Thuật toán tối ưu
Bài tập lần 1

Đoàn Trần Nguyên Tùng
MSSV: 1411352

Ngày 23 tháng 4 năm 2018

Bài 1: Cho $\Omega \subseteq \mathbb{R}^n$ là môt tập lồi, với $k \in \mathbb{N}^*,\ x_1, x_2, \ldots, x_k \in \Omega$ và $\lambda_1, \lambda_2, \ldots, \lambda_k$ và $\sum_{i=1}^k \lambda_i = 1$. Chứng minh rằng $\sum_{i=1}^k \lambda_i x_i \in \Omega$.

Giải:

• Trường hợp $k = 2$: Ta có $x_1, x_2 \in \Omega$ và $\bar{\lambda}_1, \bar{\lambda}_2 \geq 0$ với $\bar{\lambda}_1 + \bar{\lambda}_2 = 1$. Khi đó, $\bar{\lambda}_1 \in [0, 1]$ và $\bar{\lambda}_2 = 1 - \bar{\lambda}_1$. Mà do Ω là tập lồi nên

$$\bar{\lambda}_1 x_1 + \bar{\lambda}_2 x_2 = \bar{\lambda}_1 x_1 + (1 - \bar{\lambda}_1) x_2 \in \Omega \tag{1}$$

• Trường hợp $k > 2$: Giả sử với $\hat{\lambda}_1, \ldots, \hat{\lambda}_{k-1} \geq 0$ thỏa $\sum_{i=1}^{k-1} \hat{\lambda}_i = 1$ ta có

$$\sum_{i=1}^{k-1} \hat{\lambda}_i x_i \in \Omega \tag{2}$$

Lấy $\lambda_1, \ldots, \lambda_k \geq 0$ thỏa $\sum_{i=1}^k \lambda_i = 1$, ta xét

$$\sum_{i=1}^{k} \lambda_i x_i = \sum_{i=1}^{k-1} \lambda_i x_i + \lambda_k x_k \tag{3}$$

Để ý rằng nếu $\lambda_k = 1$ thì ta phải có $\lambda_i = 0$ với mọi $i = 1, \ldots, k-1$ để điều kiện $\sum_{i=1}^k \lambda_i = 1$ với $\lambda_1, \ldots, \lambda_k \geq 0$ được thỏa, khi đó (3) có thể được viết lại là

$$\sum_{i=1}^{k} \lambda_i x_i = x_k \in \Omega \tag{4}$$

Với $\lambda_k < 1$ và $\sum_{i=1}^k \lambda_i = 1$, ta có

$$\sum_{i=1}^{k-1} \lambda_i = 1 - \lambda_k \Leftrightarrow \frac{\sum_{i=1}^{k-1} \lambda_i}{1 - \lambda_k} = 1 \tag{5}$$

Sử dụng giả thiết quy nạp ở (2), ta có

$$y = \frac{\sum_{i=1}^{k-1} \lambda_i}{1 - \lambda_k} x_i \in \Omega \tag{6}$$

Lấp luận tương tự như ở trường hợp $k = 2$, ta có $\lambda_k \in [0, 1]$ và $y, x_k \in \Omega$. Mà Ω lồi nên

$$\begin{align}
\sum_{i=1}^{k} \lambda_i x_i &= \sum_{i=1}^{k-1} \lambda_i x_i + \lambda_k x_k \tag{7}\\
&= (1 - \lambda_k) \frac{\sum_{i=1}^{k-1} \lambda_i}{1 - \lambda_i} x_i + \lambda_k x_k \tag{8}\\
&= (1 - \lambda_k) y + \lambda_k x_k \tag{9}\\
&= \lambda_k x_k + (1 - \lambda_k) y \in \Omega \tag{10}
\end{align}$$

Như vậy, bằng quy nạp, ta có điều cần phải chứng minh.

Bài 2: Dùng các đặc trưng của hàm lồi, để kiểm tra xem trong các hàm sau đây, hàm số nào là hàm lồi ?

(a) $f(x) = \mathrm{e}^{\alpha x} - x$ trên miền $\mathbb{R}$

(b) $f(x) = x^q$, với $q > 1$, trên miền $\mathbb{R}_+$

(c) $f(x) = -\ln(x)$ trên miền $\mathbb{R}_+$

(d) $f(x) = x\ln(x)$ trên miền $\mathbb{R}_+$

(e) $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2 + x_1 - 2x_2$ trên miền $\mathbb{R}^2$

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

(g) $f(x_1, x_2) = \dfrac{x_1^2}{x_2}$ trên miền $\mathbb{N} \times \mathbb{R}_{++}$

(h) $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, với $0 \le \alpha \le 1$, trên miền $\mathbb{R}^2_{++}$

(i) $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3 - x_3x_1$ trên miền $\mathbb{R}^n$

Giải:
(a) $f(x) = \mathrm{e}^{\alpha x} - x$ trên miền $\mathbb{R}$
Ta có

$$\nabla f(x) = \alpha\,\mathrm{e}^{\alpha x} - 1 \tag{11}$$

và

$$\nabla^2 f(x) = \alpha^2\,\mathrm{e}^{\alpha x} \tag{12}$$

Ta thấy $\nabla^2 f(x) = \alpha^2\,\mathrm{e}^{\alpha x} \ge 0$ trên miền $\mathbb{R}$ nên f là hàm lồi.

(b) $f(x) = x^q$, với $q > 1$, trên miền $\mathbb{R}_+$
Ta có

$$\nabla f(x) = qx^{q-1} \tag{13}$$

và

$$\nabla^2 f(x) = q(q-1)x^{q-2} \tag{14}$$

Ta thấy $\nabla^2 f(x) = q(q-1)x^{q-2} \ge 0$ trên miền $\mathbb{R}_+$ nên f là hàm lồi.

(c) $f(x) = -\ln(x)$ trên miền $\mathbb{R}_+$
Ta có

$$\nabla f(x) = -\frac{1}{x} \tag{15}$$

và

$$\nabla^2 f(x) = \frac{1}{x^2} \tag{16}$$

Ta thấy $\nabla^2 f(x) = \dfrac{1}{x^2} \ge 0$ trên miền $\mathbb{R}_+$ nên f là hàm lồi.

(d) $f(x) = x\ln(x)$ trên miền $\mathbb{R}_+$
Ta có

$$\begin{align} \nabla f(x) &= x\frac{1}{x} + \ln(x) \tag{17} \\ &= 1 + \ln(x) \tag{18} \end{align}$$

và

$$\nabla^2 f(x) = \frac{1}{x} \tag{19}$$

Ta thấy $\nabla^2 f(x) = \dfrac{1}{x} \geq 0$ trên miền $\mathbb{R}_+$ nên f là hàm lồi.

(e) $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2 + x_1 - 2x_2$ trên miền $\mathbb{R}^2$
Ta có

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 - x_2 + 1 \\ 2x_2 - x_1 - 2 \end{bmatrix} \tag{20}$$

và

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \tag{21}$$

Các trị riêng của $\nabla^2 f(x_1, x_2)$ là

$$\begin{cases} \lambda_1 & = 1 \\ \lambda_2 & = 3 \end{cases} \tag{22}$$

Ta thấy các trị riêng này đều dương ngặt trên miền $\mathbb{R}^2$, do đó $\nabla^2 f(x_1, x_2)$ là ma trận xác định dương và f là hàm lồi.

(f) $f(x_1, x_2) = x_1x_2$ trên miền $\mathbb{R}^2_{++}$
Ta có

$$\nabla f(x_1, x_2) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \tag{23}$$

và

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{24}$$

Các trị riêng của $\nabla^2 f(x_1, x_2)$ là

$$\begin{cases} \lambda_1 & = -1 \\ \lambda_2 & = 1 \end{cases} \tag{25}$$

Ta thấy các trị riêng này ngược dấu, do đó f không là hàm lồi.

