Homework 3

Zhen

October 14, 2021

Problem 1

Suppose f is a convex function and $S \subset \text{dom } f$ is a convex set. Let M be the set of global minima of f over S,

$$M = \{\boldsymbol{x}^* \in S : f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}), \forall \boldsymbol{x} \in S\}$$

Show that M is a convex set.

Proof:

Apparently, $M = \{\boldsymbol{x} : f(\boldsymbol{x}) = C\}$ where C is a const and $\forall \boldsymbol{x} \in C, f(\boldsymbol{x}) \geq C$.

$\forall \boldsymbol{x}, \boldsymbol{y} \in M$ and $\theta \in (0,1)$,we have

$$C \leq f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y}) = C,$$

which means that $f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = C$.

Thus $\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} \in M$ and M is convex.

Problem 2

Let f be convex. If $f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$ for some $\boldsymbol{x}, \boldsymbol{y}$ and $\theta = \theta_0 \in (0,1)$, then it holds for the same $\boldsymbol{x}, \boldsymbol{y}$ and any $\theta \in [0,1]$.

Proof by contradiction:

Without loss of generality, assume that $f(\theta_1\boldsymbol{x} + \bar{\theta_1}\boldsymbol{y}) < \theta_1 f(\boldsymbol{x}) + \bar{\theta_1} f(\boldsymbol{y})$ where $\theta_1 \in (\theta_0, 1)$.

Then we have

$$\begin{aligned}
\frac{\theta_0}{\theta_1} f\big(\theta_1\boldsymbol{x} + \bar{\theta_1}\boldsymbol{y}\big) + \left(1 - \frac{\theta_0}{\theta_1}\right) f(\boldsymbol{y}) &< \frac{\theta_0}{\theta_1}\big(\theta_1 f(\boldsymbol{x}) + \bar{\theta_1} f(\boldsymbol{y})\big) + \left(1 - \frac{\theta_0}{\theta_1}\right) f(\boldsymbol{y}) \\
&= \theta_0 f(\boldsymbol{x}) + (1-\theta_0) f(\boldsymbol{y}) \\
&= f\big(\theta_0\boldsymbol{x} + \bar{\theta_0}\boldsymbol{y}\big) \\
&= f\left(\frac{\theta_0}{\theta_1}\big(\theta_1\boldsymbol{x} + \bar{\theta_1}\boldsymbol{y}\big) + \left(1 - \frac{\theta_0}{\theta_1}\right)\boldsymbol{y}\right)
\end{aligned}$$

Let $\theta_2 = \dfrac{\theta_0}{\theta_1}$, $\boldsymbol{x}_1 = \big(\theta_1\boldsymbol{x} + \bar{\theta_1}\boldsymbol{y}\big)$ and $\boldsymbol{x}_2 = \boldsymbol{y}$. Thus we get

$$\theta_2 f(\boldsymbol{x}_1) + \bar{\theta_2} f(\boldsymbol{x}_2) < f\big(\theta_2\boldsymbol{x}_1 + \bar{\theta_2}\boldsymbol{x}_2\big)$$

which deduce a contradiction.

Therefore, $\forall \theta \in (0,1) : \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y}) \leq f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$.

Then $f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$, $\forall \theta \in (0,1)$.

Problem 3

Determine if the following functions are convex, concave, or neither.

(a) $f(\boldsymbol{x}) = f(x_1, x_2, x_3) = x_1^2 + x_1x_3 + x_2^2 + x_2x_3 + \frac{1}{2}x_3^2$ on $\mathbb{R}^3$

Solve: Convex.

Assume $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$. By the fact that

$$f(\boldsymbol{x}) = \left(x_1 + \frac{1}{2}x_3\right)^2 + \left(x_2 + \frac{1}{2}x_3\right)^2$$

we have $\nabla^2 f = 2\boldsymbol{A} \succeq \boldsymbol{O}$. Therefore f is convex.

(b) $f(\boldsymbol{x}) = f(x_1, x_2) = (x_1x_2)^{-1}$ on $\mathbb{R}^2_{++} = \{(x_1, x_2) : x_1 > 0, x_2 > 0\}$

Solve: Convex.

Considering that: $\dfrac{\partial^2 f}{\partial x_1^2} = \dfrac{2}{x_1^3x_2}, \dfrac{\partial^2 f}{\partial x_2^2} = \dfrac{2}{x_1x_2^3}, \dfrac{\partial^2 f}{\partial x_1\partial x_2} = \dfrac{1}{x_1^2x_2^2}$

Then the eigenvalue λ_1, λ_2 of $\nabla^2 f$ satisfies that:

$$(\lambda - \frac{2}{x_1^3x_2})(\lambda - \frac{2}{x_1x_2^3}) - \frac{1}{x_1^4x_2^4} = 0,\ i = 1, 2$$

Let $f(\lambda)$ denote **LHS** of the above equation. Then $f(\lambda) = 0$ is a parabola with an axis of symmetry $\lambda = \dfrac{1}{x_1^3x_2} + \dfrac{1}{x_1x_2^3} > 0$.

And considering that $f(0) > 0$ and $f(\lambda) = 0$ is of 2 unique solutions, we have $\lambda_1, \lambda_2 > 0$. Thus $\nabla^2 f \succeq \boldsymbol{O}$ and f is convex.

(c) $f(x_1, x_2) = x_1x_2^2$ on $\mathbb{R}^2_{++}$

Solve: Neither.

$$\nabla^2 f = \begin{pmatrix} 0 & 2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}$$

Then the eigenvalue of $\nabla^2 f$ $\lambda_\pm = x_1 \pm \sqrt{x_1^2 + 4x_2^2}$. Thus f is neither convex nor concave ($\lambda_+ > 0$ while $\lambda_- < 0$).

