MA5701 Optimización no Lineal
Profesor: Alejandro Jofré
Auxiliar: Benjamín Vera Vera

Auxiliar 7

Mirror Descent
16 de mayo de 2025

P1. Dada $h : \mathbb{R}^n \to \mathbb{R}$ y considerando la divergencia de Bregmann definida por

$$D_h(x, z) = h(x) - h(z) - \boldsymbol{\nabla} h(z)^\top (x - z)$$

pruebe que se tiene la siguiente identidad:

$$\forall x, y, z \in \mathbb{R}^n : D_h(x, y) = D_h(x, z) - (x - z)^\top (\boldsymbol{\nabla} h(y) - \boldsymbol{\nabla} h(z)) + D_h(z, y)$$

P2. Sea $\|\cdot\|$ una norma cualquiera sobre $\mathbb{R}^n$ y $h : \mathbb{R}^n \to \mathbb{R}$ m-fuertemente convexa con respecto a esta norma. Sea además $f : \mathbb{R}^n \to \mathbb{R}$ convexa y L-Lipschitz con respecto a $\|\cdot\|$ tal que f posee un mínimo x^* sobre $\mathcal{X} \subseteq \mathbb{R}$. Considere la iteración de Mirror Descent dada por

$$x^{k+1} = \operatorname{argmin}_{x \in \mathcal{X}} \left\{ f(x^k) + \boldsymbol{\nabla} f(x^k)^\top (x - x^k) + \frac{1}{\alpha_k} D_h(x, x^k) \right\}$$

y sean además

$$\lambda_k := \sum_{j=0}^{k} \alpha_j, \qquad \overline{x}^k = \frac{1}{\lambda_k} \sum_{j=0}^{k} \alpha_j x^j.$$

Pruebe que para $T \geq 1$ se tiene que

$$f\left(\overline{x}^T\right) - f^* \leq \frac{D_h(x^*, x^0) + \frac{L^2}{2m} \sum_{t=0}^{T} \alpha_t^2}{\sum_{t=0}^{T} \alpha_t}$$

Indicación: Recuerde que las condiciones necesarias de primer orden para x^{k+1} vienen dadas por

$$\left[\boldsymbol{\nabla} f(x^k) + \frac{1}{\alpha_k} \boldsymbol{\nabla} h(x^{k+1}) - \frac{1}{\alpha_k} \boldsymbol{\nabla} h(x^k) \right]^\top (x - x^{k+1}) \geq 0, \quad \forall x \in \mathcal{X}$$