חזרה מושגים באלגברה לינארית

יולי סלווטסקי

<u>כפל מטריצות:</u> תהי $A=(a_{ij})$ מטריצה בגודל $n\times m$ ו - $B=(b_{ij})$ מטריצה בגודל $m\times p$. מכפלתן AB היא מטריצה בגודל $n\times p$ שאבריה מוגדרים לפי הנוסחה

$$(AB)_{ij}=\sum_{k=1}^{m}a_{ik}b_{kj}$$

כלומר האיבר בשורה ה - i ובעמודה ה - j של המכפלה AB מתקבל מהכפלת השורה ה - i במטריצה A בעמודה j במטריצה B. לדוגמה:

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{pmatrix},B=\begin{pmatrix} b_{11} & b_{12}\\ b_{21} & b_{22}\\ b_{31} & b_{23}\end{pmatrix}$$

$$AB=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{23}\\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{23}\end{pmatrix}$$

<u>שחלוף מטריצות:</u> שחלוף הוא פעולת ההחלפה בין השורות והעמודות של מטריצה נתונה. לפיכך עבור $A=(a_{ij})$ מטריצה בגודל $n\times m$

$$\left(A^T\right)_{ij}=(A)_{ji}$$

ובדוגמה שלנו

$$A^T=\begin{pmatrix} a_{11} & a_{21}\\ a_{12} & a_{22}\\ a_{13} & a_{23}\end{pmatrix}$$

<u>תכונות:</u>

א. פילוג הכפל:

$$A(B+C)=AB+AC$$
$$A(BC)=(AB)C$$

ב. שחלוף:

$$\left(A^T\right)^T = A$$
$$(A+B)^T = A^T + B^T$$
$$(AB)^T = B^T A^T$$

ג. כפל בסקלר c:

$$cA = Ac$$
$$c\,(AB) = (cA)\,B$$
$$(cA)^T = c\,(A)^T$$

ד. דטרמיננטה (נזכיר בהמשך):

$$\det\left(A^T\right) = \det(A)$$
$$\det(AB) = \det(BA) = \det(A)\det(B)$$

ה. מטריצה הפוכה (נזכיר בהמשך):

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$$

ו. נבחין כי בעוד $AB \neq BA$, כלומר כפל מטריצות הוא אינו קומוטטיבי, מכפלה סקלרית של וקטורים כן. יהיו x, y שני וקטורים באורך n, אזי

$$x^T y = x_1 y_1 + \cdots + x_n y_n$$

משום ש - $x^T y$ הוא סקלר,

$$\left(x^T y\right)^T = y^T x$$

2. מטריצות סימטריות

מטריצה ריבועית $A \in \mathbb{R}^{n\times n}$ תקרא סימטרית אם $A = A^T$, כלומר לכל i, j : $A_{ij} = A_{ji}$

<u>תכונות</u>:

א. אם A, B הן מטריצות סימטריות אזי $A+B$ גם כן סימטרית
ב. אם A, B הן מטריצות סימטריות אזי AB סימטרית אם ורק אם $AB=BA$
ג. חזקה של מטריצה סימטרית A^n היא סימטרית
ד. אם A הפיכה, אזי A^{-1} סימטרית רק אם A סימטרית

3. נורמה של וקטור

נורמה היא פונקציה המוגדרת על מרחב וקטורי ומתאימה לכל וקטור ערך ממשי כך שמתקיימים מספר תנאים. נורמה יכולה להיות מוגדרת על כל מרחב וקטורי אך התנאים מבוססים על תכונות שמתקיימות עבור אורך במרחב אוקלידי.

א. חיוביות: $\|x\| \geq 0$ ו - $\|x\| = 0$ רק עבור $x = 0$
ב. הומוגניות (בכפל בסקלר): $\|cx\| = |c| \, \|x\|$
ג. אי שוויון המשולש: $\|x+y\| \leq \|x\| + \|y\|$

<u>דוגמה</u>: נורמת L^p מוגדרת על ידי

$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}}$$

בפרט, עבור $p=1$ נקבל

$$\|x\|_1 = \left(\sum_i |x_i|^1 \right)^{\frac{1}{1}} = \sum_i |x_i|$$

עבור $p=2$ נקבל

$$\|x\|_2 = \left(\sum_i |x_i|^2 \right)^{\frac{1}{2}} = \sqrt{\sum_i |x_i|^2} = \sqrt{\sum_i (x_i^2)}$$

מקרה מיוחד הוא $p=\infty$, עבורו הנורמה נקראת נורמת המקסימום:

$$\|x\|_\infty = \max_i |x_i|$$

4. עקבה

העקבה של מטריצה היא סכום ערכי האלכסון שלה

$$\mathrm{Tr}\,(A) = \sum_k A_{k,k}$$

מתקיים:

$$\begin{aligned}
&\operatorname{Tr}(A)=\operatorname{Tr}\left(A^T\right)\\
&\operatorname{Tr}(AB)=\operatorname{Tr}(BA)\\
&\operatorname{Tr}(ABC)=\operatorname{Tr}(CBA)=\operatorname{Tr}(BCA)
\end{aligned}$$

5. פעולות אלמנטריות

דירוג מטריצה הוא הפעלה של פעולות אלמנטריות שאינן משנות את מרחב הפתרונות שלה. פעולות אלמנטריות משמשות מציאת פתרונות של מערכת משוואות ליניאריות, מציאת דרגה של מטריצה, מציאת דטרמיננטה של מטריצה ומציאת המטריצה ההופכית של מטריצות הפיכות.

