מבוא למערכות לומדות - תרגיל 2 חלק תיאורטי

רועי צרפתי 208648154

1. **טענה:** $\ker(X)=\ker(X^TX)$

הוכחה: נוכיח הכלה דו-כיוונית.

$\subseteq$: נניח כי $v\in\ker(X)$ אזי $Xv=0$

נציב ונקבל כי

$$(X^TX)v = X^T(Xv) = X^T\cdot 0 = 0$$

$\Leftarrow$ $v\in\ker(X^TX)$ ולכן $\ker(X)\subseteq\ker(X^TX)$

$\supseteq$: נניח כעת כי $v\in\ker(X^TX)$, צ"ל כי

$$(X^TX)\cdot v=0 \Rightarrow X^T(Xv)=0$$

הנחנו כי $X\neq 0$ ולכן $X^T\neq 0$ $\Leftarrow$ $Xv=0$ וזה השוויון

הדרוש, קרי $v\in\ker(X)$ $\Leftarrow$ $\ker(X^TX)\subseteq\ker(X)$

$$\Downarrow$$

$$\ker(X)=\ker(X^TX)$$

מש"ל! $\blacksquare$

2. תהי A מט' ריבועית

טענה: $Im(A^T)=\ker(A)^\perp$

הוכחה: יהי $x\in Im(A^T)$ ו-$y\in\ker(A)$ עלינו להראות

$$\langle x,y\rangle=0$$

ובכך נוכיח הכלה דו-כיוונית:

$\subseteq$: ראשית נראה, כי ניתן לכתוב $x=A^Tu$ עבור וקטור

$$x^Ty = x^TA^Tu = (Ax)^Tu$$

אך כמובן $x\in\ker(A)$ ולכן $Ax=0$

$$\Leftarrow \quad (Ax)^T u = 0$$

צ"ל כי קיבלנו ש- $\mathrm{Im}\,A^T \subseteq \ker(A)^\perp$

<u>ב</u>: כעת נבחר $z\in \mathrm{Im}(A^T)^\perp$ ונ"ל כי

$$z^T y = 0$$

אך

$$z^T(A^T u) = (Az)u = 0 \quad :$$

מכך קיבלנו כי $z\in\ker A$ $\Leftarrow$

$\Leftarrow$ $\mathrm{Im}(A^T)^\perp \subseteq \ker(A)$

$\Leftarrow$ $\mathrm{Im}(A^T) \subseteq \ker(A)^\perp$ מעבר זה הוכחנו בלינארית 2!

$\Downarrow$

קיבלנו כי $\mathrm{Im}(A^T) = \ker(A)^\perp$ <u>כנדרש</u>!

3.) יהי, $y = Xu$ תהי מערכת משוואות של הנורמליות.
ו- X מט׳ המקדמים של הבעיה.

<u>3</u>: צ"ל: למערכת יש פתרון אמ"מ $y\perp\ker(X^T)$

<u>תזכורת</u>:

נזכור שכל מכפלה ש- A של הבעיה צ׳ הדרגה של
מטריצה ולכן יתכנו 2 מצבים או יש ∞ פתרונות
או 0.

אין מספיק משוואות
או מספיק ודאי תהיה.

$\Leftarrow$ נקבל את שרשרת <u>הבאה</u>: סעיף ב׳

$y\in\ker(X^T)^\perp \iff y\in\mathrm{Im}(X) \iff$ יש ∞ פתרונות

$\iff y \perp \ker(X^T)$

$\boxtimes$

4) מצד שני, אילו X^TX הפיכה אז כידוע מקורסי אלגברה לינארית כי למערכת קיים פתרון יחיד (הרי היא הפיכה).

אחרת X^TX אינה הפיכה, צ"ל שלמערכת יש לפחות ושכן קיים יותר מפתרון אחד למערכת.

נזכר כי $\ker(X^TX) = \ker(X)$

אכן כמובן $v \in \ker(X^TX) \iff v \in \ker(X)$

ונקבל כי $v \perp \ker(X)^{\perp}$

צ"ל מכיוון שקיים $x \neq 0$ בגרעין כלומר כיוון ש- $\ker(X)^{\perp} \perp x$ אך הוכחנו כי הפתרון ה"ע גורר כי למערכת יש אינסוף פתרונות (3), כנדרש!