(g) $f(x_1, x_2) = \dfrac{x_1^2}{x_2}$ trên miền $\mathbb{R} \times \mathbb{R}_{++}$
Ta có

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2\frac{x_1}{x_2} \\ -\frac{x_1^2}{x_2^2} \end{bmatrix} \tag{26}$$

và

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_2} & -2\frac{x_1}{x_2^2} \\ -2\frac{x_1}{x_2^2} & 2\frac{x_1^2}{x_2^3} \end{bmatrix} \tag{27}$$

Các trị riêng của $\nabla^2 f(x_1, x_2)$ là

$$\begin{cases} \lambda_1 & = 0 \\ \lambda_2 & = 2\dfrac{x_1^2 + x_2^2}{x_2^3} \end{cases} \tag{28}$$

Ta thấy các trị riêng này đều dương trên miền $\mathbb{R} \times \mathbb{R}_{++}$, do đó $\nabla^2 f(x_1, x_2)$ là ma trận nửa xác định dương và f là hàm lồi.

(h) $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, với $0 \le \alpha \le 1$, trên miền $\mathbb{R}^2_{++}$.
Ta có

$$\nabla f(x_1, x_2) = \begin{bmatrix} \alpha x_1^{(\alpha-1)} x_2^{(1-\alpha)} \\ (\alpha - 1)\dfrac{x_1^{\alpha}}{x_2^{\alpha}} \end{bmatrix} \tag{29}$$

và

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \alpha(\alpha - 1)x_1^{(\alpha-2)} x_2^{(1-\alpha)} & -\alpha(\alpha - 1)\dfrac{x_1^{(\alpha-1)}}{x_2^{\alpha}} \\ -\alpha(\alpha - 1)\dfrac{x_1^{(\alpha-1)}}{x_2^{\alpha}} & \alpha(\alpha - 1)\dfrac{x_1^{\alpha}}{x_2^{(\alpha+1)}} \end{bmatrix} \tag{30}$$

Các trị riêng của $\nabla^2 f(x_1, x_2)$ là

$$\begin{cases} \lambda_1 & = 0 \\ \lambda_2 & = a(a - 1)\dfrac{x_1^{(\alpha-2)}(x_1^2 + x_2^2)}{x_2^{(\alpha+1)}} \end{cases} \tag{31}$$

Ta thấy với $0 < \alpha < 1$ thì $\lambda_2 < 0$ trên miền $\mathbb{R}^2_{++}$, do đó f không là hàm lồi.

(i) $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3 - x_3x_1$ trên miền $\mathbb{R}^n$
Ta có

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} 2x_1 - x_2 - x_3 \\ 2x_2 - x_1 - x_3 \\ 2x_3 - x_2 - x_1 \end{bmatrix} \tag{32}$$

và

$$\nabla^2 f(x_1, x_2, x_3) = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \tag{33}$$

Các trị riêng của $\nabla^2 f(x_1, x_2, x_3)$ là

$$\begin{cases} \lambda_1 & = 0 \\ \lambda_2 & = 3 \\ \lambda_3 & = 3 \end{cases} \tag{34}$$

Ta thấy các trị riêng này đều dương, do đó $\nabla^2 f(x_1, x_2)$ là ma trận nửa xác định dương và f là hàm lồi.

Bài 3: Cho hai ánh xạ $f_1, f_2 : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ là các hà lồi. Đặt

$$g(x) = \mathrm{Min}\{f_1(x), f_2(x)\} \tag{35}$$

và

$$h(x) = \mathrm{Max}\{f_1(x), f_2(x)\} \tag{36}$$

Hỏi g và h hàm nào là hàm lồi ? Vì sao ? (Nếu là hàm lồi hãy chứng minh, ngược lại hãy cho phản thí dụ)
Giải:
• Lấy $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ thỏa $f_1(x) = x^2$ và $f_2(x) = (x-2)^2$ là hai hàm lồi. Cho $x, y \in \mathbb{R}$ và $t \in [0,1]$, nếu g là hàm lồi, ta phải có

$$g(tx + (1-t)y) \le tg(x) + (1-t)g(y) \tag{37}$$

hay

$$\mathrm{Min}\{(tx+(1-t)y)^2, (tx+(1-t)y-2)^2\} \le t\,\mathrm{Min}\{x^2, (x-2)^2\} + (1-t)\,\mathrm{Min}\{y^2, (y-2)^2\} \tag{38}$$

Chọn $x = 0$, $y = 2$ và $t = \dfrac{1}{2}$, (38) có thể được viết lại là

$$\mathrm{Min}\left\{\left(\frac{1}{2}0 + \left(1-\frac{1}{2}\right)2\right)^2, \left(\frac{1}{2}0 + \left(1-\frac{1}{2}\right)2-2\right)^2\right\} \le \frac{1}{2}\,\mathrm{Min}\left\{0^2, (0-2)^2\right\} + \left(1-\frac{1}{2}\right)\mathrm{Min}\{2^2, (2-2)^2\} \tag{39}$$

$$\Leftrightarrow \mathrm{Min}\{1,1\} \le \frac{1}{2}\,\mathrm{Min}\{0,4\} + \frac{1}{2}\,\mathrm{Min}\{4,0\} \tag{40}$$

$$\Leftrightarrow 1 \le \frac{1}{2}0 + \frac{1}{2}0 \tag{41}$$

$$\Leftrightarrow 1 \le 0 \tag{42}$$

Ta thấy (42) vô lý, do đó, g không phải là hàm lồi.

• Xét $x, y \in \mathrm{dom}\, h \subset \mathbb{R}^n$ và $t \in [0,1]$, do f_1, f_2 là hàm lồi, ta có

$$\begin{cases} f_1(tx+(1-t)y) & \le tf_1(x) + (1-t)f_1(y) \\ f_2(tx+(1-t)y) & \le tf_2(x) + (1-t)f_2(y) \end{cases} \tag{43}$$

$$\Rightarrow \begin{cases} f_1(tx+(1-t)y) & \le t\,\mathrm{Max}\{f_1(x), f_2(x)\} + (1-t)\,\mathrm{Max}\{f_1(y), f_2(y)\} \\ f_2(tx+(1-t)y) & \le t\,\mathrm{Max}\{f_1(x), f_2(x)\} + (1-t)\,\mathrm{Max}\{f_1(y), f_2(y)\} \end{cases} \tag{44}$$

$$\Rightarrow \mathrm{Max}\{f_1(tx+(1-t)y), f_2(tx+(1-t)y)\} \le t\,\mathrm{Max}\{f_1(x), f_2(x)\} + (1-t)\,\mathrm{Max}\{f_1(y), f_2(y)\} \tag{45}$$

Tức là

$$h(tx + (1-t)y) \le th(x) + (1-t)h(y) \tag{46}$$

Từ đó suy ra h là hàm lồi.

Bài 4: Cho ánh xạ $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ và α là một số thực bất kì, tập mức α được định nghĩa như sau

$$L_\alpha = \{x \in \mathbb{R}^n | f(x) \le \alpha\} \tag{47}$$

(a) Chứng minh rằng nếu f là hàm lồi thì với mọi $\alpha \in \mathbb{R}$, tập mức L_α là tập lồi

(b) Mệnh đề đảo của câu (a) là đúng hay sai ? Vì sao ?

Giải:
(a) Lấy $x, y \in L_\alpha$, ta có

$$\begin{cases} f(x) & \le \alpha \\ f(y) & \le \alpha \end{cases} \tag{48}$$

Do f là hàm lồi nên với $t \in [0, 1]$, ta có

$$\begin{aligned} f(tx + (1-t)y) &\le tf(x) + (1-t)f(y) && (49) \\ &\le t\alpha + (1-t)\alpha && (50) \\ &= \alpha && (51) \end{aligned}$$

Do $f(tx + (1-t)y) \le \alpha$ nên ta suy ra được

$$tx + (1-t)y \in L_\alpha \tag{52}$$

Như vậy, L_α là tập lồi.

