.ד.1.1

א)

$$A=\begin{bmatrix}1 & 1 & 0\\ 1 & -1 & 2\end{bmatrix} \Rightarrow A^TA=\begin{bmatrix}2 & 0 & 2\\ 0 & 2 & -2\\ 2 & -2 & 4\end{bmatrix}$$

נמצא את הערכים העצמיים של A^TA.

$$P_{A^TA}(\lambda)=|A^TA-\lambda I|=\begin{vmatrix}2-\lambda & 0 & 2\\ 0 & 2-\lambda & -2\\ 2 & -2 & 4-\lambda\end{vmatrix}=(2-\lambda)\big((2-\lambda)(4-\lambda)-4\big)$$
$$+2\big(-2(2-\lambda)\big)$$

$$=(2-\lambda)\left(8-6\lambda+\lambda^2-4-4\right)=(2-\lambda)\left(\lambda^2-6\lambda\right)=\lambda(2-\lambda)(\lambda-6)$$

$$\Rightarrow \underline{\lambda_{1,2,3}=0,2,6} \Rightarrow \sigma_{1,2,3}=\underline{0,\sqrt{2},\sqrt{6}}$$

$\underline{V_0}$:
$$A^TA-\lambda I=\begin{bmatrix}2 & 0 & 2\\ 0 & 2 & -2\\ 2 & -2 & 4\end{bmatrix}$$

$$N\left(\begin{bmatrix}2 & 0 & 2\\ 0 & 2 & -2\\ 2 & -2 & 4\end{bmatrix}\right)\Rightarrow\begin{bmatrix}2 & 0 & 2\\ 0 & 2 & -2\\ 2 & -2 & 4\end{bmatrix}\xrightarrow[R_3+R_2]{R_3-R_1}\begin{bmatrix}2 & 0 & 2\\ 0 & 2 & -2\\ 0 & 0 & 0\end{bmatrix}\to\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & -1\\ 0 & 0 & 0\end{bmatrix}$$

$$\Rightarrow N(A^TA)=\operatorname{span}\left\{\begin{bmatrix}-1\\ 1\\ 1\end{bmatrix}\right\}\Rightarrow \underline{V_0=\begin{bmatrix}-1\\ 1\\ 1\end{bmatrix}}$$

$\underline{v_2}$: $A^TA-2I=\begin{bmatrix}0&0&2\\0&0&-2\\2&-2&2\end{bmatrix}\xrightarrow[/2]{R_1\Leftrightarrow R_3}\begin{bmatrix}1&-1&0\\0&0&-1\\0&0&0\end{bmatrix}$

$\Rightarrow N(A^TA-2I)=\mathrm{span}\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}\Rightarrow \underline{v_2=\begin{bmatrix}1\\1\\0\end{bmatrix}}$

$\underline{v_6}$: $A^TA-6I=\begin{bmatrix}-4&0&2\\0&-4&-2\\2&-2&-2\end{bmatrix}\xrightarrow{R_3+\frac{1}{2}R_1}\begin{bmatrix}-4&0&2\\0&-4&-2\\0&-2&-1\end{bmatrix}$

$\xrightarrow[R_3-\frac{1}{2}R_2,\ R_2/-4]{R_1/-4}\begin{bmatrix}1&0&-\frac{1}{2}\\0&1&\frac{1}{2}\\0&0&0\end{bmatrix}\Rightarrow N(A^TA-6I)=\mathrm{span}\left\{\begin{bmatrix}1/2\\-1/2\\1\end{bmatrix}\right\}$

$\Rightarrow \underline{v_6=\begin{bmatrix}1\\-1\\2\end{bmatrix}}$

$\Sigma=\begin{bmatrix}\sqrt{6}&0\\0&\sqrt{2}\end{bmatrix}\qquad V=\begin{bmatrix}1/\sqrt{6}&1/\sqrt{2}\\-1/\sqrt{6}&1/\sqrt{2}\\\sqrt{6}/3&0\end{bmatrix}$

יש כלי מצוין:
(SVD קצוריי
ויקיפדיה)

נמצא את U:

