רגרסיה- תרגול 3 :

מטריצת הטלה, המודל הלינארי והתפלגויות משותפות :

מטריצת הטלה אורתוגונלית

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

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

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

הוכחה : תרגיל

מטריצת הטלה למרחב הנפרש על ידי עמודות X:

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 :

(1) P_X סימטרית.

(2) P_X איידמפוטנטית.

(3) מתקיים שלכל $v \in R^n$: $P_Xv \in IM(X)$.

הוכחה : תרגיל

המרחב המשלים האורתוגונלי: יהי $U \subseteq V$ תת מרחב. נגדיר את תת המרחב המשלים האורתוגונלי של U באופן הבא :

$$U^\perp = \{v \in V | u^tv = 0, \forall u \in U\}$$

טענה: $U \oplus U^\perp = V$

הוכחה :

ניקח $v \in V$. אז $v = P_Uv + v - P_Uv = P_Uv + (I - P_U)v$. צד שמאל ב-U . בתרגיל תראו כי לכל וקטור $u \in U$ ההטלה של u לתת המרחב U היא u עצמו, ומכאן שלכל וקטור $u \in U$ נקבל $u^T(I - P_U)v = u^Tv - u^TP_Uv = u^Tv - u^TP_U^Tv = u^Tv - (P_Uu)^Tv = u^Tv - u^Tv = 0$ ומכאן שצד ימין ב-U^T (כאשר כמובן $U \cap U^T = \{0\}$ ולכן הסכום ישר.

מצד שני- כיוון ש-U ו-U^T מוכלים ב-V כי תתי מרחבים שלו, אז גם הסכום הישר שלהם מוכל ב-V כי הוא מרחב.

תכונות חשובות-אולי הכי חשובות בקורס(!!!)-של מטריצת הטלה:

Proposition 4. *Let $\boldsymbol{X}$ be an $n \times m$ matrix and assume that it has linearly independent columns (i.e., full column rank; remember that this implies $m \leq n$). Then the projection matrix $\boldsymbol{P}_{\boldsymbol{X}}$ has the following properties.*

1. $\boldsymbol{P}_X$ *is symmetric*
2. $\boldsymbol{P}_X$ *is idempotent,* $\boldsymbol{P}_X^2 = \boldsymbol{P}_X$
3. $\boldsymbol{P}_X \boldsymbol{X} = \boldsymbol{X}$
4. $\boldsymbol{X}^\top (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}}) = \boldsymbol{0} \in \mathbb{R}^{m\times n}$
5. $\boldsymbol{P}_X v \in \operatorname{Im}(X)$ *for all* $v \in \mathbb{R}^n$
6. *If* $m = n$ *and* $\boldsymbol{X}$ *is invertible, then* $\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{I}$
7. $(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}}) \boldsymbol{v} \in \operatorname{Im}(\boldsymbol{X})^\perp$ *for all* $\boldsymbol{v} \in \mathbb{R}^n$
8. *If* $w \in \operatorname{Im}(\boldsymbol{X})$*, then* $\boldsymbol{P}_{\boldsymbol{X}} \boldsymbol{w} = \boldsymbol{w}$
9. *If* $\boldsymbol{w} \in \operatorname{Im}(\boldsymbol{X})^\perp$*, then* $\boldsymbol{P}_X \boldsymbol{w} = \boldsymbol{0}$
10. *If* $\boldsymbol{Z}$ *is another* $n \times m$ *matrix s.t.* $\operatorname{Im}(\boldsymbol{Z}) = \operatorname{Im}(\boldsymbol{X})$*, then* $\boldsymbol{P}_{\boldsymbol{Z}} = \boldsymbol{P}_{\boldsymbol{X}}$*. This means that* $\boldsymbol{P}_{\boldsymbol{X}}$ *depends on* $\boldsymbol{X}$ *only through the span of its columns. Hence, for an arbitrary linear space* M*, we can define the projection matrix* $\boldsymbol{P}_M$ *onto* M *(an explicit form for* $\boldsymbol{P}_M$ *can be obtained by taking any basis of* M *and stacking its elements as columns in a matrix* $\boldsymbol{X}$*, then forming* $\boldsymbol{P}_{\boldsymbol{X}} := \boldsymbol{X}\left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top\Big)$
11. *If* L *and* M *are two subspaces with* $L \subseteq M$*, then* $\boldsymbol{P}_M \boldsymbol{P}_L = \boldsymbol{P}_L \boldsymbol{P}_M = \boldsymbol{P}_L$.

