318442241 יאיר שטכן

IML חלק תאורטי:

שאלה 1:

כיוון ראשון: יהי $v \in \ker(X)$ $\Rightarrow$ $Xv = 0$ $\Rightarrow$ $X^TXv = 0$

$\Rightarrow$ $v \in \ker(X^TX)$

כיוון שני: יהי $v \in \ker(X^TX)$ $\Rightarrow$ $X^TXv = 0$ $\Rightarrow$ $v^TX^TXv = 0$

$\Rightarrow$ $(v^TX^T)Xv = 0 = (Xv)^TX = 0$ $\Rightarrow$ $\langle Xv, Xv \rangle = 0$

$\Rightarrow$ $Xv = 0$ $\Rightarrow$ $v \in \ker(X)$.

שאלה 2:

$$v \in \ker(A) \iff Av = 0 \iff \langle Av, w \rangle = 0 \quad \text{לכל } w \in V$$

$$\iff \langle v, A^Tw \rangle = 0 \quad w \in V \iff v \perp A^Tw$$

$$\iff v \in \operatorname{Im}(A^T)^\perp$$

ולפי המשלימים האורת' נקבל את הנדרש.

שאלה 3:

למערכת קיים פתרון יחיד אמ"מ היא הפיכה

אחרת יש אינסוף פתרונות או שלא קיים פתרון

לכן, עבור מערכת $Xw = y$ יש אינסוף פתרונות אמ"מ קיים

$u \in \mathbb{R}^n$ כך ש- $Xu = y$ $\iff$ $y \in Im(X)$

משאלה קודמת, $y \in Im(X^T)^\perp$ אמ"מ $y \in Im(X)$

לכן $y \in Im(X)$ אמ"מ $y \perp ker(X^T)$

שאלה 4:

נניח ש- X^TX הפיכה ← קיימת $(X^TX)^{-1}$ ההופכית. נקבל:

$$(X^TX)^{-1} \cdot X^TXw = (X^TX^{-1}) \cdot X^Ty \Rightarrow w = (X^TX)^{-1}X^Ty$$

צ"ל קיים w כך שלמערכת יש פתרון יחיד.

כעת נניח ש- X^TX לא הפיכה.

משאלה קודמת למערכת $X^TXw = X^Ty$ יש אינסוף פתרונות

אמ"מ $X^Ty \perp ker(X^TX) \underset{\text{שאלה 1}}{\iff} X^Ty \perp ker(X)$.

לכן, עבור $w \in ker(X)$ נקבל: $\langle X^Ty | w \rangle = y^T\underbrace{Xw}_{\in ker(X)} = 0$

← קיבלנו ש- $X^Ty \perp ker(X)$.

שאלה 5:

א) לכל $i \in [n]$ $v \in \mathbb{R}^d$ בסיס אורתוגונלי.

$$P=\sum_{i=1}^{n} v_i \otimes v_i^T = \begin{pmatrix} v_1^1 v_1^1 & \cdots & v_1^1 v_d^1 \\ \vdots & & \vdots \\ v_d^1 v_1^1 & \cdots & v_d^1 v_d^1 \end{pmatrix} + \cdots +$$

$$\cdots + \begin{pmatrix} v_1^n v_1^n & \cdots & v_1^n v_d^n \\ \vdots & & \vdots \\ v_d^n v_1^n & \cdots & v_d^n v_d^n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n v_1^i v_1^i & \cdots & \sum_{i=1}^n v_1^i v_d^i \\ \vdots & & \vdots \\ \sum_{i=1}^n v_d^i v_1^i & \cdots & \sum_{i=1}^n v_d^i v_d^i \end{pmatrix}$$

כל אחד מהמחוברים זו מטריצה סימטרית ולכן גם החיבור שלהם.

ב)

עבור $i\in[k]$ נראה ש- $Pv_i=v_i$

$$Pv_i=\sum_{j=1}^{n}(v_j v_j^T)v_i = \sum_{j=1}^{n} v_j \langle v_j | v_i\rangle = v_i \langle v_i | v_i \rangle$$

$$= v_i \|v_i\|^2 \underset{\|v_i\|=1}{=} v_i$$

שונה מאפס רק כאשר $j=i$ כי הם אורתונורמליים.