$u_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1&0\\1&-1&2\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{bmatrix}$

$=\begin{bmatrix}1/\sqrt{2}&1/\sqrt{2}&0\\1/\sqrt{2}&-1/\sqrt{2}&\sqrt{2}\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{bmatrix}=\underline{\underline{\begin{bmatrix}1\\0\end{bmatrix}}}$

$$u_6 = \frac{1}{\sqrt6}\begin{bmatrix}1 & 1 & 0\\ 1 & -1 & 2\end{bmatrix}\begin{pmatrix}1/\sqrt6\\ -1/\sqrt6\\ \sqrt6/3\end{pmatrix} = \begin{bmatrix}1/\sqrt6 & 1/\sqrt6 & 0\\ 1/\sqrt6 & -1/\sqrt6 & 2/\sqrt6\end{bmatrix}\begin{pmatrix}1/\sqrt6\\ -1/\sqrt6\\ \frac{\sqrt6}{3}\end{pmatrix}$$

$$= \begin{bmatrix}0\\ \frac16 + \frac16 + \frac23\end{bmatrix} = \begin{bmatrix}0\\ 1\end{bmatrix}$$

$$U = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$ ובסך הכל קיבלנו

ובסך הכל Σ

$$U\Sigma V^T = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{bmatrix}\sqrt6 & 0\\ 0 & \sqrt2\end{bmatrix}\begin{pmatrix}1/\sqrt6 & -\frac{1}{\sqrt6} & \frac{\sqrt6}{3}\\ 1/\sqrt2 & 1/\sqrt2 & 0\end{pmatrix}$$

ובמטריצות נשים לב כי

$$U\Sigma V^T = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & -1 & 2\\ 1 & 1 & 0\end{bmatrix} = \begin{bmatrix}1 & 1 & 0\\ 1 & -1 & 2\end{bmatrix} = A$$

כמצופה.

(*) SVD בצורה הלא קומפקטית יהיה:

$$U\Sigma V^T = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\sqrt6 & 0 & 0\\ 0 & \sqrt2 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt6 & -\frac{1}{\sqrt6} & \frac{\sqrt6}{3}\\ 1/\sqrt2 & 1/\sqrt2 & 0\\ -\frac{1}{\sqrt3} & 1/\sqrt3 & \frac{1}{\sqrt3}\end{bmatrix}$$

ב)

$v\in\mathbb{R}^n,\ u\in\mathbb{R}^m,\ v=\begin{bmatrix}v_1\\ \vdots\\ v_n\end{bmatrix},\ u=\begin{bmatrix}u_1\\ \vdots\\ u_m\end{bmatrix}$

$$v\otimes u = v\cdot u^T=\begin{bmatrix} v_1\cdot u_1 & v_1\cdot u_2 & \cdots & v_1\cdot u_m\\ v_2\cdot u_1 & v_2\cdot u_2 & \cdots & v_2\cdot u_m\\ \vdots & \vdots & \ddots & \vdots\\ v_n\cdot u_1 & v_n\cdot u_2 & \cdots & v_n\cdot u_m\end{bmatrix}$$

$$v\cdot u^T=\begin{bmatrix} — & v_1\cdot u^T & — \\ — & v_2\cdot u^T & — \\ & \vdots & \\ — & v_n\cdot u^T & — \end{bmatrix} \Rightarrow$$

כל השורות הן
מכפלות בסקלר של
u^T, ולכן תלויות לינארית.

$$v\cdot u^T=\begin{bmatrix} | & | & & | \\ v\cdot u_1 & v\cdot u_2 & \cdots & v\cdot u_m \\ | & | & & | \end{bmatrix}$$

באופן דומה
וגם כל העמודות
של $v\cdot u^T$ הן כפל בסקלר
של הוקטור v ולכן תלויות לינארית.

קיבלנו כי כל השורות (וכן העמודות) תלויות
לינארית, ולכן $A=v\cdot u^T$ מקיימת $\underline{rank(A)=1}$

3

$$\langle x,u_j\rangle \overset{(I)}{=} \left\langle \sum_{i=1}^{n} a_i \cdot u_i , u_j \right\rangle \overset{(II)}{=} \sum_{i=1}^{n} a_i \langle u_i , u_j \rangle \overset{(III)}{=} a_j$$

(I) הצגת u_i (II) לינאריות במשתנה הראשון של מכפלה פנימית

(III) אורתונורמליות הבסיס.