(b) Giả sử L_α là tập lồi với mọi $\alpha \in \mathbb{R}$, ta cần kiểm tra xem f có phải là hàm lồi hay không.
Xét $f : \mathbb{R} \to \mathbb{R}_+$ với $f(x) = \sqrt{|x|}$. Với $\alpha \in \mathbb{R}_+$, ta có

$$\begin{aligned} L_\alpha &= \{x \in \mathbb{R} | \sqrt{|x|} \le \alpha\} && (53) \\ &= \{x \in \mathbb{R} | |x| \le \alpha^2\} && (54) \\ &= \{x \in \mathbb{R} | -\alpha^2 \le x \le \alpha^2\} && (55) \\ &= \left[-\alpha^2, \alpha^2\right] && (56) \end{aligned}$$

Ta thấy L_α là tập lồi.
Lấy $\bar{x}, \bar{y} \in \mathbb{R}$ và $t \in [0, 1]$. Nếu f là hàm lồi thì ta phải có

$$f(t\bar{x} + (1-t)\bar{y}) \le tf(\bar{x}) + (1-t)f(\bar{y}) \tag{57}$$

Tức là với f đã chọn, ta phải có

$$\sqrt{|t\bar{x} + (1-t)\bar{y}|} \le t\sqrt{|\bar{x}|} + (1-t)\sqrt{|\bar{y}|} \tag{58}$$

Chọn $\bar{x} = 4$, $\bar{y} = 16$ và $t = \dfrac{7}{12}$, ta có

$$t\bar{x} + (1-t)\bar{y} = \frac{7}{12}4 + \left(1 - \frac{7}{12}\right)16 = 9 \tag{59}$$

và

$$\begin{cases} f(\bar{x}) & = \sqrt{|4|} = 2 \\ f(\bar{y}) & = \sqrt{|16|} = 4 \\ f(t\bar{x} + (1-t)\bar{y}) & = \sqrt{|9|} = 3 \\ tf(\bar{x}) + (1-t)f(\bar{y}) & = \dfrac{7}{12}2 + \dfrac{5}{12}4 = \dfrac{17}{6} \end{cases} \tag{60}$$

Ta thấy

$$f(t\bar{x} + (1-t)\bar{y}) = 3 = \frac{18}{6} > \frac{17}{6} = tf(\bar{x}) + (1-t)f(\bar{y}) \tag{61}$$

Điều này mâu thuẫn với (57). Do đó f như đã chọn không là hàm lồi, mặc dù tập mức L_α lồi với mọi $\alpha \in \mathbb{R}_+$.

Bài 5: (Jensen's inequality) Cho ánh xạ $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ là ánh xạ lồi, với $k \in \mathbb{N}^*, x_1, x_2, \ldots, x_k \in$ dom f và $\lambda_1, \lambda_2, \ldots, \lambda_k \geq 0$ và $\sum_{i=1}^k \lambda_i = 1$, chứng minh rằng

$$f\left(\sum_{i=1}^k \lambda_i x_i\right) \leq \sum_{i=1}^k \lambda_i f(x_i) \tag{62}$$

Áp dụng: bằng cách vận dụng tính lồi của hàm số $f(x) = -\ln(x)$.

(a) (Cauchy's inequality) Cho $a_1, a_2, \ldots, a_m \in \mathbb{R}_+$, ta có

$$\frac{a_1 + a_2 + \cdots + a_m}{m} \geq \sqrt[m]{a_1 a_2 \ldots a_m} \tag{63}$$

(b) (Holder's inequality) Cho $x, y \in \mathbb{R}^n$, $p > 1$ và $\dfrac{1}{p} + \dfrac{1}{q} = 1$, ta có

$$\sum_{i=1}^n x_i y_i \leq \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^p\right)^{\frac{1}{q}} \tag{64}$$

Giải:

• Trường hợp $k = 2$: Ta có $\lambda_1 \in [0, 1]$ và $\lambda_2 = 1 - \lambda_1$, do f là hàm lồi nên ta có

$$\begin{aligned} f(\lambda_1 x_1 + \lambda_2 x_2) &= f(\lambda_1 x_1 + (1 - \lambda_1) x_2) && (65) \\ &\leq \lambda_1 f(x_1) + (1 - \lambda_1) f(x_2) && (66) \\ &= \lambda_1 f(x_1) + \lambda_2 f(x_2) && (67) \end{aligned}$$

• Trường hợp $k > 2$: Giả sử với $\hat{\lambda}_1, \ldots, \hat{\lambda}_{k-1} \geq 0$ thỏa $\sum_{i=1}^{k-1} \hat{\lambda}_i = 1$ ta có

$$f\left(\sum_{i=1}^{k-1} \hat{\lambda}_i x_i\right) \leq \sum_{i=1}^{k-1} \hat{\lambda}_i f(x_i) \tag{68}$$

Xét $\lambda_1, \ldots, \lambda_k \geq 0$ thỏa $\sum_{i=1}^k \lambda_i = 1$, ta có

$$\begin{aligned} &\sum_{i=1}^k \lambda_i = 1 && (69) \\ \Leftrightarrow &\sum_{i=1}^{k-1} \lambda_i = 1 - \lambda_k && (70) \end{aligned}$$

Giả sử $\lambda_k = 1$, khi đó ta phải có $\lambda_1, \ldots, \lambda_{k-1} = 0$. Lúc này ta có

$$f\left(\sum_{i=1}^k \lambda_i x_i\right) = f(x_k) \leq \sum_{i=1}^k \lambda_i f(x_i) = f(x_k) \tag{71}$$

Giả sử $\lambda_k < 1$, (70) có thể viết lại thành

$$\frac{\sum_{i=1}^{k-1} \lambda_i}{1 - \lambda_k} = 1 \tag{72}$$

Khi đó ta có thể viết

$$\sum_{i=1}^k \lambda_i x_i = (1 - \lambda_k) \frac{\sum_{i=1}^{k-1} \lambda_i x_i}{1 - \lambda_k} + \lambda_k x_k \tag{73}$$

Sử dụng kết quả **Bài 1** ta có

$$\frac{\sum_{i=1}^{k-1} \lambda_i x_i}{1-\lambda_k} \in \operatorname{dom} f \tag{74}$$

Do f là hàm lồi, ta có

$$\begin{align}
f\left(\sum_{i=1}^{k} \lambda_i x_i\right) &= f\left((1-\lambda_k)\frac{\sum_{i=1}^{k-1} \lambda_i x_i}{1-\lambda_k} + \lambda_k x_k\right) \tag{75}\\
&\le (1-\lambda_k) f\left(\frac{\sum_{i=1}^{k-1} \lambda_i x_i}{1-\lambda_k}\right) + \lambda_k f(x_k) \tag{76}\\
&\le (1-\lambda_k)\frac{\sum_{i=1}^{k-1} \lambda_i f(x_i)}{1-\lambda_k} + \lambda_k f(x_k) \tag{77}\\
&= \sum_{i=1}^{k} \lambda_i f(x_i) \tag{78}
\end{align}$$

Trong đó, dòng (77) có được từ (76) do giải thiết quy nạp (68) và (72).
Như vậy, bằng quy nạp ta có điều cần phải chứng minh.

