Q1

Monday, November 15, 2021 10:01 AM

שאלה 1

בשאלה זו נשלים את הוכחת השקילות בין ההצגה של רגרסיה לינארית פשוטה לבין פרספטרון לינארי.
נניח שנתונות לנו נקודות $\{(x^\mu, y^\mu)\}_{\mu=1}^{P}$, ואנו רוצים לתאר את הקשר בין x ל y כפונקציה לינארית $\hat{y} = ax + b$.
למדתם בקורס סטטיסטיקה או שיטות מחקר (וכן ראינו בתרגול) שהערכים שממזערים את השגיאה הריבועית (סכום ריבועי ההפרשים בין $\hat{y}^\mu$ ל y^μ) הם:

$$b^* = \bar{y} - a^* \bar{x}$$

$$a^* = \frac{\mathrm{Cov}[x,y]}{\mathrm{Var}[x]} = r_{xy}\frac{s_y}{s_x}$$

כאשר $\bar{x}$ ו $\bar{y}$ הם הממוצעים. r_{xy} הוא מקדם המתאם של פרסון. s_y, s_x הם סטיות התקן.

הניחו כעת שכל דוגמא חד־מימדית ממופה לוקטור דו־מימדי באופן הבא: $x^\mu \mapsto \begin{bmatrix} x^\mu \\ 1 \end{bmatrix}$ (וערכי ה y^μ לא משתנים).

הראו כי הפתרון שמתקבל בשיטת היפוך המטריצה שלמדתם בשיעור, $w^* = C^{-1}u$ זהה לפתרון המוצג לעיל במונחי a^*, b^*. כלומר, הוכיחו כי מתקיים $w_1^* = a^*$ ו $w_2^* = b^*$. (כאשר הכוכבית מסמלת את הפתרון האופטימלי - הממזער את השגיאה)

אנו צריכים להוכיח בשאלה זו כי $w_1^* = a^*$! $w_2^* = b^*$

אנו יודעים כי: לכל $i,j \in \{1,2\}$ $C_{ij} = \frac{1}{p}\sum_{\mu=1}^{p} x_i^\mu x_j^\mu$

נחשב את C:

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{p}\sum_{\mu=1}^{p} x_1^{\mu^2} & \frac{1}{p}\sum_{\mu=1}^{p} x_1^\mu \cdot 1 \\ \frac{1}{p}\sum_{\mu=1}^{p} 1\cdot x_1^\mu & \frac{1}{p}\sum_{\mu} 1^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{p}\sum_{\mu=1}^{p} x_1^{\mu^2} & \frac{1}{p}\sum_{\mu=1}^{p} x_1^\mu \\ \frac{1}{p}\sum_{\mu=1}^{p} x_1^\mu & \frac{1}{p}\cdot p \end{bmatrix} = \begin{bmatrix} \overline{x^2} & \bar{x} \\ \bar{x} & 1 \end{bmatrix}$$

כעת נחשב את ההפכי: C^{-1}

$$C^{-1} = \frac{1}{c_{11}\cdot c_{22} - c_{12}^2}\begin{bmatrix} c_{22} & -c_{12} \\ -c_{12} & c_{11} \end{bmatrix} = \frac{1}{\overline{x^2} - \bar{x}^2}\cdot\begin{bmatrix} 1 & -\bar{x} \\ -\bar{x} & \overline{x^2} \end{bmatrix}$$

כעת, נחשב את $\vec{u}$:

$$\vec{u} = \begin{bmatrix} \frac{1}{p}\cdot\sum_{\mu=1}^{p} x^\mu\cdot y^\mu \\ \frac{1}{p}\cdot\sum_{\mu=1}^{p} 1\cdot y^\mu \end{bmatrix} = \frac{1}{p}\begin{bmatrix} \sum_{\mu=1}^{p} x^\mu\cdot y^\mu \\ \sum_{\mu=1}^{p} y^\mu \end{bmatrix}$$

ראינו בכיתה כי בשביל למצוא את $w^* = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ עלינו לפתור $\vec{w} = C^{-1}\vec{u}$. ולכן:

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{\overline{x^2} - \bar{x}^2}\begin{bmatrix} 1 & -\bar{x} \\ -\bar{x} & \overline{x^2} \end{bmatrix}\cdot\frac{1}{p}\begin{bmatrix} \sum_{\mu=1}^{p} x^\mu y^\mu \\ \sum_{\mu=1}^{p} y^\mu \end{bmatrix} = \frac{1}{p(\overline{x^2} - \bar{x}^2)}\begin{bmatrix} \sum_{\mu=1}^{p} x^\mu y^\mu - \bar{x}\sum_{\mu=1}^{p} y^\mu \\ -\bar{x}\sum_{\mu=1}^{p} x^\mu\cdot y^\mu + \overline{x^2}\sum_{\mu=1}^{p} y^\mu \end{bmatrix}$$

