Solution to Homework 2

Problem 1. Which of the following sets are convex?

(a) A slab, i.e., a set of the form $\{\mathbf{x} \in \mathbb{R}^n | \alpha \leq \mathbf{a}^{\mathrm{T}}\mathbf{x} \leq \beta\}$.

(b) The set of points closer to a given point than a given set, i.e., $\{\mathbf{x} | \|\mathbf{x} - \mathbf{x}_0\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_2 \text{ for all } \mathbf{y} \in S\}$ where $S \subseteq \mathbb{R}^n$.

(c) The set of points closer to one set than another, i.e., $\{\mathbf{x} | \mathbf{dist}(\mathbf{x}, S) \leq \mathbf{dist}(\mathbf{x}, T)\}$ where $S, T \subseteq \mathbb{R}^n$, and $\mathbf{dist}(\mathbf{x}, S) = \inf\{\|\mathbf{x} - \mathbf{z}\|_2 | \mathbf{z} \in S\}$.

(d) The set of points whose distance to $\mathbf{a}$ does not exceed a fixed fraction θ of the distance to $\mathbf{b}$, i.e., the set $\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\|_2 \leq \theta \|\mathbf{x} - \mathbf{b}\|_2\}$ ($\mathbf{a} \neq \mathbf{b}$ and $0 \leq \theta \leq 1$).

Solution.

(a) A slab is convex. If $\alpha \leq \mathbf{a}^{\mathrm{T}}\mathbf{x} \leq \beta, \alpha \leq \mathbf{a}^{\mathrm{T}}\mathbf{y} \leq \beta$ and $0 \leq \theta \leq 1$, then it's obvious that $\alpha \leq \mathbf{a}^{\mathrm{T}}(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \beta$. Alternatively, you can regard a slab as an intersection of two half-spaces.

(b) The set of points closer to a given point than a given set is convex. In fact, the condition $\|\mathbf{x} - \mathbf{x}_0\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_2$ for all $\mathbf{y} \in S$ is equivalent to $(\mathbf{y} - \mathbf{x}_0)^{\mathrm{T}}\mathbf{x} \leq \frac{1}{2}(\|\mathbf{y}\|_2^2 - \|\mathbf{x}_0\|_2^2)$ for any $\mathbf{y} \in S$. Therefore, the set can be rewritten as

$$\cap_{\mathbf{y} \in S} \{\mathbf{x} | (\mathbf{y} - \mathbf{x}_0)^{\mathrm{T}}\mathbf{x} \leq \frac{1}{2}(\|\mathbf{y}\|_2^2 - \|\mathbf{x}_0\|_2^2)\},$$

which is an intersection of convex sets.

(c) The set of points closer to one set than another may be non-convex. For example, we take $S = \{-2, 2\}, T = \{0\}$, then the set is $(-\infty, -1] \cup [1, \infty)$, which is non-convex.

(d) The set of points whose distance to $\mathbf{a}$ does not exceed a fixed fraction θ of the distance to $\mathbf{b}$ is convex. When $\theta = 0$ or $\theta = 1$, the conclusion is trivial. When $\theta \in (0, 1)$, the condition $\|\mathbf{x} - \mathbf{a}\|_2 \leq \theta \|\mathbf{x} - \mathbf{b}\|_2\}$ is equivalent to

$$\left\| \mathbf{x} - \frac{\mathbf{a} - \theta^2 \mathbf{b}}{1 - \theta^2} \right\|_2^2 \leq \frac{1}{1 - \theta^2} \left(\frac{\|\mathbf{a} - \theta\mathbf{b}\|_2^2}{1 - \theta^2} - \|\mathbf{a}\|_2^2 + \theta^2 \|\mathbf{b}\|_2^2 \right).$$

Therefore, the set is a ball, which is convex.

Problem 2. Judge which of the following functions are (strict) convex.

(a) $f(x) = e^x - 1$.

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

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

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

Solution.

(a) Note that $f''(x) = e^x > 0$, thus $f(x)$ is strictly convex.

(b) 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.

(c) 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.

(d) 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 3. Prove that $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only of for every $\mathbf{x} \neq \mathbf{y} \in \mathrm{dom} f$, 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 4. Prove that if f is a convex function, then for all $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$, and a_1, a_2 and $a_3 \in (0,1)$ such that $a_1 + a_2 + a_3 = 1$, we have

$$\langle \nabla f(\mathbf{x}_3), a_1\mathbf{x}_1 + a_2\mathbf{x}_2 - (1-a_3)\mathbf{x}_3 \rangle \leq a_1 f(\mathbf{x}_1) + a_2 f(\mathbf{x}_2) - (1-a_3) f(\mathbf{x}_3).$$

Solution. By the first-order condition and the Jensen inequality, we have

$$\begin{aligned}\langle \nabla f(\mathbf{x}_3), a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + a_3\mathbf{x}_3 - \mathbf{x}_3 \rangle &\leq f(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + a_3\mathbf{x}_3) - f(\mathbf{x}_3) \\ &\leq a_1 f(\mathbf{x}_1) + a_2 f(\mathbf{x}_2) + a_3 f(\mathbf{x}_3) - f(\mathbf{x}_3).\end{aligned}$$