(a) Ta có hàm số $f : \mathbb{R}_+ \to \mathbb{R}$ với $f(x) = -\ln(x)$ là hàm lồi theo kết quả câu (c) **Bài 2**.
Với $a_1, a_2, \ldots, a_m \in \mathbb{R}_+$ và các hệ số $\lambda_i = \dfrac{1}{m}$, áp dụng bất đẳng thức Jensen như đề cho, ta có

$$\begin{align}
-\ln\left(\sum_{i=1}^{m} \frac{1}{m} a_i\right) &\le -\sum_{i=1}^{m} \frac{1}{m} \ln(a_i) \tag{79}\\
\Leftrightarrow \ln\left(\sum_{i=1}^{m} \frac{1}{m} a_i\right) &\ge \sum_{i=1}^{m} \frac{1}{m} \ln(a_i) \tag{80}
\end{align}$$

Do hàm số mũ e^x đồng biến trên $\mathbb{R}_+$ nên ta có

$$\begin{align}
\mathrm{e}^{\ln\left(\sum_{i=1}^{m} \frac{1}{m} a_i\right)} &\ge \mathrm{e}^{\sum_{i=1}^{m} \frac{1}{m} \ln(a_i)} \tag{81}\\
\Rightarrow \sum_{i=1}^{m} \frac{a_i}{m} &\ge \mathrm{e}^{\frac{\ln(a_1)}{m}} \mathrm{e}^{\frac{\ln(a_2)}{m}} \ldots \mathrm{e}^{\frac{\ln(a_1 m)}{m}} \tag{82}\\
\Rightarrow \sum_{i=1}^{m} \frac{a_i}{m} &\ge a_1^{\frac{1}{m}} a_2^{\frac{1}{m}} \ldots a_m^{\frac{1}{m}} \tag{83}\\
\Rightarrow \sum_{i=1}^{m} \frac{a_i}{m} &\ge \sqrt[m]{a_1 a_2 \ldots a_m} \tag{84}
\end{align}$$

Vậy, ta có điều cần chứng minh.

(b) Ta có hàm số $f : \mathbb{R}_+ \to \mathbb{R}$ với $f(x) = x^q$ với $q > 1$ là hàm lồi theo kết quả câu (b) **Bài 2**.
Với $a_1, a_2, \ldots, a_n \in \mathbb{R}_+$ cùng các hệ số $\lambda_i > 0$ thỏa $\sum_{i=1}^{n} \lambda_i = 1$, áp dụng bất đẳng thức Jensen như đề cho, ta có

$$\left(\sum_{i=1}^{n} \lambda_i a_i\right)^q \le \sum_{i=1}^{n} \lambda_i a_i^q \tag{85}$$

- <u>Trường hợp không tồn tại h sao cho $x_h = 0$</u>

Giả sử $\lambda_i > 0$ với mọi $i = 1, \ldots, n$, đặt $a_i = \dfrac{|x_i||y_i|}{\lambda_i}$, (85) trở thành

$$\left(\sum_{i=1}^{n} \lambda_i \frac{|x_i||y_i|}{\lambda_i}\right)^q \leq \sum_{i=1}^{n} \lambda_i \left(\frac{|x_i||y_i|}{\lambda_i}\right)^q \tag{86}$$

$$\Leftrightarrow \left(\sum_{i=1}^{n} |x_i||y_i|\right)^q \leq \sum_{i=1}^{n} \lambda_i^{1-q} |x_i|^q |y_i|^q \tag{87}$$

Đặt $\lambda_i = \dfrac{|x_i|^p}{\sum_{k=1}^{n} |x_k|^p}$ và do $x_i \neq 0$ với mọi $i = 1, \ldots, n$ nên $\lambda_i > 0$. (87) trở thành

$$\left(\sum_{i=1}^{n} |x_i||y_i|\right)^q \leq \sum_{i=1}^{n} \left(\frac{|x_i|^p}{\sum_{k=1}^{n} |x_k|^p}\right)^{1-q} |x_i|^q |y_i|^q \tag{88}$$

$$\Rightarrow \left(\sum_{i=1}^{n} |x_i||y_i|\right)^q \leq \frac{1}{\left(\sum_{k=1}^{n} |x_k|^p\right)^{1-q}} \sum_{i=1}^{n} \left(|x_i|^p\right)^{1-q} |x_i|^q |y_i|^q \tag{89}$$

$$\Rightarrow \left(\sum_{i=1}^{n} |x_i||y_i|\right)^q \leq \frac{1}{\left(\sum_{k=1}^{n} |x_k|^p\right)^{1-q}} \sum_{i=1}^{n} |x_i|^{p-pq} |x_i|^q |y_i|^q \tag{90}$$

$$\Rightarrow \left(\sum_{i=1}^{n} |x_i||y_i|\right)^q \leq \frac{1}{\left(\sum_{k=1}^{n} |x_k|^p\right)^{1-q}} \sum_{i=1}^{n} |x_i|^{p+q-pq} |y_i|^q \tag{91}$$

Do $\dfrac{1}{p} + \dfrac{1}{q} = 1$ nên $p + q = pq$, khi đó (91) trở thành

$$\left(\sum_{i=1}^{n} |x_i||y_i|\right)^q \leq \frac{1}{\left(\sum_{k=1}^{n} |x_k|^p\right)^{1-q}} \sum_{i=1}^{n} |y_i|^q \tag{92}$$

$$\Rightarrow \left(\sum_{i=1}^{n} |x_i||y_i|\right)^q \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{q-1} \sum_{i=1}^{n} |y_i|^q \tag{93}$$

$$\Rightarrow \sum_{i=1}^{n} |x_i||y_i| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{q-1}{q}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}} \tag{94}$$

$$\Rightarrow \sum_{i=1}^{n} |x_i||y_i| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{1-\frac{1}{q}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}} \tag{95}$$

$$\Rightarrow \sum_{i=1}^{n} |x_i||y_i| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}} \tag{96}$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}} \tag{97}$$

• Trường hợp tồn tại h sao cho $x_h = 0$

Lúc này, ta có thể xắp xếp lại thứ tự của $x_1, \ldots, x_n$ sao cho $x_1, \ldots, x_m \neq 0$ và $x_{m+1}, \ldots, x_n = 0$.

Khi đó ta có

$$
\begin{aligned}
\sum_{i=1}^{n} x_i y_i &= \sum_{i=1}^{m} x_i y_i + \sum_{j=m+1}^{n} x_j y_j && (98)\\
&= \sum_{i=1}^{m} x_i y_i && (99)\\
&\leq \left(\sum_{k=1}^{m} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{m} |y_i|^q\right)^{\frac{1}{q}} && (100)\\
&\leq \left(\sum_{k=1}^{m} |x_k|^p + \sum_{j=m+1}^{n} |x_h|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{m} |y_i|^q + \sum_{j=m+1}^{n} |y_j|^q\right)^{\frac{1}{q}} && (101)\\
&= \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}} && (102)
\end{aligned}
$$

Trong đó dòng (100) có được từ việc áp dụng (97).
Như vậy, ta có điều phải chứng minh.

Bài 6: Sử dụng điều kiện cần cấp 1 để tìm các điểm tới hạn của các hàm số sau.