ולכן:

$$w_1 = \frac{\sum_{\mu=1}^{p} x^\mu y^\mu - \bar{x}\sum_{\mu=1}^{p} y^\mu}{p(\overline{x^2} - \bar{x}^2)} = \frac{\frac{1}{p}\sum_{\mu=1}^{p} x^\mu y^\mu - \bar{x}\frac{1}{p}\sum_{\mu=1}^{p} y^\mu}{\overline{x^2} - \bar{x}^2} = \frac{\overline{x\cdot y} - \bar{x}\cdot\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\mathrm{cov}(X,Y)}{\mathrm{var}(X)} = a^*$$

$$w_2 = \frac{-\bar{x}\sum_{\mu=1}^{p} x^\mu y^\mu + \overline{x^2}\sum_{\mu=1}^{p} y^\mu}{p(\overline{x^2} - \bar{x}^2)} = \frac{-\bar{x}\frac{1}{p}\sum_{\mu=1}^{p} x^\mu y^\mu + \overline{x^2}\frac{1}{p}\sum_{\mu=1}^{p} y^\mu}{\overline{x^2} - \bar{x}^2} = \frac{-\overline{x\cdot y}\cdot\bar{x} + \overline{x^2}\cdot\bar{y}}{\overline{x^2} - \bar{x}^2} =$$

$$\frac{-\overline{x\cdot y}\cdot\bar{x} + \bar{x}^2\cdot\bar{y} - \bar{x}^2\cdot\bar{y} + \overline{x^2}\,\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\bar{x}(-\overline{x\cdot y} + \bar{x}\cdot\bar{y}) + \bar{y}(-\bar{x}^2 + \overline{x^2})}{\overline{x^2} - \bar{x}^2} =$$

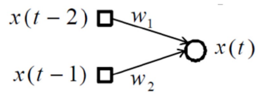
$$\frac{\bar{y}(-\bar{x}^2 + \overline{x^2})}{\overline{x^2} - \bar{x}^2} - \bar{x}\,\frac{\overline{x\cdot y} - \bar{x}\cdot\bar{y}}{\overline{x^2} - \bar{x}^2} = \bar{y} - \bar{x}w_2 = \bar{y} - \bar{x}a^* = b^*$$

Q2

10:45 AM Monday, November 15, 2021

שאלה 2

נתון פרספטרון לינארי המנסה לחזות את ערכו של סיגנל מחזורי חד־מימדי בזמן t, באמצעות הערכים בשני צעדי הזמן הקודמים $t-1$ ו $t-2$. כלומר, הפלט של הפרספטרון בזמן t הוא $x(t)=w_1x(t-2)+w_2x(t-1)$, כמתואר בציור:



בסעיפים הבאים, נתונים שלושה סוגים של הפלט הרצוי מהפרספטרון כפונקציה של הזמן. שימו לב שהסיגנל מחזורי, ונניח שהוא נמשך זמן ממושך. עליכם למצוא בכל סעיף את w_1, w_2 אשר ממזערים את השגיאה הריבועית בין הפלט הרצוי לפלט המצוי, וכן לחשב את ערכה של שגיאת ההכללה.

1. הסיגנל הרצוי: $1,1,1\ldots$ כלומר $y(1)=1$, $y(2)=1$, $y(3)=1$, …
2. הסיגנל הרצוי: $1,2,1,2,1\ldots$ כלומר $y(1)=1$, $y(2)=2$, $y(3)=1$, …
3. הסיגנל הרצוי: $0,1,2,0,1,2\ldots$ כלומר $y(1)=0$, $y(2)=1$, $y(3)=2$, …

מכך שהשגיאה הריבועית היא אי שלילית (סכום של ריבועים בריבוע) נובע כי אם נמצא משקולות שהשגיאה עבורן היא 0 הן בבחירה ממזערות את השגיאה הריבועית ועונה על השאלה. (ככלל הספציפי).

① $w=\begin{bmatrix}0\\1\end{bmatrix}$ מקיים את התנאים.

$\forall t\ \ x(t)=w_1x(t-1)+w_2x(t-2)=x(t-1)+0x(t-2)=x(t-1)=1$

נקבל לכן כי עבור t והפלט המתקבל ושגיאת ההכללה היא 0.

② $w=\begin{bmatrix}1\\0\end{bmatrix}$ מקיים את התנאים.

