

# 国家税务总局徐州市税务局稽查局

Convex Optimization HW2 王奕博 2100011025

1. a. interior:  $\emptyset$  closure:  $\emptyset$  boundary:  $\emptyset$   
b. interior:  $\mathbb{R}^n$  closure:  $\mathbb{R}^n$  boundary:  $\emptyset$   
c. interior:  $(0, 1) \cup (2, 3) \cup (4, 5)$  closure:  $[0, 1] \cup [2, 3] \cup [4, 5]$   
boundary:  $\{0, 1, 2, 3, 4, 5\}$   
d. interior:  $\{(x, y)^T \mid x > 0, y > 0\}$  closure:  $\{(x, y)^T \mid x \geq 0, y \geq 0\}$   
boundary:  $\{(x, y)^T \mid (x=0, y>0) \text{ 或 } (x>0, y=0)\}$   
e. 由于  $C_5^C = \dots \cup (-1, 0) \cup (0, 1) \cup \dots$  为开集  
故  $(C_5^C)^\circ = C_5^C$ , closure  $= ((C_5^C)^\circ)^C = C_5 = \{k \mid k \in \mathbb{Z}\}$   
interior  $= \emptyset$ . boundary  $= \overline{C} / C^\circ = C_5 = \{k \mid k \in \mathbb{Z}\}$   
f. 其为离散点列, 故 interior  $= \emptyset$   
而  $C_6$  不为闭 (因为  $\lim_{k \rightarrow \infty} \frac{1}{k} = 0$ , 0 不在  $C_6$  内), 而  $C_6 \cup \{0\}$  为闭  
又因为 显然  $C_6 \cup \{0\}$  为含  $C_6$  的最小闭集, 故:  
closure  $= \{0\} \cup \{\frac{1}{k} \mid k \in \mathbb{Z}\}$   
boundary  $= \text{closure} / \text{interior} = \{0\} \cup \{\frac{1}{k} \mid k \in \mathbb{Z}\}$   
g. 其为离散点列, 故 interior  $= \emptyset$   
而  $\sin k, k \in \mathbb{Z}$  在  $[-1, 1]$  上稠密, 故可知:  
closure  $= \left\{ \begin{bmatrix} 1/k \\ \sin k \end{bmatrix}, k \in \mathbb{Z} \right\} \cup \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix}, y \in [-1, 1] \right\}$   
boundary  $= \text{closure} = \mathcal{I}$

4. (b) 若  $x \in (A \cap B)^\circ$ , 则  $\exists \varepsilon, s.t. B_\varepsilon(x) \subseteq A$  且  $B_\varepsilon(x) \subseteq B$

故  $x \in A^\circ$  且  $x \in B^\circ$ ,  $x \in A^\circ \cap B^\circ$

若  $x \in A^\circ$  且  $x \in B^\circ$ , 则  $\exists \varepsilon_1, s.t. B_{\varepsilon_1}(x) \subseteq A, \exists \varepsilon_2, s.t. B_{\varepsilon_2}(x) \subseteq B$

则  $\exists \varepsilon = \min(\varepsilon_1, \varepsilon_2), s.t. B_\varepsilon(x) \subseteq A \cap B \therefore x \in (A \cap B)^\circ$

$$\therefore (A \cap B)^\circ = A^\circ \cap B^\circ$$

若  $x \in A^\circ$  或  $x \in B^\circ$ , 则  $\{ \exists \varepsilon_1, s.t. B_{\varepsilon_1}(x) \subseteq A \}$  或  $\{ \exists \varepsilon_2, s.t. B_{\varepsilon_2}(x) \subseteq B \}$  成立  
则  $x \in A^\circ \cup B^\circ$  - 定成立.

$$\therefore \text{故: } (A \cup B)^\circ \supseteq A^\circ \cup B^\circ$$

令  $A = [0, 1], B = [1, 2]$ . 则  $A^\circ \cup B^\circ = (0, 1) \cup (1, 2)$ ,  $(A \cup B)^\circ = (0, 2) \neq A^\circ \cup B^\circ$ .

