תרגיל כיתה 4 – העתקת מאפיינים

שאלה 1

נתונות הדוגמאות הבאות של עצמים שמיוצגחם עם שני מאפיינים :

$$\mathbf{x}_1 = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad \mathbf{x}_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{x}_5 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \mathbf{x}_6 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

א) חשב את העתקת PCA עבור הנתונים הבאים :

ב) העתק את הנתונים לכאלה עם מאפיין יחיד. מהי השגיאה הריבוית הממוצעת ?

פתרון שאלה 1

נסמן ב- m_x את וקטור התוחלת של הנתונים, וב- C_x מטריצת הקווריאנס שלהם.

$$\mathbf{m}_\mathbf{x} = \frac{1}{M}\sum_{k=1}^{M}\mathbf{x}_k$$

$$\mathbf{C}_\mathbf{x} = \frac{1}{M}\sum_{k=1}^{M}\mathbf{x}_k\mathbf{x}_k^T - \mathbf{m}_\mathbf{x}\mathbf{m}_\mathbf{x}{}^T$$

$$\mathbf{m}_\mathbf{x} = \frac{1}{6}\begin{pmatrix} -2-1+0+0+1+2 \\ 0+2+3+1+2+4 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\mathbf{C}_\mathbf{x} = \frac{1}{6}\begin{pmatrix} 4+1+0+0+1+4 & 0-2+0+0+2+8 \\ 0-2+0+0+2+8 & 0+4+9+1+4+16 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 5/3 & 4/3 \\ 4/3 & 5/3 \end{pmatrix}$$

העתקת PCA היא אז :

$$\mathbf{y} = \mathbf{A}(\mathbf{x} - \mathbf{m}_\mathbf{x})$$

כאשר השורות של המטריצה A הן הוקטורים העצמיים של מטריצת הקווריאנס.
הוקטורים העצמיים של מטריצת הקווריאנס מחושבים לפי :

$$\mathbf{C}_\mathbf{x}\mathbf{e}_i = \lambda_i\mathbf{e}_i \quad \Leftrightarrow \quad \begin{pmatrix} 5/3 & 4/3 \\ 4/3 & 5/3 \end{pmatrix}\begin{pmatrix} \mathbf{e}_{i1} \\ \mathbf{e}_{i2} \end{pmatrix} = \lambda_i\begin{pmatrix} \mathbf{e}_{i1} \\ \mathbf{e}_{i2} \end{pmatrix}$$

והערכים העצמיים מחושבים מתוך

$$\text{ig } |\mathbf{C}_\mathbf{x} - \lambda\mathbf{I}| = 0,$$

$$\begin{vmatrix} 5/3-\lambda & 4/3 \\ 4/3 & 5/3-\lambda \end{vmatrix} = \left(\frac{5}{3}\right)^2 - 2\cdot\frac{5}{3}\cdot\lambda + \lambda^2 - \left(\frac{4}{3}\right)^2 = \lambda^2 - \frac{10}{3}\lambda + 1 = 0 \Rightarrow \lambda_1 = 3 \text{ and } \lambda_2 = 1/3.$$

והוקטורים העצמיים המתאימים הם :

$$\Rightarrow \quad \mathbf{e}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

ומטריצת ההעתקה של PCA היא

$$\mathbf{A} = \begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

ב) ההעתקה משני מאפיינים למאפיין אחד היא על ידי שימוש בוקטור עצמי יחיד, זה שמתאים לערך העצמי הגבוה יותר :

$$\mathbf{A} = \mathbf{e}_1^T = (1/\sqrt{2} \quad 1/\sqrt{2})$$

השגיאה הריבועית הממוצעת היא אז :

is $R = \lambda_2 = 1/3.$

הערכים של המאפיין היחיד עבור הנתונים הם :

$$\mathbf{y}_1 = \mathbf{A}(\mathbf{x} - \mathbf{m_x}) = \mathbf{e}_1^T(\mathbf{x}_1 - \mathbf{m_x}) = (1/\sqrt{2} \quad 1/\sqrt{2}) \begin{pmatrix} -2-0 \\ 0-2 \end{pmatrix} = -2\sqrt{2}$$

$$\mathbf{y}_2 = -\sqrt{2}/2,\ \mathbf{y}_3 = \sqrt{2}/2,\ \mathbf{y}_4 = -\sqrt{2}/2,\ \mathbf{y}_5 = \sqrt{2}/2,\ \mathbf{y}_6 = 2\sqrt{2}.$$