Proposition 6. *We have*

1. $\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}} = P_{Im(\boldsymbol{X})^\perp}$
2. *if* L *and* M *are two subspaces of* $\mathbb{R}^n$ *with* $L \subseteq M$*, then* $\boldsymbol{P}_M - \boldsymbol{P}_L = \boldsymbol{P}_{M \cap L^\perp}$

Proposition 7. *Let* $\boldsymbol{Q}$ *be an* $n \times n$ *matrix of rank* $m \leq n$ *which is symmetric and idempotent,* $\boldsymbol{Q}^\top = \boldsymbol{Q}, \boldsymbol{Q}^2 = \boldsymbol{Q}$*. Then* $\boldsymbol{Q} = \boldsymbol{P}_M$ *where* $M := \operatorname{Im}(\boldsymbol{Q})$.

Proof. Exercise. □

הוכחות ברשימות השיעור.

שאלה

1. הניחו כי A מטריצה ריבועית. הוכיחו כי $IM(A^T) = Ker(A)^{\perp}$.
2. טענה (ללא הוכחה): מטריצה A אינה לכסינה אס"ם קיים לפחות ע"ע אחד של A, λ_i עבורו $(A-\lambda_i)^k = 0, (A-\lambda_i)^m \neq 0, \forall k \geq 2, k \geq m$. השתמשו בטענה זו והראו כי אם A מטריצה סימטרית אז היא ניתנת ללכסון.
3. השתמשו בתוצאות אלו כדי להראות כי $Q = I - P_X$ היא מטריצת הטלה למרחב המשלים האורתוגונלי של $colspace(X)$ וכתבו במפורש את הפירוק הספקטרלי של Q במונחי הו"ע של P_X.
4. הסיקו מכך ש-$\hat{\beta}_{OLS}$ הוא הממזער של $||Y - X\beta||^2$ וכן שככל שדרגת X גדולה יותר, נורמה זו הולכת וקטנה.

פתרון:

1. נניח $v \in IM(A^T)$. אז $v = A^T w$ עבור $w \in R^n$. ניקח $u \in Ker(A)$. נקבל: $v^T u = w^T A u = 0$ ולכן $v \in Ker(A)^{\perp}$. באופן דומה מראים את ההכלה בכיוון ההפוך.

2. נניח ש-A אינה ניתנת ללכסון. בהתאם לטענה, זה קורה אס"ם קיים ע"ע λ המאפס לפחות פעמיים את הפולינום האופייני (בה"כ שני האחרונים):

$$\Pi_{i=1}^{n}(t_i - \lambda_i) = (t_i - \lambda_n)^2 \Pi_i^{n-2}(t_i - \lambda_i)$$

וגם מתקיים $(A - \lambda I)v \neq 0$ (אחרת v היה ו"ע) ואילו $(A-\lambda I)(A-\lambda I)v = (A-\lambda I)^2 v = 0$.

כעת כי סימטרית: $v^T(A-\lambda I)^2 v = ||(A-\lambda I)v||^2 = 0$. מתכונות של נורמה זה קורה אס"ם $(A-\lambda I)v = 0$ בסתירה.

3. ראשית שימו לב שזו מטריצת הטלה. מהשאלה הקודמת P_X סימטרית ולכן $IM(P_X) = Ker(P_X)^{\perp}$. ולהפך. כמו כן מהשאלה הקודמת, אם ניקח וקטור $v \in IM(X)$ אז $Qv \in Ker(P_X)$ וזהו המרחב המשלים האורתוגונלי ל-$IM(P_X) = IM(X)$. כיוון שניתנת ללכסון, לכל ו"ע u:

$$(I - P_X)u = u - \lambda u = (1-\lambda)u$$

לכן u הוא ו"ע של Q עם ע"ע $1-\lambda$. כאשר $\lambda \in \{0,1\}$

4. $||Y - X\beta||^2 = ||Y - P_X Y + P_X Y - X\beta||^2 = ||Y - P_X Y||^2 + 2|(Y - P_X Y)^T(P_X Y - X\beta)| +$
$$||P_X Y - X\beta||^2 \geq ||Y - P_X Y||^2 = ||(I - P_X)Y||^2 = ||QY||^2$$

