Solution to Homework 2

Total 30 points

Problem 1. (6 points) Judge which of the following functions are (strict) convex.

(a) $f(x_1, x_2) = x_1 x_2$, $x_1 > 0, x_2 > 0$.

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

(c) $f(x_1, x_2) = x_1^2/x_2$, $x_2 > 0$.

Solution.

(a) Note that

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which is neither positive semi-definite nor negative semi-definite. Thus, $f(x_1, x_2)$ is not convex or concave.

(b) Note that

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

Since $2x_1^{-3}x_2^{-1} > 0$ and $\det(\nabla^2 f(\mathbf{x})) = 3x_1^{-4}x_2^{-4} > 0$ if $x_1 > 0$ and $x_2 > 0$, we immediately know that $\nabla^2 f(\mathbf{x})$ is positive definite and $f(x_1, x_2)$ is strictly convex.

(c) Note that

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} 2x_2^{-1} & -2x_1x_2^{-2} \\ -2x_1x_2^{-2} & 2x_1^2x_2^{-3} \end{pmatrix}.$$

Since $2x_2^{-1} > 0$ if $x_2 > 0$ and $\det(\nabla^2 f(\mathbf{x})) = 0$, we know $\nabla^2 f(\mathbf{x})$ is positive semi-definite, and thus $f(x_1, x_2)$ is convex but not strictly convex.

Problem 2. (6 points) Prove that $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for all $\mathbf{x}, \mathbf{y} \in \text{dom } f$ and $\mathbf{x} \neq \mathbf{y}$, the function $g(t) = f(t\mathbf{x} + (1-t)\mathbf{y})$ is a convex function on $[0, 1]$.

$\Rightarrow$: If f is convex, then for $t_1, t_2 \in [0, 1]$ and $\lambda \in [0, 1]$, we have

$$\begin{aligned} g(\lambda t_1 + (1-\lambda)t_2) =& f([\lambda t_1 + (1-\lambda)t_2]\mathbf{x} + [1 - (\lambda t_1 + (1-\lambda)t_2)]\mathbf{y}) \\ =& f(\lambda(t_1\mathbf{x} + (1-t_1\mathbf{y})) + (1-\lambda)(t_2\mathbf{x} + (1-t_2)\mathbf{y}) \\ \leq& \lambda f(t_1\mathbf{x} + (1-t_1\mathbf{y})) + (1-\lambda)f(t_2\mathbf{x} + (1-t_2)\mathbf{y}) \\ =& \lambda g(t_1) + (1-\lambda)g(t_2). \end{aligned}$$

Therefore, $g(t)$ is a convex function on $[0, 1]$.

$\Leftarrow$: If $g(t) = f(t\mathbf{x} + (1-t)\mathbf{y})$ is convex on $[0,1]$, then for $\lambda \in [0,1]$ we have

$$f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) = g(\lambda) = g(\lambda \cdot 1+(1-\lambda)\cdot 0) \leq \lambda g(1)+(1-\lambda)g(0) = \lambda f(\mathbf{x})+(1-\lambda)f(\mathbf{y}).$$

Therefore, we conclude f is also convex.

Problem 3. (6 points) Prove that for any convex function $f : \mathbb{R}^n \to \mathbb{R}$, the Bregman distance $B_f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y}\rangle$ is convex in $\mathbf{x}$ but not necessarily in $\mathbf{y}$. **Solution.** As $B_f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y}\rangle$, then we have

$$\begin{aligned}
&B_f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \mathbf{y}) \\
=&f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 - \mathbf{y}\rangle \\
\leq&\theta f(\mathbf{x}_1) + (1-\theta)f(\mathbf{x}_2) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 - \mathbf{y}\rangle \\
=&\theta(f(\mathbf{x}_1) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \mathbf{x}_1 - \mathbf{y}\rangle) + (1-\theta)(f(\mathbf{x}_2) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \mathbf{x}_2 - \mathbf{y}\rangle) \\
=&\theta B_f(\mathbf{x}_1, \mathbf{y}) + (1-\theta)B_f(\mathbf{x}_2, \mathbf{y})
\end{aligned}$$

thus, $B_f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$.

