הטכניון, מכון טכנולוגי לישראל

הפקולטה להנדסת חשמל

שיטות חישוביות באופטימיזציה

046197

סמסטר אביב תשפ"א 2021

100

מגישים:

הדס בירן	300916285
עידו יחזקאל	204397368

שיטות חישוביות באופטימיזציה -

תרגיל בית 1

מגישים: הדס בירן, ת.ז. 300916285

עידו

תרגיל 1

1. לא נכון, מכיוון שאם M היא מטריצת האפס אז $\langle x,x\rangle_M=0$ גם אם $x\neq 0$.

2. נכון. נוכיח לפי ההגדרה:

1) סימטריה:

$$\langle x,y\rangle_Q = x^TQy \underset{\substack{\uparrow\\ \text{הכפל}\\ \text{הוא סקלר}}}{=} (x^TQy)^T = y^TQ^Tx \underset{\substack{\uparrow\\ Q\succ 0\\ \text{ולכן סימטרית}}}{=} y^TQx = \langle y,x\rangle_Q$$

2) אדיטיביות:

$$\langle x+y,z\rangle_Q = (x+y)^TQz = x^TQz + y^TQz = \langle x,z\rangle_Q + \langle y,z\rangle_Q$$

3) הומוגניות:

$$\langle \lambda x,y\rangle_Q = \lambda x^TQy = \lambda\langle x,y\rangle_Q$$

4) חיוביות מוגדרת:

$$\langle x,x\rangle_Q = x^TQx \underset{Q\succ 0}{>} 0 \quad \forall x\neq 0$$

ועבור $x=0$:

$$\langle x,x\rangle_Q = x^TQx \underset{\substack{\uparrow\\ x=\vec{0}}}{=} \vec{0}\cdot Q\cdot\vec{0} = 0$$

תרגיל 2

1. אם $A \succ 0$ אזי היא הפיכה, ו $A^{-1} \succ 0$.

הוכחה

תהי $A \in \mathbb{R}^{n\times n}$ מטריצה סימטרית המקיימת $A \succ 0$.
לכן, הערכים עצמיים של A כולם חיוביים ממש,
ולכן הדטרמיננטה של A, $\det(A)$, אשר שווה למכפלת
הע"ע חיוביים ממש, בפרט לא אפס ולכן A הפיכה. $\Leftarrow \det(A) \neq 0$

בנוסף, נראה כי A^{-1} סימטרית:

$$AA^{-1} = I \Leftrightarrow (AA^{-1})^T = I$$

$$\Rightarrow (A^{-1})^T \cdot A^T = I$$

A סימטרית $\Leftrightarrow (A^{-1})^T \cdot A = I \quad / \cdot A^{-1}$

$$\Rightarrow (A^{-1})^T \cdot \underbrace{A \cdot A^{-1}}_{I} = A^{-1}$$

$$\Rightarrow (A^{-1})^T = A^{-1}$$

ומתכונות מטריצות הפיכות הע"ע של A^{-1} הינם $\frac{1}{\lambda_i}$ הע"ע של A ולכן גם הם חיוביים ממש.
ומשפט בהרצאה $A^{-1} \succ 0$ כי כל הע"ע של חיוביים ממש.

---

2. תהי $A \in \mathbb{R}^{n\times n}$ סימטרית ו- $A \succeq 0$ והפיכה, אזי $A \succ 0$

הוכחה.

$A \succeq 0 \Rightarrow$ כל ע"ע אי-שליליים, A הפיכה $\Rightarrow \det(A) \neq 0$
ומכיוון ש- $\det(A) = \prod \lambda_i$, λ_i ע"ע של A אזי $\lambda_i \neq 0$
לכל i ולכן כל ע"ע של A חיוביים ממש ומתקיים
$A \succ 0$.

③ נראה כי $A=B^TB$ עבור $B\in\mathbb{R}^{m\times n}$ כלשהי:

צ"ל $A\succeq 0$:

הוכחה: יהי $x\in\mathbb{R}^n$ כלשהו.

$$x^TAx = x^TB^TBx$$

ומתקיים לפי כי $\underbrace{Bx}_{m\times n\ \ n\times 1}\in\mathbb{R}^m$ ונסמן כ- y

$$\Rightarrow x^TB^TBx = (Bx)^TBx = y^Ty = \|y\|_2^2 \geq 0$$

נורמת L_2

ומתקיים לנו כי הנורמה אי שלילית ולכן $A\succeq 0$ (מתכונות הנורמה).