ואם ניקח $X\hat{\beta} = P_X Y$ נקבל את החסם.

$$||QY||^2 = Y^T Q^2 Y = Y^T Q Y = \sum_{(i=1)}^{(n-r)} Y^T U_i U_i^T Y = \sum ||U_i Y||^2$$

וזה סכום של איברים אי שליליים ולכן גדל במספר הנסכמים.

התפלגויות רב מימדיות:

יהיו $Z_1, \dots, Z_n$ משתנים מקריים המפולגים במשותף (ללא שום הנחות נוספות) בהתפלגות $f_{Z_1,\dots,Z_n}(z_1, \dots, z_n)$. נגדיר $(Z_1, \dots Z_n) = Z \in R^n$ ונגדיר $E(Z) = (E[Z_1], \dots, E[Z_n])$.

באותו האופן, אם A היא **מטריצה מקרית**, כלומר A_{ij} הוא מ"מ: $E[A] = \begin{pmatrix} E[A_{11}] & \dots & E[A_{1m}] \\ \dots & \dots & \dots \\ E[A_{n1}] & \dots & E[A_{nm}] \end{pmatrix}$

תכונות של תוחלות של מטריצות (ובפרט וקטורים) מקריים:

עבור Z, W מטריצות מקריות ו-A, B מטריצות קבועות, לא מקריות (דטרמיניסטיות):

1. $\mathbb{E}[\boldsymbol{Z} + \boldsymbol{W}] = \mathbb{E}[\boldsymbol{Z}] + \mathbb{E}[\boldsymbol{W}]$
2. $\mathbb{E}[\boldsymbol{AZB}] = \boldsymbol{A}\mathbb{E}[\boldsymbol{Z}]\boldsymbol{B}$

3. $\mathbb{E}[AU + C] = A\mathbb{E}[U] + C$ (from 1+2)

עבור זוג וקטורים מקריים Z,W נגדיר את מטריצת השונות המשותפת בין Z,W:

$$Cov(Z,W) \coloneqq E([Z - E(Z)][W - E(W)]^T)$$

ובאופן דומה את מטריצת השונויות של הוקטור Z:

$$Var(Z) := Cov(Z,Z) \coloneqq E([Z - E(Z)][Z - E(Z)]^T)$$

טענה: $Var(Z)_{ij} = cov(Z_i, Z_j) = cov(Z_j, Z_i) = Var(Z)_{ji}$

"הוכחה":

$$Var(Z)_{ij} = E([Z - E(Z)][Z - E(Z)]^T)_{ij}$$
$$= E\big([Z - E(Z)]_i [Z - E(Z)]_j\big) = E([Z_i - E(Z_i)][Z_j - E(Z_j)]) = cov(Z_i, Z_j)$$

בתרגיל תוכיחו את התכונות הבאות:

Properties of covariance matrix. $\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{R}$ random vectors; $\boldsymbol{a}$ fixed vector. Then the following properties hold:

1. $\text{cov}(\boldsymbol{Z}, \boldsymbol{W}) = \text{cov}(\boldsymbol{W}, \boldsymbol{Z})^\top$
2. $\text{cov}(\boldsymbol{Z} + \boldsymbol{R}, \boldsymbol{W}) = \text{cov}(\boldsymbol{Z}, \boldsymbol{W}) + \text{cov}(\boldsymbol{R}, \boldsymbol{W})$
3. $\text{cov}(A\boldsymbol{Z}, B\boldsymbol{W}) = A\,\text{cov}(\boldsymbol{Z}, \boldsymbol{W})B^\top$
4. $\text{cov}(A\boldsymbol{Z}) = A\,\text{cov}(\boldsymbol{Z})A^\top$ (from 3)
5. $V\left(\boldsymbol{a}^\top \boldsymbol{Z}\right) = \boldsymbol{a}^\top \text{cov}(\boldsymbol{Z})\boldsymbol{a}$ (from 4)
6. $\text{cov}(\boldsymbol{Z})$ is a nonnegative definite matrix (from 5)

<u>שאלה:</u>

יהיו המשתנים המקריים הבאים: $X \sim Ber(p), Y \sim Ber(q)$ כך שמתקיים: $M \coloneqq XY \sim Ber(r)$.

