<u>רגרסיה ומודלים סטטיסטיים- תרגול 2</u>

מטריצת מעבר מבסיס לבסיס ומטריצה מייצגת העתקה:

טענת עזר: תהי $u \in R^m, A \in R^{n\times m}$. אז $v = Au$ הוא קומבינציה לינארית של עמודות A, כלומר $v \in colspace(A)$. הוכחה:

$$v_i = \sum_{j=1}^{m} A_{ij}u_j = \sum_{j=1}^{m} [A^j]_i u_j$$

ובאותו האופן:

$$v = \sum_{j=1}^{m} A^j u_j$$

<u>הגדרות:</u>

1. יהי $v \in V$ ו-$B = \{b_1, \dots, b_n\}$ בסיס **סדור** ל-V, ולכן $v = \sum_{i=1}^{n} k_i b_i = Bk$ וכן זוהי הקומבינציה הלינארית היחידה המקיימת את שוויון זה (פתרון יחיד למערכת המשוואות). נגדיר את וקטור הקואורדינאטות של v ביחס לבסיס B: $[v]_B = (k_1, \dots, k_n)^T$.

הערה: שימו לב שאם B הוא הבסיס הסטנדרטי אז $[v]_B = v$. ניתן להסתכל על כך באופן הבא: תחת הבסיס הסטנדרטי "מערכת הצירים" היא זו המוכרת לנו- הקרטזית, ולכן הקואורדינאטות מוגדרות בדיוק באופן שבו אנו מכירים. תחת בסיס אחר, ניתן לחשוב על כך כשינוי של מערכת צירים לכזו בעלת נקודות ייחוס שונות מראשית הצירים, ולכן נצטרך לייצג את הוקטור v תחת מערכת הצירים החדשה.

דוגמה:

$V = \mathbb{R}^2$ ו $v = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, $\mathcal{B} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$.

אז על מערכת הצירים: $v = a_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$

מכאן, $[v]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. כי מתקיים ש- $v = -3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$

לחילופין, הפעולה זהה ל- $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ שהיא למעשה פתרון המטריצה: $\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 5 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right]$

הפתרון היחיד היינו $\begin{bmatrix} -3 \\ 5 \end{bmatrix} = [v]_{\mathcal{B}}$

2. מטריצת מעבר מבסיס לבסיס: יהיו B, C בסיסים סדורים ל-V, $v \in V$, $[v]_B$. מטריצת המעבר מבסיס B לבסיס C היא המטריצה $[I]_C^B$ (היחידה) המקיימת: $[I]_C^B [v]_B = [v]_C$.[1] במילים- המטריצה שהכפלה משמאל בה של וקטור הקואורדינאטות של v לפי הבסיס B, תתן את וקטור הקואורדינאטות של v לפי הבסיס C.

<u>טענה 1:</u> עמודות המטריצה הן וקטורי הקואורדינטות של וקטורי הבסיס B לפי הבסיס C. הוכחה:

ראשית, נשים לב שכיוון ש-B, C בסיסים ומהגדרת וקטור הקואורדינטות: $[v]_B = B^{-1}v$, $[v]_c =$ $C^{-1}v$, לכן:

$$[I]_C^B [v]_B = [v]_C \Leftrightarrow [I]_C^B B^{-1}v = C^{-1}v \Leftrightarrow [I]_C^B B^{-1}v = C^{-1}BB^{-1}v \Leftrightarrow [I]_C^B = C^{-1}B \Leftrightarrow [I]_{C \cdot j}^B = C^{-1}b_j = [b_j]_C$$

[1] הערה: בחלק מהמקורות היא נקראת דווקא " מטריצת המעבר מבסיס B לבסיס C". נתייחס אליה כפי שמוגדר כאן.

טענה 2: $[I]_C^B = \left([I]_B^C\right)^{-1}$. ההוכחה נובעת ישירות מהטענה הקודמת. שימו לב שמכך נוכל להסיק שאם B מטריצה ריבועית מדרגה מלאה- כלומר עמודותיה מהוות בסיס, אז B^{-1} היא מטריצת מעבר מהבסיס הסטנדרטי לבסיס שהוא העמודות של B.

טענה 3: **לכל** העתקה לינארית $T: V \to W$, ובסיסים C, B: $V = span(B)$, $W = span(C)$, קיימת מטריצה יחידה $[T]_C^B$ המקיימת שלכל וקטור $v \in V$: $[T]_C^B[v]_B = [T(v)]_C$. זוהי המטריצה המייצגת את ההעתקה T לפי בסיס B בתחום ו-C בטווח.
שימו לב שמטריצת מעבר היא מקרה פרטי בו T היא העתקת הזהות.

