# פתרון תרגיל מספר 7 ־ כלים מתמטיים

**שם:** מיכאל גרינבאום, **ת.ז:** 211747639

4 בדצמבר 2020

1. **פתרון:**

(א) **צ"ל:** $\lambda_{k+t-1}\left(A+B\right)\le\lambda_{k}\left(A\right)+\lambda_{t}\left(B\right)$ וגם $\lambda_{i+j-n}\left(A+B\right)\ge\lambda_{i}\left(A\right)+\lambda_{j}\left(B\right)$

**הוכחה:**

נזכר שראינו כי $\lambda_{k}\left(A\right)=\max_{F;\dim F=k-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax$, נשתמש בכך ונשים לב כי

$$\begin{aligned}\lambda_{i+j-n}\left(A+B\right) & =\max_{F;\dim F=i+j-n-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}\left(A+B\right)x\\
 & =\max_{F;\dim F=i+j-n-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax+x^{T}Bx\\
 & \overset{\star}{\ge}\left(\max_{F;\dim F=i+j-n-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax\right)+\left(\max_{F;\dim F=i+j-n-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Bx\right)\\
 & \overset{\star\star}{\ge}\left(\max_{F;\dim F=i-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax\right)+\left(\max_{F;\dim F=j-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Bx\right)\\
 & =\lambda_{i}\left(A\right)+\lambda_{j}\left(B\right)
\end{aligned}$$

נשים לב ש־ $\star$ נכון כי $\min a+b\ge\min a+\min b$,

נשים לב ש־ $\star\star$ נכון כי $i,j\le n$ ולכן $j-n,i-n<0$ ולכן אם נסמן ב־ $F=\text{argmax}_{F;\dim F=i-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax$, נוכל להגדיר $F'\subseteq F$ תת מרחב כך ש־ $\dim F'=i+j-n-1$ ולקבל כי

$$\begin{aligned}\max_{F;\dim F=i+j-n-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax\ge\min_{\|x\|=1\wedge x\in F'^{\perp}}x^{T}Ax & \ge\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax\\
 & =\max_{F;\dim F=i-1}\min_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax
\end{aligned}$$

באופן דומה ל־ $B,j$.

עתה נזכר שראינו כי $\lambda_{k}\left(A\right)=\min_{F;\dim F=n-k}\max_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax$, נשתמש בכך ונשים לב כי

$$\begin{aligned}\lambda_{k+t-1}\left(A+B\right) & =\min_{F;\dim F=n-k-t+1}\max_{\|x\|=1\wedge x\in F^{\perp}}x^{T}\left(A+B\right)x\\
 & =\min_{F;\dim F=n-k-t+1}\max_{\|x\|=1\wedge x\in F^{\perp}}\left[x^{T}Ax+x^{T}Bx\right]\\
 & \overset{\star}{\le}\left[\min_{F;\dim F=n-k-t+1}\max_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax\right]+\left[\min_{F;\dim F=n-k-t+1}\max_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Bx\right]\\
 & \overset{\star\star}{\le}\left[\min_{F;\dim F=n-k}\max_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax\right]+\left[\min_{F;\dim F=n-t}\max_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Bx\right]\\
 & =\lambda_{k}\left(A\right)+\lambda_{t}\left(B\right)
\end{aligned}$$

נשים לב ש־ $\star$ נכון כי $\max a+b\le\max a+\max b$,

נשים לב ש־ $\star\star$ נכון כי $k,t\ge1$ ולכן $k-1,t-1\ge0$ ולכן אם נסמן ב־ $F=\text{argmin}_{F;\dim F=n-k}\max_{\|x\|=1\wedge x\in F^{\perp}}x^{T}Ax$,

נוכל להגדיר $F' \subseteq F$ תת מרחב כך ש־ $\dim F' = n-k-t+1$ ולקבל כי

$$\min_{F;\dim F=n-k-t+1}\max_{\|x\|=1\wedge x\in F^{\perp}} x^TAx \le \max_{\|x\|=1\wedge x\in F'^{\perp}} x^TAx \le \max_{\|x\|=1\wedge x\in F^{\perp}} x^TAx$$
$$= \min_{F;\dim F=k-1}\max_{\|x\|=1\wedge x\in F^{\perp}} x^TAx$$

באופן דומה ל־ $B,t$.
והראנו את הנדרש.

מ.ש.ל.א.☺

(ב) **צ"ל:** $\lambda_k(A)+\lambda_n(B)\le\lambda_k(A+B)\le\lambda_k(A)+\lambda_1(B)$
**הוכחה:**
נשתמש בסעיף הקודם ונקבל כי

$$\lambda_k(A+B)=\lambda_{k+1-1}(A+B)\le\lambda_k(A)+\lambda_1(B)$$

וגם

$$\lambda_k(A+B)=\lambda_{k+n-n}(A+B)\ge\lambda_k(A)+\lambda_n(B)$$

כנדרש.

