✓ 1 (א)

$$\left.\begin{array}{l}\forall A\in\mathbb{R}^{m\times n},\ A=[a_{ij}]_{m\times n}\\ \forall B\in\mathbb{R}^{n\times p},\ B=[b_{jk}]_{n\times p}\end{array}\right\} \Rightarrow (AB)_{ik}=\sum_{j=1}^{n}a_{ij}b_{jk} \Rightarrow AB=\left[\sum_{j=1}^{n}a_{ij}b_{jk}\right]_{m\times p} \Rightarrow \left[(AB)^T\right]_{ki}=\left[(AB)\right]_{ik}=\sum_{j=1}^{n}a_{ij}b_{jk}$$

$$\left.\begin{array}{l}(A^T)_{ji}=a_{ij}\\ (B^T)_{kj}=a_{jk}\end{array}\right\} \Rightarrow (B^TA^T)_{ki}=\sum_{j=1}^{n}(B^T)_{kj}(A^T)_{ji}=\sum_{j=1}^{n}b_{jk}a_{ij}=\sum_{j=1}^{n}a_{ij}b_{jk}$$

$\Leftrightarrow$

ובמיוחד עבור $n=m=p$.

✓ 2

Let $C=A+B$ then $C_{ij}=a_{ij}+b_{ij} \Rightarrow (C^T)_{ji}=\left[(A+B)^T\right]_{ji}=(A+B)_{ij}=(C)_{ij}=a_{ij}+b_{ij}$

$$(A^T+B^T)_{ji}=(A^T)_{ji}+(B^T)_{ji}=a_{ij}+b_{ij}$$

$\Leftrightarrow$

× 3

$$\left.\begin{array}{l}A=\begin{pmatrix}1&0\\0&2\end{pmatrix}\Rightarrow tr(A)=3\\ B=\begin{pmatrix}3&0\\0&4\end{pmatrix}\Rightarrow tr(B)=7\end{array}\right\} \Rightarrow tr(A)tr(B)=21\neq tr(AB)=tr\begin{pmatrix}3&0\\0&8\end{pmatrix}=11$$

✓ 4

$$tr(AB)=\sum_{i=1}^{n}(ab)_{ii}=\sum_{i=1}^{n}a_{ii}b_{ii}=\sum_{j=1}^{n}b_{ii}a_{ii}=\sum_{i=1}^{n}(ba)_{ii}=tr(BA)$$

5

$$tr(A+B)=\sum_{i=1}^{n}(a+b)_{ii}=\sum_{i=1}^{n}a_{ii}+b_{ii}=\sum_{i=1}^{n}a_{ii}+\sum_{i=1}^{n}b_{ii}=tr(A)+tr(B)$$

✓ 6

$$\left.\begin{array}{l}(AB)(B^{-1}A^{-1})=ABB^{-1}A^{-1}=AIA^{-1}=AA^{-1}=I\\ (B^{-1}A^{-1})(AB)=B^{-1}A^{-1}AB=B^{-1}IB=B^{-1}B=I\end{array}\right\} (AB)(B^{-1}A^{-1})=(B^{-1}A^{-1})(AB)\Rightarrow (AB)^{-1}=(B^{-1}A^{-1})$$

× 7

$$A=I,\ B=-I \Rightarrow \left\{\begin{array}{l}A+B=[0] \Rightarrow (A+B)^{-1} \text{ Does not Exist}\\ A^{-1}+B^{-1}=I-I=[0]\end{array}\right\} \Rightarrow [0]=\text{undefined}$$

8

$$\forall A\in\mathbb{R}^{m\times n},\ \bar{u},\bar{v}\in \ker(A)\subseteq\mathbb{R}^n,\ c\in\mathbb{R}:\left\{\begin{array}{l}A\bar{0}=\bar{0}\Rightarrow\bar{0}\in\ker(A)\\ A\bar{u}=\bar{0}\text{ and }A\bar{v}=\bar{0}\Rightarrow A(\bar{u}+\bar{v})=A\bar{u}+A\bar{v}=\bar{0}+\bar{0}=\bar{0}\in\ker(A)\\ A\bar{u}=0\Rightarrow A(c\bar{u})=cA\bar{u}=c\bar{0}=\bar{0}\in\ker(A)\end{array}\right.$$

