8.3 Smooth Convex Optimization

Definition 8.1 ***(L-smooth)***
f is L-smooth (with a constant $L > 0$) on $\mathcal{X}$ if f is continuously differentiable and $\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2$ for all $x, y \in \mathcal{X}$.

Proposition 8.2 *The followings are equivalent.*
(a) f is convex and L-smooth.
(b) $0 \leq f(y) - f(x) - \nabla f(x)^T(y - x) \leq \frac{L}{2}\|x - y\|_2^2$ for all $x, y \in \mathcal{X}$.
(c) $f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|_2^2$ for all $x, y \in \mathcal{X}$.
(d) $\{\nabla f(x) - \nabla f(y)\}^T(x - y) \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|_2^2$ for all $x, y \in \mathcal{X}$.

Proof: (a) $\Rightarrow$ (b) By the fundamental theorem of calculus,

$$\begin{aligned}
f(y) - f(x) - \nabla f(x)^T(y-x) &= \int_0^1 \nabla f(x+t(y-x))^T(y-x)\,dt - \nabla f(x)^T(y-x) \\
&= \int_0^1 [\nabla f(x+t(y-x)) - \nabla f(x)]^T(y-x)\,dt \\
&\leq \int_0^1 \|\nabla f(x+t(y-x)) - \nabla f(x)\|_2 \|y-x\|_2\,dt \quad \text{(by Cauchy-Schwarz inequality)} \\
&\leq \int_0^1 L\|t(y-x)\|_2\|y-x\|_2\,dt \quad \text{(by L-smoothness of } f\text{)} \\
&= L\|(y-x)\|_2^2 \int_0^1 t\,dt \\
&= \frac{L}{2}\|(y-x)\|_2^2
\end{aligned}$$

(b) $\Rightarrow$ (c) Let $z = y + \frac{1}{L}(\nabla f(x) - \nabla f(y))$

$$\begin{aligned}
f(y) - f(x) &= f(y) - f(z) + f(z) - f(x) \\
&\geq -\nabla f(y)^T(z-y) - \frac{L}{2}\|y-z\|_2^2 + \nabla f(x)^T(z-x) \\
&= \nabla f(x)^T(y-x) - \{\nabla f(x) - \nabla f(y)\}^T(y-z) - \frac{L}{2}\|y-z\|_2^2 \quad \text{(by plugging in z)} \\
&= \nabla f(x)^T(y-x) + \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|_2^2 - \frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|_2^2 \\
&= \nabla f(x)^T(y-x) + \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|_2^2
\end{aligned}$$

(c) $\Rightarrow$ (d) Suppose that

$$\begin{aligned}
f(y) &\geq f(x) + \nabla f(x)^T(y-x) + \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|_2^2 \\
f(x) &\geq f(y) + \nabla f(y)^T(x-y) + \frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|_2^2
\end{aligned}$$

By summing up two inequalities, we can obtain

$$[\nabla f(x) - \nabla f(y)]^T(x-y) \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|_2^2$$

(d) $\Rightarrow$ (a) Suppose that $[\nabla f(x) - \nabla f(y)]^T(x-y) \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|_2^2$.

By Cauchy-Schwarz inequality,

$$\|\nabla f(x) - \nabla f(y)\|_2\|x-y\|_2 \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|_2^2$$

which implies f is L-smooth. The convexity of f is due to the following claim.
Therefore, (a), (b), (c), and (d) are equivalent. $\blacksquare$