[A סימטרית כי $A^T=(B^TB)^T=B^T(B^T)^T=B^TB=A$].

דוגמה: עבור B כך ש- $A\succeq 0$ וגם $A\nsucc 0$

$$B=\begin{pmatrix}0&0\\0&1\end{pmatrix} \; ; \; B^T=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

$$A=\begin{pmatrix}0&0\\0&1\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

$$x=\begin{pmatrix}1\\0\end{pmatrix}$$

$$(1\ 0)\begin{pmatrix}0&0\\0&1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=(0\ 0)\begin{pmatrix}1\\0\end{pmatrix}=0$$

כלומר מצאנו וקטור $x=\begin{pmatrix}1\\0\end{pmatrix}\neq 0$ כך ש- $x^TAx=0$

ולכן $A\succeq 0$.

4. תהי $A \in \mathbb{R}^{n\times n}$ סימטרית כלשהי.

צ"ל אם קיים $A_{ii} < 0$ עבור $1 \le i \le n$

אזי $A \not\succeq 0$.

<u>הוכחה:</u>

נה"כ $A_{ii} < 0$. נבחר את $x \neq 0$ כך ש-

$x = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ כלומר $x_i = 1$

ו-$x_j = 0$ $j \neq i$; $1 \le j \le n$.

$$(0,0,\dots,1,\dots,0)\begin{pmatrix} a_{11} & & a_{1n} \\ & \ddots & \\ a_{n1} & & a_{nn} \end{pmatrix}\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} =$$

$$(a_{i1}, a_{i2}, \dots, a_{ii}, \dots, a_{in})\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_{ii} < 0$$

כלומר, מצאנו $x \neq 0$ כך ש- $x^T A x < 0$

ובפרט $A \not\succeq 0$.

<u>זה לא תנאי מספיק -</u>

<u>דוגמא:</u> $A = \begin{pmatrix} 1 & -50 \\ -50 & 1 \end{pmatrix}$ ושים לב כי $a_{11}, a_{22} > 0$

וניקח $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$(1\ 1)\begin{pmatrix} 1 & -50 \\ 50 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = (-49\ \ -49)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = -98 < 0$$

מצאנו A עם איברי אלכסון חיוביים ממש אשר לא חיובית מוגדרת ולכן התנאי רק תנאי הכרחי.

Scanned with CamScanner

תרגיל 3

$f:U\to\mathbb{R}$ דיפרנציאבילית ברציפות. $U\subseteq\mathbb{R}^n$.

$d\neq 0 \equiv$ כיוון ירידה של הפונקציה בנקודה x אם קיים $T>0$ כך שלכל $0<t<T$ מתקיים $f(x+td)<f(x)$.

1. צ"ל: $f'(x;d)<0 \Rightarrow d$ כיוון ירידה של f ב-x.

הוכחה:

לכל $x\in U$:

$$f'(x;d)=\lim_{t\to 0^+}\frac{f(x+td)-f(x)}{t}\underset{\text{נתון}}{<}0$$

נניח בשלילה ש-d אינו כיוון ירידה של f ב-x, כלומר לא קיים T קטן מספיק כך שלכל $0<t<T$ מתקיים $f(x+td)<f(x)$.

כלומר, לכל T_0 קיים $0<t_0<T_0$ כך ש- $f(x+td)>f(x)$.
אבל אז ניתן לקחת T_1 $(T_1<t_0)$ כך שקיים $0<t_1<T_1$ כך ש- $f(x+td)>f(x)$.
ואפשר להמשיך באופן הזה וכך שלקבל סדרה הולכת ומתכנסת ל-0: $\{t_0,t_1,t_2,\dots\}$ ושעבור t בה ששייך לסדרה מתקיים $f(x+td)>f(x)$. אבל אז (לפי הגדרת היינה לגבול) $\lim_{t\to 0^+}\frac{f(x+td)-f(x)}{t}>0$ וזה מנוגד לנתון. לכן ההנחה בשלילה הייתה שגויה
ו-d הוא כיוון ירידה של f ב-x. **מ.ש.ל.**

2. צ"ל: $\nabla f(x)\neq 0 \Rightarrow d=-\nabla f(x)$ כיוון ירידה של f ב-x.