ב.) 1) לפי הגדרה של ρ מוגדרת ע"י מכפלה חיצונית
ולכן בהינתן בסיס של $(v_1,\dots,v_k)$ הטלה כזו ρ מוגדרת
כסכום φ של כאשר σ של היא $v_i \otimes v_i$ $1\le i\le k$.

מכך נשים לב כי
$\rho_{i,j} = \sum_{t=1}^{k} v_{i,t} \cdot v_{j,t}$ כאשר $v_{i,t}$ הינה
הקואורדינטה ה-i בוקטור v_t $1\le t\le k$, $1\le i,j\le d$.
$\Leftarrow$ נשים לב כי $\rho_{i,j} = \rho_{j,i}$, דבר נובע מהסימטריה
φ בפעולת הכפל.
קרי, ρ סימטרית כנדרש!

ד) נוכיח כי הע"ע של ρ הינם $1,0$ ל-d בבסיס המתאים ש
שנרצה אותו למלא. נרחיב את הבסיס $B=(v_1,\dots,v_k)$
לבסיס לגמרי ע"י תהליך גרם-שמידט כך שנקבל
$B'=(v_1,\dots,v_k,u_{k+1},\dots,u_d)$ ונראה מהו הע"ע של ρ לכל וקטור
בבסיס.

והי $V=\mathrm{span}(v_1,\dots,v_k)$, $W=\mathrm{span}(u_{k+1},\dots,u_d)$

נשים לב כי לכל $v\in\mathbb{R}^d$ ניתן לרשום אותו כסכום
φ $u\in V$ ו-$w\in W$ וקיים σ

$$\rho(u+w)=\rho u+\rho w$$

$\Leftarrow$

$$\rho u = \sum_{i=1}^{k} v_i v_i^T \cdot u \Rightarrow u=\alpha_1 v_1+\dots+\alpha_k v_k$$

$$\Rightarrow \sum_{i=1}^{k} v_i \sum_{j=1}^{k} v_i^T \cdot \alpha_j v_j = \sum_{i=1}^{k} v_i \sum_{j=1}^{k} \alpha_j \delta_{ij} =$$

$$= \sum_{i=1}^{k} \alpha_i v_i = u$$

⇐ כי לכל $u \in V$ מתקיים כי $Pu = u$, וקרי $\underline{\lambda = 1}$

ועבור $u \in W$

$$Pu = \sum_{i=1}^{k} v_i v_i^T \sum_{j=k+1}^{d} \beta_j u_j = 0 \Rightarrow \lambda = 0$$

כל הוקטורים הם ניצבים זה לזה.

⇐ מכיוון ש- $\mathbb{R}^d = V \oplus W$ אזי אלו הערכים היחידים ונקבל כי ע"ע של P הם 0,1 כנדרש!

III. ע"מ להוכיח כי $\lambda = 0,1$ הראנו כי לכל $v \in V$ $Pv = 1v$ וקרי $Pv = v$ (ויזוהו בהגדרות ססים ע"ע).

IV.

$$P^2 = P \cdot P = \sum_{i=1}^{k} v_i v_i^T \cdot \sum_{j=1}^{k} v_j v_j^T =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{k} v_i \underline{v_i^T v_j} v_j^T = \sum_{i=1}^{k} \sum_{j=1}^{k} v_i \delta_{ij} v_j^T =$$

ע"פ

בהגדרת ע"פ של ססס אז.

בזכות 6 הכפולות מתאפס.

$$= \sum_{i=1}^{k} v_i v_i^T = P$$

V. ושים לב:

$$(I-P)P = IP - P \cdot P =$$

$$= P - P^2$$

אך כפי ב-IV הוכחנו כי $P = P^2$ ולכן $P - P^2 = 0$

וקרי $(I-P)P = 0$ כנדרש! ∎

⇐) נניח כי X^TX היא הפיכה. נ"ע קיימת ההופכית $[X^TX]^{-1}$ בתרגיל ראינו כי

$$w^* = (X^TX)^{-1}X^Ty$$

⇐ כלומר, פתרונות שגם מגיעים לאותה תשובה, נראה כי $(X^TX)^{-1}X^T = X^+$ ⇐ מה שיוכיח כי הם אותו פתרון.

נזכר כי:

$$X = U\Sigma V^T$$
$$X^+ = V\Sigma^+U^T$$

כמו כן, מכיוון ש- X^TX הפיכה אז אנו יודעים כי ערכי $\sigma_i \neq 0$, ולכן $\Sigma^+ = \Sigma^{-1}$ שכבר יודעים לנו כעת הפיכה.

