrictamente convexa en M sii $\forall x, y \in M, \forall \alpha \in (0,1)$ se verifica que:
$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y)$$

Prop: Sea $M \subset \mathbb{R}^n, M \neq \emptyset$ abierto convexo, $f : M \to \mathbb{R}$, $f \in \mathcal{C}^2$. Entonces, f es convexa en M sii $\forall x \in M, y^T H_f(x) y \geqslant 0$, para cualquier $y \in \mathbb{R}^n$.

Condición de Slater

Sea el problema

$$\min \ f(x_1, ..., x_n)$$
$$s.t. \ g_i(x_1, ..., x_n) = 0$$
$$h_j(x_1, ..., x_n) \leq 0$$

donde las f, h_j son funciones convexas y las g_i son funciones afines (lineales o constantes). Si existe un punto $\vec{x_0}$ **estrictamente factible**, es decir que $g_i(\vec{x_0}) = 0, h_j(\vec{x_0}) < 0 \ \forall i, j$ entonces:

1. Hay dualidad fuerte: óptimo del dual coincide con óptimo del primal
2. Las condiciones KKT son necesarias y suficientes

Programación Lineal

Un problema de programación lineal es donde tanto la función objetivo como las funciones que definen las restricciones son lineales. La forma general es:

$$
\begin{aligned}
opt\ & c_1x_1 + ... + c_nx_n \\
s.t.\ & a_{11}x_1 + ... + a_{1n}x_n \leqslant b_1 \\
& \vdots \\
& a_{m1}x_1 + ... + a_{mn}x_n \leqslant b_m \\
& d_{11}x_1 + ... + d_{1n}x_n \geqslant e_1 \\
& \vdots \\
& d_{r1}x_1 + ... + d_{rn}x_n \geqslant e_r \\
& g_{11}x_1 + ... + g_{1n}x_n = b_1 \\
& \vdots \\
& g_{s1}x_1 + ... + g_{sn}x_n = b_s
\end{aligned}
$$

Propiedades de la Programación Lineal

Definición: Sea $S \subset \mathbb{R}^n$ convexo $S \neq \emptyset$, $\bar{x^*} \in S$ es un punto extremo de S si $\bar{x^*}$ no puede expresarse como combinación lineal convexa de puntos de S distintos de él.

1. Es un problema convexo ya sea de minimización o maximización.
2. La solución óptima, si existe, es global.
3. Nunca existen óptimos locales que no sean globales.
4. Tiene 0, 1 o infinitas soluciones.

Forma estándar: Se busca hallar min $\bar{c}\bar{x}$, s.t. $A\bar{x} = \bar{b}; \bar{x} \geqslant 0$, donde $A \in \mathbb{R}^{m \times n}$, $m < n$, $rg(A) = m$.

Métodos típicos de solución: Algoritmo Simplex y revised Simplex, Interior Point (IPM).

Programación cuadrática

En el caso de que la función objetivo sea una cuadrática convexa:

$$\min_{x \in \mathbb{R}^d} \ \frac{1}{2}x^T Q x + c^T x$$
$$s.t.: Ax \leqslant b$$

donde $A \in \mathbb{R}^{m \times d}, b \in \mathbb{R}^m, c \in \mathbb{R}^d$.

La matriz $Q \in \mathbb{R}^{d \times d}$ es simétrica y definida positiva, y por lo tanto la función objetivo es convexa. El lagrangiano está dado por:

$$\mathscr{L}(x, \lambda) = \frac{1}{2}x^T Q x + c^T x + \lambda^T (Ax - b)$$

Para obtener $D(\lambda) = \min_x \mathscr{L}(x, \lambda)$ buscamos x aplicando $\nabla \mathscr{L}(x) = 0$:

$$\nabla \mathscr{L}(x) = Qx + (c + A^T \lambda) = 0 \Rightarrow x = -Q^{-1}(c + A^T \lambda)$$

Y sustituyendo obtenemos el objetivo del dual:

$$D(\lambda) = -\frac{1}{2}(c + A^T \lambda) - Q^{-1}(c + A^T \lambda) - \lambda^T b.$$