运筹与优化 Homework 2

12112627 李乐平

1.(a) Denote $\sum_{k=1}^{n} e^{x_k}$ as s.

Then

$$\nabla f(x) = \frac{1}{s}(e^{x_1}, e^{x_2}, \cdots, e^{x_n})$$

$$\nabla^2 f(x) =$$

$$\frac{1}{s^2}\begin{bmatrix} (s-e^{x_1})e^{x_1} & -e^{x_1}e^{x_2} & \cdots & -e^{x_1}e^{x_n} \\ -e^{x_2}e^{x_1} & (s-e^{x_2})e^{x_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ -e^{x_n}e^{x_1} & \cdots & \cdots & (s-e^{x_n})e^{x_n} \end{bmatrix}$$

$$= \frac{1}{s}\operatorname{diag}(e^{x_1}, e^{x_2}, \cdots, e^{x_n}) - \frac{1}{s^2}\begin{bmatrix} e^{x_1} \\ e^{x_2} \\ \vdots \\ e^{x_n} \end{bmatrix}\begin{bmatrix} e^{x_1} & e^{x_2} & \cdots & e^{x_n} \end{bmatrix}$$

For $\forall v = [v_1\ v_2 \cdots v_n]^T$.

$$v^T\nabla^2 f(x) v = \frac{1}{s}\sum_{k=1}^{n} v_k^2 e^{x_k} - \frac{1}{s^2}\left(\sum_{k=1}^{n} v_k e^{x_k}\right)^2$$

$$= \frac{1}{s^2}\left(\|a\|_2^2\|b\|_2^2 - |a^Tb|^2\right) \ge 0, \text{ where}$$

$$a = [\sqrt{e^{x_1}}\ \sqrt{e^{x_2}} \cdots \sqrt{e^{x_n}}]^T,\ b = [v_1\sqrt{e^{x_1}} \cdots v_n\sqrt{e^{x_n}}]^T.$$

$\therefore$ $\nabla^2 f(x) \succeq 0$. $\forall x \in \operatorname{dom} f$.

$\therefore$ f is convex (2^{nd}-order condition).

(b) $$\frac{df(x)}{dx} = \sum_{k=1}^{n}\frac{\partial h(g(x))}{\partial g_k(x)}\cdot\frac{dg_k(x)}{dx}$$

$$\frac{d^2 f_2(x)}{dx^2} = \sum_{k=1}^{n}\left(\sum_{i=1}^{n}\frac{\partial^2 h(g(x))}{\partial g_i(x)\partial g_k(x)}\cdot\frac{dg_i(x)}{dx}\cdot\frac{dg_k(x)}{dx} + \frac{d^2 g_k(x)}{dx^2}\cdot\frac{\partial h(g(x))}{\partial g_k(x)}\right)$$

$$= \left[\frac{dg_1(x)}{dx} \cdots \frac{dg_n(x)}{dx}\right](\nabla^2 h)\circ g(x)\begin{bmatrix} \frac{dg_1(x)}{dx} \\ \vdots \\ \frac{dg_n(x)}{dx} \end{bmatrix} + \frac{d^2 g_k(x)}{dx^2}\cdot\frac{\partial h(g(x))}{\partial g_k(x)} \ge 0.$$

This is because

$$\begin{cases} h \text{ is convex} \to \forall v.\ v^T\nabla^2 h v \ge 0 \\ \text{each entry of } h \text{ is nondecreasing} \to \dfrac{\partial h(g(x))}{\partial g_k(x)} \ge 0 \\ g \text{ is convex} \to \dfrac{d^2 g_k(x)}{dx^2} \ge 0 \end{cases}$$

$\therefore$ f is convex. (2^{nd} order-condition)

2. Proof. $\because x_1 < x_2 < x_3$. $f(x)$ is convex.

$\therefore$ $x_2 = \theta x_1 + (1-\theta)x_3$, $\theta \in (0,1)$.

$f(x_2) \le \theta f(x_1) + (1-\theta)f(x_3)$

Then $\theta f(x_2) - \theta f(x_1) \le (1-\theta)f(x_3) - (1-\theta)f(x_2)$

$$\Rightarrow \frac{f(x_2)-f(x_1)}{1-\theta} \le \frac{f(x_3)-f(x_2)}{\theta}$$

$$\Rightarrow \frac{f(x_2)-f(x_1)}{(1-\theta)(x_3-x_1)} \le \frac{f(x_3)-f(x_2)}{\theta(x_3-x_1)}$$

$$\Rightarrow \frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2}$$, Q.E.D.

3(a) Proof.

"$\Rightarrow$": $\because$ f is quasiconvex

$\therefore$ $f(\lambda x + (1-\lambda)y) \le \max\{f(x), f(y)\}$. $\forall x, y$, $\lambda \in [0,1]$.

Then for the level set $\{x \mid f(x) \le a\}$, for

$\forall c, b \in \{x \mid f(x) \le a\}$. $\forall \theta \in (0,1)$. (Suppose not empty).

$f(\theta c + (1-\theta)b) \le \max\{f(c), f(b)\} \le a$

$\therefore$ $\theta c + (1-\theta)b \in \{x \mid f(x) \le a\} \Rightarrow$ Level set is convex. $\forall a$.

"$\Leftarrow$": $\because$ $\{x \mid f(x) \le a\}$ is convex for ~~any~~ every $a \in \mathbb{R}$.

$\therefore$ For $\forall c, b$, $\theta \in (0,1)$. $\theta c + (1-\theta)b \in \{x \mid f(x) \le \max\{f(c), f(b)\}\}$

$f(\theta c + (1-\theta)b) \le \max\{f(c), f(b)\} \Rightarrow f(x)$ is quasiconvex.

Q.E.D.

(b). Suppose f is a convex function. i.e. $\forall x, y \in \operatorname{dom} f$

$\forall \theta \in (0,1)$ $f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y) \le \max\{f(x), f(y)\}$.

$\therefore$ f is also quasiconvex.

However, the converse is not necessarily true.

For example. $f(x) = \sqrt{|x|}$ is quasiconvex but not convex.

4. Proof.

$\partial I_C(x) = \{g \mid g^T(y-x) \le I_C(y) - I_C(x)\}$. If $x \notin C$, $\partial I_C(x) = \emptyset$.

Then for $\forall g \in \partial I_C(x)$, (implies $x \in C$) and $y \in C$.

$0 = f(y) - f(x) \ge g^T(y-x)$ $\therefore$ $g^T x \ge g^T y$, $g \in N_C(x)$

$\therefore$ $\partial I_C(x) \subseteq N_C(x)$

Conversely. $\forall g \in N_C(x)$, $x \in C$. then $g^T x \ge g^T y$. $\forall y \in C$.

$\because$ $x, y \in C$. $\therefore$ $I_C(x) = I_C(y) = 0$. $\therefore$ $I_C(y) \ge I_C(x) + g^T(y-x)$

$\therefore$ $g \in \partial I_C(x)$. $\Rightarrow N_C(x) \subseteq \partial I_C(x)$.

$\because$ $\partial I_C(x) \subseteq N_C(x)$. $N_C(x) \subseteq \partial I_C(x)$ $(x \in C)$.

$\therefore$ $\partial I_C(x) = N_C(x)$. $\forall x \in C$.