⇐

$$X^TX = (U\Sigma V^T)^T(U\Sigma V^T) =$$
$$= V\Sigma^TU^TU\Sigma V^T = V\Sigma^2V^T$$

⇐ זהו פירוק SVD גם כן ולכן לפי מה שראינו, ההופכי פירוק זה

$$(X^TX)^{-1} = V(\Sigma^2)^{-1}V^T$$

⇐

$$(X^TX)^{-1}X^T = V(\Sigma^2)^{-1}V^TV\Sigma U^T =$$
$$= V(\Sigma^2)^{-1}\Sigma U^T$$

ונשים לב כי

$$(\Sigma^2)^{-1} = \begin{pmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_d^2 \end{pmatrix}$$

ומכיוון ש- $(\Sigma^2)^{-1}$ ו- Σ אלכסוניות מתקיים כי

$$(\Sigma^2)^{-1}\cdot\Sigma = \begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_d \end{pmatrix} = \Sigma^{-1} = \Sigma^+$$

$\Leftarrow$

$$(X^TX)^{-1}X^T = V\Sigma^+U^T = X^+$$

לפי <u>הגדרה</u>.

קרי קיבלנו כי גם במקרה זה הפתרון שהצענו הינו X^+, <u>כנדרש</u>!

$\blacksquare$

7.) נראה כי X^TX הפיכה אמ"מ $span(z_1,\dots,z_m)=\mathbb{R}^d$.

בתרגיל 4 ראינו כי דרגת המט' x היא דרגת תת-המרחב הנפרש ע"י הדגימות, קרי $span\{z_1,\dots,z_m\}$, ועכשיו גם ל-X^TX מכיוון שראינו שדרגה שווה לדרגה של x (והכל בתרגיל 4, ראינו כי הדרגות שוות ולכן מספר העמודות נק' כאן).

כן גם לא וזכרו דרגתה של את הן הפיכה.

?

שוויון ו הוסבר

$$\text{הפיכה } X^TX \iff Rank(X^TX)=d \overset{!}{=} Rank(X) \iff$$

מה דרך בה הגדרנו את דרגת הט'!

$$\iff span(z_1,\dots,z_m)=\mathbb{R}^d$$

ז"א כי קיבלנו ש-

$$\text{הפיכה } X^TX \iff span(z_1,\dots,z_m)=\mathbb{R}^d$$

<u>כנדרש</u>!

$\blacksquare$

ה.) עלינו להראות כעת כי בכל מקרה הפתרון שמצאנו $w^1 = X^+y$ הינו המינימלי, קרי לכל פתרון $\bar{w}$ מתקיים $\|\bar{w}\| \geq \|w^1\|$.

<u>הוכחה</u>: נזכר כי כבר מצאנו את w^1 בתרגיל, ראינו כי כאשר פתרנו את המשוואות הנורמליות בתבנית compact SVD אז קיבלנו תחת הפירוק כי

$$\Sigma = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}, \quad S = diag(\sigma_1, \dots, \sigma_\rho)$$

נ"ז כי וקטור הפתרון w^1 בקורדינטות $1, \dots, \rho$ וקבוע באופן יחודי לפי הערכים הסינגולריים של X, בעוד כשאר ושאר מהפירוק שלה גילינו את w^1 ככלל 0 X^+y. ⇐ נזכיר גם ש-$\bar{w}$ פתרון ל- $X^TXw = X^Ty$, לקיים את הנדרש מכיוון ש- $\bar{w} = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^Ty$.

מכך גם $\bar{w}$ פתרון אנו יודעים כי $w^1_i = \bar{w}_i$ כאשר $1 \leq i \leq \rho$ מהפירוק שראינו בתרגיל. ויתר ה $\rho+1, \dots, d+1$ ערכים חופשיים.

⇐ $$\|\bar{w}\|^2 = \sum_{i=1}^{d+1} \bar{w}_i^2 =$$

$$= \sum_{i=1}^{r} \bar{w}_i^2 + \sum_{i=\rho+1}^{d+1} \bar{w}_i^2 \geq \sum_{i=1}^{\rho} \bar{w}_i^2 + 0 = \sum_{i=1}^{\rho} {w^1_i}^2 =$$

הורדנו כמות מחוברים אי-שליליים!

$$= \|w^1\|$$

⇐ $\|\bar{w}\| \geq \|w^1\|$ <u>כנדרש</u>!

∎