Solution to Homework 3

Total 20 points

Problem 1. (4 points) Judge whether the following functions are smooth.

(a) $f(x) = \sin x$.

(b) $f(\mathbf{x}) = \|\mathbf{x}\|_1$, $\mathbf{x} \in \mathbb{R}^d$.

Solution.

(a) Since $|f''(x)| = |\sin x| \leq 1$, $f(x)$ is 1-smooth.

(b) $f(\mathbf{x})$ is not smooth since it is not differentiable.

Problem 2. (4 points) Judge whether the following functions are strongly convex.

(a) $f(\mathbf{x}) = \sum_{i=1}^{m} (\mathbf{a}_i^\top \mathbf{x} - b_i)^2$, $\mathbf{a}_i, \mathbf{x} \in \mathbb{R}^d$, $m > d$.

(b) $f(x_1, x_2) = 1/(x_1 x_2)$, $x_1 > 0, x_2 > 0$.

Solution.

(a) Let $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_m]$, then $\nabla^2 f(\mathbf{x}) = \mathbf{A}\mathbf{A}^\top$. If $\mathbf{A}\mathbf{A}^\top$ is singular, then $f(\mathbf{x})$ is not strongly convex. If $\mathbf{A}\mathbf{A}^\top$ is non-singular, we suppose its minimum eigenvalue is $\lambda_{\min}$. Then $f(\mathbf{x})$ is $\lambda_{\min}$-smooth.

(b) Note that

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} 2x_1^{-3}x_2^{-1} & x_1^{-2}x_2^{-2} \\ x_1^{-2}x_2^{-2} & 2x_1^{-1}x_2^{-3} \end{pmatrix}.$$

When $x_1, x_2 \to \infty$, $\nabla^2 f(\mathbf{x}) \to \mathbf{0}$. Thus $f(x_1, x_2)$ is not strongly convex.

Problem 3. (12 points)

(a) Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is α-strongly convex and β-smooth for some $\beta > \alpha$. Show that $h(\mathbf{x}) = f(\mathbf{x}) - \frac{\alpha}{2}\|\mathbf{x}\|^2$ is $(\beta - \alpha)$-smooth.

(b) Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is α-strongly convex and $g : \mathbb{R}^d \to \mathbb{R}$ is β-smooth. Prove that the function $h(x) = f(x) - g(x)$ is convex if $\alpha \geq \beta$. Is the converse true?

(c) Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is μ-strongly convex and L-smooth. Show that

$$\langle \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq \frac{\mu L}{\mu + L} \|\mathbf{x} - \mathbf{y}\|_2^2 + \frac{1}{\mu + L} \|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\|_2^2.$$

(hint: by the conclusion of (a), $h(\mathbf{x}) = f(\mathbf{x}) - \frac{\mu}{2}\|\mathbf{x}\|^2$ is $(L - \mu)$-smooth and convex.)

Solution.

(a) Since $f(\mathbf{x})$ is α-strongly convex, we know that $h(\mathbf{x})$ is convex. Notice that $\nabla h(\mathbf{x}) = \nabla f(\mathbf{x}) - \alpha \mathbf{x}$. Thus we can get

$$\begin{aligned}
&h(\mathbf{y}) - h(\mathbf{x}) - \langle \nabla h(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \\
=&f(\mathbf{y}) - f(\mathbf{x}) + \frac{\alpha}{2}(\|\mathbf{x}\|_2^2 - \|\mathbf{y}\|_2^2) - \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \alpha \langle \mathbf{x}, \mathbf{y} - \mathbf{x} \rangle \\
=&f(\mathbf{y}) - f(\mathbf{x}) - \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle - \frac{\alpha}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 \\
\leq&\frac{\beta - \alpha}{2} \|\mathbf{x} - \mathbf{y}\|_2^2,
\end{aligned}$$

where the last inequality comes from the fact that $f(\mathbf{x})$ is β-smooth. Thus $h(\mathbf{x})$ is $(\beta - \alpha)$-smooth.

(b) We have that

$$\begin{aligned}
&h(\mathbf{y}) - h(\mathbf{x}) - \langle \nabla h(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \\
=&(f(\mathbf{y}) - f(\mathbf{x}) - \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle) - (g(\mathbf{y}) - g(\mathbf{x}) - \langle \nabla g(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle) \\
\geq&\frac{\alpha}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 - \frac{\beta}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 \geq 0
\end{aligned}$$

where in the last line we used the α-strong convexity condition on f and the β-smoothness of g. The converse is false. A simple counter example is $f(\mathbf{x}) = g(\mathbf{x}) = \mathbf{x}^\top \mathbf{Q} \mathbf{x}$ where $\mathbf{Q} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, Then $f(\mathbf{x})$ is 1-strongly convex and $g(\mathbf{x})$ is 2-smooth. We find that $h(\mathbf{x}) = 0$ is convex but $1 < 2$.

(c) By the conclusion of (a), $h(\mathbf{x}) = f(\mathbf{x}) - \frac{\mu}{2}\|\mathbf{x}\|^2$ is $(L - \mu)$-smooth and convex, i.e.,

$$\langle \nabla h(\mathbf{x}) - \nabla h(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq \frac{1}{L - \mu} \|\nabla h(\mathbf{x}) - \nabla h(\mathbf{y})\|^2.$$

Since $\nabla h(\mathbf{x}) = \nabla f(\mathbf{x}) - \mu \mathbf{x}$, we have

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

which indicates

$$\langle \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq \frac{\mu L}{\mu + L} \|\mathbf{x} - \mathbf{y}\|_2^2 + \frac{1}{\mu + L} \|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\|_2^2.$$

Remark: For the solution to the bonus homework in Lecture 4, please refer to Theorem 2.1.5 in the book "Lectures on Convex Optimization" by Nesterov.