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

Examen

Tiempo: 3:00
10 de julio de 2025

Considérese el problema de encontrar un minimizador para:

$$\min_x f(x)$$
$$g_j(x) \leq 0, \quad j \in \{1, \dots, m\}$$

en que $f : \mathbb{R}^d \to \mathbb{R}$ tiene la forma

$$f(x) = \frac{1}{n} \sum_{i=1}^{n} f_i(x)$$

con $(f_i)_i, (g_j)_j \in \mathcal{C}^1(\mathbb{R}^d)$. Suponga que para $0 < \mu \leq L$, las $(f_i)_i$ son convexas L-suaves y que f es μ-fuertemente convexa. Suponga también que las $(g_j)_j$ son afines dadas por $g_j(x) = a_j^\top x + b_j$ ($a_j \in \mathbb{R}^d, b_j \in \mathbb{R}$) de tal modo que el conjunto factible

$$\mathcal{X} = \left\{x \in \mathbb{R}^d : g_j(x) \leq 0, j \in \{1, \dots, m\}\right\}$$

tiene interior no vacío en $\mathbb{R}^d$.

Sea la función barrera logarítmica relajada $B : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ dada por

$$B(z, \delta) = \begin{cases} -\delta \log(-z), & z \leq -\delta \\ \frac{1}{2}\left(\frac{(z+2\delta)^2}{\delta} - \delta\right) - \delta \log(\delta), & z \geq -\delta. \end{cases}$$

Dado $k \in \mathbb{N}$ y $\delta_k \geq \delta_\infty > 0$, consideremos el problema irrestricto

$$\min_{x \in \mathbb{R}^d} \left\{ f(x) + \frac{1}{m} \sum_{j=1}^{m} B(g_j(x), \delta_k) \right\}$$

con minimizador que denotamos por $x^*(\delta_k)$. Dado $\delta_\infty > 0$, consideremos el siguiente esquema de gradiente estocástico con penalización:

$$x^{k+1} = x^k - \gamma(\nabla f_{i_k}(x_k) + \nabla B(g_{j_k}(x_k), \delta_k)) \tag{1}$$

en que $x_0 \in \mathbb{R}^d$ e $i_k \in \{1, \dots, n\}, j_k \in \{1, \dots, m\}$ se escogen de manera uniforme e independiente entre iteraciones, δ_k representa un paso positivo y δ_k es parámetro de barrera en la iteración k. El objetivo es probar que si se toma $\{\gamma_k\}_{k\geq 0}$ tal que

$$\sum_{k=0}^{\infty} \gamma_k = \infty, \qquad \sum_{k=0}^{\infty} \gamma_k^2 < \infty$$

y $\{\delta_k\}_{k\geq 0}$ dada por $\delta_k = \delta_\infty + \varepsilon_k$ en que $\{\varepsilon_k\}_{k\geq 0}$ es tal que

$$\sum_{k=0}^{\infty} \gamma_k \varepsilon_k < \infty,$$

entonces para casi cualquier $\{x_k\}_{k\geq 0}$ dado por (1) se cumple que $\lim_{k\to\infty} x_k = x^*(\delta_\infty)$. Para esto, procedemos como sigue:

P1. Sea $\tilde{x}_k = x_k - x^*(\delta_\infty)$.

a) Pruebe que

$$\begin{aligned}\mathbb{E}_k\left[\|\tilde{x}_{k+1}\|^2\right] &= \|\tilde{x}_k\|^2 + \gamma_k^2\mathbb{E}_k\left[\|\nabla\Phi_{i_k,j_k}(x_k)\|^2\right] + \gamma_k^2\mathbb{E}_k\left[\|\nabla C_{j_k}(x_k,\delta_k)\|^2\right] \\ &\quad - 2\gamma_k\mathbb{E}_k[\nabla\Phi_{i_k,j_k}(x_k)^\top\tilde{x}_k] + 2\gamma_k\mathbb{E}_k\left[\nabla C_{j_k}(x,k,\delta_k)^\top\tilde{x}_k\right] \\ &\quad - 2\gamma_k^2\mathbb{E}_k\left[\nabla\Phi_{i_k,j_k}(x_k)^\top\nabla C_{j_k}(x_k,\delta_k)\right]\end{aligned}$$

en que

$$\begin{aligned}\nabla\Phi_{i_k,j_k}(x_k) &= \nabla f_{i_k}(x_k) + \nabla B(g_{j_k}(x_k),\delta_\infty) \\ \nabla C_{j_k}(g(x_k),\delta_k) &= \nabla B(g_{j_k}(x_k),\delta_\infty) - \nabla B(g_{j_k}(x_k),\delta_k)\end{aligned}$$

b) Sea $\Phi : \mathbb{R}^d \to \mathbb{R}$ dada por

$$\Phi(x_k) = \frac{1}{n}\sum_{i=1}^{n} f_i(x_k) + \frac{1}{m}\sum_{j=1}^{m} B(g_j(x_k),\delta_\infty).$$

