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

Auxiliar 10

Descenso de Gradiente Estocástico
4 de julio de 2025

P1. Sea $f : \mathbb{R}^n \to \mathbb{R}$ fuertemente convexa de parámetro m y L-suave. Consideremos un algoritmo que consiste en búsquedas de línea exactas a lo largo de direcciones aleatorias. El esquema para obtener la siguiente iteración x^+ desde x es el siguiente:

- Escoger $v \sim \mathcal{N}(0, \sigma^2 I)$ independiente de las iteraciones previas.
- Encontrar $t_{\text{min}} = \operatorname{argmin}_t f(x + tv)$
- Definir $x^+ = x + t_{\text{min}} v$

El objetivo es probar que $\mathbb{E}[f(x^T) - f(x^*)] \leq \varepsilon$ siempre que

$$T \geq \frac{CnL}{m} \log\left(\frac{f(x^0) - f(x^*)}{\varepsilon}\right)$$

para algún $C > 0$. Para ello, proceda como sigue:

a) Pruebe que, dado $v \in \mathbb{R}^n, t > 0$ se tiene que

$$f(x + tv) \leq f(x) + \boldsymbol{\nabla} f^\top v + \frac{L}{2} t^2 \|v\|^2.$$

b) Para v fijo, minimice sobre t ambos lados y obtenga una cota para $f(x + t_{\text{min}} v)$.

c) Tomando esperanza a ambos lados sobre v, utilice que $\mathbb{E}[v_j^2 / \|v\|^2] = \frac{1}{n}$ además de la cota conocida

$$\|\boldsymbol{\nabla} f(x)\|^2 \geq 2m(f(x) - f(x^*))$$

para probar que

$$\mathbb{E}_v[f(x^+) - f(x^*)] \leq \left(1 - \frac{m}{nL}\right)(f(x) - f(x^*)).$$

d) Concluya.