(a) $f(x_1, x_2) = x_1^2 + 3x_2^2 - 4x_1 + 8x_2$

(b) $f(x_1, x_2, x_3) = 2x_1^2 + x_1x_2 + x_2^2 + x_2x_3 + x_3^2 - 6x_1 - 7x_2 - 8x_3 + 9$

(c) $f(x_1, x_2) = (x_1x_2 - x_1 - 1)^2 + (x_2^2 - 1)^2$

(d) $f(x_1, x_2, x_3) = x_1x_2x_3\, \mathrm{e}^{-x_1-x_2-x_3}$

(e) $f(x_1, x_2) = \dfrac{1}{x_1x_2}x_1 + x_2$ trên miền $\mathbb{R}^2_{++}$

Giải:
(a) $f(x_1, x_2) = x_1^2 + 3x_2^2 - 4x_1 + 8x_2$
Ta có

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 - 4 \\ 6x_2 + 8 \end{bmatrix} \tag{103}$$

Điểm $(\bar{x}_1, \bar{x}_2)$ là điểm tới hạn của f khi

$$\nabla f(\bar{x}_1, \bar{x}_2) = 0 \tag{104}$$

Tức là

$$\begin{cases} 2\bar{x}_1 - 4 &= 0 \\ 6\bar{x}_2 + 8 &= 0 \end{cases} \tag{105}$$

Giải hệ tuyến tính này ta được

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ \frac{4}{3} \end{bmatrix} \tag{106}$$

(b) $f(x_1, x_2, x_3) = 2x_1^2 + x_1x_2 + x_2^2 + x_2x_3 + x_3^2 - 6x_1 - 7x_2 - 8x_3 + 9$
Ta có

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} 4x_1 + x_2 - 6 \\ x_1 + 2x_2 + x_3 - 7 \\ x_2 + 2x_3 - 8 \end{bmatrix} \tag{107}$$

Điểm $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ là điểm tới hạn của f khi

$$\nabla f(\bar{x}_1, \bar{x}_2, \bar{x}_3) = 0 \tag{108}$$

Tức là

$$\begin{cases} 4\bar{x}_1 + \bar{x}_2 - 6 &= 0 \\ \bar{x}_1 + 2\bar{x}_2 + \bar{x}_3 - 7 &= 0 \\ \bar{x}_2 + 2\bar{x}_3 - 8 &= 0 \end{cases} \tag{109}$$

Giải hệ tuyến tính này ta được

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -26 \\ 17 \end{bmatrix} \tag{110}$$

(c) $f(x_1, x_2) = (x_1x_2 - x_1 - 1)^2 + (x_2^2 - 1)^2$
Ta có

$$\nabla f(x_1, x_2) = \begin{bmatrix} -2(x_2 - 1)(x_1 - x_1x_2 + 1) \\ 4x_2(x_2^2 - 1) - 2x_1(x_1 - x_1x_2 + 1) \end{bmatrix} \tag{111}$$

Điểm $(\bar{x}_1, \bar{x}_2)$ là điểm tới hạn của f khi

$$\nabla f(\bar{x}_1, \bar{x}_2) = 0 \tag{112}$$

Tức là

$$\begin{cases} -2(\bar{x}_2 - 1)(\bar{x}_1 - \bar{x}_1\bar{x}_2 + 1) & = 0 \\ 4\bar{x}_2(\bar{x}_2^2 - 1) - 2\bar{x}_1(\bar{x}_1 - \bar{x}_1\bar{x}_2 + 1) & = 0 \end{cases} \tag{113}$$

Từ dòng 1 của (113), ta có

$$\left[\begin{aligned} &\bar{x}_2 - 1 && = 0 \\ &\bar{x}_1 - \bar{x}_1\bar{x}_2 + 1 && = 0 \end{aligned}\right. \tag{114}$$

• <u>Trường hợp 1:</u> $\bar{x}_2 - 1 = 0 \Leftrightarrow \bar{x}_2 = 1$, dòng 2 của (113) trở thành

$$\begin{aligned} \bar{x}_1(\bar{x}_1 - \bar{x}_1\bar{x}_2 + 1) &= \bar{x}_1(\bar{x}_1 - \bar{x}_1 + 1) && (115) \\ &= \bar{x}_1 && (116) \\ &= 0 && (117) \end{aligned}$$

Vậy

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{118}$$

là một điểm tới hạn của f.

• <u>Trường hợp 2</u> $\bar{x}_1 - \bar{x}_1\bar{x}_2 + 1 = 0 \Leftrightarrow \bar{x}_1 = \dfrac{1}{\bar{x}_2 - 1}$ và $\bar{x}_2 \neq 1$, dòng 2 của (113) trở thành

$$\bar{x}_2(\bar{x}_2^2 - 1) = 0 \tag{119}$$

Khi đó ta có

$$\left[\begin{aligned} &\bar{x}_2 && = 0 \text{ và } \bar{x}_1 = -1 \\ &\bar{x}_2 && = -1 \text{ và } \bar{x}_1 = -\frac{1}{2} \end{aligned}\right. \tag{120}$$

Vậy

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \tag{121}$$

và

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -\dfrac{1}{2} \end{bmatrix} \tag{122}$$

cũng là các điểm tới hạn của f.

(d) $f(x_1, x_2, x_3) = x_1x_2x_3\, \mathrm{e}^{-x_1 - x_2 - x_3}$
Ta có

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} (x_2x_3 - x_1x_2x_3)\, \mathrm{e}^{-x_1-x_2-x_3} \\ (x_1x_3 - x_1x_2x_3)\, \mathrm{e}^{-x_1-x_2-x_3} \\ (x_1x_2 - x_1x_2x_3)\, \mathrm{e}^{-x_1-x_2-x_3} \end{bmatrix} \tag{123}$$

Điểm $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ là điểm tới hạn của f khi

$$\nabla f(\bar{x}_1, \bar{x}_2, \bar{x}_3) = 0 \tag{124}$$

Tức là

$$\begin{cases} (\bar{x}_2\bar{x}_3 - \bar{x}_1\bar{x}_2\bar{x}_3)\,\mathrm{e}^{-\bar{x}_1-\bar{x}_2-\bar{x}_3} &= 0 \\ (\bar{x}_1\bar{x}_3 - \bar{x}_1\bar{x}_2\bar{x}_3)\,\mathrm{e}^{-\bar{x}_1-\bar{x}_2-\bar{x}_3} &= 0 \\ (\bar{x}_1\bar{x}_2 - \bar{x}_1\bar{x}_2\bar{x}_3)\,\mathrm{e}^{-\bar{x}_1-\bar{x}_2-\bar{x}_3} &= 0 \end{cases} \tag{125}$$

$$\Leftrightarrow \begin{cases} \bar{x}_2\bar{x}_3 - \bar{x}_1\bar{x}_2\bar{x}_3 &= 0 \\ \bar{x}_1\bar{x}_3 - \bar{x}_1\bar{x}_2\bar{x}_3 &= 0 \\ \bar{x}_1\bar{x}_2 - \bar{x}_1\bar{x}_2\bar{x}_3 &= 0 \end{cases} \tag{126}$$

$$\Leftrightarrow \begin{cases} \bar{x}_2\bar{x}_3(1-\bar{x}_1) &= 0 \\ \bar{x}_1\bar{x}_3(1-\bar{x}_2) &= 0 \\ \bar{x}_1\bar{x}_2(1-\bar{x}_3) &= 0 \end{cases} \tag{127}$$

Từ dòng 1 của (127), ta có

$$\left[\begin{aligned} \bar{x}_1 &= 1 \\ \bar{x}_2 &= 0 \\ \bar{x}_3 &= 0 \end{aligned}\right. \tag{128}$$

- <u>Trường hợp 1:</u> $\bar{x}_1 = 1$, (127) trở thành

$$\begin{cases} \bar{x}_3(1-\bar{x}_2) &= 0 \\ \bar{x}_2(1-\bar{x}_3) &= 0 \end{cases} \tag{129}$$

Từ dòng 1 của (129), ta có

$$\left[\begin{aligned} \bar{x}_2 &= 1 \\ \bar{x}_3 &= 0 \end{aligned}\right. \tag{130}$$

Thay $\bar{x}_2 = 1$ vào dòng 2 của (129), ta có

$$\bar{x}_3 = 1 \tag{131}$$

Thay $\bar{x}_3 = 0$ vào dòng 2 của (129), ta có

$$\bar{x}_2 = 0 \tag{132}$$

Vậy

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{133}$$

và

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{134}$$

là các điểm tới hạn của f.