טענה 4: אם $[I]_C^B$ מטריצת מעבר מבסיס B לבסיס C, ו-$T: V \to V$ היא העתקה לינארית, אז:

$$[T]_C = [I]_C^B[T]_B[I]_B^C$$

הוכחה:

לכל $v \in V$ מתקיים: $[T]_C[v]_C = [T(v)]_C = [I]_C^B[T(v)]_B = [I]_C^B[T]_B[v]_B = [I]_C^B[T]_B[I]_B^C[v]_C$.
כיוון שקיים יצוג יחיד
$[v]_C \to v$ מתקבלת הטענה.

Consider the linear transformation F on $\mathbf{R}^2$ defined by $F(x,y) = (5x - y,\ 2x + y)$ and the following bases of $\mathbf{R}^2$:

$$E = \{e_1, e_2\} = \{(1,0),\ (0,1)\} \qquad \text{and} \qquad S = \{u_1, u_2\} = \{(1,4),\ (2,7)\}$$

(c) Find the matrix B that represents F in the basis S.

Method 2. By Theorem 6.7, $B = P^{-1}AP$. Thus,

$$B = P^{-1}AP = \begin{bmatrix} -7 & 2 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} 5 & -1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ -2 & 1 \end{bmatrix}$$

הערה: כאן $P = [I]_E^S$, כלומר מטריצת המעבר מהבסיס בייצוג S לבסיס הסטנדרטי.

אם לא נאמר אחרת, מטריצה נתונה A מייצגת לפי בסיס סטנדרטי (בתחום ובטווח).

בייצוג זה מתקיים:

$$Ker(T) = Ker(A) = \{v \in V | Av = 0\}$$
$$IM(T) = IM(A) = colspace(A) = \{w \in W | Av = w\}$$

ונסיק כי דרך למציאת בסיס למרחב הגרעין של העתקה תהיה מציאת בסיס למרחב הפתרונות של מערכת המשוואות ההומוגנית $Av = 0$.
ואילו כדי למצוא בסיס למרחב התמונה של ההעתקה, נדרג את A^T לתצורה קנונית (לדוגמא), ונבחר את השורות שאינן מתאפסות.

מהשקילות הנ״ל וממשפט המימד להעתקות לינאריות נוכל להסיק כי לכל מטריצה $A \in R^{n \times m}$, כך ש- $rank(A) = r \le \min(n,m)$ מתקיים: $\dim(Ker(A)) = n - r$, $\dim(IM(A)) = r$.

תזכורת:

1 וקטורים וערכים עצמיים, פולינום אופייני והפירוק הספקטראלי

1.1 הגדרות

1. תהי המטריצה **הריבועית** $A \in \mathbb{R}^{n\times n}$ ויהיו $\lambda \in \mathbb{R}$ סקלאר ו- $x \in \mathbb{R}^n$ וקטור. אם מתקיים השוויון $Ax = \lambda x$ אזי x הוא **וקטור עצמי** של Aו- λ הוא **ערך עצמי** של A.

2. $\det(A - \lambda I)$ מוגדר להיות **הפולינום האופייני** של A, עבור I מטריצת היחידה מאותו מימד של A. באמצעות הפולינום האופייני מחשבים את הע״ע של המטריצה, מאחר ועבור פתרונות לא טריוויאלים (וע שאינם 0) הפתרון הוא יחיד, משמע A היא מטריצה סינגולרית, כלומר $\det(A - \lambda I) = 0$.

3. תהי המטריצה $A \in \mathbb{R}^{n\times n}$מטריצה סימטרית. אזי ניתן להציג את A כך: $A = U\Lambda U^T$ וההצגה הזו נקראת **הפירוק הספקטרלי של** A, כאשר:

- $\Lambda \in \mathbb{R}^{n\times n}$ היא מטריצה אלכסונית, שעל האלכסון שלה נמצאים הע״ע של A.
- $U \in \mathbb{R}^{n\times n}$ היא מטריצה המורכבת מהוקטורים העצמיים של A. העמודה ה j היא הוקטור העצמי המתאים לערך העצמי λ_j.
- כמו כן מתקיים ש $UU^T = U^TU = I$ כלומר U היא מטריצה אורתונורמלית.