(a) (a) 利用 (b) 的结论,

$$\begin{aligned} \overline{A \cup B} &= [[A \cup B]^c]^\circ = [[A^c \cap B^c]]^\circ \xrightarrow{\text{上问结论}} [[A^c]^\circ \cap [B^c]^\circ]^\circ \\ &= [[A^c]^\circ]^\circ \cup [[B^c]^\circ]^\circ = \bar{A} \cup \bar{B}. \text{ 得证.} \end{aligned}$$

$$\overline{A \cap B} = [[A \cap B]^c]^\circ = [[A^c \cup B^c]]^\circ$$

而  $[A^c \cup B^c]^\circ \supseteq [A^c]^\circ \cup [B^c]^\circ$

$$\begin{aligned} \text{故 } [[A^c \cup B^c]]^\circ &\subseteq [[A^c]^\circ \cup [B^c]^\circ]^\circ = [[A^c]^\circ]^\circ \cap [[B^c]^\circ]^\circ \\ &= \bar{A} \cap \bar{B}. \text{ 得证.} \end{aligned}$$

令  $A = [0, 1], B = [1, 2]$

且  $A \cap B = \emptyset, \overline{A \cap B} = \emptyset$

$$\bar{A} = [0, 1], \bar{B} = [1, 2], \bar{A} \cap \bar{B} = \{1\} \neq \overline{A \cap B}$$

2. " $\Rightarrow$ "  $x \in \partial C \Rightarrow x \in \bar{C}$  且  $x \notin C^\circ$

$x \notin C^\circ$  知,  $\forall \varepsilon > 0, \exists z \notin C$ , s.t.  $\|z - x\|_2 \leq \varepsilon$

$x \in \bar{C}$  知,  $x \notin (C^\circ)^\circ$ . 故  $\forall \varepsilon > 0, \exists y \notin C^\circ$ , s.t.  $\|y - x\|_2 \leq \varepsilon$   
而  $y \notin C^\circ \Leftrightarrow y \in C$ . 即证.

" $\Leftarrow$ "

$\forall \varepsilon > 0, \exists z \notin C$ , s.t.  $\|z - x\|_2 \leq \varepsilon$  知,  $x \notin C^\circ$

$\forall \varepsilon > 0, \exists y \in C$ , 即  $y \notin C^\circ$ , s.t.  $\|y - x\|_2 \leq \varepsilon$ , 知,  $x \notin (C^\circ)^\circ$

故  $x \in ((C^\circ)^\circ)^\circ$  且  $x \notin C^\circ$

即  $x \in ((C^\circ)^\circ)^\circ / C^\circ = \partial C$

3. 以下使用两个分析学结论:

Lemma①  $C$  is closed  $\Leftrightarrow \mathbb{R}^n / C$  is open; vice versa. (开闭互补, 经典结论)

Lemma②  $C$  is open  $\Leftrightarrow C = C^\circ$  (由是定义)

[1]  $C$  is closed  $\Leftrightarrow \partial C \subseteq C$

Proof: " $\Rightarrow$ ":  $C$  is closed  $\Rightarrow C^\circ$  is open  $\Rightarrow (C^\circ)^\circ = C^\circ$

$\Rightarrow C = ((C^\circ)^\circ)^\circ$ , 而  $C^\circ \subseteq C$  显然

故  $\partial C = ((C^\circ)^\circ)^\circ / C^\circ \subseteq C$ .

" $\Leftarrow$ ":  $\partial C \subseteq C \Rightarrow ((C^\circ)^\circ)^\circ / C^\circ \subseteq C$ , 而  $C^\circ \subseteq C$  显然

$\Rightarrow ((C^\circ)^\circ)^\circ \subseteq C$

$\Rightarrow C^\circ \subseteq (C^\circ)^\circ$  (利用  $A \subseteq B \Leftrightarrow \mathbb{R}^n / A \supseteq \mathbb{R}^n / B$ )

而  $(C^\circ)^\circ \subseteq C^\circ$  显然,  $\therefore C^\circ = (C^\circ)^\circ$ , 故  $C^\circ$  为开

$\therefore C$  闭.

[2]  $C$  is open  $\Leftrightarrow C \cup \partial C = \emptyset$

Proof: 注意到  $C$  is open  $\Leftrightarrow C^\circ$  is closed,  $C \cup \partial C = \emptyset \Leftrightarrow \partial C \subseteq C^\circ$

而 Lemma 3:  $\partial C = \partial C^\circ$ . 证明:  $\partial C = [C^\circ]^\circ / C^\circ, \partial C^\circ = [C^\circ]^\circ / [C^\circ]^\circ$

而  $A^\circ / B^\circ = B / A$ , 放  $\partial C^\circ = [C^\circ]^\circ / C^\circ = \partial C$ . 得证.

