שאלה 1- אלגברה לינארית:

א. יהיו $A,B\in R^{n\times n}$. תזכורת: $tr(A)\coloneqq\sum_{i=1}^{n}A_{ii}$
הוכיחו או הפריכו את התכונות הבאות:

1) $(AB)^T=B^TA^T$.

2) $(A+B)^T=A^T+B^T$

3) $tr(AB)=tr(A)tr(B)$

4) $tr(AB)=tr(BA)$

5) $tr(A+B)=tr(A)+tr(B)$

אם A,B הפיכות:

6) $(AB)^{-1}=B^{-1}A^{-1}$

7) $(A+B)^{-1}=A^{-1}+B^{-1}$

ב. הוכיחו כי $Ker(A)=\{v|v\in R^n, Av=0\}$ ו-$IM(A)=\{w|w\in R^n,\exists v\in R^n|\ Av=w\}$ הם תתי מרחבים לינאריים.

ג. עבור $A\in R^{n\times p}$ הוכיחו כי $IM(A^T)=Ker(A)^{\perp}$.

שאלה 2- מטריצת הטלה:

מטריצה איידמפוטנטית היא מטריצה $A\in R^{n\times n}$ שדרגתה r המקיימת $A=A^2$.

מטריצה סימטרית ואיידמפוטנטית נקראת **מטריצת הטלה אורתוגונלית.**

א. הוכיחו כי הע״ע של מטריצת הטלה הם 1, בריבוי כדרגת המטריצה, ו-0 בריבוי השווה למימד של גרעין המטריצה.

ב. תהי $X\in R^{n\times p}$ מטריצה מדרגה מלאה ונגדיר $P_X=X(X^TX)^{-1}X^T$. הראו כי P_X היא מטריצת הטלה למרחב הנפרש על ידי העמודות של X. כלומר:

P_X סימטרית, איידמפוטנטית ומתקיים שלכל $v\in R^n$: $P_Xv\in IM(X)$.

ג. מצאו את $trace(P_X)$.

שאלה 3- יישומים של ליכסון אורתוגונלי:

<u>חלק 1</u>

תהי $A \in R^{n\times p}$ מטריצה כלשהי.

א. הראו של- AA^T ול- A^TA יש את אותם הע"ע. (תזכורת: בתרגיל הקודם הוכחתם שהם גם אי שליליים).

ב. השתמשו בכך כדי להראות שניתן לכתוב כל מטריצה $A = USV^T$ עבור U ו- V מטריצות ריבועיות מהממדים המתאימים ו- $S \in R^{n\times p}$ כך ש- $S_{ii} \geq 0$ וכן $S_{ij} = S_{ji} = 0$ לכל $i \neq j$. פירוק זה נקרא פירוק ה- SVD של A.

הדרכה:

א. כתבו $AA^T = U\Lambda U^T$. מדוע קיים הפירוק הזה? כעת על פי הסעיף הקודם, בחרו את V המתאימה. מדוע היא אורתוגונלית?

ב. השתמשו בכך וכתבו $I = V^TV$ כדי לקבל $U\Lambda U^T = USV^T(USV^T)^T$. מי אלו איברי S במונחי איברי המטריצה Λ? רמז: נמקו מדוע איברי Λ אי שליליים.

ג. איך ניתן לבטא את העמודה ה- i של U באמצעות העמודה ה- i של V, A ואיברי S? הסיקו כי ניתן לכתוב $AV = US$ ומשם מצאו את הפירוק של A.

אם אתם משתמשים בהדרכה עליכם לענות על השאלות בדרך.

<u>חלק 2</u>

א. בעבור מטריצה $X \in R^{n\times p}$ נגדיר:

הגדרה ראשונה:[1]

$$PC_1^{var} = \underset{\substack{w\in R^p \\ ||w||=1}}{\text{argmax}} \sum_{i=1}^{n} \left|w^T x_i\right|^2 = \underset{\substack{w\in R^p \\ ||w||=1}}{\text{argmax}} \left||Xw|\right|^2 = \underset{\substack{w\in R^p \\ ||w||=1}}{\text{argmax}}\, w^T X^T X w$$

הגדרה שניה - שגיאת ריבועים פחותים מינימאלית

$$\mathrm{PC}_1^{\mathrm{LS}} = \underset{\substack{w\in\mathbb{R}^p \\ \|w\|_2=1}}{\arg\min} \sum_{i=1}^{n} \mathrm{dist}(x_i, w)^2,$$

כאשר

$$\mathrm{dist}(x_i, w) = \|x_i - P_w(x_i)\|_2$$

ו $P_w(x)$ הוא ההיטל האורתוגונלי של הנק' x על תת המרחב הנפרש על ידי הוקטור w.

הניחו כי $\lambda_1 > \lambda_2 > \cdots > \lambda_p$. הראו ש- $PC_1^{var} = v_1$ כאשר u_1 הוא העמודה הראשונה של U מהפירוק הספקטרלי: $X^TX = U\Lambda U^T$.

ב. הראו כי עבור כל וקטור w כך ש $\|w\|_2 = 1$ מתקיים

$$\sum_{i=1}^{n} \mathrm{dist}(x_i, w)^2 = \|X\|_{\mathrm{F}}^2 - \sum_{i=1}^{n} |w^T x_i|^2.$$

והסיקו כי $PC_1^{LS} = PC_1^{var}$. (הוא נקרא "המרכיב הראשי הראשון" של A, והתהליך למציאתו נקרא PCA עליו תרחיבו בהמשך התואר).

---

[1] בהמשך נראה שתחת הנחות מסויימות, מדובר בוקטור המנורמל ששונותו היא הגבוהה ביותר מבין כל הוקטורים המכפילים מימין את מטריצת הנתונים.

הערה: פירוק ה-SVD שימושי מאוד בסטטיסטיקה, עיבוד תמונה ולמידת מכונה, ואף בבעיות רגרסיה כאשר עמודות X תלויות לינארית.
תוכלו לקרוא על חלק מהשימושים כאן ו-כאן.

שאלה 4- הקדמה למודל הלינארי

יהי Y משתנה מקרי כלשהו עם תוחלת ושונות סופיים.

א. נגדיר את הפונקציה - $M(a) = E[(Y\text{-}a)^2]$.
לכל a , הראו כי M מקבלת מינימום בערך $\hat{a} = E[Y]$.

ב. X ו-Y הם משתנים מקריים בעלי תוחלות ושונויות סופיות. לכל פונקציה f נגדיר – $MSE = E[(Y\text{-}f(x))^2]$. היעזרו בסעיף א׳ כדי להסיק ש-MSE ממוזער עבור $\widehat{f(x)} = E[Y|X = x]$.

<u>הדרכה:</u>
השתמשו בנוסחת התוחלת השלמה וכתבו – $MSE = E[g(x)]$, כאשר מתקיים –
$$g(x) = E[(Y\text{-}f(x))^2|X = x]$$
והפעילו את סעיף א׳ על Y <u>בהינתן</u> X = x.