$\Rightarrow \ker(A)\subseteq\mathbb{R}^n$ is a subspace

$$\forall\ \bar{y}_1=A\bar{x}_1,\ \bar{y}_2=A\bar{x}_2\in Im(A),\ c\in\mathbb{R}:\left\{\begin{array}{l}A\bar{0}=\bar{0}\in Im(A)\\ \bar{y}_1+\bar{y}_2=A\bar{x}_1+A\bar{x}_2=A(\bar{x}_1+\bar{x}_2)\Rightarrow\bar{y}_1+\bar{y}_2\in Im(A)\\ c\bar{y}_1=cA\bar{x}_1=A(c\bar{x}_1)\Rightarrow c\bar{y}_1\in Im(A)\end{array}\right.$$

$$\left.\begin{array}{l} \forall \bar{y} \in Im(A^T)\ \exists \bar{z} \in \mathbb{R}^n \text{ s.t. } \bar{y} = A^T\bar{z} \\ \forall \bar{x} \in ker(A),\ A\bar{x} = 0 \end{array}\right\} \bar{y}\cdot\bar{x} = (A^T\bar{z})\bar{x} = \bar{z}\,(A\bar{x}) = \bar{z}\,\bar{0} = \bar{0} \Rightarrow \bar{y} \perp \bar{x} \Rightarrow \bar{y} \in ker(A)^\perp$$

$$\Rightarrow im(A^T) \subseteq ker(A)^\perp$$

$$dim(ker(A)) + dim(Im(A^T)) = dim(ker(A)) + dim(ker(A)^\perp) = p \quad \Rightarrow dim(Im(A^T)) = dim(ker(A)^\perp)$$

$$\left.\begin{array}{l} dim(Im(A^T)) = dim(ker(A)^\perp) \\ im(A^T) \subseteq ker(A)^\perp \end{array}\right\} \Rightarrow im(A^T) = ker(A)^\perp$$

: 2 fica

1c

$$\forall P \in \mathbb{R}^{n\times n},\ P = P^2 = P^T,\ \forall \lambda \in \mathbb{R},\ \bar{v} \neq \bar{0} : P\bar{v} = \lambda\bar{v} \Rightarrow P^2\bar{v} = P\lambda\bar{v} \Rightarrow P\bar{v} = \lambda(P\bar{v}) \Rightarrow P\bar{v} = \lambda(\lambda\bar{v}) \Rightarrow P\bar{v} = \lambda^2\bar{v}$$
$$\Rightarrow \lambda\bar{v} = \lambda^2\bar{v} \Rightarrow \bar{v}(\lambda^2 - \lambda) = 0 \Rightarrow \lambda = 0, 1$$

$\lambda = 1$:

$P\bar{v} = \bar{v} \Rightarrow (P - I)\bar{v} = \bar{0} \Rightarrow \bar{v} \in ker(P - I)$

To Prove: $ker(P-I) = Im(P) \Rightarrow dim(ker(P-I)) = dim(Im(P))$

$\forall \bar{v} \in Im(P)\ \exists \bar{x} \in \mathbb{R}^n$ s.t. $\bar{v} = P\bar{x}$ then: $\quad P\bar{v} = P(P\bar{x}) = P^2\bar{x} = P\bar{x} = \bar{v} \Rightarrow v \in ker(P-I)$

$\forall \bar{v} \in ker(P-I),\ P\bar{v} = \bar{v} \Rightarrow \bar{v} \in Im(P)$

$ker(P-I) = Im(P) \Rightarrow dim(ker(P-I)) = dim(Im(P)) \Rightarrow$ mult of $\lambda = 1$ $= rank(P)$

$\lambda = 0$:

Eigenspace of $\lambda = 0$ $= \{\bar{v} \in \mathbb{R}^n \mid P\bar{v} = 0\} = ker(P) \Rightarrow$ mult of $\lambda = 0$ $= dim(ker(P))$