4

a

$\|x\|_1 = 3+4+1+2 = \underline{10}$ $\qquad$ $\|x\|_\infty = \max_i |x_i| = \underline{4}$

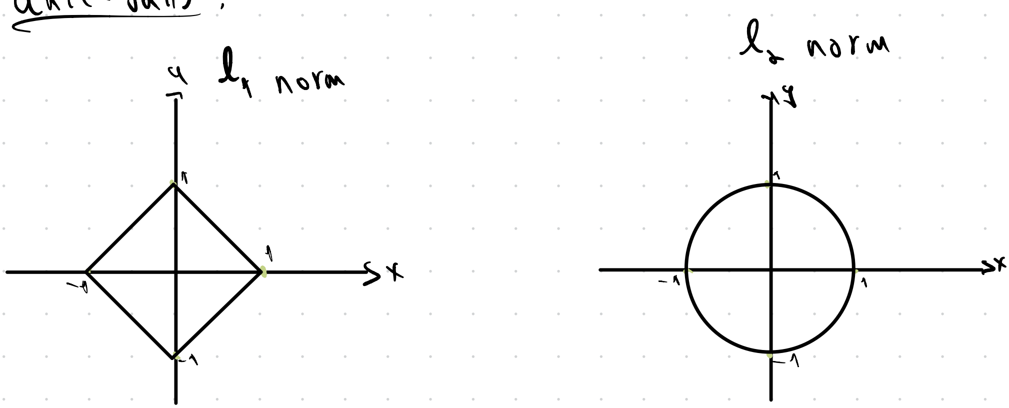
$\|x\|_2 = \sqrt{9+16+1+4} = \underline{\sqrt{30}}$

הערך הקטן ביותר מתקבל ב- $\|x\|_\infty$, שכן הנורמה שלו היא

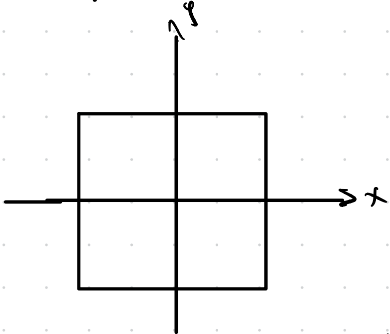
רק ערך יחיד מתוך $\{|x_i|\}_{i=1}^{n}$ לכן קטן מסכום מוחלטי

הערכים שלו ב- $\|x\|_1$, $\|x\|_2$

unit-balls:



l_∞ norm



נורמת l_1 norm משקפת את כך ש- $l_1(v)=|x|+|y|$

ולמעשה "מעניש" וקטורים בעלי ערכים גדולים - צריך ע"מ למזער מכפלה לפי נורמה זו, נרצה שכמה מהמשקלים הגדולים יהיו מתכנסים לערך כלשהו.

נורמת l_2 norm לוקחת בריבוע את הערכים, ולכן לא מעניש כמו מקודם, אלא מעניש את המרחק בצורה "אוקלידית".

נורמת l_∞ norm האינסופית משקפת את כל שהאיבר הגדול ביותר

מבחינת מבחינה זו הוא היחיד המשפיע ביותר מכולם, אם

בעולם שבו כל הצירים ל-ד שונים הערך השני ($\mathbb{R}^2$ -ב)

לא משפיע על הנורמה. ולכן יש שימוש בעיקר במקרים מיוחדים.