תרגיל:

א. תהי $A := X^TX = U\Lambda U^T \in R^{n\times n}$, כאשר $\Lambda = diag(\lambda_1, \dots, \lambda_n)$ $= \mathrm{U^TAU}$ ועמודות $X \in R^{n\times p}$ בת״ל. הניחו שכל הע״ע של A שונים זה מזה.
הסבירו מדוע Λ מייצגת את ההעתקה:

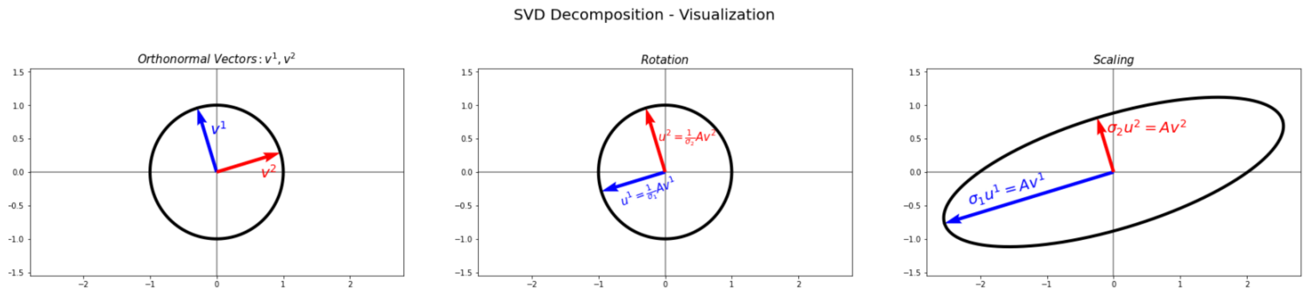
$$T: R^p \to R^p, T(v) = Av$$

ביחס לבסיס אורתונורמלי הנפרש על ידי עמודות U והסבירו כיצד זה מסתדר עם יחידות הייצוג של מטריצה מייצגת העתקה ועם כך שהמטריצה Λ מייצגת את העתקה

$$f: R^p \to R^p, f(v) = \left(\lambda_1 \cdot v_1, \dots, \lambda_p \cdot v_p\right)^T$$

ביחס לבסיס הסטנדרטי.

ב. פרשנות גיאומטרית לכפל במטריצה: הראו שהעתקה $A: V \to W$ יכולה להיכתב כסיבוב, מתיחה וסיבוב של $v \in V$ כאשר מיוצגת בבסיס הסטנדרטי.



פתרון:

א. לכל וקטור $v \in R^n$ מתקיים כי $v = [v]_E$. לכן:

$$[A]_E[v]_E = [A(v)]_E = Av = U\Lambda U^T v = U\Lambda U^{-1} v = [I]_E^U [\Lambda]_U [I]_U^E [v]_E$$

כאשר השתמשנו בטענה 2 וטענה 4. כלומר Λ מייצגת את ההעתקה A ביחס לבסיס האורתונורמלי שהוא עמודות U.

זה מסתדר כיוון שיחידות הייצוג (איזומורפיזם בין המטריצה המייצגת ובין ההעתקה) היא ביחס לבסיס סדור מסוים.

באותו האופן, A מייצגת את ההעתקה f ביחס לבסיס האורתונורמלי שהוא שורות U.

ב. כאשר מיוצגת בבסיס הסנדרטי, $[A]_E = A = U\Lambda U^T$

ראשית נבין מה היא הפעולה $v \to U^T v$:

ניקח $v_1, v_2 \in V$ ונגדיר $\widetilde{v_1} \coloneqq U^T v_1, \widetilde{v_2} \coloneqq U^T v_2$ אזי :

$$||\widetilde{v_1}||^2 = ||U^T v_1||^2 = v_1^T U U^T v_1 = ||v_1||^2$$

כלומר הכפלה ב-U לא משנה את הגודל של הוקטור.

באותו האופן, נקבל כי :

$$v_1^T v_2 = \widetilde{v_1}^T \widetilde{v_2}$$

נזכור כי עבור הזווית שבין הוקטורים θ מתקיים :

$$Cos(\theta) = \frac{v_1^T v_2}{||v_1|| \cdot ||v_2||} = Cos(\tilde{\theta})$$

כלומר הפעולה משמרת גודל וזווית- לכן מדובר בסיבוב או שיקוף (או שניהם).

כעת, כיוון שמכפלת וקטור במטריצה אלכסונית היא כמו מכפלת כל אחת מהקואורדינאטות של הוקטור באיבר המתאים על האלכסון, נקבל כי כי מדובר במתיחת הקואורדינטה ה-j ב-λ_j.

לבסוף, כיוון $UU^T = U^T U = I$ ומתחילת הסעיף, נקבל שזהו סיבוב בחזרה לנקודות המוצא, לאחר המתיחה.