א. מצאו את וקטור התוחלות ומטריצת השונויות של הוקטור המקרי $Z = (X, Y, M)^T$

ב. נגדיר את ההעתקה: $A: R^3 \to R$; $A(v) = 2v_1 - 3v_2 + 4v_3 + 7$. האם ההעתקה לינארית? על בסיס הסעיף הקודם חשבו את וקטור התוחלות ומטריצת השונויות של $A(Z)$.

ג. כעת הניחו ש-X,Y ב"ת. חשבו את ההסתברות $P(X = 1, Y = 1, M = 1)$.

פתרון:

א. עבור וקטור התוחלות אנו צריכים לחשב את התוחלת של כל אחת מ-3 הכניסות:

$$E(Z)^t = \big(E(X), E(Y), E(M)\big)^T = (p, q, r)^T$$

בעבור מטריצת השונויות: כפי שראינו, עלינו לחשב כל אחת מהשונויות המשותפות:

$$cov(X,X) = p(1-p), cov(X,Y) = E(XY) - E(X)E(Y) = r - pq, cov(X,M)$$
$$= E(X^2Y) - E(X)E(XY) = E(XY) - E(X)E(XY) = r - pr$$

ובאופן דומה לשאר הכניסות.

ב. זו איננה העתקה ליניארית כיוון שהוכחנו שכל העתקה ליניארית מקיימת $T(0) = 0$. שימו לב שנוכל לכתוב את A כך:

$$A(v) = Bv + 7$$

עבור $B = (2 - 3\ \ 4)$. נקבל:

$$E(AZ) = AE(Z) = (2 - 3\ \ 4)\begin{pmatrix} p \\ q \\ r \end{pmatrix} + 7$$

$$Var(AZ) = Var(BZ) = BVar(Z)B^T = (2\ \ -3\ \ 4)\begin{pmatrix} p(1-p) & r - pq & r(1-p) \\ r - pq & q(1-q) & r(1-q) \\ r(1-p) & r(1-q) & r(1-r) \end{pmatrix}\begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}$$

ג. במקרה כזה $P(X = 1, Y = 1, M = 1) = P(X = 1, Y = 1) = P(X = 1)P(Y = 1) = pq = r$

יהיו $Z, W \in R^p$ וקטורים מקריים. הראו שהבאים שקולים:

(1) $\forall v \in R^p, Var(v^T Z) \geq Var(v^T W)$

(2) $B \coloneqq Var(Z) - Var(W)$ היא מטריצה חיובית למחצה.

(3) קיימת המטריצה $B^{\frac{1}{2}}$.

המודל הליניארי:

נתונים: $(\boldsymbol{X}_i, Y_i), \quad i = 1,\ldots,n$

מודל ליניארי:

$$Y_i = \sum_{j=0}^{p} \beta_j X_{ij} + \epsilon_i, \qquad \mathrm{Cov}(\epsilon_i, \epsilon_{i'}) = \begin{cases} \sigma^2, & i = i' \\ 0, & i \neq i' \end{cases}$$

כאשר $\boldsymbol{X_i} = (1, X_{i1}, \ldots, X_{ip})^\top$ הם ממימד $p + 1$,

וכאשר $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^\top$ ו-σ^2 הם קבועים לא ידועים

אפשר לקבל ייצוג קומפקטי בעזרת כתיב מטריצות. נסמן:

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \qquad \boldsymbol{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{bmatrix}, \qquad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}, \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

אז את המודל הליניארי אפשר לכתוב:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}, \qquad \mathbb{E}[\boldsymbol{\epsilon}] = \boldsymbol{0}, \quad \mathrm{cov}[\boldsymbol{\epsilon}] = \sigma^2 \boldsymbol{I}$$