הפעלת הפעולות הבאות אינה משנה את מרחב הפתרונות אך משנות את הדטרמיננטה באופן הבא:

א. החלפה בין שתי שורות $R_i \leftrightarrow R_j$ משנה את סימן הדטרמיננטה.

ב. כפל שורה בקבוע שאינו 0, $R_i \leftarrow cR_i$, משנה את הדטרמינטה בכפל באותו קבוע.

ג. הוספה של שורה לשורה אחרת לאחר שהוכפלה בקבוע $R_i \leftarrow cR_j + R_i$ לא משנה את הדטרמיננטה.

<u>פתרון משוואות:</u> תהי $A \in \mathbb{R}^{m\times n}$ מטריצה שערכיה ידועים, $b \in \mathbb{R}^m$ וקטור שערכיו ידועים ו־$x \in \mathbb{R}^n$ וקטור של משתנים x_i שערכיהם לא ידועים. מערכת המשוואות

$$\begin{aligned}
A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1\\
A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2\\
&\vdots\\
A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m
\end{aligned}$$

ניתנת לרישום קומפקטי באמצעות כפל מטריצות על ידי $Ax=b$. בהמשך הקורס, אנו נשתמש ברישום זה למציאת הפתרון לרגרסיה לינארית, כאשר בתפקיד המטריצה שערכיה ידועים תהיה מטריצת הנתונים X ואנו נחפש את ערכי המקדמים β שכאן הם מיוצגים על ידי הוקטור x. שימו לב לחילוף התפקידים בסימון.

6. מטריצה צמודה

<u>הגדרה</u>: מטריצה צמודה של מטריצה A מגודל $m\times m$ מסומנת $\operatorname{adj}(A)$ ומוגדרת כך:

$$\operatorname{adj}(A)_{ij} = (-1)^{\mathrm{i+j}}\left|M_{j,i}\right|$$

כאשר $M_{j,i}$ הוא המינור של האיבר במקום ה - j,i במטריצה A, כלומר המטריצה המתקבלת כאשר מוחקים את שורה j ועמודה i במטריצה A.

7. דטרמיננטה

דטרמיננטה של מטריצה ריבועית היא סקלר התלוי ברכיבי המטריצה, ושווה לאפס אם ורק אם המטריצה אינה הפיכה.

<u>הגדרה:</u> הדטרמיננטה של מטריצה בגודל $n \times n$ מוגדרת על פי הנוסחה הבאה:

$$|A| = \det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \dots A_{n,\sigma(n)}$$

<u>חישוב הדטרמיננטה:</u>

א. דירוג המטריצה: נדרג את המטריצה על ידי הפעלה של פעולות אלמנטריות עד שמגיעים למטריצה משולשית. הדטרמיננטה של מטריצה משולשית שווה למכפלת איברי האלכסון הראשי שלה.

ב. פיתוח לפי מינורים:

$$|A| = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} |M_{i,j}|$$

כאשר $M_{i,j}$ הוא המינור של האיבר במקום ה - i,j במטריצה A, כלומר המטריצה המתקבלת כאשר מוחקים את שורה i ועמודה j במטריצה A.

<u>דוגמה:</u>

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{vmatrix} = a \begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b \begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c \begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d \begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix}$$

<u>דטרמיננטה של מטריצה ריבועית:</u>

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

<u>דטרמיננטה של מטריצה מגודל 3×3:</u>

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$

$$= aei + bfg + cdh - ceg - bdi - afh$$

8. מטריצה הופכית

הגדרה: נתבון במטריצה ריבועית A מגודל $m \times m$. אם קיימת מטריצה A^{-1} כל ש- $AA^{-1} = A^{-1}A = I_m$, אז המטריצה A נקראית הפיכה והמטריצה A^{-1} היא המטריצה ההופכית שלה.

שיטות למציאת המטריצה ההופכית:

א. הפעלה של פעולות אלמנטריות על המטריצה A עד שמקבלים את I.