<u>שאלה</u>

תהי $A \in R^{n\times n}$ מטריצה לכסינה (לאו דווקא סימטרית). הראו כי :

$$\det(A) = \Pi_{i=1}^n \lambda_i$$

$$tr(A) = \sum_{i=1}^n \lambda_i$$

הוכחה :

$$\det(A) = \det(U\Lambda U^{-1}) = \det(U)\det(\Lambda)\det(U^{-1}) = \det(U)\det(\Lambda)\det(U^{-1})$$
$$= \det(UU^{-1})\, det\,(\Lambda) = \det(I)\det(diag[\lambda_1, \dots, \lambda_n]) = \Pi_{i=1}^n \lambda_i$$

הגדרה של tr נמצאת בתרגול הקודם (סכום האיברים על האלכסון). תרגיל:

הראו ש-$tr(AB) = tr(BA)$

ואז : $\sum_i \lambda_i = \mathrm{tr}(\lambda) = tr(U^{-1}U\Lambda) = tr(U\Lambda U^{-1}) = tr(A)$.

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:

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:

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 $\boldsymbol{P_X}$ 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. $\boldsymbol{X}^\top(\boldsymbol{I} - \boldsymbol{P_X}) = \boldsymbol{0} \in \mathbb{R}^{m\times n}$
5. $\boldsymbol{P}_X v \in \mathrm{Im}(X)$ *for all* $v \in \mathbb{R}^n$
6. *If* $m = n$ *and* $\boldsymbol{X}$ *is invertible, then* $\boldsymbol{P_X} = \boldsymbol{I}$
7. $(\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{v} \in \mathrm{Im}(\boldsymbol{X})^\perp$ *for all* $\boldsymbol{v} \in \mathbb{R}^n$
8. *If* $w \in \mathrm{Im}(\boldsymbol{X})$*, then* $\boldsymbol{P_X w} = \boldsymbol{w}$
9. *If* $\boldsymbol{w} \in \mathrm{Im}(\boldsymbol{X})^\perp$*, then* $\boldsymbol{P}_X\boldsymbol{w} = \boldsymbol{0}$
10. *If* $\boldsymbol{Z}$ *is another* $n \times m$ *matrix s.t.* $\mathrm{Im}(\boldsymbol{Z}) = \mathrm{Im}(\boldsymbol{X})$*, then* $\boldsymbol{P_Z} = \boldsymbol{P_X}$*. This means that* $\boldsymbol{P_X}$ *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* $\boldsymbol{P_X} := \boldsymbol{X}\left(\boldsymbol{X}^\top\boldsymbol{X}\right)^{-1}\boldsymbol{X}^\top$*)*
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. $\boldsymbol{I} - \boldsymbol{P_X} = P_{Im(\boldsymbol{X})^\perp}$
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 $M := \mathrm{Im}(\boldsymbol{Q})$.*

Proof. Exercise. □

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

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

טענה: $U \oplus U^\perp = V$

שאלה

1. הניחו כי A מטריצה ריבועית. הוכיחו כי $IM(A^T) = Ker(A)^\perp$.
2. טענה (ללא הוכחה): מטריצה A אינה לכסינה אם״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

<u>שאלה</u>

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

(

נגזרות של וקטורים ומטריצות

א. גרדיאנט של פונקציה מרובת משתנים $f\left(x_1,\ldots,x_m\right)$

$$\nabla f=\frac{\partial f}{\partial x}=\begin{pmatrix}\frac{\partial f}{\partial x_1}\\ \vdots\\ \frac{\partial f}{\partial x_m}\end{pmatrix}$$

ב. נגזרת של וקטור $y=\left(y_1,\ldots y_m\right)^T$ לפי סקלר x:

$$\frac{\partial y}{\partial x}=\begin{pmatrix}\frac{\partial y_1}{\partial x}\\ \vdots\\ \frac{\partial y_m}{\partial x}\end{pmatrix}$$

ג. נגזרת של וקטור $y=\left(y_1,\ldots y_m\right)^T$ לפי וקטור $x=\left(x_1,\ldots x_n\right)^T$:

$$\frac{\partial y}{\partial x}=\begin{pmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n}\\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n}\end{pmatrix}$$

ד. נגזרת של מכפלה סקלרית של וקטורים:

$$\frac{\partial}{\partial x}\left(b^T x\right)=\frac{\partial}{\partial x}\left(x^T b\right)=b$$

$$\frac{\partial}{\partial x}\left(x^T x\right)=2x$$

ה. נגזרת של כפל מטריצה בוקטור לפי הוקטור:

$$\frac{\partial}{\partial x}\left(Ax\right)=A$$
$$\frac{\partial}{\partial x}\left(x^T Ax\right)=\left(A^T+A\right)x$$

כאשר את המשוואה האחרונה מקבלים על ידי

$$\frac{\partial}{\partial x_k}\left(x^T A x\right) = \frac{\partial}{\partial x_k}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j = \sum_{i=1}^{n} a_{ik}x_i + \sum_{j=1}^{n} a_{kj}x_j$$

לפיכך, עבור מטריצה סימטרית A נקבל

$$\frac{\partial}{\partial x}\left(x^T A x\right) = 2Ax$$