故  $C \cup \partial C = \emptyset \Leftrightarrow \partial C^\circ \subseteq C^\circ$  而上面已证  $C^\circ$  is closed  $\Leftrightarrow \partial C^\circ \subseteq C^\circ$

故:  $C$  is open  $\Leftrightarrow C \cup \partial C = \emptyset$

# 国家税务总局徐州市税务局稽查局

5. 令  $x_n = 4 e^{-1.5^n}$

$$\begin{aligned} \text{(a)} \quad \lim_{n \rightarrow \infty} \frac{x_{n+1}}{x_n} &= \lim_{n \rightarrow \infty} \exp(\ln x_{n+1} - 1.5 \ln x_n) \\ &= \lim_{n \rightarrow \infty} \exp \left\{ \ln 4 - 1.5^{n+1} - 1.5 (\ln 4 - 1.5^n) \right\} \\ &= \lim_{n \rightarrow \infty} \exp \{-0.5 \ln 4\} = 4^{-0.5} = 0.5 \end{aligned}$$

且  $\lim_{n \rightarrow \infty} x_n = 0$ , 放得证.

6. a.  $\lim_{n \rightarrow \infty} x_n = 0$ , 且  $\lim_{n \rightarrow \infty} \left| \frac{x_{n+1}}{x_n} \right| = 0.5$ , 故  $p=1$ , rate constant = 0.5  
 CP is rate of convergence)

b.  $\lim_{n \rightarrow \infty} x_n = 1$ , 且  $\lim_{n \rightarrow \infty} \left| \frac{x_{n+1}-1}{x_n-1} \right| = 0.01$ , 故  $p=1$ , rate constant = 0.01

c.  $\lim_{n \rightarrow \infty} x_n = 0$ , 且  $\lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|^2} = \frac{2^{-2}}{2^{-2k-2}} = 1$ , 故  $p=2$ , rate constant = 1

d.  $\lim_{n \rightarrow \infty} x_n = 0$ ,  $\lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|} = \lim_{n \rightarrow \infty} 2^{n-1} = 0$  故  $p=1$ , rate constant = 0  
 (Superlinear)

e.  $\lim_{n \rightarrow \infty} x_n = 0$ , 但  $\lim_{n \rightarrow \infty} \frac{|x_{n+1}-1|}{|x_n-1|^r}$  不存在, 故不存在  $p$  & rate constant

7. (a)  $\frac{d \|\vec{x}\|_p^p}{d \vec{x}} = \frac{d(x_1^p + \dots + x_n^p)}{d \vec{x}} = \begin{bmatrix} p x_1^{p-1} \\ \vdots \\ p x_n^{p-1} \end{bmatrix}$ . 故  $\frac{df}{d\vec{x}} = \frac{d\|x\|_p^p}{d\|x\|_p^p} \cdot \frac{d\|x\|_p^p}{dx}$

$$\begin{aligned} \text{(b)} \quad \frac{df}{dx} &= (a^T x) \frac{d b^T x}{dx} + (b^T x) \cdot \frac{d a^T x}{dx} \\ &= (a^T x) \vec{b} + (b^T x) \vec{a}. \end{aligned}$$

$$\begin{aligned} &= \frac{1}{p \|x\|_p^{p-1}} \begin{bmatrix} p x_1^{p-1} \\ \vdots \\ p x_n^{p-1} \end{bmatrix} \\ &= \frac{1}{\|x\|_p^{p-1}} \begin{bmatrix} x_1^{p-1} \\ \vdots \\ x_n^{p-1} \end{bmatrix}. \end{aligned}$$

(c) 设  $Ax = y$ ,

$$\frac{df}{dy} = \frac{d}{dy} \frac{1}{2} \|y - b\|_2^2 = \vec{y} - \vec{b}, \text{ 即 } \hat{f}(y + A\delta x) \approx \hat{f}(y) + \langle \frac{df}{dy}, \delta y \rangle$$

$$\therefore f(x + \Delta x) \approx f(x) + \langle Ax - b, A\Delta x \rangle$$

$$= f(x) + (Ax - b)^T \cdot A\Delta x = f(x) + \langle A^T(Ax - b), \Delta x \rangle \therefore \frac{df}{dx} = A^T(Ax - b)$$

d. 设  $\tilde{f}(y) = u^T g(y) = u_1 g_{(1)} + \dots + u_n g_{(n)}$

$$\therefore \frac{d\tilde{f}}{dy} = \begin{bmatrix} u_1 g'_1 \\ \vdots \\ u_n g'_n \end{bmatrix},$$