- <u>Trường hợp 2:</u> $\bar{x}_2 = 0$, (127) trở thành

$$\bar{x}_1\bar{x}_3 = 0 \tag{135}$$

Từ (135), ta có

$$\left[\begin{aligned} \bar{x}_1 &= 0 \\ \bar{x}_3 &= 0 \end{aligned}\right. \tag{136}$$

Vậy

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \bar{x}_3 \end{bmatrix} \tag{137}$$

với hệ số tự do $\bar{x}_3$ và

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ 0 \\ 0 \end{bmatrix} \tag{138}$$

với hệ số tự do $\bar{x}_1$ là các điểm tới hạn của f.
• Trường hợp 3: $\bar{x}_3 = 0$, (127) trở thành

$$\bar{x}_1\bar{x}_2 = 0 \tag{139}$$

Từ (139), ta có

$$\left[\begin{aligned} \bar{x}_1 &= 0 \\ \bar{x}_2 &= 0 \end{aligned}\right. \tag{140}$$

Vậy

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{x}_2 \\ 0 \end{bmatrix} \tag{141}$$

với hệ số tự do $\bar{x}_2$ và

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ 0 \\ 0 \end{bmatrix} \tag{142}$$

với hệ số tự do $\bar{x}_1$ (trùng với một họ các điểm tới hạn tìm được ở Trường hợp 2) là các điểm tới hạn của f.

(e) $f(x_1, x_2) = \dfrac{1}{x_1x_2}x_1 + x_2$ trên miền $\mathbb{R}^2_{++}$
Ta có

$$\nabla f(x_1, x_2) = \begin{bmatrix} 1 - \dfrac{1}{x_1^2x_2} \\ 1 - \dfrac{1}{x_1x_2^2} \end{bmatrix} \tag{143}$$

Điểm $(\bar{x}_1, \bar{x}_2)$ là điểm tới hạn của f khi

$$\nabla f(\bar{x}_1, \bar{x}_2) = 0 \tag{144}$$

Tức là

$$\begin{cases} 1 - \dfrac{1}{\bar{x}_1^2\bar{x}_2} &= 0 \\ 1 - \dfrac{1}{\bar{x}_1\bar{x}_2^2} &= 0 \end{cases} \tag{145}$$

$$\Leftrightarrow \begin{cases} \bar{x}_1^2\bar{x}_2 &= 1 \\ \bar{x}_1\bar{x}_2^2 &= 1 \end{cases} \tag{146}$$

Từ đó ta suy ra được $\bar{x}_1, \bar{x}_2 \neq 0$ và

$$\bar{x}_1^2 \bar{x}_2 = \bar{x}_1 \bar{x}_2^2 \Leftrightarrow \bar{x}_1 \bar{x}_2 (\bar{x}_1 - \bar{x}_2) = 0 \tag{147}$$

Tức là ta có

$$\bar{x}_1 = \bar{x}_2 \tag{148}$$

Thế vào (146), ta có

$$\bar{x}_1^3 = 1 \tag{149}$$

Vậy, ta có thể giải ra được

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{150}$$

là một điểm tới hạn của f.

Bài 7: Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$ là hàm lồi. Xét bài toán cực tiểu không ràng buộc sau

$$(P): \ \ \mathrm{Min}_{x\in\mathbb{R}^n} f(x) \tag{151}$$

Chứng minh các tính chất sau

(a) $\bar{x}$ là cực tiểu địa phương của $(P) \Leftrightarrow \nabla f(\bar{x}) = 0$

(b) $\bar{x}$ là cực tiểu địa phương của $(P) \Leftrightarrow \bar{x}$ là cực tiểu toàn cục của (P)

(c) Tập các điểm cực tiểu của bài toán (P) là tập lồi.

(d) Nếu f là hàm lồi chặt thì bài toán (P) có duy nhất một cực tiểu.

Giải:
(a)
($\Rightarrow$) Cho $\bar{x}$ là một cực tiểu địa phương của f, khi đó có một tập mở $U_{\bar{x}}$ sao cho

$$f(\bar{x}) \leq f(x), \ \ \forall x \in U_{\bar{x}} \tag{152}$$

Theo định nghĩa, ta có

$$\lim_{h\to 0} \frac{f(\bar{x}+h) - f(\bar{x}) - \langle \nabla f(\bar{x}), h\rangle}{\|h\|} = 0 \tag{153}$$

Đặt $h = te$ với $0 \neq t \in \mathbb{R}$ và $0 \neq e \in \mathbb{R}^n$ là một vector khác 0 cố định, ta có

$$\lim_{t\to 0} \frac{f(\bar{x}+te) - f(\bar{x}) - t\langle \nabla f(\bar{x}), e\rangle}{\|h\|} = 0 \tag{154}$$

Với $t > 0$ đủ gần 0 sao cho $\bar{x} + te \in U_{\bar{x}}$, ta có

$$f(\bar{x}+te) - f(\bar{x}) \geq 0 \tag{155}$$

Khi đó, do ta có

$$\lim_{t\to 0^+} \frac{f(\bar{x}+te) - f(\bar{x}) - t\langle \nabla f(\bar{x}), e\rangle}{|t|\|e\|} = 0 \tag{156}$$

nên ta có

$$\begin{cases} t\langle \nabla f(\bar{x}), e\rangle \geq 0 \\ t > 0 \end{cases} \Rightarrow \langle \nabla f(\bar{x}), e\rangle \geq 0 \tag{157}$$

Bên cạnh đó, ta cũng có

$$\lim_{t\to 0^-} \frac{f(\bar{x}+te) - f(\bar{x}) - t\langle \nabla f(\bar{x}), e\rangle}{|t|\|e\|} = 0 \tag{158}$$

nên ta có

$$\begin{cases} t\langle \nabla f(\bar{x}), e\rangle \geq 0 \\ t < 0 \end{cases} \Rightarrow \langle \nabla f(\bar{x}), e\rangle \leq 0 \tag{159}$$

Từ (156) và (159) ta suy ra được

$$\langle \nabla f(\bar{x}), e\rangle = 0 \tag{160}$$

Mà do vector $e \neq 0$ nên ta phải có

$$\nabla f(\bar{x}) = 0 \tag{161}$$

tức là điều cần chứng minh.

($\Leftarrow$) và (b) Ta có $f : \mathbb{R}^n \to \mathbb{R}$ là hàm lồi và cho $\nabla f(\bar{x}) = 0$, ta cần chứng minh $\bar{x} \in \mathbb{R}^n$ là cực tiểu địa phương (toàn cục) của f.

$$\lim_{h \to 0} \frac{f(\bar{x}+h) - f(\bar{x}) - \langle \nabla f(\bar{x}), h \rangle}{\|h\|} = 0 \tag{162}$$

Cho $x, y \in \mathbb{R}^n$, $t \in [0, 1]$ và $z = tx + (1-t)y$. Do f lồi nên ta có

$$\begin{aligned} f(z) &= f(tx + (1-t)y) && (163) \\ &\leq tf(x) + (1-t)f(y) && (164) \end{aligned}$$

Từ đó ta có

$$\begin{aligned} f(z) - f(y) &= tf(x) + (1-t)f(y) - f(y) && (165) \\ &= tf(x) - tf(y) && (166) \end{aligned}$$

Mà theo lý thuyết về đạo hàm theo hướng, ta có

$$\begin{aligned} \nabla f(y)^\intercal (x-y) &= \lim_{t \to 0} \frac{f(x + t(x-y)) - f(x)}{t} && (167) \\ &= \lim_{t \to 0} \frac{f(z) - f(x)}{t} && (168) \end{aligned}$$

Sử dụng (166), ta có

$$\begin{aligned} \nabla f(y)^\intercal (x-y) &= \lim_{t \to 0} \frac{f(z) - f(x)}{t} && (169) \\ &\leq f(x) - f(y) && (170) \end{aligned}$$

Đặt $y = \bar{x}$, ta có

$$\nabla f(\bar{x})^\intercal (x - \bar{x}) \leq f(x) - f(\bar{x}) \tag{171}$$

Mà do $\nabla f(\bar{x}) = 0$ nên ta có

$$f(\bar{x}) \leq f(x) \tag{172}$$

Do $x \in \mathbb{R}^n$ bất kì nên $\bar{x}$ là cực tiểu toàn cục của f.

