ISyE8813
Subgradient Method for Strongly Convex NCO

Renato D.C. Monteiro

January 17, 2024

1 Problem Description and Assumptions

Consider the set optimization problem

$$f_* = \min\{(f + \delta_X)(x) : x \in \mathbb{R}^n\} = \min\{f(x) : x \in X\}$$

Assumptions:

A.1) $X \subset \mathbb{R}^n$ is nonempty closed convex and, for some $\mu > 0$, function f is μ-strongly convex on X

A.2) there exists $M \geq 0$ with the following property: for every $x \in X$, there exists $s(x) \in \partial f(x)$ such that $\|s(x)\| \leq M$

A.3) optimal solution set X_* is nonempty, and hence $f_* \in \mathbb{R}$

Subgradient Method (SM)

---

0) $x_0 \in X$ is given

1) For $k = 0, 1, 2, \ldots$, do

   - choose stepsize $\lambda_k > 0$ and set $s_k = s(x_k)$
   - compute
   $$x_{k+1} = \mathrm{Proj}_X(x_k - \lambda_k s_k)$$

---

2 Analysis

2.1 Variable stepsize case

This subsection considers the case where the sequence of stepsizes $\{\lambda_k\}$ is chosen as

$$\lambda_k = \frac{2}{\mu(k+1)} \quad \forall k \geq 0. \tag{1}$$

Proposition 2.1 *Assume that* (1) *holds and, for any* $K \geq 0$*, define*

$$\theta_K^* = \min_{k=0,\ldots,K} [f(x_k) - f(x_*)].$$

Then, for every $x_* \in X_*$*, we have*

$$\theta_K^* + \frac{\mu(K+1)}{2(K+2)} \|x_{K+1} - x_*\|^2 \leq \frac{2M^2}{\mu(K+2)}$$

As a consequence, if $K \geq 4M^2/(\mu\varepsilon)$ *then* $\theta_K^* \leq \varepsilon$*.*

The proof of the above result requires a few technical lemmas.

Lemma 2.2 *For every* $k \geq 0$ *and* $x \in X$*, have*

$$2\lambda_k[f(x_k) - f(x)] \leq (1 - \lambda_k\mu)\|x_k - x\|^2 - \|x_{k+1} - x\|^2 + \lambda_k^2 \|s_k\|^2$$

Proof: Let $x \in X$ be given. Have

$$\begin{aligned} \|x_{k+1} - x\|^2 &= \|\text{Proj}_X(x_k - \lambda_k s_k) - x\|^2 \\ &= \|\text{Proj}_X(x_k - \lambda_k s_k) - \text{Proj}_X(x)\|^2 \\ &\leq \|x_k - \lambda_k s_k - x\|^2 \\ &= \|x_k - x\|^2 + \lambda_k^2 \|s_k\|^2 + 2\lambda_k \langle s_k, x - x_k \rangle \end{aligned}$$

The result now follows from the fact that

$$f(x) - f(x_k) \geq \langle s_k, x - x_k \rangle + \frac{\mu}{2}\|x - x_k\|^2$$

■

Lemma 2.3 *Assume that* (1) *holds. Then, for every* $k \geq 1$ *and* $x \in \mathbb{R}^n$*, we have:*

$$(k+1)[f(x_k) - f(x)] \leq \frac{\mu k^2}{4}\|x_k - x\|^2 - \frac{\mu(k+1)^2}{4}\|x_{k+1} - x\|^2 + \frac{M^2}{\mu}$$

Proof: It follows from Lemma 2.2 and Assumption A.2 that

$$2\lambda_k[f(x_k) - f(x)] \leq (1 - \lambda_k\mu)\|x_k - x\|^2 - \|x_{k+1} - x\|^2 + \lambda_k^2 M^2$$

Now using (1), we have

$$\begin{aligned} \frac{4[f(x_k) - f(x)]}{\mu(k+1)} &\leq \left(1 - \frac{2}{k+1}\right)\|x_k - x\|^2 - \|x_{k+1} - x\|^2 + \frac{4M^2}{\mu^2(k+1)^2} \\ &\leq \left(\frac{k-1}{k+1}\right)\|x_k - x\|^2 - \|x_{k+1} - x\|^2 + \frac{4M^2}{\mu^2(k+1)^2} \end{aligned}$$