מ.ש.ל.ב.☺

(ג) **צ"ל:** $\lambda_{k+r}(A)\le\lambda_k(A+B)\le\lambda_{k-r}(A)$
**הוכחה:**
נשתמש בסעיף הקודם ונקבל כי

$$\lambda_k(A+B)=\lambda_{(k-r)+(r+1)-1}(A+B)\le\lambda_{k-r}(A)+\lambda_{r+1}(B)=\lambda_{k-r}(A)+0=\lambda_{k-r}(A)$$

וגם

$$\lambda_k(A+B)=\lambda_{(k+r)+(n-r)-n}(A+B)\le\lambda_{k=r}(A)+\lambda_{n-r}(B)=\lambda_{k=r}(A)+0=\lambda_{k+r}(A)$$

כנדרש.

מ.ש.ל.ג.☺

2. **פתרון:**

(א) **צ"ל:** $\mathbf{Rank}\left(\tilde{A}\right)=2\mathbf{Rank}(A)$
**הוכחה:**
תחילה נראה כי $\text{Rank}\left(\tilde{A}\right)\le 2\text{Rank}(A)$,
נניח בשלילה ש־ $\text{Rank}\left(\tilde{A}\right)>2\text{Rank}(A)$, אז קיימות שורות $i_1,\dots,i_{\text{Rank}(\tilde{A})}$ שהן בלתי תלויות.
נסמן $C=\left\{i_1,\dots,i_{\text{Rank}(\tilde{A})}\right\}$, נשים לב כי $|C|>2\text{Rank}(A)$ וגם כל אחת מהשורות בטווח $1,\dots,2\text{Rank}(A)$,
לכן ממשפט שובך היונים, לפחות חצי מאיברי $C$ קטנים שווים ל־ $\text{Rank}(A)$ או שלפחות חצי מאיברי $C$ גדולות מ־ $\text{Rank}(A)$.

i. אם לפחות חצי מאיברי $C$ קטנים שווים ל־ $\text{Rank}(A)$, נסמנם ב־ $j_1,\dots,j_k$ כאשר $k\ge\frac{|C|}{2}>\text{Rank}(A)$,
נשים לב שכל השורות האלה הן שורות של המטריצה $A$ המקורית ולכן גם ב־ $A$ השורות $j_1,\dots,j_k$ הן בלתי תלויות
ולכן $\text{Rank}(A)\ge k$ בסתירה לכך ש־ $k>\text{Rank}(A)$.

ii. אחרת לפחות חצי מאיברי $C$ גדולים מ־ $\text{Rank}(A)$, נסמנם ב־ $j_1,\dots,j_k$ כאשר $k\ge\frac{|C|}{2}>\text{Rank}(A)$,
נשים לב שכל השורות האלה הן שורות של המטריצה $A^T$ המקורית ולכן גם ב־ $A^T$ השורות $j_1-n,\dots,j_k-n$
הן בלתי תלויות ולכן $\text{Rank}(A)=\text{Rank}\left(A^T\right)\ge k$ בסתירה לכך ש־ $k>\text{Rank}(A)$.

כלומר הראנו כי $\boxed{\text{Rank}\left(\tilde{A}\right) \leq 2\text{Rank}\left(A\right)}$,

עתה נסמן ב־ $i_1, \ldots, i_{\text{Rank}(A)}$ את אינדקסי העמודות בלתי תלויות ב־ $A$ וב־ $j_1, \ldots, j_{\text{Rank}(A)}$ את אינדקסי עמודות בלתי תלויות של $A^T$.

נסמן ב־ $A_l$ את העמודה ה־ $l$ של $A$ ובאופן דומה ב־ $A_l^T$ את העמודה ה־ $l$ של $A^T$ ובאופן דומה את העמודה ה־ $l$ של $\tilde{A}$ בתור $\tilde{A}_l$.

נראה כי העמודות ה־ $i_1, \ldots, i_{\text{Rank}(A)}, j_1, \ldots, j_{\text{Rank}(A)}$ של $\tilde{A}$ בתל.