Muestre que

$$\mathbb{E}_k\left[\nabla\Phi_{i_k,j_k}(x_k)^\top\tilde{x}_k\right] = \nabla\Phi(x_k)^\top\tilde{x}_k$$

y utilice la fuerte convexidad de f para obtener que

$$-2\gamma_k\nabla\Phi(x_k)^\top\tilde{x}_k \le -2\gamma_k\Phi_0(x_k) - \mu\gamma_k\|\tilde{x}_k\|^2$$

en que $\Phi_0(u) = \Phi(u) - \Phi(x^*(\delta_\infty))$.

c) Utilizando lo anterior, concluya que

$$\begin{aligned}\mathbb{E}_k\left[\|\tilde{x}_k\|^2\right] &\le \|\tilde{x}_k\|^2 + 2\gamma_k^2\mathbb{E}_k\left[\|\nabla\Phi_{i_k,j_k}(x_k)\|^2\right] + 2\gamma_k^2\mathbb{E}_k\left[\|\nabla C_{j_k}(x_k,\delta_k)\|^2\right] \\ &\quad - \mu\gamma_k\|\tilde{x}_k\|^2 - 2\gamma_k\Phi_0(x_k) + 2\gamma_k\mathbb{E}_k[\|\nabla C_{j_k}(x_k,\delta_k)\|\|\tilde{x}_k\|].\end{aligned}$$

Indicación: Recuerde que $2u^\top v \le 2\|u\|\|v\| \le \|u\|^2 + \|v\|^2$.

P2. Se puede probar (no lo haga) que existen constantes positivas $\hat{c},\bar{a},\bar{\bar{a}},\hat{a},\bar{b},\bar{\bar{b}}$ tales que

$$\begin{aligned}\mathbb{E}_k[\|\nabla C_{j_k}(x_k,\delta_k)\|\|\tilde{x}_k\|] &\le \|\tilde{x}_k\|^2\left(\bar{\bar{a}}\frac{\varepsilon_k}{\delta_k\delta_\infty} + \hat{c}\varepsilon_k\bar{a} + \bar{b}\frac{\varepsilon_k}{\delta_k\delta_\infty}\right) + \frac{\hat{c}\varepsilon_k}{4}\bar{a} + \frac{1}{4}\bar{b}\frac{\varepsilon_k}{\delta_k\delta_\infty} \\ \mathbb{E}_k\left[\|\nabla C_{j_k}(x_k,\delta_k)\|^2\right] &\le 3\hat{a}\frac{\varepsilon_k^2\|\tilde{x}_k\|^2}{\delta_k^2\delta_\infty^2} + 3\hat{c}^2\varepsilon_k^2\bar{\bar{a}} + \frac{3\bar{\bar{b}}\varepsilon_k^2}{\delta_k^2\delta_\infty^2} \\ \mathbb{E}_k\left[\|\nabla\Phi_{i_k,j_k}(x_k)\|^2\right] &\le 4\hat{L}\Phi_0(x_k) + 2\sigma_\Phi\end{aligned}$$

con $\hat{L} > 0, \sigma_\Phi \ge 0$.

a) Muestre que, para ξ_k, r_k a definir:

$$\mathbb{E}_k\left[\|\tilde{x}_{k+1}\|^2\right] \le (1-\gamma_k(\mu-\xi_k))\|\tilde{x}_k\|^2 - 2\gamma_k(1-4\gamma_k\hat{L})\Phi_0(x_k) + \gamma_k\varepsilon_k r_k + 4\gamma_k^2\sigma_\Phi$$

b) Utilizando lo encontrado, pruebe que existe $k_0 \ge 0$ tal que para $k \ge k_0$:

$$\mu - \xi_k > 0, \qquad \gamma_k(1-4\gamma_k\hat{L}) \ge 0.$$

Concluya que para $k \ge k_0$,

$$\mathbb{E}_k\left[\|\tilde{x}_{k+1}\|^2\right] \le (1-\gamma_k(\mu-\xi_k))\|\tilde{x}_k\|^2 + \gamma_k\varepsilon_k r_k + 4\gamma_k^2\sigma_\Phi$$

y que las iteraciones $\tilde{x}_k$ están acotadas casi seguramente.