דוגמה:

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\left(\begin{array}{cccc|cccc} 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R_1 \leftrightarrow R_3]{R_2 \leftrightarrow R_4} \left(\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 & 0 & 0 \end{array}\right)$$

$$\xrightarrow[R_3 \leftarrow \frac{1}{3}R_3]{R_1 \leftarrow \frac{1}{2}R_1} \left(\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & \frac{1}{3} & 0 & 0 \end{array}\right)$$

ב. שימוש במטריצה צמודה

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

תחילה נמצא את $\det(A)$. לשם כך נדרג את המטריצה עד לקבלת מטריצה משולשית. בדוגמה שלנו, זהו בעצם הצעד הראשון שביצענו בשיטה הקודמת. לכן, $\det(A) = 2 \cdot 1 \cdot 1 \cdot 3 = 6$ ונותר לנו למצוא את $\operatorname{adj}(A)$:

$$\operatorname{adj}(A)_{1,2} = (-1)^{1+2} \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 0$$

$$\operatorname{adj}(A)_{1,3} = (-1)^{1+3} \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{vmatrix} = 3$$

ממשיכים באותו אופן לכל (i,j) ומקבלים:

$$A^{-1}=\frac{1}{6}\begin{pmatrix}0&0&3&0\\0&0&0&6\\6&0&0&0\\0&2&0&0\end{pmatrix}=\begin{pmatrix}0&0&\frac{1}{2}&0\\0&0&0&1\\1&0&0&0\\0&\frac{1}{3}&0&0\end{pmatrix}$$

<u>פתרון משוואות באמצעות מטריצה הופכית:</u> נניח כי נתונה לנו מערכת המשוואות $Ax=b$ ו ־ A היא מטריצה הפיכה. אזי

$$\begin{aligned}&A^{-1}Ax=A^{-1}b\\&Ix=A^{-1}b\\&x=A^{-1}b\end{aligned}$$

9. וקטורים וערכים עצמיים

<u>הגדרה</u>: תהי A מגודל $m\times n$. λ נקרא ערך עצמי של A אם $\det(A-\lambda I)=0$.
תנאי זה שקול לתנאי שקיים $x\in\mathbb{R}^m$ כך ש־ $(A-\lambda I)x=0$, כלומר $Ax=\lambda x$.
x הוא הוקטור העצמי המתאים לערך העצמי λ.

<u>טענה</u>: תהי A מטריצה סימטרית מגודל $m\times m$ ויהיו λ_1 ו־ λ_2 שני ערכים עצמיים שונים של A המתאימים לוקטורים העצמיים x_1,x_2. אזי $x_1\perp x_2$.
<u>הוכחה</u>: **תרגיל**

10. פירוק ספקטראלי

נזכיר כאן תכונה של מטריצות סימטריות ללא הוכחה:

כל מטריצה סימטרית A מגודל $m \times m$ ניתן לרשום באופן $A = U\Omega U^T$ כאשר Ω היא מטריצה אלכסונית שאברי האלכסון שלה הם הערכים העצמיים של A:

$$\Omega = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{pmatrix}$$

ו־U היא מטריצת הוקטורים העצמיים בהתאמה:

$$U = \begin{pmatrix} | & & | \\ u_1 & \ldots & u_m \\ | & & | \end{pmatrix}$$

11. מטריצה חיובית וחיובית למחצה

<u>הגדרה:</u> מטריצה סימטרית A נקראית חיובית לחלוטין (Positive Definite) אם לכל $x \neq 0$, $x^T Ax > 0$.

מטריצה סימטרית A נקראית חיובית למחצה (Positive Semidefinite) אם לכל x, $x^T Ax \geq 0$.

<u>טענה:</u> A מטריצה סימטרית מגודל $m \times m$ ויהיו $\lambda_1, \ldots, \lambda_m$ ערכיה העצמיים. אזי

א. A חיובית לחלוטין אם ורק אם לכל k, $\lambda_k > 0$.

ב. A חיובית למחצה אם ורק אם לכל k, $\lambda_k \geq 0$.

<u>הוכחה:</u>

ב. תרגיל.

12. שורש סימטרי של מטריצה חיובית למחצה:

ראינו קודם כי ניתן לייצג מטריצה סימטרית וחיובית למחצה A מגודל $m \times m$ על ידי הפירוק הספקטרלי שלה. כעת, נגדיר $B = U\Omega^{\frac{1}{2}}U^T$, כאשר

$$\Omega^{\frac{1}{2}} = \begin{pmatrix} \lambda_1^{\frac{1}{2}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m^{\frac{1}{2}} \end{pmatrix}$$

נבחין כי

$$\begin{aligned} B^2 &= \left(U\Omega^{\frac{1}{2}}U^T\right)\left(U\Omega^{\frac{1}{2}}U^T\right) \\ &= U\Omega^{\frac{1}{2}}U^TU\Omega^{\frac{1}{2}}U^T = U\Omega^{\frac{1}{2}}I\Omega^{\frac{1}{2}}U^T \\ &= U\Omega U^T = A \end{aligned}$$

כלומר, B הוא שורש של המטריצה A.

13. 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 $y = (y_1, \dots y_m)^T$ לפי וקטור $x = (x_1, \dots x_n)^T$:

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

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:

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

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

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

ו. נגזרת מכפלה של מטריצה בשני וקטורים שונים לפי המטריצה:

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