(d) $f(x_1, x_2) = x_1x_2^{-1/2}$ on $\mathbb{R}^2_{++}$

Solve: Neither.

Considering that: $\dfrac{\partial^2 f}{\partial x_1^2} = 0,\ \dfrac{\partial^2 f}{\partial x_1\partial x_2} = -\dfrac{1}{2x_2^{3/2}},\ \dfrac{\partial^2 f}{\partial x_2^2} = \dfrac{3x_1}{4x_2^{5/2}}.$

Then the eigenvalue of $\nabla^2 f$

$$\lambda_\pm = \frac{1}{2}\left(\frac{3x_1}{4x_2^{5/2}} \pm \sqrt{\frac{9x_1^2}{16x_2^5} + \frac{1}{x_2^3}}\right)$$

Thus f is neither convex nor concave ($\lambda_+ > 0$ while $\lambda_- < 0$).

(e) $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ on $\mathbb{R}^2_{++}$

Solve: Concave.

Considering that

$$\nabla^2 f = -\alpha(1-\alpha)x_1^{\alpha-2}x_2^{-1-\alpha}\begin{bmatrix} x_2^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{bmatrix} \triangleq C(x_1, x_2, \alpha)\boldsymbol{A}$$

By the fact that $C(x_1, x_2, \alpha) < 0 \wedge \boldsymbol{A} \succeq \boldsymbol{O}$, $\nabla^2 f \preceq \boldsymbol{O}$ and f is concave.

Problem 4

Suppose $f_i : \mathbb{R} \to \mathbb{R}, i = 1, 2$, are strictly convex functions. Show that $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x_1, x_2) = f_1(x_1) + f_2(x_2)$ is strictly convex over $\mathbb{R}^2$, and in particular $f(x_1, x_2) = x_1^2 + x_2^4$ is strictly convex.

Proof:

$\forall \boldsymbol{x} = (x_1, x_2), \boldsymbol{y} = (y_1, y_2) \in \mathbb{R}^2$ and $\forall \theta \in (0, 1)$,

$$\begin{aligned} f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) &= f(\theta x_1 + \bar{\theta} y_1, \theta x_2 + \bar{\theta} y_2) \\ &= f_1(\theta x_1 + \bar{\theta} y_1) + f_2(\theta x_2 + \bar{\theta} y_2) \\ &\leq \theta f_1(x_1) + \bar{\theta} f_1(y_1) + \theta f_2(x_2) + \bar{\theta} f_2(y_2) \\ &= \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y}) \end{aligned}$$

Therefore f is strictly convex.

Trivially, $f_1(x_1) = x_1^2$ and $f_2(x_2) = x_2^4$ are strictly conevx. Then $f(x_1, x_2) = f_1(x_1) + f_2(x_2)$ is strictly convex.

Problem 5

Let $f : C \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function defined on a nonempty open convex set C. Show that f is convex if and only if

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \geq 0, \forall \boldsymbol{x}, \boldsymbol{y} \in C$$

Proof:

$$\begin{aligned} & f \text{ is convex} \\ \iff & \forall \text{direction } \boldsymbol{d}, \forall \boldsymbol{x} \in C, g(t) = f(\boldsymbol{x} + t\boldsymbol{d}) \text{ is convex on } \{t : \boldsymbol{x} + t\boldsymbol{d} \in C\} \\ \iff & \forall \text{direction } \boldsymbol{d}, \forall \boldsymbol{x} \in C, g'(t) = \nabla f(\boldsymbol{x} + t\boldsymbol{d})^T \boldsymbol{d} \text{ is increasing on } \{t : \boldsymbol{x} + t\boldsymbol{d} \in C\} \\ \iff & \forall \text{direction } \boldsymbol{d}, \forall \boldsymbol{x} \in C, \forall h, s \in \{t : \boldsymbol{x} + t\boldsymbol{d} \in C\}, [g'(h) - g'(s)]\,(h - s) \geq 0 \\ \iff & \forall \text{direction } \boldsymbol{d}, \forall \boldsymbol{x} \in C, \forall h, s \in \{t : \boldsymbol{x} + t\boldsymbol{d} \in C\}, \left[\nabla f(\boldsymbol{x} + h\boldsymbol{d})^T \boldsymbol{d} - f(\boldsymbol{x} + s\boldsymbol{d})^T \boldsymbol{d}\right](h - s) \geq 0 \\ \iff & \forall \text{direction } \boldsymbol{d}, \forall \boldsymbol{x} \in C, \forall h, s \in \{t : \boldsymbol{x} + t\boldsymbol{d} \in C\}, \left[\nabla f(\boldsymbol{x} + h\boldsymbol{d})^T - f(\boldsymbol{x} + s\boldsymbol{d})^T\right][(\boldsymbol{x} + h\boldsymbol{d}) - (\boldsymbol{x} + s\boldsymbol{d})] \geq 0 \\ \iff & \forall \boldsymbol{x}, \boldsymbol{y} \in C, \left[\nabla f(\boldsymbol{x})^T - \nabla f(\boldsymbol{y})^T\right](\boldsymbol{x} - \boldsymbol{y}) \geq 0 \text{ (Note that: } \exists s, \boldsymbol{d},\ \boldsymbol{x} = \boldsymbol{x} + 0\boldsymbol{d}, \boldsymbol{y} = \boldsymbol{x} + t\boldsymbol{d}) \\ \iff & \forall \boldsymbol{x}, \boldsymbol{y} \in C, \langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \geq 0 \end{aligned}$$