$\forall t\ \ x(t)=w_1x(t-2)+w_2x(t-1)=2x(t-2)=x(t-2)$

כך מתקיים:

$$\sum_t \left(x(t)-\left\langle w \middle| \begin{bmatrix}x(t-2)\\x(t-1)\end{bmatrix}\right\rangle\right)^2=\sum_t\left(x(t)-\left\langle \begin{bmatrix}1\\0\end{bmatrix} \middle| \begin{bmatrix}x(t-2)\\x(t-1)\end{bmatrix}\right\rangle\right)^2=$$

$$\sum_t (x(t)-x(t-2))^2=0$$

כנדרש

③ בהסתברות $\frac{1}{3}$, יש אחת מהאפשרויות בשוויון מתקיימת:

$$\begin{bmatrix}x(t-2)\\x(t-1)\end{bmatrix}=\begin{cases}\begin{bmatrix}0\\1\end{bmatrix} & x(t)=2 \text{ or } \\ \begin{bmatrix}1\\2\end{bmatrix} & x(t)=0 \text{ or } \\ \begin{bmatrix}2\\0\end{bmatrix} & x(t)=1\end{cases}$$

נחשב את מטריצה C:

$$C=\begin{bmatrix}x(t-2)\\x(t-1)\end{bmatrix}\begin{bmatrix}x(t-2)\\x(t-1)\end{bmatrix}^T=\frac{1}{3}\begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}0&1\end{bmatrix}+\frac{1}{3}\begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}1&2\end{bmatrix}+\frac{1}{3}\begin{bmatrix}2\\0\end{bmatrix}\begin{bmatrix}2&0\end{bmatrix}=$$

$$\frac{1}{3}\begin{bmatrix}0&0\\0&1\end{bmatrix}+\frac{1}{3}\begin{bmatrix}1&2\\2&4\end{bmatrix}+\frac{1}{3}\begin{bmatrix}4&0\\0&0\end{bmatrix}=\frac{1}{3}\begin{bmatrix}5&2\\2&5\end{bmatrix}=\begin{bmatrix}\frac{5}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{5}{3}\end{bmatrix}$$

כעת נחשב את C^{-1}

$$C^{-1}=\frac{1}{\left(\frac{5}{3}\right)^2-\left(\frac{2}{3}\right)^2}\cdot\begin{bmatrix}\frac{5}{3}&-\frac{2}{3}\\-\frac{2}{3}&\frac{5}{3}\end{bmatrix}=\frac{1}{\frac{25}{9}-\frac{4}{9}}\cdot\frac{1}{3}\begin{bmatrix}5&-2\\-2&5\end{bmatrix}=\frac{3}{21}\cdot\begin{bmatrix}5&-2\\-2&5\end{bmatrix}=\frac{1}{7}\begin{bmatrix}5&-2\\-2&5\end{bmatrix}$$

נחשב את W!

$(\frac{5}{3})^2 - (\frac{2}{3})^2 \quad [3 \quad 3]$

נחשב את M:

$$M = \begin{bmatrix} x(t)x(t-2) \\ x(t)x(t-1) \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2\cdot 0 \\ 2\cdot 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1\cdot 2 \\ 1\cdot 0 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 0\cdot 1 \\ 0\cdot 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

ולכן:

$$W = C^{-1}\cdot \vec{M} = \frac{1}{7}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}\cdot \frac{1}{3}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \frac{1}{21}\begin{bmatrix} 2 & 2 \end{bmatrix}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix} = \frac{1}{21}\begin{bmatrix} 10-4 \\ -4+10 \end{bmatrix} = \frac{1}{21}\begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ \frac{2}{7} \end{bmatrix}$$

כעת נחשב את שגיאת ההכללה. נחשב את השגיאה הריבועית לכל מקרה:

$x(t)=2$: $\begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$\left(y - w^T \begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix}\right)^2 = \left(2 - \frac{2}{7}\cdot 0 - \frac{2}{7}\cdot 1\right) = \left(2 - \frac{2}{7}\right)^2 = \left(\frac{12}{7}\right)^2 = \frac{144}{49}$$

$x(t)=1$: $\begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$$\left(y - W^T \begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix}\right)^2 = \left(1 - \frac{2}{7}\cdot 2 - \frac{2}{7}\cdot 0\right) = \left(1 - \frac{4}{7}\right)^2 = \left(\frac{3}{7}\right)^2 = \frac{9}{49}$$

$x(t)=0$: $\begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$$\left(y - w^T \begin{bmatrix} x(t-2) \\ x(t-1) \end{bmatrix}\right)^2 = \left(0 - \frac{2}{7}\cdot 1 - \frac{2}{7}\cdot 2\right) = \left(-\frac{6}{7}\right)^2 = \frac{36}{49}$$

ולכן, שגיאת ההכללה היא:

$$E_g = \frac{1}{2}\left(\frac{1}{3}\cdot\frac{144}{49} + \frac{1}{3}\cdot\frac{9}{49} + \frac{1}{3}\cdot\frac{36}{49}\right) = \frac{1}{6}\cdot\frac{1}{49}(144+9+36) = \frac{189}{6\cdot 49} = \frac{9}{14}$$