פתרון לשאלה 2

א) התוחלת של וקטורי המאפיינים בשתי הקבוצות משוערת לפי

$$\underline{\mu}_i = \frac{1}{5}\sum_{x\in C_i} \underline{x} \qquad i=1,2$$

ומתקבל

$$\underline{\mu}_1^T=(0.4,\ 0.8) \qquad \underline{\mu}_2^T=(-0.6,\ -1)$$

ב) שערוך מטריצת הקווריאנס הוא לפי

$$\hat{\Sigma}_i = \frac{1}{n_i}\sum_{x\in C_i} (\underline{x}-\underline{\mu}_i)(\underline{x}-\underline{\mu}_i)^T \qquad i=1,2$$

מאחר שמטריצת הקווריאנס זהה בשתי הקבוצות, נשערך אותה כממוצע שתי מטריצות הקווריאנס:

$$\hat{\Sigma} = \frac{1}{2}(\hat{\Sigma}_1+\hat{\Sigma}_2) = \frac{1}{10}\left[\sum_{x\in C_1}(\underline{x}-\underline{\mu}_1)(\underline{x}-\underline{\mu}_1)^T + \sum_{x\in C_2}(\underline{x}-\underline{\mu}_2)(\underline{x}-\underline{\mu}_2)^T\right]$$

$$\hat{\Sigma} = \begin{pmatrix} 0.54 & -0.35 \\ -0.35 & 0.68 \end{pmatrix}$$

ג) את הערכים העצמיים של $\hat{\Sigma}$ נמצא על ידי פתרון:

$$\det(\hat{\Sigma}-\lambda I) = 0 = \begin{vmatrix} 0.54-\lambda & -0.35 \\ -0.35 & 0.68-\lambda \end{vmatrix}$$

או

$$(0.54-\lambda)(0.68-\lambda) - 0.35^2 = 0$$

$$\lambda_1 = 0.967 \qquad \lambda_2 = 0.253$$

ג) את הוקטורים העצמיים מוצאים על ידי פתרון

$$\hat{\Sigma}_x \underline{a}_i = \lambda_i \underline{a}_i$$

עבור הערך העצמי $\lambda_1 = 0.967$ מתקבל

$$(\hat{\Sigma}_x - \lambda_1 I)\underline{a}_1 = 0$$

$$\begin{pmatrix} 0.54-\lambda_1 & -0.35 \\ -0.35 & 0.68-\lambda_1 \end{pmatrix}\begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} = 0$$

$$\left.\begin{aligned} -0.427\,a_{11} - 0.35\,a_{12} &= 0 \\ -0.35\,a_{11} - 0.287\,a_{12} &= 0 \end{aligned}\right\}$$

שתי המשוואות הן תלויות ודי לפתור אחת מהן. לפתרון משוואה אחת בשני נעלמים, נבחר ערך לאחד, למשל $a_{12}=1$ ואז

$$a_{11} = 0.35/-0.427 = -0.82$$

כלומר $\underline{a}_1^T = (-0.82,\ 1)$

כל הוקטורים העצמיים האחרים המתאימים ל-λ_1 הם כפל קבוע של a_1^T לעיל.

באופן דומה מתקבל עבור הערך העצמי $\lambda_2 = 0.253$ הוקטור העצמי

$$\underline{a}_2 = (1\ ,\ 0.82)^T$$

אחר נרמול, מתקבלים הוקטורים העצמיים

$$a_1^T = (-0.82,\ 1)^T/(1+0.82^2)^{1/2} = (-0.634\ ,\ 0.7733)^T$$
$$a_2^T = (1\ ,\ 0.82)^T/(1+0.82^2)^{1/2} = (0.7733\ ,\ 0.634)^T$$

הוקטורים העצמיים שהתקבלו הם אורתוגונליים, כפי שצריך
במקרה של מטריצה סימטרית ($\hat{\Sigma}_x$). כלומר:

$$\underline{a}_1^T \cdot \underline{a}_2 = 0$$

מטריצת ההעתקה, בה נערוך את וקטורי האופייניים הנתונים לפי
סדר קטנות (משמאל לימין סדר קטנות):

$$A = \begin{pmatrix} \underline{a}_1 & \underline{a}_2 \end{pmatrix} = \begin{pmatrix} -0.634 & 0.7733 \\ 0.7733 & 0.634 \end{pmatrix}$$

ה) אפשר לראות שמטריצת הקווריאנס של הוקטור המוצע y
היא אלכסונית בשני הרכיבים. בדרך הראשונה נציג דוגמא.

$$\Sigma_y = A \Sigma_x A^T$$

$$\Sigma_y = \begin{pmatrix} -0.634 & 0.7733 \\ 0.7733 & 0.634 \end{pmatrix} \begin{pmatrix} 0.54 & -0.35 \\ -0.35 & 0.68 \end{pmatrix} \begin{pmatrix} -0.634 & 0.7733 \\ 0.7733 & 0.634 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.967 & 0 \\ 0 & 0.253 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \sigma_{y_1}^2 & 0 \\ 0 & \sigma_{y_2}^2 \end{pmatrix}$$

כלומר, השונות של מאפיין y_1 היא λ_1 ושל מאפיין y_2 היא λ_2.