יהיו $a_1, \ldots, a_{\text{Rank}(A)}, b_1, \ldots, b_{\text{Rank}(A)} \in \mathbb{R}$ כך ש־ $0 = \sum_{l=1}^{\text{Rank}(A)} a_l \cdot \tilde{A}_{i_l} + \sum_{l=1}^{\text{Rank}(A)} b_l \cdot \tilde{A}_{j_l}$,

נשים לב כי

$$0 = \sum_{l=1}^{\text{Rank}(A)} a_l \cdot \tilde{A}_{i_l} + \sum_{l=1}^{\text{Rank}(A)} b_l \cdot \tilde{A}_{j_l} = \sum_{l=1}^{\text{Rank}(A)} a_l \cdot \left[\tilde{A} \cdot e_{i_l+n}\right] + \sum_{l=1}^{\text{Rank}(A)} b_l \cdot \left[\tilde{A} \cdot e_{j_l}\right]$$
$$\sum_{l=1}^{\text{Rank}(A)} a_l \cdot \begin{bmatrix} A_{i_l} \\ 0 \end{bmatrix} + \sum_{l=1}^{\text{Rank}(A)} b_l \cdot \begin{bmatrix} 0 \\ A_{j_l}^T \end{bmatrix} = \begin{bmatrix} \sum_{l=1}^{\text{Rank}(A)} a_l \cdot A_{i_l} \\ \sum_{l=1}^{\text{Rank}(A)} b_l \cdot A_{j_l}^T \end{bmatrix}$$

נשים לב ש־ $A_{i_1}, \ldots, A_{i_{\text{Rank}(A)}}$ בלתי תלויות ולכן $a_1 = \cdots = a_{\text{Rank}(A)} = 0$,

נשים לב ש־ $A_{j_1}^T, \ldots, A_{j_{\text{Rank}(A)}}^T$ בלתי תלויות ולכן $b_1 = \cdots = b_{\text{Rank}(A)} = 0$,

כלומר כל המקדמים הם 0 ולכן העמודות בלתי תלויות ולכן

$$\text{Rank}\left(\tilde{A}\right) \geq \left|\left\{i_1, \ldots, i_{\text{Rank}(A)}, j_1, \ldots, j_{\text{Rank}(A)}\right\}\right| = 2\text{Rank}\left(A\right)$$

משילוב 2 האי שוויונות שהוכחנו, נקבל כי $\boxed{\text{Rank}\left(\tilde{A}\right) = 2\text{Rank}\left(A\right)}$, כנדרש

מ.ש.ל.א.☺

(ב) **צ"ל:** הערכים העצמיים שאינם 0 של $\tilde{A}$ הם בדיוק $\pm\sigma_1(A), \ldots, \pm\sigma_r(A)$

**הוכחה:**

נסמן ב־ $v_1, \ldots, v_r$ את הוקטורים העצמיים המתאימים לערכים העצמיים שאינם 0 במטריצה $A \cdot A^T$,

נשים לב שהערכים העצמיים שאינם 0 במטריצה $A \cdot A^T$ לפי משפט ה־ $SVD$ הם בדיוק $\sigma_1^2(A), \ldots, \ldots \sigma_r^2(A)$.

בלי הגבלת הכלליות נגדיר את $v_i$ להיות הוקטור העצמי המתאים ל־ $\sigma_i^2(A)$.

יהי $1 \leq i \leq r$, נשים לב כי

$$\tilde{A} \cdot \begin{bmatrix} v_i \\ \frac{1}{\sigma_i} A^T v_i \end{bmatrix} = \begin{bmatrix} A \cdot \left(\frac{1}{\sigma_i} A^T v_i\right) \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_i} A \cdot A^T v_i \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_i}\left(A \cdot A^T\right) v_i \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_i} \sigma_i^2 \cdot v_i \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} \sigma_i \cdot v_i \\ A^T \cdot v_i \end{bmatrix}$$
$$= \sigma_i \cdot \begin{bmatrix} v_i \\ \frac{1}{\sigma_i} A^T v_i \end{bmatrix}$$

כלומר $\sigma_i$ הוא ערך עצמי של $A$.

נשים לב כי

$$\tilde{A} \cdot \begin{bmatrix} v_i \\ \frac{-1}{\sigma_i} A^T v_i \end{bmatrix} = \begin{bmatrix} A \cdot \left(\frac{-1}{\sigma_i} A^T v_i\right) \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sigma_i} A \cdot A^T v_i \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sigma_i}\left(A \cdot A^T\right) v_i \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sigma_i} \sigma_i^2 \cdot v_i \\ A^T \cdot v_i \end{bmatrix} = \begin{bmatrix} -\sigma_i \cdot v_i \\ A^T \cdot v_i \end{bmatrix}$$
$$= -\sigma_i \cdot \begin{bmatrix} v_i \\ \frac{-1}{\sigma_i} A^T v_i \end{bmatrix}$$

כלומר $-\sigma_i$ הוא ערך עצמי של $A$.

