Coursework (3) for *Introductory Lectures on Optimization*

Zhou Nan
3220102535

November 6, 2024

Excercise 1. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is L-smooth ($C_L^{1,1}$), and μ-PL, that is

$$\mu\text{-PL:} \ \ \frac{1}{2}\|\nabla f(\boldsymbol{x})\|_2^2 \geq \mu\left(f(\boldsymbol{x}) - f(\boldsymbol{x}^*)\right),$$

then GD iterates with step size $h_k = 1/L$ converge linearly, i.e.

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) \leq \left(1 - \frac{\mu}{L}\right)^k \left(f(\boldsymbol{x}_0) - f(\boldsymbol{x}^*)\right).$$

Proof of Excercise 1: According to the lemma 9, we have
for any $\boldsymbol{x}, \boldsymbol{y}$ from $\mathbb{R}^n$,

$$|f(\boldsymbol{x}) - f(\boldsymbol{y}) - \langle \nabla f(\boldsymbol{x}), \boldsymbol{x} - \boldsymbol{y}\rangle| \leq \frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_2^2.$$

Since $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \frac{1}{L}\nabla f(\boldsymbol{x}_k)$, then

$$\begin{aligned}
&|f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) - \langle \nabla f(\boldsymbol{x}_k), \boldsymbol{x}_{k+1} - \boldsymbol{x}_k\rangle| \leq \frac{L}{2}\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|_2^2. \\
\Rightarrow &|f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) + \frac{1}{L}\|\nabla f(\boldsymbol{x}_k)\|_2^2| \leq \frac{L}{2}\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|_2^2. \\
\Rightarrow &|f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) + \frac{1}{L}\|\nabla f(\boldsymbol{x}_k)\|_2^2| \leq \frac{L}{2}\left(\frac{1}{L}\right)^2\|\nabla f(\boldsymbol{x}_k)\|_2^2. \\
\Rightarrow &-\frac{3}{2L}\|\nabla f(\boldsymbol{x}_k)\|_2^2 \leq f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) \leq -\frac{1}{2L}\|\nabla f(\boldsymbol{x}_k)\|_2^2.
\end{aligned}$$

Since the function f is μ-PL, then

$$\frac{1}{2}\|\nabla f(\boldsymbol{x}_k)\|_2^2 \geq \mu\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)\right).$$

Therefore,

$$\begin{aligned}
&L(f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k)) + \mu(f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)) \\
\leq &-\frac{1}{2}\|\nabla f(\boldsymbol{x}_k)\|_2^2 + \frac{1}{2}\|\nabla f(\boldsymbol{x}_k)\|_2^2 = 0.
\end{aligned}$$

Then,

$$\begin{aligned}
&f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) \leq -\frac{\mu}{L}(f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)) \\
&f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k) + f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) \leq (1 - \frac{\mu}{L})(f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)). \\
&f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}^*) \leq (1 - \frac{\mu}{L})(f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)).
\end{aligned}$$

Therefore,

$$\begin{aligned}
f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}^*) &\leq (1 - \frac{\mu}{L})(f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)) \\
f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) &\leq (1 - \frac{\mu}{L})(f(\boldsymbol{x}_{k-1}) - f(\boldsymbol{x}^*)) \\
\Rightarrow f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}^*) &\leq (1 - \frac{\mu}{L})^2(f(\boldsymbol{x}_{k-1}) - f(\boldsymbol{x}^*)). \\
\Rightarrow f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) &\leq (1 - \frac{\mu}{L})^k(f(\boldsymbol{x}_0) - f(\boldsymbol{x}^*)).
\end{aligned}$$

□