סיכום רגרסיה לינארית

14 בינואר 2026

גיא יער־און

קיבלנו שאלה טכנית ברגרסיה לינארית. כיצד פועלים? לפי השלבים הבאים:
נניח כי יש לנו n תצפיות וk משתנים מסבירים (למשל: גיל, גובה, משקל).

1. סדר את הנקודות בצורה של (x_i, y_i)
2. מטריצת המשתנים המסבירים (X):

א. זו מטריצה בגודל $n \times (k+1)$:
ב. k העמודות הראשונות ־ ערכי המשתנים
ג. העמודה האחרונה ־ עמודה של אחדות, מייצגת את החיתוך b כך שלא תמיד $f(0) = 0$.

3. וקטור המטרה (Y): וקטור עמודה בגודל $n \times 1$ שמכיל את הערכים שאנחנו רוצים לחזות: $(y_1, ..., y_n)$
4. וקטור המקדמים (W): הוקטור אותו אנחנו מחפשים הוא יכיל את השיפועים ובשורה האחרונה את השיפוע כלומר:

$$W = \begin{pmatrix} w_1 \\ w_2 \\ .. \\ w_k \\ b \end{pmatrix}$$

5. חשב את המטריצה $X^T X$. זו תהיה תמיד מטריצה בגודל $(k+1)(k+1)$. עבור רגרסיה פשוטה $(k = 1)$ נקבל כי:

$$X^T X = \begin{pmatrix} \sum_i x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix}$$

עבור רגרסיה $k = 2$ כלומר מטריצה 3×3 נקבל באשר המשתנים המסבירים הם x_i, z_i:

$$X^T X = \begin{pmatrix} \sum x_i^2 & \sum x_i z_i & \sum x_i \\ \sum x_i z_i & \sum z_i^2 & \sum z_i \\ \sum x_i & \sum z_i & n \end{pmatrix}$$

6. חשב את המטריצה ההופכית: $(X^T X)^{-1}$. חשוב להעזר בנוסחה:

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

7. חשב את המכפלה של X^TY
8. כפול לפי הנוסחה:

$$W = (X^TX)^{-1}X^TY$$

התוצאה היא וקטור, השיפועים בשורות הראשונות ובשורה האחרונה יהיה b.
9. כתוב לבסוף את התוצאה $\hat{y} = w_1x_1 + ... + w_kx_k + b$
10. הערכת איכות הקשר:

מתי? כשישאלו אותנו מהו טיב ההתאמה , או כמה השונות מוסברת או מהי עוצמת הקשר הלינארי? נרצה לדעת כמה הקו באמת מסביר את הנתונים. לשם כך נחשב שלושה סכומים:

א. $SS_{TOT} = \sum(y_i - \overline{y_i})^2$ - כמה ערכי הy המקוריים רחוקים מהממוצע שלהם.

ב. $SS_{REG} = \sum(\hat{y}_i - \overline{y_i})^2$ - כמה הניבויים של הy מהקו רחוקים מהממוצע.

ג.$RSS = SS_{ERR} = \sum(y_i - \hat{y}_i)^2$ - כמה שונות לא הוסברה, המרחק בין הנתונים האמיתיים לבין הקו.

באשר $\overline{y}$ הוא הממוצע של ערכי הy, וכן $\hat{y_i}$ הוא המשערך לאחר שהצבנו בנוסחת הרגרסיה.
חשב את הערך הבא:

$$R^2 = \frac{SS_{REG}}{SS_{TOT}}$$

אם $R^2 = 0.8$ משמעות הדבר כי המודל מסביר 80% מהשינויים בy. ערך זה הוא שקול למתאם המתאם של פירסון (וב2 משתנים יתן גם את אותה התוצאה) ומבצע מעין הכללה שלו. כלומר ברגרסיה פשוטה $R^2 = r^2$.

11. בדיקת מובהקות במבחן t:

א. בתחילה, נרצה להעריך כמה רעש יש במערכת. לשם כך נקח את הRSS שחישבנו ונחלק אותו במס' דרגות החופש:

$$\hat{\sigma^2} = \frac{RSS}{n-k}$$

מדובר בערך סקלרי. כעת נכפול את הערך הנ"ל במטריצה $(X^TX)^{-1}$ שכבר חישבנו. קיבלנו מטריצה חדשה: היא נקראת מטריצת השנויות והקו וריאנס של המקדמים ונסמנה בC:

$$C = \hat{\sigma^2} \times (X^TX)^{-1}$$

איפה נמצא הSE (סטיית התקן)? על האלכסון הראשי של המטריצה. הSE יהיה השורש של האיברים הללו. כלומר $SE(w_1) = \sqrt{C_{11}}$ וכן הלאה.

ב. השלב האחרון: לכל מקדם מחשבים את:

$$\frac{\hat{w}}{SE(w)}$$

משווים לערך הקריטי בטבלת t: הולכים לטבלת התפלגות t עם רמת מובהקות ומקדם דרגות חופש $n-k$.

אם הערך שחישבנו גדול מהערך בטבלה - דוחים את השערת האפס. המשתנה מובהק. אחרת: לא דוחים.