(c) Nếu f có một điểm cực tiểu (toàn cục) thì ta thấy tập các điểm cực tiểu của f là hàm lồi nên ta giả sử f có nhiều hơn một cực tiểu.
Lấy $x, y \in \mathbb{R}^n$ là hai cực tiểu của f. Đặt $\alpha = f(x) = f(y)$ là giá trị cực tiểu của f.
Do f là hàm lồi nên với $t \in [0, 1]$, ta có

$$\begin{aligned} f(tx + (1-t)y) &\leq tf(x) + (1-t)f(y) && (173) \\ &= t\alpha + (1-t)\alpha && (174) \\ &= \alpha && (175) \end{aligned}$$

Mà do α là giá trị cực tiểu của f nên ta có $f(tx + (1-t)y) = \alpha$ và $tx + (1-t)y$ cũng là một điểm cực tiểu của f. Do đó, tập các điểm cực tiểu của f là tập lồi.

(d) Giả sử $x, y \in \mathbb{R}^n$ là hai cực tiểu của f ứng với giá trị cực tiểu $f(x) = f(y) = \alpha$.
Do f là hàm lồi ngặt nên với $t \in [0, 1]$, ta có

$$\begin{aligned} f(tx + (1-t)y) &< tf(x) + (1-t)f(y) && (176) \\ &= t\alpha + (1-t)\alpha && (177) \\ &= \alpha && (178) \end{aligned}$$

Điều này mâu thuẫn với việc x và y là cực tiểu. Do đó, cực tiểu của hàm lồi ngặt (nếu tồn tại) là duy nhất.

Bài 8: Biện luận theo tham số m số điểm cực tiểu của bài toán sau

$$(P): \quad \text{Min}_{(x,y)\in\mathbb{R}^2} \frac{3}{2}(x^2+y^2)+(1+m)xy-x-y+4 \tag{179}$$

Giải: Đặt ánh xạ $f:\mathbb{R}^n \to \mathbb{R}$ với

$$f(x,y)=\frac{3}{2}(x^2+y^2)+(1+m)xy-x-y+4 \tag{180}$$

Tức là bài toán (P) có thể viết là

$$(P): \quad \text{Min}_{(x,y)\in\mathbb{R}^2} f(x,y) \tag{181}$$

Ta tính được

$$\nabla f(x,y)=\begin{bmatrix} 3x+(1+m)y-1 \\ (1+m)x+3y-1 \end{bmatrix}=\begin{bmatrix} 3 & 1+m \\ 1+m & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}-\begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{182}$$

và

$$\nabla^2 f(x,y)=\begin{bmatrix} 3 & 1+m \\ 1+m & 3 \end{bmatrix} \tag{183}$$

Các trị riêng của $\nabla^2 f(x,y)$ là

$$\begin{cases} \lambda_1 & = 2-m \\ \lambda_2 & = m+4 \end{cases} \tag{184}$$

Nghiêm của phương trình $\nabla f(\bar{x},\bar{y})=0$ là

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}=\begin{bmatrix} \dfrac{1}{m+4} \\ \dfrac{1}{m+4} \end{bmatrix} \tag{185}$$

Tức là nếu $(\bar{x},\bar{y})$ tồn tại $(m\neq -4)$ thì đó là điểm tới hạn duy nhất của f.

• <u>Trường hợp 1:</u> $-4<m<2$. Khi đó

$$\begin{cases} \lambda_1 & > 0 \\ \lambda_2 & > 0 \end{cases} \tag{186}$$

Tức là f là hàm lồi chặt trên miền $\mathbb{R}^2$. Khi đó, $(\bar{x},\bar{y})$ là điểm tới hạn duy nhất của f và cũng là cực tiểu duy nhất của f.

• <u>Trường hợp 2:</u> $m=-4$. Khi đó

$$\begin{cases} \lambda_1 & = 6 > 0 \\ \lambda_2 & = 0 \geq 0 \end{cases} \tag{187}$$

Tức là f là hàm lồi không chặt trên miền $\mathbb{R}^2$. Nhưng do $m=-4$ nên f không có điểm tới hạn nào và do đó f không có cực tiểu.

• <u>Trường hợp 3:</u> $m=2$. Khi đó

$$\begin{cases} \lambda_1 & = 0 > 0 \\ \lambda_2 & = 6 \geq 0 \end{cases} \tag{188}$$

Tức là f là hàm lồi không chặt trên miền $\mathbb{R}^2$. Ta thấy $(\bar{x}, \bar{y}) = \left(\frac{1}{6}, \frac{1}{6}\right)$ là điểm tới hạn duy nhất của f và do đó cũng là cực tiểu duy nhất.

• Trường hợp 4: $m < -4$. Khi đó

$$\begin{cases} \lambda_1 & = 6 > 0 \\ \lambda_2 & < 0 \end{cases} \tag{189}$$

Ta thấy $(\bar{x}, \bar{y}) = \left(\frac{1}{m+4}, \frac{1}{m+4}\right)$ là điểm tới hạn duy nhất của f và do 2 trị riêng của $\nabla^2 f(\bar{x}, \bar{y})$ trái dấu nên đây là điểm yên ngựa. Do đó, f không có cực tiểu.

• Trường hợp 5: $m > 2$. Khi đó

$$\begin{cases} \lambda_1 & < 0 \\ \lambda_2 & = 6 > 0 \end{cases} \tag{190}$$

Ta thấy $(\bar{x}, \bar{y}) = \left(\frac{1}{m+4}, \frac{1}{m+4}\right)$ là điểm tới hạn duy nhất của f và do 2 trị riêng của $\nabla^2 f(\bar{x}, \bar{y})$ trái dấu nên đây là điểm yên ngựa. Do đó, f không có cực tiểu.

Bài 9: Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$ và $\bar{x}, d \in \mathbb{R}^n$. Chứng minh các tính chất sau

(a) Nếu $\nabla f(\bar{x})^\intercal d < 0$ thì d là hướng giảm của f tại $\bar{x}$.

(b) Nếu f là lồi thì: d là hướng giảm của f tại $\bar{x} \Leftrightarrow \nabla f(\bar{x})^\intercal d < 0$.

Giải:
(a) $d \in \mathbb{R}^n$ là một hướng giảm của f tại $\bar{x}$ nếu tìm được $\epsilon > 0$ sao cho

$$f(\bar{x} + td) < f(\bar{x}), \ \ \forall t \in (0, \epsilon) \tag{191}$$

Cho $\nabla f(\bar{x})^\intercal d < 0$, theo lý thuyết về đạo hàm theo hướng, ta có

$$\nabla f(\bar{x})^\intercal d = \lim_{t \to 0^+} \frac{f(\bar{x} + td) - f(\bar{x})}{t} < 0 \tag{192}$$

Do đó, với t đủ nhỏ tức là tồn tại $\epsilon > 0$ sao cho $t \in (0, \epsilon)$ thì

$$f(\bar{x} + td) < f(\bar{x}) \tag{193}$$

hay

$$f(\bar{x} + td) - f(\bar{x}) < 0 \tag{194}$$

Như vậy, d thỏa $\nabla f(\bar{x})^\intercal d < 0$ là một hướng giảm của f tại $\bar{x}$.