Let $f : \mathbb{R} \to \mathbb{R}$, $f(t) = t^3$ then $B_f(x, y) = x^3 - y^3 - 3y^2(x-y)$, $\frac{\partial B_f(x,y)}{\partial y} = -6y^2 - 6xy$, $\frac{\partial^2 B_f(x,y)}{\partial y^2} = 12y - 6x$ is not positive semidefinite, then $B_f(\mathbf{x}, \mathbf{y})$ is not necessarily convex in $\mathbf{y}$.

Problem 4. (6 points) Compute the subdifferentials of the following functions

(a) $f(\mathbf{x}) = \|\mathbf{x}\|_2$.

(b) Given a closed convex set $\mathcal{C}$, define

$$f(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \in \mathcal{C} \\ +\infty & \text{otherwise.} \end{cases}$$

Solution.

(a)

$$\partial f(\mathbf{x}) = \begin{cases} \frac{\mathbf{x}}{||\mathbf{x}||_2} & \text{if } \mathbf{x} \neq 0 \\ \{\mathbf{g}|\ ||\mathbf{g}||_2 \leq 1\} & \text{if } \mathbf{x} = 0 \end{cases}$$

(b)

$$\partial f(\mathbf{x}) = \begin{cases} \emptyset & \text{if } \mathbf{x} \notin \mathcal{C} \\ \{\mathbf{g}|\ \langle \mathbf{g}, \mathbf{y} - \mathbf{x}\rangle \leq 0, \forall \mathbf{y} \in \mathcal{C}\} & \text{if } \mathbf{x} \in \partial\mathcal{C} \\ 0 & \text{if } \mathbf{x} \in \mathcal{C}^\circ \end{cases}$$

Problem 5. (6 points) If function f is convex, Show that $\partial f(\mathbf{x}) \neq \emptyset$ for all $\mathbf{x} \in (dom\ f)^\circ$. **Solution.** Notice that $(\mathbf{x}, f(\mathbf{x}))$ is on the boundary of epi f. The hyperplane supporting theorem say there exists $(\mathbf{a}, b)$ with $\mathbf{a} \neq \mathbf{0}$ such that

$$\left\langle \begin{bmatrix} \mathbf{a} \\ b \end{bmatrix}, \begin{bmatrix} \mathbf{y} - \mathbf{x} \\ t - f(\mathbf{x}) \end{bmatrix} \right\rangle \leq 0$$

for any $(\mathbf{y}, t) \in \text{epi } f$, which means

$$S \triangleq \langle \mathbf{a}, \mathbf{y} - \mathbf{x} \rangle + b(t - f(\mathbf{x})) \leq 0.$$

We can conclude $b \leq 0$, otherwise, let $t \to +\infty$, then S goes to $+\infty$.

Since $\mathbf{x}$ is in the interior, we can find some $\epsilon > 0$ such that $\mathbf{y} = \mathbf{x} + \epsilon \mathbf{a} \in \text{dom } f$, which leads to $S = \epsilon \|\mathbf{a}\|_2^2 + b(t - f(\mathbf{x}))$. Let $t > f(\mathbf{x})$, then we know $b \neq 0$. Hence we can say $b < 0$. Thus, $\langle \mathbf{a}/b, \mathbf{y} - \mathbf{x} \rangle + (t - f(\mathbf{x})) \geq 0$, i.e., $t \geq f(\mathbf{x}) + \langle -\mathbf{a}/b, \mathbf{y} - \mathbf{x} \rangle$.

Take $t = f(\mathbf{y})$ means $\mathbf{g} = -\mathbf{a}/b$ is a subgradient at $\mathbf{x}$.