הוכחה:

באופן אינטואיטיבי, אם $\nabla f(x)\neq 0$ אז x אינה נקודת מינימום/מקסימום והגרדיאנט $\nabla f(x)$ הוא השיפוע באותה נקודה, ולכן הנגטיב שלו יצביע בדיוק בכיוון הירידה.

נוכיח את זה:

f דיפרנציאבילית ברציפות על U, אז
$\forall x\in U$, $d\in\mathbb{R}^n$, $f'(x;d)=\nabla f(x)^T d$
(לפי משפט 4 מתרגול 1).

נציב $d=-\nabla f(x)$ ונקבל:

$$f'(x;d) = -\nabla f(x)^T \nabla f(x)$$
$$= -\sum_{i=1}^{n} \nabla f(x)_i^2 < 0$$

ובהכרח, אז $\nabla f(x) \neq \vec{0}$
קיים $1 \leq i \leq n$ שעבורו
$\nabla f(x)_i \neq 0$.

לכן לפי 1 d הוא כיוון ירידה של f ב-x.

<u>מ.ש.ל</u>

3. דוגמא:

$$f(x_1,x_2) = x_1^2 \quad , \quad x^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad d_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} , \quad d_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

f דיפרנציאבילית ברציפות על $\mathbb{R}^2$.

$$\nabla f(x_1,x_2) = \begin{pmatrix} \frac{\partial f(x_1,x_2)}{\partial x_1} \\ \frac{\partial f(x_1,x_2)}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 0 \end{pmatrix}$$

$$f'(x_1,x_2;d_1) = \nabla f(x_1,x_2)^T \Big|_{x^{(0)}} \cdot d_1$$
$$= \begin{pmatrix} 2 \\ 0 \end{pmatrix}^T \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -2 < 0$$

$\Leftarrow$ d_1 הוא כיוון ירידה של $f(x_1,x_2)$ ב-$x^{(0)}$.

$$f'(x_1,x_2;d_2) = \nabla f(x_1,x_2) \Big|_{x^{(0)}} \cdot d_2$$
$$= \begin{pmatrix} 2 \\ 0 \end{pmatrix}^T \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

וניתן לראות:

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + t \cdot d_2\right) = f(x_1 + t\cdot 0, x_2 + t \cdot 1)$$
$$= x_1^2$$
$$f(x_1,x_2) = x_1^2$$

$\Longleftarrow$ $f(x+t\cdot d_2) = f(x)$ לכן d_2 <u>לא</u> כיוון ירידה של f ב-x.

תרגיל 4 -

יהי $C \subseteq \mathbb{R}^m$ תחום קמור ויהי $A \in \mathbb{R}^{m \times n}$

נסמן:

$$B_1 = \{ x \in \mathbb{R}^n \mid x = A^T z \ , \ z \in C \}$$

צ"ל B_1 קמור.

יהיו $x_1, x_2 \in B_1$ כך ש- $x_1 = A^T z_1 \ , \ z_1 \in C$

$x_2 = A^T z_2 \ , \ z_2 \in C$

כמו כן מקמירות של C ידוע כי

$z_3 = t z_1 + (1-t) z_2 \in C \quad \forall \ 0 \le t \le 1.$

נראה כי: $\forall 0 \le t \le 1 \quad t x_1 + (1-t) x_2 \in B_1$

$$t x_1 + (1-t) x_2 = t A^T z_1 + (1-t) A^T z_2 = A^T (t z_1 + (1-t) z_2)$$

דיסטריביוטיביות של מכפלת מטריצה

$= A^T \cdot z_3 \ , \ z_3 \in C$ ✓

ולכן לפי הגדרת קמירות B_1 קמור.