עבור $i\in\{k+1,\dots,d\}$ נקבל מכך שהם אורת' ש: $Pv_i=0$

סה"כ נקבל שקיים בסיס $v_1,\dots,v_d$ בסיס מו"ע עם ע"ע $0,1$.

ג)

יהי $v\in V$, v הוא צ"ל של הבסיס האורתונורמלי $v_1,\dots,v_k$

כך ש $v=\sum_{i=1}^{k}\alpha_i v_i$

מסעיף קודם הם ו"ע נקבל $Pv=P(\sum_{i=1}^{k}\alpha_i v_i)=\sum_{i=1}^{k}\alpha_i \underbrace{Pv_i}_{v_i}=v$

ג)

נשתמש ב- EVD, צ"ל קיימת מטריצה U אורתוגונלית

כשהעמודות שלה וע"ע של P, D אלכסונית כאשר על האלכסון הע"ע של P

מתקיים $P = UDU^T$

$$P^2 = (UDU^T)^2 = UD\underbrace{U^TU}_{I}DU^T = UDU^T = P$$

ד)

$$(I-P)P = IP - P^2 = P - P^2 = 0$$

סעיף קודם $P = P^2$

שאלה 6:

נרצה להראות ש- $(X^TX)^{-1}X^T = X^{\perp}$

נשתמש ב- SVD על X ונקבל:

$$X^TX = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T\underbrace{U^TU}_{I}\Sigma V^T = V\Sigma^T\Sigma V^T$$

כעת:

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

וסה"כ:

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

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

שאלה 7:

X^TX הפיכה אמ"מ $\ker(X^TX)=0$ אמ"מ (שאלה 1) $\ker(X)=0$

וזה אמ"מ $\dim(\mathrm{Im}(X))=d$ אמ"מ $\dim(\mathrm{col}(X))=d$

וזה אמ"מ גם $\dim(\mathrm{row}(X))=d$ $\Longleftrightarrow$ $\mathbb{R}^d \ni \mathrm{Span}\{x_1,\dots,x_m\}$

שאלה 8:

נסתכל על פתרון w של המשוואה $\underset{w\in\mathbb{R}^d}{\mathrm{argmin}}\|Xw-y\|^2$

ויהי $X=U\Sigma V^T$ פירוק SVD של X.

$$\|Xw-y\|^2=\|U\Sigma V^Tw-y\|^2=\left\|U\Sigma\begin{bmatrix}V_1^Tw\\V_2^Tw\end{bmatrix}-UU^Ty\right\|^2=$$

$$=\left\|U\left(\Sigma\begin{bmatrix}V_1^Tw\\V_2^Tw\end{bmatrix}-U^Ty\right)\right\|^2=\left\|\begin{pmatrix}\Sigma&0\\0&0\end{pmatrix}\begin{bmatrix}V_1^Tw\\V_2^Tw\end{bmatrix}-\underbrace{[V_1V_2]}_{=U}{}^Ty\right\|^2$$

Σ הערך שאינם אפסים

$$=\left\|\begin{bmatrix}\Sigma_1V_1^Tw\\0\end{bmatrix}-\begin{bmatrix}V_1^Ty\\0\end{bmatrix}-\begin{bmatrix}0\\V_2^Ty\end{bmatrix}\right\|^2=\left\|\begin{bmatrix}\Sigma_1V_1^Tw-V_1^Ty\\0\end{bmatrix}\right\|^2+\left\|\begin{pmatrix}0\\V_2^Ty\end{pmatrix}\right\|^2$$

כעת גם שווה אפס אמ"מ $\Sigma_1V_1^Tw-V_1^Ty=0$

$$\Longleftrightarrow \quad V_1^Tw=(\Sigma_1)^{-1}V_1^Ty$$

ועבור $w=X^+y$ נקבל $V_1^TX^+y=(\Sigma_1)^{-1}V_1^Ty$

$$V_2^TX^+y=V_2^T(V\Sigma^+V^T)y=0$$

לכן נקבל שבפתרון $\tilde{w}$ מתקיים ש- $\|w\| \leq \|\tilde{w}\|$