(b)
($\Leftarrow$) Như ở câu (a)
($\Rightarrow$) Cho $f : \mathbb{R}^n \to \mathbb{R}$ lồi và $d \in \mathbb{R}^n$ là một hương giảm của f tại $\bar{x}$. Khi đó ta có $\epsilon > 0$ sao cho

$$f(\bar{x} + td) < f(\bar{x}), \ \ \forall t \in (0, \epsilon) \tag{195}$$

Do f là hàm lồi nên với mọi $x, y \in \mathbb{R}^n$, ta có

$$f(x) \geq f(y) + \nabla f(y)^\intercal (x - y) \tag{196}$$

Chọn $x = \bar{x} + td$ với $t \in (0, \epsilon)$ và $y = \bar{x}$, ta có

$$\begin{align}
& f(\bar{x} + td) \geq f(\bar{x}) + \nabla f(\bar{x})^\intercal (\bar{x} + td - \bar{x}) \tag{197}\\
\Leftrightarrow & f(\bar{x} + td) \geq f(\bar{x}) + t \nabla f(\bar{x})^\intercal d \tag{198}\\
\Leftrightarrow & f(\bar{x} + td) - f(\bar{x}) \geq t \nabla f(\bar{x})^\intercal d \tag{199}\\
& \tag{200}
\end{align}$$

Mà ta có

$$\begin{cases} t > 0 \\ f(\bar{x} + td) - f(\bar{x}) < 0 \end{cases} \tag{201}$$

nên

$$\nabla f(\bar{x})^\intercal d < 0 \tag{202}$$

Bài 10: Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$ và $\bar{x}, d \in \mathbb{R}^n$. Chứng minh rằng nếu d khác 0 và $\|\nabla f(\bar{x}) + d\|^2 \leq \|\nabla f(\bar{x})\|^2$ thì d là hướng giảm của f tại $\bar{x}$.
Giải:
Với các giả thiết đề cho, ta có

$$\|\nabla f(\bar{x}) + d\|^2 \leq \|\nabla f(\bar{x})\|^2 \tag{203}$$
$$\Leftrightarrow (\nabla f(\bar{x}) + d)^\intercal (\nabla f(\bar{x}) + d) \leq \nabla f(\bar{x})^\intercal \nabla f(\bar{x}) \tag{204}$$
$$\Leftrightarrow \nabla f(\bar{x})^\intercal \nabla f(\bar{x}) + \nabla f(\bar{x})^\intercal d + d^\intercal \nabla f(\bar{x}) + d^\intercal d \leq \nabla f(\bar{x})^\intercal \nabla f(\bar{x}) \tag{205}$$
$$\Leftrightarrow 2\nabla f(\bar{x})^\intercal d \leq -d^\intercal d \tag{206}$$

Mà do $d \in \mathbb{R}^n$ khác 0 nên

$$d^\intercal d > 0 \tag{207}$$

và

$$-d^\intercal d < 0 \tag{208}$$

Từ đó, (206) dẫn đến

$$\nabla f(\bar{x})^\intercal d < 0 \tag{209}$$

Do đó, d là một hướng giảm của f tại $\bar{x}$

Bài 11: Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$ và $\bar{x}, y \in \mathbb{R}^n$, biết rằng $f(y) < f(\bar{x})$. Chứng minh rằng $d = y - \bar{x}$ là hướng giảm của f tại $\bar{x}$.
Giải:
Do f là hàm lồi nên với mọi $x, y \in \mathbb{R}^n$, ta có

$$f(y) \geq f(x) + \nabla f(x)^\intercal (y - x) \tag{210}$$

Chọn $x = \bar{x}$, ta có

$$f(y) \geq f(\bar{x}) + \nabla f(\bar{x})^\intercal (y - \bar{x}) \tag{211}$$
$$\Leftrightarrow f(y) \geq f(\bar{x}) + \nabla f(\bar{x})^\intercal d \tag{212}$$
$$\Leftrightarrow f(y) - f(\bar{x}) \geq \nabla f(\bar{x})^\intercal d \tag{213}$$

Mà do $f(y) < f(\bar{x})$ nên từ (213) ta có

$$\nabla f(\bar{x})^\intercal d \leq f(y) - f(\bar{x}) < 0 \tag{214}$$

Do đó, $d = y - \bar{x}$ là một hướng giảm của f tại $\bar{x}$.

Bài 12: Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$, hãy tìm một hướng giảm d của f tại $\bar{x}$ trong các trường hợp sau.

(a) $f(x, y) = x^2 + y^2 - xy - x + 2y - 3$ và $\bar{x} = (0, 0)$

(b) $f(x, y) = 2x^2 + y^2 - 2xy + 2x^3 + x^4$ và $\bar{x} = (-1, 0)$

(c) $f(x, y) = \dfrac{1}{2}(x - 2y)^2 + x^4$ và $\bar{x} = (2, 1)$

(d) $f(x, y, z) = x^2 + y^2 + z^2 - xy - xz + 2x - 4y - 2z$ và $\bar{x} = (0, 0, 1)$

Giải:
Ta có

$$\nabla f(\bar{x})^\intercal(-\nabla f(\bar{x})) = -\|\nabla f(\bar{x})\|^2 \tag{215}$$

Do đó, nếu $\nabla f(\bar{x}) \neq 0$ thì $d = -\nabla f(\bar{x})$ là một hướng giảm của f tại $\bar{x}$
(a) $f(x, y) = x^2 + y^2 - xy - x + 2y - 3$ và $\bar{x} = (0, 0)$
Ta tính được

$$\nabla f(x, y) = \begin{bmatrix} 2x - y - 1 \\ -x + 2y + 2 \end{bmatrix} \tag{216}$$

Do đó

$$\nabla f(\bar{x}) = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \tag{217}$$

Vậy

$$d = -\nabla f(\bar{x}) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \tag{218}$$

là một hướng giảm của f tại $\bar{x}$.

(b) $f(x, y) = 2x^2 + y^2 - 2xy + 2x^3 + x^4$ và $\bar{x} = (-1, 0)$
Ta tính được

$$\nabla f(x, y) = \begin{bmatrix} 4x^3 + 6x^2 + 4x - 2y \\ -2 * x + 2y \end{bmatrix} \tag{219}$$

Do đó

$$\nabla f(\bar{x}) = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \tag{220}$$

Vậy

$$d = -\nabla f(\bar{x}) = \begin{bmatrix} 2 \\ -2 \end{bmatrix} \tag{221}$$

là một hướng giảm của f tại $\bar{x}$.

(c) $f(x, y) = \dfrac{1}{2}(x - 2y)^2 + x^4$ và $\bar{x} = (2, 1)$
Ta tính được

$$\nabla f(x, y) = \begin{bmatrix} 4x^3 + x - 2y \\ -2 * x + 4y \end{bmatrix} \tag{222}$$

Do đó

$$\nabla f(\bar{x}) = \begin{bmatrix} 32 \\ 0 \end{bmatrix} \tag{223}$$

Vậy

$$d = -\nabla f(\bar{x}) = \begin{bmatrix} -32 \\ 0 \end{bmatrix} \tag{224}$$

là một hướng giảm của f tại $\bar{x}$.

(d) $f(x, y, z) = x^2 + y^2 + z^2 - xy - xz + 2x - 4y - 2z$ và $\bar{x} = (0, 0, 1)$
Ta tính được

$$\nabla f(x, y, z) = \begin{bmatrix} 2x - y - z + 2 \\ -x + 2y - 4 \\ -x + 2z - 2 \end{bmatrix} \tag{225}$$

Do đó

$$\nabla f(\bar{x}) = \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix} \tag{226}$$

Vậy

$$d = -\nabla f(\bar{x}) = \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} \tag{227}$$

là một hướng giảm của f tại $\bar{x}$.