נסמן: $B_2 = \{ x \in \mathbb{R}^n \mid Ax \in C \}$

יהיו $x_1, x_2 \in B_2$ שלהם ולכן $Ax_2 \in C$ וגם $Ax_1 \in C$

נראה כי $t x_1 + (1-t) x_2 \in B_2 \quad \forall \ 0 \le t \le 1$

כלומר ש- $A(t x_1 + (1-t) x_2) \in C$

$$A(t x_1 + (1-t) x_2) = t \underbrace{A x_1}_{\in C} + (1-t) \underbrace{A x_2}_{\in C} \in C$$ ✓

מקמירות של C. ולכן מהגדרת קמירות B_2 קמור.

תרגיל 5-

1. תהי $S \subseteq \mathbb{R}^n$ ו- $x \in \mathbb{R}^n$ נקודת צבר של x.
אזי לכל $\varepsilon>0$ הכדור הפתוח $B(x,\varepsilon)$
מכיל אינסוף נקודות מ- S

הוכחה-

נניח בשלילה כי קיים $\exists\ \varepsilon>0$ כך שהכדור הפתוח
$B(x,\varepsilon)$ מכיל <u>מספר סופי</u> של נקודות מ- S.
נסמן אותם ב- $\{y_i\}_{i=1}^{n}$ כך ש- $x\notin\{y_i\}$.

נגדיר $\delta=\min_{y_i}\|y_i-x\|$, ושים לב ש- δ
מוגדר על פני מינימום איברים סופי ולכן מוגדר היטב.

כעת נסתכל על הכדור $\tilde{B}(x,\varepsilon-\delta)$ ונשים לב כי
$\varepsilon-\delta<\varepsilon$ כי $\delta\neq 0$ $(x\notin\{y_i\})$ ו- $\varepsilon-\delta>0$
בגלל שהכדור $B(x,\varepsilon)$ פתוח.
- מניסוח $B(x,\varepsilon-\delta)$ $\{y_i\}\notin B(x,\varepsilon-\delta)$ $\forall\, 1\le i\le n$
- ו- $\tilde{B}(x,\varepsilon-\delta)\subseteq B(x,\varepsilon)$ ולכן של נקודה ב-
כדור קטן מוכל בגדול ומכאן שבניית $\tilde{B}$
אינו מכיל נקודות נוספות ב- S (אחרת היו ב- B)
ומכיוון ש- x אינה נקודת צבר כי מצאנו
$0<\varepsilon-\delta=\tilde{\varepsilon}$ כך שלא מכיל נקודה ב- S וזו
סתירה להנחה כי מספר נקודות סופי!

② S סגורה אמ"מ מכילה את כל נקודות הצבר שלה.

הוכחה-

$S\subseteq\mathbb{R}^n$ קבוצה סגורה $\Rightarrow$ S מכילה את כל נקודות הצבר שלה.

נניח בשלילה כי קיימת $x\in\mathbb{R}^n$ נקודת צבר כך ש- $x\notin S$.

אזי בפרט $x\in S^c$ ולכן קיים $\varepsilon>0$ שעבורו $B(x,\varepsilon)\in S^c$, כלומר כל הנקודות בכדור הפתוח מוכלות ב- S^c. $\Rightarrow$ קיים $\varepsilon>0$ כך שעבורו $B(x,\varepsilon)$ לא מכיל נקודה אחת מ- S ולכן x לא נקודת צבר $\Rightarrow$ מצאנו סתירה, S מכילה את x.

S מכילה את כל נקודות הצבר שלה $\Rightarrow$ S סגורה-

יהי $x\in S^c$, נקודה בקבוצה המשלימה אזי x אינה נקודת צבר של S, (אחרת היתה מוכלה בה).
ולכן קיים $\varepsilon>0$ כך ש- $B(x,\varepsilon)$ אינו מכיל אף נקודה מ- S (בפרט סביבה זו מוכלת במשלים.
ולכן x שייכת ל- S^c כלומר $x\in S^c$ נקודת פנים.
$\Rightarrow$ כל הנקודות $x\in S^c$ שייכות לפנים שלו
ולכן S^c קבוצה פתוחה $\Rightarrow$ S סגורה.

