Coursework (3) for *Introductory Lectures on Optimization*

Hanxuan Li
3220106039

Nov. 23, 2023

Excercise 1. Proof that μ-Strongly convex functions are μ-Error Bounds.

Proof of Excercise 1: Suppose f is μ convex, then for $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^n$,we have:

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{\mu}{2} \|\boldsymbol{y} - \boldsymbol{x}\|^2 \tag{1}$$

Thus we have:

$$\min_{\boldsymbol{y}} \{f(\boldsymbol{y})\} \geq \min_{\boldsymbol{y}} \left\{ f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{\mu}{2} \|\boldsymbol{y} - \boldsymbol{x}\|^2 \right\} \tag{2}$$

LHS is $f(\boldsymbol{x}^*)$,And the minumum of RHS can be solved by$\nabla RHS = \nabla f(\boldsymbol{x}) + \mu(\boldsymbol{y} - \boldsymbol{x}) = 0$

Thus we have the optimal $\hat{\boldsymbol{y}} = \boldsymbol{x} - \frac{1}{\mu} \nabla f(\boldsymbol{x})$, so we have:

$$f(\boldsymbol{x}^*) \geq f(\boldsymbol{x}) - \frac{1}{2\mu} \|\nabla f(\boldsymbol{x})\|^2 \tag{3}$$

That is:

$$\|\nabla f(\boldsymbol{x})\|^2 \geq 2\mu(f(\boldsymbol{x}) - f(\boldsymbol{x}^*)) \tag{4}$$

And by the definition of strong convex function and $\nabla f(\boldsymbol{x}^*) = 0$ we have:

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) + \langle \nabla f(\boldsymbol{x}^*), \boldsymbol{x} - \boldsymbol{x}^* \rangle + \frac{\mu}{2} \|\boldsymbol{x} - \boldsymbol{x}^*\|^2 \tag{5}$$

$$= f(\boldsymbol{x}^*) + \frac{\mu}{2} \|\boldsymbol{x} - \boldsymbol{x}^*\|^2 \tag{6}$$

Substitute (6) into (4) we have:

$$\|\nabla f(\boldsymbol{x})\|^2 \geq \mu^2 \|\boldsymbol{x} - \boldsymbol{x}^*\|^2 \tag{7}$$

That is $\|\nabla f(\boldsymbol{x})\| \geq \mu \|\boldsymbol{x} - \boldsymbol{x}^*\|$, which is exactly $\mu-$Error Bounds. □

Excercise 2. Let f be continuously differentiable. Proof that both conditions below, holding for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, are equivalent to inclusion: $\mathcal{S}^1_\mu(\mathbb{R}^n)$

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}),\ \boldsymbol{x} - \boldsymbol{y} \rangle \geq \mu \left\| \boldsymbol{x} - \boldsymbol{y} \right\|^2, \tag{8}$$

and

$$\begin{aligned} \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \geq\ & f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \\ & + \alpha(1-\alpha)\frac{\mu}{2} \left\| \boldsymbol{x} - \boldsymbol{y} \right\|^2. \end{aligned} \tag{9}$$

Proof of Excercise 2: Consider the definition of strong convex function:

$$f \in \mathcal{S}_\mu^1 \iff \exists \mu > 0, s.t. \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{1}{2}\mu \|\boldsymbol{y} - \boldsymbol{x}\|^2 \tag{10}$$

By transformation, we can obtain:

$$f(\boldsymbol{y}) - \frac{\mu}{2}\|\boldsymbol{y}\|^2 \geq f(\boldsymbol{x}) - \frac{\mu}{2}\|\boldsymbol{x}\|^2 + \langle \nabla f(\boldsymbol{x}) - \mu \boldsymbol{x}, \boldsymbol{y} - \boldsymbol{x} \rangle \tag{11}$$

By the definition of convex function, we have $f(\boldsymbol{x}) - \frac{\mu}{2}\|x\|^2$ is convex.

Recall that f is convex $\iff$ $\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \geq 0$ So

$$\langle (\nabla f(\boldsymbol{x}) - \mu \boldsymbol{x}) - (\nabla f(\boldsymbol{y}) - \mu \boldsymbol{y}), \boldsymbol{y} - \boldsymbol{x} \rangle \geq 0 \tag{12}$$

That is, $\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \geq \mu \|\boldsymbol{x} - \boldsymbol{y}\|^2$

The proof above is euqivalent.

For the second condition, recall that the equivalent definition of the convex function:

$$For \alpha \in [0,1], \forall \boldsymbol{x}, \boldsymbol{y}, \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \geq f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \tag{13}$$

In this case, $f(\boldsymbol{x}) - \frac{\mu}{2}$ is convex, So:

$$\alpha(f(\boldsymbol{x}) - \frac{\mu}{2}\|\boldsymbol{x}\|^2) + (1-\alpha)(f(\boldsymbol{y}) - \frac{\mu}{2}\|\boldsymbol{y}\|^2) \geq f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) - \frac{\mu}{2}\|\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}\|^2 \tag{14}$$

By transformation, we can obtain the result:

$$\begin{aligned} \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) &\geq f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) + \alpha \frac{\mu}{2}\|\boldsymbol{x}\|^2 + (1-\alpha)\frac{\mu}{2}\|\boldsymbol{y}\|^2 - \frac{\mu}{2}\|\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}\|^2 && (15) \\ &= f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) + \alpha(1-\alpha)\frac{\mu}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2 && (16) \end{aligned}$$

The proof above is equivalent as well.

$\square$