Solution to Homework 6

Total 20 points

Problem 1. (5 points) If function f is μ-strongly convex, and $\mathbf{g}$ is a subgradient of f at $\mathbf{x}$. Show that for any $\mathbf{y} \in dom\ f$,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 .$$

Solution. Let $h(\mathbf{x}) = f(\mathbf{x}) - \frac{\mu}{2} \|\mathbf{x}\|_2^2$. Then $h(\mathbf{x})$ is convex and $\mathbf{g} - \mu\mathbf{x}$ is a subgradient of h at $\mathbf{x}$. Thus we have

$$h(\mathbf{y}) \geq h(\mathbf{x}) + \langle \mathbf{g} - \mu\mathbf{x}, \mathbf{y} - \mathbf{x} \rangle,$$

which means

$$\begin{aligned} f(\mathbf{y}) \geq & f(\mathbf{x}) + \frac{\mu}{2}(\|\mathbf{y}\|_2^2 - \|\mathbf{x}\|_2^2) + \langle \mathbf{g} - \mu\mathbf{x}, \mathbf{y} - \mathbf{x} \rangle \\ = & f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 . \end{aligned}$$

Problem 2. (5 points) Suppose f is convex and G-Lipschitz continuous over the constraint $\mathcal{C}$, which is bounded and convex with diameter $D > 0$. If we run projected subgradient descent method for T rounds with $\eta_t = \frac{D}{G\sqrt{T}}$, then we have

$$f(\bar{\mathbf{x}}_t) - f^* \leq \frac{DG}{\sqrt{T}},$$

where $\mathbf{x}_t = \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_t$.

Solution. In the class, we have shown

$$f(\bar{\mathbf{x}}_t) - f^* \leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|_2^2 + \sum_{k=0}^{t} \eta_k^2 \|\mathbf{g}_k\|^2}{2\sum_{k=0}^{t} \eta_k}.$$

Since $\|\mathbf{x}_0 - \mathbf{x}^*\|_2 \leq D$ and $\|\mathbf{g}_k\| \leq G$, we have

$$f(\bar{\mathbf{x}}_t) - f^* \leq \frac{DG}{\sqrt{T}}.$$

Problem 3. (10 points) Let f be μ-strongly convex and G-Lipschitz continuous over the constraint $\mathcal{C}$. Let $\eta_t = \frac{2}{\mu(t+1)}$ and $\bar{\mathbf{x}}_t = \sum_{k=1}^{t} \frac{2k}{t(t+1)} \mathbf{x}_k$. Prove that the projected subgradient descent obeys

(a)

$$f(\bar{\mathbf{x}}_t) - f^* \leq \frac{2G^2}{\mu(t+1)};$$

(b)

$$\|\bar{\mathbf{x}}_t - \mathbf{x}^*\|_2 \leq \frac{2G}{\mu\sqrt{t+1}}.$$

Solution.

(a) In the class, we have shown that

$$\sum_{k=0}^{t} k(f(\mathbf{x}_k) - f^*) \leq \frac{tG^2}{\mu}.$$

By Jensen inequality, we have

$$\sum_{k=0}^{t} k(f(\mathbf{x}_k) - f^*) = \frac{t(t+1)}{2}\left(\sum_{k=1}^{t} \frac{2k}{t(t+1)} f(\mathbf{x}_k) - f^*\right) \geq \frac{t(t+1)}{2}(f(\bar{\mathbf{x}}_t) - f^*).$$

Thus we can get

$$f(\bar{\mathbf{x}}_t) - f^* \leq \frac{2G^2}{\mu(t+1)}.$$

(b) By strong convexity and (a), we have

$$\frac{\mu}{2}\|\bar{\mathbf{x}}_t - \mathbf{x}^*\|_2^2 \leq \langle \nabla f(\mathbf{x}^*), \bar{\mathbf{x}}_t - \mathbf{x}^* \rangle + \frac{\mu}{2}\|\bar{\mathbf{x}}_t - \mathbf{x}^*\|_2^2 \leq f(\bar{\mathbf{x}}_t) - f^* \leq \frac{2G^2}{\mu(t+1)}.$$