כמו כן: $\hat{\boldsymbol{Y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}, \qquad \boldsymbol{e} = \boldsymbol{Y} - \hat{\boldsymbol{Y}}$

(הערה: אם לא נציין אחרת, $\hat{\boldsymbol{\beta}}$ זה אומד הריבועים הפחותים)

שאלה

לפניכם מספר טענות. התאימו לכל טענה האם מדובר בהנחה או בתוצאה מתמטית:

1. $\hat{\beta} = argmin_\beta ||Y - X\beta||^2$
2. $X\beta = E(Y|X)$
3. $0 = E(e_i)$
4. $0 = E(\bar{e})$
5. $X^T e = 0$
6. $Cov(Y) = \sigma^2 I$

פתרון:

1. תוצאה מתמטית. הראינו במספר דרכים. זהו פתרון בעיית אופטימיזציה.
2. הנחה: $Y = X\beta + \epsilon$ וכן $E(\epsilon) = EE(\epsilon|X) = E(0) = 0$.
3. הנחה:
$E(e_i) = E\left(Y_i - X_i^T\hat{\beta}\right) = E(Y_i) - E(X[X^TX]^{-1}X_i^TY) = X_i^T\beta - X[X^TX]^{-1}X_i^TX\beta = X_i^T\beta - X_i^T\beta = 0$
4. תוצאה: מתנאי סדר ראשון מתקבלת המשוואה הנורמלית הראשונה: $\sum_{i=1}^n e_i = 0$
5. תוצאה: $X^Te = X^T(Y - P_XY) = X^T(I - P_X)Y = 0$
6. הנחה: מצויין במפורש למעלה.

שאלה:

לפניכם מתוארים מספר מקרים. עבור כל אחד מהם פרטו את ההתפלגויות של $X, Y, Y|X, \epsilon$ וכתבו אילו הנחות של המודל הליניארי כל אחד מהם מקיים:

1. $X_i \in R^p$ **קבועים מראש**, $Y_i = X_i^T\beta + \epsilon_i$ כאשר $\epsilon_i \sim N(0,1)$ ב״ת.
2. $X_i \in R^p$ מ״מ נורמליים סטנדרטיים וב״ת, $Y_i = X_i^T\beta + \epsilon_i$ כאשר $\epsilon_i \sim N(0,1)$ ב״ת.
3. $X_i \in R^1$ מ״מ נורמליים סטנדרטיים וב״ת, $Y_i = X_i^2\beta + \epsilon_i$ כאשר $\epsilon_i \sim U(-1,1)$ ב״ת.
4. $X_{it} \in R^p$ **קבועים מראש**- כאשר X_{itj} היא הדגימה של האדם ה- i בתקופה ה-t,
$Y_{it} = X_{it}\beta + \epsilon_{it}$ כאשר $\epsilon_{it} \sim N(0,1)$ והנדגמים ב״ת.

פתרון:

1. כל ההנחות מתקיימות.
2. כנ"ל.
3. כנ"ל. באופן כללי, נוכל להגדיר את X כמטריצת $Vandermonde$:

$$\begin{bmatrix} \mathbf{1} & x_0 & x_0^2 & \dots & x_0^n \\ \mathbf{1} & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{1} & x_n & x_n^2 & \dots & x_n^n \end{bmatrix}$$

כאשר במקרה כזה האמידה של וקטור המקדמים תהיה בעצם אמידה של מקדמי פולינום.

4. כאן סביר להניח שהנחת חוסר המתאם בין $\epsilon_{it_k}, \epsilon_{it_m}$ אינה מתקיימת. דוגמה למודל שכזה:

$$Y_i = X_i\beta + \epsilon_i + \epsilon_{i+1} \cdot \mathbf{1}_{\{i\ is\ odd\}}$$

שאלה

1. יהי $0 \neq v \in R^n$. הראו כי $\frac{vv^T}{||v||^2}$ היא מטריצת הטלה אורתוגונלית. מה דרגת המטריצה?
2. יהיו $Y_1, \dots, Y_n \sim (\mu, \sigma^2)$ מ"מ ב"ת ש"ה, כאשר שני הפרמטרים לא ידועים. הראו כי $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \bar{Y})^2$ הוא אומד ח"ה ל-σ^2.
3. הניחו כעת כי $Y_1, \dots, Y_n \sim N(\mu, \sigma^2)$. הוכיחו את התוצאה, שראינו בעבר $(n-1)S_n^2 \sim \sigma^2\chi_{n-1}^2$.