אם S קבוצה סגורה והסדרה $x_i\in S$, $i=1,2,\dots$ מתכנסת ל-x, $\lim_{i\to\infty}\|x_i-x\|=0$ אזי $x\in S$.

הוכחה-

לפי הגדרת הגבול $\forall\varepsilon>0, \exists N: \|x_N-x\|<\varepsilon \Rightarrow x_N\in B(x,\varepsilon)$

ולכן $\forall\varepsilon>0$ הכדור הפתוח $B(x,\varepsilon)$ מכיל נקודה $x_i\in S$, $\exists i$, x נקודת צבר של S, S סגורה $x\in S$.

Scanned with CamScanner

תרגיל 6-

יהי $C \subseteq \mathbb{R}^n$ תחום קמור וסגור ויהי $x_0 \notin C$

נתונה בעיית האופטימיזציה:

$$\min_{y\in C} \|y-x_0\|$$

כאשר $\|\cdot\|$ נורמה מושרית מכפ"פ:

נניח כי קיים פתרון $y_0 \in C$ לבעיה. הוכח כי הוא יחיד.

הוכחה- נראה בהוכחה בשלילה.

נניח כי קיימים 2 פתרונות- $y_0, z_0 \in C$

כך ש-

$$\|y_0-x_0\|^2 = \|z_0-x_0\|^2 = \left(\min_{y\in C} \|y-x_0\|\right)^2 = \delta^2$$

תזכורת- מקמירות & תחום C מתקיים $(y_0+z_0)\in C$ כי לכל $0 \le t \le 1$ $ty_0+(1-t)z_0 \in C$ בפרט $t=\frac{1}{2}$

ולכן:

$$\underbrace{\left\|\tfrac{1}{2}y_0+\tfrac{1}{2}z_0-x_0\right\|^2}_{\ge\delta^2} = \left\|\tfrac{1}{2}y_0-\tfrac{1}{2}x_0+\tfrac{1}{2}z_0-\tfrac{1}{2}x_0\right\|^2$$

$$= \frac{1}{4}\|y_0-x_0+z_0-x_0\|^2$$

$$= \frac{1}{4}\|y_0-x_0\|^2+\frac{1}{4}\|z_0-x_0\|^2+\frac{1}{4}\cdot 2\langle y_0-x_0, z_0-x_0\rangle$$

$$= \frac{1}{4}\delta^2+\frac{1}{4}\delta^2+\frac{1}{2}\langle x_0-y_0, x_0-z_0\rangle = \frac{1}{2}\delta^2+\frac{1}{2}\langle x_0-y_0, x_0-z_0\rangle$$

$$\ge \delta^2 \quad \Rightarrow \quad \frac{1}{2}\langle x_0-y_0, x_0-z_0\rangle \ge \frac{1}{2}\delta^2$$

גדול שווה למינימום.

$$\Rightarrow \langle x_0-y_0, x_0-z_0\rangle \ge \delta^2$$

$$\|y_0 - z_0\|^2 = \|y_0 - x_0 + x_0 - z_0\|^2$$

$$= \|y_0 - x_0\|^2 + \|x_0 - z_0\|^2 + 2\langle y_0 - x_0, x_0 - z_0\rangle$$

$$= \|y_0 - x_0\|^2 + \|z_0 - x_0\|^2 - 2\langle x_0 - y_0, x_0 - z_0\rangle$$

$$= \delta^2 + \delta^2 - 2\langle x_0 - y_0, x_0 - z_0\rangle$$

$$\le 2\delta^2 - 2\delta^2 \le 0$$

הנחת חלק שלי

מכיוון ש- $0 \le \|\cdot\|$ (מחיוביות) הרי

$$0 \le \|y_0 - z_0\|^2 \le 0 \Rightarrow \|y_0 - z_0\|^2 = 0$$

$\|x\| = 0$ אמ"מ $x = 0$ ולכן $y_0 - z_0 = 0 \Rightarrow y_0 = z_0$

<u>לכן ההנחה כי יש 2 פתרונות אופטימליים לא נכונה</u>

<u>וקיים פתרון אופטימלי יחיד.</u>