Rank - Nullity: $rank(P) + dim(ker(P)) = n$ = All mult of all $\lambda = 0, 1$

$$P_x^2 = [X(X^TX)^{-1}X^T][X(X^TX)^{-1}X^T] = X\underbrace{[(X^TX)^{-1}X^TX]}_{=I}(X^TX)^{-1}X^T = X(X^TX)^{-1}X^T = P_x \qquad : 2$$

$$P_x^T = [X(X^TX)^{-1}X^T]^T = (X^T)^T[(X^TX)^{-1}]^T(X)^T = X[\underbrace{(X^TX)}_{X^TX = (X^TX)^T}{}^T]^{-1}X^T = X(X^TX)^{-1}X^T = P_x$$

$$\forall v \in \mathbb{R}^n : P_x\bar{v} = X(X^TX)^{-1}X^T\bar{v} = X[(X^TX)^{-1}X^T\bar{v}] = X\bar{\beta} \quad (\bar{\beta} \in \mathbb{R}^p)$$

$$\Rightarrow P_x\bar{v} = X\bar{\beta} \in Im(X)$$

$$tr(P_x) = tr(X(X^TX)^{-1}X^T) = tr((X^TX)^{-1}X^TX) = tr(I_p) = \sum_{i=1}^{p} 1 = p = rank(X) \qquad 2$$

$AA^T \in \mathbb{R}^{n\times n}$
$A^TA \in \mathbb{R}^{p\times p}$

שאלה 3

ג

$\forall \bar v \in \mathbb{R}^p, v\neq 0, u\in\mathbb{R}^n, u=A\bar v:$
$(A^TA)\bar v = \lambda v \Rightarrow AA^TA\bar v = \lambda A\bar v \Rightarrow AA^T\bar u = \lambda \bar u$

לכן, נקבל את אותו λ, כל עוד $\bar u,\bar v \neq 0$ נקבל הטענה כי ערכים של ע"ע

ב $A^TA\in\mathbb{R}^{p\times p}$ סימטרית, אז יש לה בסיס עצמי אורתונורמלי $A^TA = V\Lambda V^T$ עבור $V\in\mathbb{R}^{p\times p}$ אורתוגונלית

ו- $\Lambda = Diag(\lambda_1,\dots\lambda_r,0,0)$ עבור $Rank(A)=r$ ע"ע של A.

ונגדיר $\Sigma = Diag(\sqrt{\lambda_1},\dots\sqrt{\lambda_r},0,0) \in \mathbb{R}^{n\times p}$

מכיוון שאותם ע"ע ל AA^T ול A^TA, נקבל כמו ע, להגדיר $u\in\mathbb{R}^{n\times n}$ כך שהוא אורתוגונלית ומתקיים

$AA^T=U\Lambda U^T$, ו $r=rank(A)$ עמודות הראשונות פורסות את $Im(A)$

אז $u_i = \frac{AV_i}{\sqrt{\lambda_i}}$ עבור $i=1\dots r$ $\Leftarrow$ $AV_i=\sqrt{\lambda_i}u_i$ $\Leftarrow$ $AV=U\Sigma$ $\Leftarrow$ $A=U\Sigma V^T$ והשלמנו את הפירוק

שאלה 4

א

$$m(a)=E(Y-a)^2 \Rightarrow \frac{d}{da}m(a)=\frac{d}{da}E(Y-a)^2 = E[-2(Y-a)] \Rightarrow E[Y]=\hat a$$

ב

$$MSE=E(Y-f(x))^2 \Rightarrow \frac{\partial}{\partial f(x)}MSE = \frac{\partial}{\partial f(x)}E(Y-f(x))^2 = E\left(\frac{\partial}{\partial f(x)}(Y-f(x))^2\right) = E(-2(Y-f(x))) = 0 \Rightarrow$$

$$E(Y-f(x)) = E(Y-f(x)|x) = E(Y|x)-f(x) \Rightarrow E(Y|x)=f(x)$$