כלומר הראנו של־ $\tilde{A}$ יש ערכים עצמיים $\pm\sigma_1(A), \ldots, \pm\sigma_r(A)$,

עתה מהסעיף הקודם ראינו ש־ $\text{Rank}(A) = 2r$ ולכן יש $n + m - 2r$ וקטורים עצמיים שמתאימים לערך העצמי 0.

כלומר מצאנו בסיס של וקטורים עצמיים $2r + n + m - 2r = n + m$ ולכן הערכים העצמיים הם בדיוק $\pm\sigma_1(A), \ldots, \pm\sigma_r(A), 0^{n+m-2r}$, כנדרש.

מ.ש.ל.ב.☺

3. **צ"ל:** לא בהכרח $\|A\|_{2,2}=\max\left\{\lambda_1,|\lambda_n|\right\}$

**הוכחה:**

נסתכל על $A=\begin{bmatrix}0&3\\2&0\end{bmatrix}$, נשים לב שהפולינום האופייני הוא

$$\lambda\cdot\lambda-(-2)\cdot(-3)=\lambda^2-6=\left(\lambda-\sqrt{6}\right)\cdot\left(\lambda+\sqrt{6}\right)$$

כלומר $A$ לכסינה (כי לכל ערך עצמי יש לפחות וקטור עצמי 1 וזה יוצר בסיס של וקטורים עצמיים)
ונשים לב כי

$$\|A\|_{2,2}=\max_{\|x\|_2=1}\|Ax\|_2\geq\|A\cdot e_2\|_2=\left\|\begin{bmatrix}3\\0\end{bmatrix}\right\|_2=3>\sqrt{6}=\max\left\{\lambda_1,|\lambda_2|\right\}$$

כנדרש בשאלה.

מ.ש.ל.☺

4. **צ"ל:** $\mathrm{Trace}\,(BABA)\leq\mathrm{Trace}\left(B^2A^2\right)$

**הוכחה:**

נשים לב כי

$$\begin{aligned}
|\mathrm{Trace}\,(BABA)|&=\left|\mathrm{Trace}\left(BA\,(AB)^T\right)\right|=\langle BA,AB\rangle^2\\
&\overset{\text{Cauchy-Schwarz inequality}}{\leq}\langle BA,BA\rangle\cdot\langle AB,AB\rangle\\
&=\sqrt{\mathrm{Trace}\left(BA\cdot(BA)^T\right)\cdot\mathrm{Trace}\left(AB\cdot(AB)^T\right)}\\
&=\sqrt{\mathrm{Trace}\,(BA\cdot A^TB^T)\cdot\mathrm{Trace}\,(ABB^TA^T)}\\
&\overset{A,B\text{ symmetric}}{=}\sqrt{\mathrm{Trace}\,(BA^2B)\cdot\mathrm{Trace}\,(AB^2A)}\\
&\overset{\mathrm{Trace}(CD)\cdot\mathrm{Trace}(DC)}{=}\sqrt{\mathrm{Trace}\,(A^2B\cdot B)\cdot\mathrm{Trace}\,(B^2A\cdot A)}\\
&=\mathrm{Trace}\left(A^2B^2\right)\cdot\mathrm{Trace}\left(B^2A^2\right)\\
&\overset{\mathrm{Trace}(CD)\cdot\mathrm{Trace}(DC)}{=}\sqrt{\mathrm{Trace}\,(B^2A^2)\cdot\mathrm{Trace}\,(B^2A^2)}=\left|\mathrm{Trace}\left(B^2A^2\right)\right|
\end{aligned}$$

ועתה נשים לב כי

$$\begin{aligned}
\mathrm{Trace}\left(B^2A^2\right)&=\mathrm{Trace}\,(B\,(BAA))=\mathrm{Trace}\,(BAAB)=\mathrm{Trace}\left(BAA^TB^T\right)\\
&=\mathrm{Trace}\left(BA\,(BA)^T\right)
\end{aligned}$$

ונשים לב כי $BA\,(BA)^T$ זאת מטריצה $PSD$ כי היא מכפלה של מטריצה בטרנפוז שלה ולכן

$$\mathrm{Trace}\left(B^2A^2\right)=\mathrm{Trace}\left(BA\,(BA)^T\right)\geq0$$

ולכן נקבל כי

$$\mathrm{Trace}\,(BABA)\leq|\mathrm{Trace}\,(BABA)|\leq\left|\mathrm{Trace}\left(B^2A^2\right)\right|=\mathrm{Trace}\left(B^2A^2\right)$$

כלומר הראנו כי $\boxed{\mathrm{Trace}\,(BABA)\leq\mathrm{Trace}\left(B^2A^2\right)}$, כנדרש

מ.ש.ל.☺