b

1.1.2

$$h(\sigma) = \frac{1}{2}\|f(\sigma) - y\|^2 = \underbrace{\frac{1}{2}\|f(\sigma)\|^2 - f(\sigma)^T y + \frac{1}{2}\|y\|^2}_{(*)}$$

$$\boxed{\frac{\partial h}{\partial \sigma_i} = \frac{\partial h}{\partial f} \cdot \frac{\partial f}{\partial \sigma_i}} = \frac{\partial}{\partial f}\left((*)\right) \cdot \frac{\partial f}{\partial \sigma_i} = (f(\sigma) - y)^T \cdot \frac{\partial f}{\partial \sigma_i}$$

$$\Rightarrow \nabla h(\sigma) = \begin{bmatrix} - & \frac{\partial h}{\partial \sigma_1} & - \\ - & \frac{\partial h}{\partial \sigma_2} & - \\ & \vdots & \\ - & \frac{\partial h}{\partial \sigma_d} & - \end{bmatrix} = \boxed{(f(\sigma) - y)^T \cdot J_\sigma(f)}$$

c

$i \neq j$:
$$\frac{\partial S_i}{\partial x_j} = \frac{0 \cdot \sum_{l=1}^{k} e^{x_l} - e^{x_i} e^{x_j}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} = \frac{-e^{x_i}}{\sum_{l=1}^{k} e^{x_l}} \cdot \frac{e^{x_j}}{\sum_{l=1}^{k} e^{x_l}} \leftarrow \text{[illegible]}$$

$$= -S(x)_i \cdot S(x)_j$$

$i = j$:
$$\frac{\partial S_i}{\partial x_i} = \frac{e^{x_i} \sum_{l=1}^{k} e^{x_l} - e^{x_i} e^{x_i}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} = \frac{e^{x_i}}{\sum_{l=1}^{k} e^{x_l}} \cdot \frac{\sum_{l=1}^{k} e^{x_l} - e^{x_i}}{\sum_{l=1}^{k} e^{x_l}}$$

$$= S(x)_i (1 - S(x)_i)$$

$$\left[\nabla S(x)_i\right]_j = S(x)_i \cdot \left(\delta_{ij} - S(x)_j\right)$$ נקבל ש-

וסך הכל נקבל שהיעקוביאן יהיה:

$$\boxed{J_x(S) = \mathrm{diag}(S) - SS^T}$$

1.2

א)

(a) $v \in \ker(X) \Rightarrow Xv = 0 \Rightarrow X^T X v = X^T 0 = \underline{0}$

$$\ker(X) \subseteq \ker(X^T X)$$

$v \in \ker(X^T X) \Rightarrow X^T X v = 0 \Rightarrow v^T X^T X v = v^T 0 \Rightarrow \|Xv\|^2 = 0$

$\Rightarrow \underline{Xv = 0} \Rightarrow$

$$\ker(X^T X) \subseteq \ker(X)$$

$$\underline{\ker(X) = \ker(X^T X)}$$ וסך הכל

(b) $v \in Im(A^T) \Rightarrow \exists_{x \in \dim(A)[0]} x$ s.t. $A^T x = v$.

$Aa = 0 \Rightarrow 0 = (Aa)^T x = a^T A^T x = a^T v \Rightarrow \underline{v \perp a}$ $\Leftarrow a \in \ker(A)$ כי

ונכון ל- $a \in \ker(A)$ אז $v \in \ker(A)^\perp$, וסך הכל $\underline{Im(A^T) \subseteq \ker(A)^\perp}$

$$x \notin Im(A^T) \Rightarrow x \in Im(A^T)^\perp \Rightarrow \exists_{y \in Im(A^T)^\perp} y \text{ s.t. } \underline{\langle x, y\rangle \neq 0}$$

$$\Rightarrow \langle y, A^T A y\rangle = 0 \Rightarrow \langle y, A^T A y\rangle = \langle Ay, Ay\rangle = \underline{\|Ay\|^2 = 0}$$

y אורתוגונלי לכל $v \in Im(A^T)$ ובפרט $AA^Ty \in Im(A^T)$

$$y \in Ker(A) \Leftarrow Ay = 0$$

[illegible] $\Leftarrow$ $x \notin Im(A^T)$ $\Leftarrow$ קיים $y \in Ker(A)$ כך
$\langle x, y\rangle \neq 0$ ש[illegible]
ולכן $\underline{x \notin Ker(A)^\perp}$

וגם $\Leftarrow$

$$x \notin Im(A^T) \Rightarrow x \notin Ker(A)^\perp$$

$$\Downarrow$$

$$x \in Ker(A)^\perp \Rightarrow x \in Im(A^T)$$

$$\Downarrow$$

$$\underline{Ker(A)^\perp \subseteq Im(A^T)}$$

הראינו הכלה דו-כיוונית, ולכן $\underline{\underline{Ker(A)^\perp = Im(A^T)}}$ סיים

(c) מכיוון ש- X לא הפיך, אזי $Xw = y$

ולכן מערכת בעלת ∞ פתרונות או 0 פתרונות.