Multiplying the above inequality by $(k+1)^2\mu/4$, we have

$$(k+1)[f(x_k) - f(x)] \leq \frac{(k+1)(k-1)\mu}{4}\|x_k - x\|^2 - \frac{(k+1)^2\mu}{4}\|x_{k+1} - x\|^2 + \frac{M^2}{\mu}$$

and hence that the conclusion of the lemma holds. ■

Lemma 2.4 *Assume that* (1) *holds and, for any* $K \geq 0$, *define*

$$\theta_K(x) = \min_{k=0,\ldots,K} [f(x_k) - f(x)].$$

Then,

$$\theta_K(x) + \frac{\mu(K+1)}{2(K+2)}\|x_{K+1} - x\|^2 \leq \frac{2M^2}{\mu(K+2)}$$

Proof: Let $K \geq 0$ be given. Then, by Lemma 2.3, we have

$$(k+1)\Theta_K(x) \leq \frac{\mu k^2}{4}\|x_k - x\|^2 - \frac{\mu(k+1)^2}{4}\|x_{k+1} - x\|^2 + \frac{M^2}{\mu} \quad k = 0, \ldots, K$$

Adding these inequalities, we have

$$\frac{(K+2)(K+1)}{2}\Theta_K(x) \leq -\frac{\mu(K+1)^2}{4}\|x_{K+1} - x\|^2 + \frac{(K+1)M^2}{\mu}.$$

The conclusion of the lemma now follows by dividing the above inequality by $(K+2)(K+1)/2$. ■

The proof of Proposition 2.1 follows immediately from Lemma 2.4 with $x = x_*$.

2.2 Constant stepsize case

By Lemma 2.2 with $x = x_*$ and $\lambda_k = \lambda$ for all $k \geq 1$, have

$$2\lambda[f(x_k) - f_*] \leq (1 - \lambda\mu)\|x_k - x_*\|^2 - \|x_{k+1} - x_*\|^2 + \lambda^2 M^2$$

Hence,

$$2\lambda\theta_K^* \leq (1 - \lambda\mu)\|x_k - x_*\|^2 - \|x_{k+1} - x_*\|^2 + \lambda^2 M^2 \quad k = 0, \ldots, K.$$

Dividing this expression by $(1 - \lambda\mu)^{k+1}$, we have

$$\frac{2\lambda\theta_K}{(1-\lambda\mu)^{k+1}} \leq \tau_k - \tau_{k+1} + \frac{\lambda^2 M^2}{(1-\lambda\mu)^{k+1}} \quad k = 0, \ldots, K$$

where

$$\tau_k := \frac{\|x_k - x_*\|^2}{(1-\lambda\mu)^k}$$

Adding the above inequalities, we have

$$2\lambda\theta_K \sum_{k=0}^{K} \beta^{k+1} \leq \tau_0 - \tau_{K+1} + \lambda^2 M^2 \sum_{k=0}^{K} \beta^{k+1}$$

where $\beta = (1 - \lambda\mu)^{-1}$. Since $\tau_0 = d_0^2$, we have

$$\theta_K \leq \frac{d_0^2 - \tau_{K+1}}{2\lambda \sum_{k=0}^{K} \beta^{k+1}} + \frac{\lambda M^2}{2} \leq \frac{d_0^2}{2\lambda \sum_{k=0}^{K} \beta^{k+1}} + \frac{\lambda M^2}{2}$$

Now,

$$\sum_{k=0}^{K} \beta^{k+1} = \beta \frac{\beta^{K+1} - 1}{\beta - 1}$$

So

$$\theta_K \le \frac{(\beta-1)d_0^2}{2\beta\lambda[\beta^{K+1}-1]} + \frac{\lambda M^2}{2}$$

So, need to choose K such that

$$\frac{(\beta-1)d_0^2}{2\beta\lambda[\beta^{K+1}-1]} \le \frac{\varepsilon}{2}$$

or

$$\beta^{K+1} \ge 1 + \frac{(\beta-1)d_0^2}{\beta\lambda\varepsilon}$$

or

$$(K+1)\log\beta \ge \log\left(1 + \frac{(\beta-1)d_0^2}{\beta\lambda\varepsilon}\right)$$

Now

$$\log\beta = -\log(1-\lambda\mu) \ge \lambda\mu$$

and hence, it suffices to choose K such that

$$K\lambda\mu \ge \log\left(1 + \frac{(\beta-1)d_0^2}{\beta\lambda\varepsilon}\right)$$

or

$$K \ge \frac{1}{\lambda\mu}\log\left(1 + \frac{(\beta-1)d_0^2}{\beta\lambda\varepsilon}\right)$$

The final complexity is obtained with $\lambda = \varepsilon/M^2$, i.e.,

$$\mathcal{O}\left(\frac{M^2}{\varepsilon\mu}\log\left(1 + \frac{(\beta-1)d_0^2}{\beta\lambda\varepsilon}\right)\right)$$