而由上问, 易知:  $\frac{df}{dx} = R^T \frac{d\tilde{f}}{dy} = R^T \begin{bmatrix} u_1 g'_1(y) \\ \vdots \\ u_n g'_n(y) \end{bmatrix}$

e. 在一阶小量意义下, 考虑:

$$\begin{aligned} f(x+AX) &= \| (x^T + \Delta x^T) A (x + \Delta x) \|_F^2 \rightarrow \text{抛弃二阶小量} \\ &= \sum_{i=1}^m \sum_{j=1}^n [(x^T A x)_{ij} + (\Delta x^T A x)_{ij} + (x^T A \Delta x)_{ij}]^2 \\ &\approx \sum_{i=1}^m \sum_{j=1}^n (x^T A x)_{ij}^2 + 2(x^T A x)_{ij} (\Delta x^T A x + x^T A \Delta x)_{ij} \rightarrow \text{抛弃二阶小量} \\ &= \| x^T A x \|_F^2 + 2 \sum_{i=1}^m \sum_{j=1}^n (x^T A x)_{ij} (\Delta x^T A x + x^T A \Delta x)_{ij} \quad \sum_{i=1}^m \sum_{j=1}^n A_{ij} B_{ij} = \text{tr}(A^T B) \\ &= \| x^T A x \|_F^2 + 2 \text{tr}((x^T A x)^T (\Delta x^T A x + x^T A \Delta x)) \\ &= \| x^T A x \|_F^2 + 2 \text{tr}(x^T A x \cdot A x^T \cdot A x) + 2 \text{tr}(x^T A x \cdot x^T A \Delta x) \\ &= \| x^T A x \|_F^2 + \text{tr}(2 x^T A \Delta x \cdot x^T A x + 2 x^T A x x^T A \Delta x) \\ &= \| x^T A x \|_F^2 + \text{tr}(4 (A x x^T A x)^T A x) \quad \text{利用 } \text{tr}(AB) = \text{tr}(BA) \\ \text{故: } \frac{\partial f}{\partial x} &= 4 A x x^T A x. \end{aligned}$$

f. 依然在一阶小量下考虑, 部分过程与上面相同,

$$\begin{aligned} f(x+AX) &\approx \| \text{diag}(x^T A x) + 2 \text{diag}(x^T A A x) \|_F^2 \\ &= \sum_{i=1}^m [(x^T A x)_{ii} + 2(x^T A A x)_{ii}]^2 \approx \sum_{i=1}^m (x^T A x)_{ii}^2 + 4(x^T A x)_{ii} \cdot (x^T A A x)_{ii} \\ &= \| \text{diag}(x^T A x) \|_F^2 + 4 \text{tr}(\text{diag}(x^T A x) \cdot \text{diag}(x^T A A x)) \end{aligned}$$

而易知,  $\text{tr}(\text{diag}(A) \cdot \text{diag}(B)) = \text{tr}(\text{diag}(A) \cdot B)$

故  $f(x+AX) \approx f(x) + \text{tr}(4 \text{diag}(x^T A x) \cdot x^T A A x)$ , 由此,  $\frac{df}{dx} = 4 A x \text{diag}(x^T A x)$

g.  $f(x+AX) = \text{tr}(A(x+AX)B) = \text{tr}(AxB) + \text{tr}(A\Delta x B)$

$$\frac{\text{tr}(AB) = \text{tr}(BA)}{\text{tr}(AxB) + \text{tr}(B\Delta x B)}, \text{ 故 } \frac{df}{dx} = (BA)^T = A^T B^T$$

# 国家税务总局徐州市税务局稽查局

HW2 修正:

7. 的 (a) (d) 为 vector form:

(a):  $\nabla f = \|x\|_p^{1-p} \cdot \vec{x}^{p-1}$ , 其中 定义  $\vec{x}^a = \begin{bmatrix} x_1^a \\ \vdots \\ x_n^a \end{bmatrix}$ .  
或:  $\nabla f = \|x\|_p^{1-p} \cdot [\text{diag}(\vec{x})]^{p-1} \cdot \vec{1}$

(d)

$$\frac{df}{dx} = \nabla f = R^T \cdot (u \odot g(R\vec{x}))$$

其中  $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \odot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} a_1 b_1 \\ \vdots \\ a_n b_n \end{bmatrix}$

或  $\nabla f = R^T \text{diag}(w \cdot g(R\vec{x}))$