בנוסף, אם $y \in Im(X)$ אזי יש למערכת לפחות פתרון אחד,

ולכן יש למערכת פתרון אמ"מ לפחות אחד נמצא $y \in Im(X)$

וסה"כ קל $y \in Im(X) \iff$ קיים למערכת פתרון.

נקבל:

- $y \perp ker(X^T) \Leftarrow y \in ker(X^T)^{\perp} \overset{(b)}{\Leftarrow} y \in Im(X) \Leftarrow$ יש ∞ פתרונות
- יש ∞ פתרונות $\Leftarrow y \in Im(X) \overset{(b)}{\Leftarrow} y \in ker(X^T)^{\perp} \Leftarrow y \perp ker(X^T)$

(d)

X לא הפיכה $\Rightarrow$ ∞ פתרונות למערכת $\overset{(c)}{\iff} X^T y \perp ker(X^T X)$

$\overset{(b)}{\iff} X^T y \perp ker(X). \longrightarrow$ ניקח את כ' $v \in ker(X)$ אזי

כי $0 = \langle v, X^T y\rangle = \langle vX, y\rangle = \langle 0, y\rangle$

ולכן $v \perp X^T y$, ומכאן $X^T y \perp ker(X)$

וסה"כ X לא-הפיכה $\Leftarrow$ יש ∞ פתרונות למערכת

במקרה שבו X הפיכה אזי:

$X^T X w = X^T y \iff Xw = y \Rightarrow w = X^{-1} y$

וקיים למערכת פתרון יחיד

9.2.2

אנחנו יודעים כי $X^+=[X^TX]^{-1}X^T$, במקרה בו X^TX הפיך

$$[X^TX]^{-1}X^T=\left[(U\Sigma V^T)^T U\Sigma V^T\right]^{-1}(U\Sigma V^T)^T$$

$$=\left[V\Sigma^T \underbrace{U^T U}_{I}\Sigma V^T\right]^{-1}(V\Sigma^T U^T)=\left[V\Sigma^T\Sigma V^T\right]^{-1}(V\Sigma^T U^T)$$

כי U היא
מטריצת וקטורים
אורתונורמלים

$$=\underbrace{V^{T^{-1}}}\Sigma^{-1}\Sigma^{T^{-1}}\cancel{V^{-1}}\cancel{V}\Sigma^T U^T=V(\Sigma^T\Sigma)^{-1}\Sigma^T U^T$$

V, בהתאם
הפיכה ואורתו
נורמלית

נשים לב כי $\Sigma^T\Sigma$ היא מטריצה 5x6

ומתקיים $\Sigma^T\Sigma_{ii}=(\Sigma_{ii})^2=\sigma_i^2$ כמו כן Σ אלכסונית

ובהתאם מתקיים $(\Sigma^T\Sigma)_{ij}=0$ לכל $j\neq i$.

לכן, $(\Sigma^T\Sigma)^{-1}_{ii}=\frac{1}{\sigma_i^2}$, $(\Sigma^T\Sigma)_{ij}=0$ לכל $i\neq j$.

ולכן ל-Σ $(\Sigma^T\Sigma)^{-1}\cdot\Sigma^T=:D$ מתקיים

כי $D_{ii}=\frac{1}{\sigma_i}$ ו- $D_{ij}=0$ לכל $i\neq j$, כלומר

D אלכסונית 6xm.

וסך הכל קיבלנו $D=\Sigma^{+}$ ו- $(X^TX)^{-1}X^T = V\Sigma^{+}U^{-1}$

כנדרש