פתרון: בתרגול הבא/תרגיל

1. $(\frac{vv^T}{||v||^2})^T = \frac{1}{||v||^2} v^{T^T} v^T = \frac{vv^T}{||v||^2}$

$$\left(\frac{vv^T}{||v||^2}\right)^2 = \frac{vv^T}{||v||^2}\frac{vv^T}{||v||^2} = \frac{||v||^2}{||v||^4} vv^T = \frac{vv^T}{||v||^2}$$

דרגת המטריצה היא 1 (אפשר לדרג ולראות, או לשים לב שמדובר ב-P_X עבור $X = v$.

2. תרגיל (שפתרתם הרבה פעמים בעבר).
3. ראשית, ניתן לכתוב כל משתנה נורמלי כ-$\mu + \sigma Z_i$ עבור $Z_i \sim N(0,1)$:

$$\sum_{i=1}^n (Z_i - \bar{Z})^2 = \left|\left|\left(I - \frac{1}{n} \cdot 1_n 1_n^T\right) Z\right|\right|^2 = Z^T\left(I - \frac{1}{n} \cdot 1_n 1_n^T\right) Z$$

כאשר $Z = (Z_1, \dots, Z_n)^T$, לקחנו את v מסעיף א' להיות הוקטור 1_n וכיוון שזו מטריצת הטלה אז גם $(I - \frac{1}{n} \cdot 1_n 1_n^T)$ המטילה למרחב המשלים שמימדו $n - r = n - 1$. כמו כן סימטרית ואיידמפוטנטית ומכאן המעבר האחרון. כעת כיוון שמטריצת הטלה אז ניתנת ללכסון כך שיששן $n-1$ ע"ע שהם 1 וע"ע אחד שהוא 0. לכן ממשפט הפירוק הספקטרלי:

$$Z^T \sum_{i=1}^n \lambda_i U_i U_i^T Z = \sum_{i=1}^{n-1} \left(U_i^T Z\right)^T \left(U_i^T Z\right) = \sum_{i=1}^{n-1} \left(U_i^T Z\right)^2$$

נתבונן בהתפלגות של $U_i^T Z = \sum_j U_i^j Z_j$. זו קומבינציה לינארית של מ"מ נורמליים ולכן מתפלגת נורמלית. התוחלת היא: $\sum_j U_i^j E(Z_j) = 0$ והשונות היא: $\sum_j {U_i^j}^2 = ||U_i||^2 = 1$. כלומר התפלגות נורמלית סטנדרטית, כמו במקור. כיוון שזו פונקציה של מ"מ ב"ת אז גם הפונקציות ב"ת. מכאן שנותרנו עם סכום ריבועים של $n-1$ מ"מ נורמליים סטנדרטיים וב"ת אשר בהגדרה

(עד שתראו זאת בהסתברות לסטטיסטיקאים) מפולגים χ^2_{n-1}.
כעת כל שנותר הוא להכפיל ב-σ לקבל את המבוקש.

שאלה

2.נתון $Z \in \mathbb{R}^m$ וקטור מקרי. הראו כי מתקיים ש

$$\mathbb{E}\left(||Z||^2\right) = tr(\mathbb{E}[ZZ^T])$$

הסיקו מכך כי אם $\mathbb{E}[Z] = 0$ אזי מתקיים כי

$$\mathbb{E}\left(||Z||^2\right) = tr(cov[Z])$$

פרטו והצדיקו כל שלב בהוכחה.

הוכחה:

1. $E\left(||Z||^2\right) = E(Z^T Z) = E(tr[Z^T Z]) = E(tr(ZZ^T))$

2. $E\left(||Z||^2\right) = tr\left(E(ZZ^T)\right) = tr(E\left(Z - E(Z)\right)\left(E\left(Z - E(Z)\right)^T\right) = tr(Cov(Z))$