Coursework (2) for *Introductory Lectures on Optimization*

Your name
Your ID

Oct. 8, 2022

Excercise 1. Prove that

$$0 \le f(\boldsymbol{y}) - f(\boldsymbol{x}) - \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle \le \frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \tag{1}$$

holds if we have

$$0 \le \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) - f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \le \alpha(1-\alpha)\frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2. \tag{2}$$

Proof of Excercise 1: Let $\beta = 1 - \alpha$. So inequality 2 can be rewritten as:

$$\begin{aligned} &0 \le \beta(f(\boldsymbol{y}) - f(\boldsymbol{x})) - [f(\boldsymbol{x} + \beta(\boldsymbol{y} - \boldsymbol{x}) - f(\boldsymbol{x})] \le \beta(1-\beta)\frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \\ &\Rightarrow 0 \le f(\boldsymbol{y}) - f(\boldsymbol{x}) - \frac{f(\boldsymbol{x} + \beta(\boldsymbol{y} - \boldsymbol{x}) - f(\boldsymbol{x})}{\beta} \le (1-\beta)\frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \end{aligned} \tag{3}$$

Let $\beta \to 0$ in inequality 3, we can get inequality 1. □

Excercise 2. Prove that

$$f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{1}{2L}\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|^2 \le f(\boldsymbol{y}) \tag{4}$$

holds if we have

$$\alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \ge f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) + \frac{\alpha(1-\alpha)}{2L}\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|^2. \tag{5}$$

Proof of Excercise 2: Let $\beta = 1 - \alpha$. So inequality 5 can be rewritten as:

$$\begin{aligned} &(1-\beta) f(\boldsymbol{x}) + \beta f(\boldsymbol{y}) \ge f(\boldsymbol{x} + \beta(\boldsymbol{y} - \boldsymbol{x})) + \frac{\beta(1-\beta)}{2L}\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|^2 \\ &\Rightarrow \beta f(\boldsymbol{y}) \ge \beta f(\boldsymbol{x}) + [f(\boldsymbol{x} + \beta(\boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{x})] + \frac{\beta(1-\beta)}{2L}\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|^2 \\ &\Rightarrow f(\boldsymbol{y}) \ge f(\boldsymbol{x}) + \frac{f(\boldsymbol{x} + \beta(\boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{x})}{\beta} + \frac{(1-\beta)}{2L}\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|^2 \end{aligned} \tag{6}$$

Let $\beta \to 0$ in inequality 6, we can get inequality 4. □

Excercise 3. Let f be continuously differentiable. Prove that both conditions below, holding for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, are equivalent to inclusion $\mathcal{S}^1_\mu(\mathbb{R}^n)$

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}),\ \boldsymbol{x} - \boldsymbol{y} \rangle \ge \mu\|\boldsymbol{x} - \boldsymbol{y}\|^2, \tag{7}$$

$$\alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \ge f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) + \alpha(1-\alpha)\frac{\mu}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2. \tag{8}$$

Proof of Excercise 3: For $f \in \mathcal{S}_{\mu}^{1}(\mathbb{R}^n)$, it should satisfies

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{1}{2}\mu \|\boldsymbol{y} - \boldsymbol{x}\|^2. \tag{9}$$

Let's consider inequality 7. Here we define $\boldsymbol{x}_\tau = \boldsymbol{x} + \tau(\boldsymbol{y} - \boldsymbol{x})$

$$\begin{aligned}
f(\boldsymbol{y}) &= f(\boldsymbol{x}) + \int_0^1 \langle \nabla f(\boldsymbol{x} + \tau(\boldsymbol{y} - \boldsymbol{x})),\ \boldsymbol{y} - \boldsymbol{x} \rangle \mathrm{d}\tau \\
&= f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle + \int_0^1 \langle \nabla f(\boldsymbol{x} + \tau(\boldsymbol{y} - \boldsymbol{x})) - \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle \mathrm{d}\tau \\
&= f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle + \int_0^1 \frac{1}{\tau} \langle \nabla f(\boldsymbol{x}_\tau) - \nabla f(\boldsymbol{x}),\ \boldsymbol{x}_\tau - \boldsymbol{x} \rangle \mathrm{d}\tau \\
&\overset{(7)}{\geq} f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle + \int_0^1 \frac{1}{\tau} \mu \|\boldsymbol{x}_\tau - \boldsymbol{x}\|^2 \mathrm{d}\tau \\
&= f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle + \int_0^1 \mu\tau \|\boldsymbol{y} - \boldsymbol{x}\|^2 \mathrm{d}\tau \\
&= f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{1}{2}\mu \|\boldsymbol{y} - \boldsymbol{x}\|^2.
\end{aligned} \tag{10}$$

Let's consider inequality 8 and define $\beta = 1 - \alpha$. Therefore, inequality 8 can be rewritten as

$$\begin{aligned}
\beta(f(\boldsymbol{y}) - f(\boldsymbol{x})) - [f(\boldsymbol{x} + \beta(\boldsymbol{y} - \boldsymbol{x}) - f(\boldsymbol{x})] &\leq \beta(1-\beta)\frac{\mu}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \\
\Rightarrow f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \frac{f(\boldsymbol{x} + \beta(\boldsymbol{y} - \boldsymbol{x}) - f(\boldsymbol{x})}{\beta} &+ (1-\beta)\frac{\mu}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2
\end{aligned} \tag{11}$$

Let $\beta \to 0$ in inequality 11, we can get inequality 9. $\square$