תוצאה דומה מתקבלת גם ללא הצגת מספרים.

משוואת הוקטור/ערך עצמי היא

$$\Sigma_x \underline{a}_i = \lambda_i \underline{a}_i$$

ואם מסדרים את המשוואות לכל i מקבלים

$$\Sigma_x A = A \Lambda$$

כאשר Λ היא מטריצה אלכסונית עם ערכי λ_i על האלכסון.

אחר כפל ב-A^{-1} של שני האגפים מקבל

$$A^{-1}\Sigma_x A = A^{-1}A\Lambda = \Lambda$$

מאחר שעמודות A הן אורתונורמליות, מתקיים $A^T = A^{-1}$

ולכן
$$\Sigma_y = A^T \Sigma_x A = \Lambda$$

ובמסקנה: המשתנים של וקטור המשתנים החדש $\underline{y}$ הם חסרי קורלציה, והשונויות שלהם הן הערכים העצמיים של מטריצת הקווריאנס המקורית Σ_x.

1) ההטלה של וקטורי הדוגמאות על ציר $\underline{a}_1$ מתקבלת לפי

$$y_1 = \underline{a}_1^T \underline{x} = a_{11}x_1 + a_{12}x_2 = -0.634x_1 + 0.7733x_2$$

ובאופן דומה, ההטלה על ציר $\underline{a}_2$ של הדוגמאות מתקבלת לפי

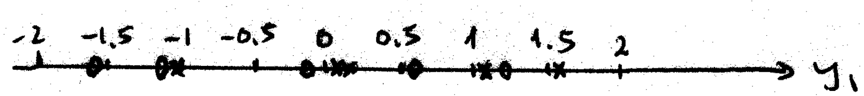
$$y_2 = a_2^T \underline{x} = a_{21}x_1 + a_{22}x_2 = 0.7733x_1 + 0.634x_2$$

הקבוצה הראשונה מתקבלים ערכי y_1:

$$\{1.09,\ 1.5466,\ 0.1393,\ 0.07,\ -1.02\}$$

ובקבוצה השניה ערכי y_1 הם:

$$\{1.268,\ -0.07,\ 0.526,\ -1.09,\ -1.5466\}$$



x - קבוצה אחת

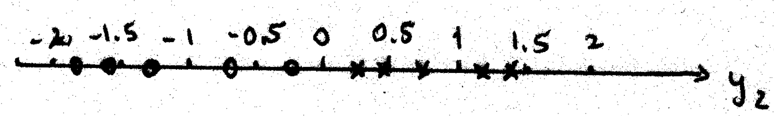
o - קבוצה שניה

בקבוצה הראשונה ערכי y_2 הם:

$$\{0.2474,\ 1.268,\ 1.407,\ 0.704,\ 0.456\}$$

בקבוצה השניה ערכי y_2 הם:

$$\{-1.5466,\ -0.704,\ -1.724,\ -0.2474,\ -1.268\}$$



x - קבוצה אחת

o - קבוצה שניה

בגרף של פילוג התוצאות של מאפיין y_1 אין הפרדה בין הקבוצות. לעומת זאת, ההפרדה בין הקבוצות עבור מאפיין y_2 די טובה. תוצאה זו עולה בקנה אחד עם ההנחיות הכלליות: בוחרים את המאפיינים עם הערכים העצמיים (שונות) הקטנים ביותר.

ג) יש לבחור במאפיין y_2 כדי לשמור על הפרדה טובה בין הקבוצות. יש לבחור במאפיין y_1 כדי לשמור על שגיאה מינימלית ביחס לנתונים המקוריים.

ד) מדד ההפרדה עבור מאפיין יחיד ושתי קבוצות הוא

$$J = \frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}$$

μ_1 ו-μ_2 הן הממוצעים של המאפיין בשתי הקבוצות, ו-σ_1^2, σ_2^2 הן השונויות של שתי הקבוצות.

בסע' ג: מבקשים לבחור אחד משני המאפיינים המוגדרים - x_1 ו-x_2 -
זה שלגביו מדד ההפרדה גדול יותר בין שתי הקבוצות.

עבור המאפיין x_1:

$$\hat{\mu}_1 = \frac{-0.5+0+1+0.5+1}{5} = 0.4$$

$$\hat{\mu}_2 = \frac{-2-0.5-1+0.5+0}{5} = -0.6$$

$$\hat{\sigma}_1^2 = \frac{1}{5}\sum(x_1-0.4)^2 = \frac{0.9^2+0.4^2+0.6^2+0.1^2+0.6^2}{5} = 0.34$$

$$\hat{\sigma}_2^2 = \frac{1}{5}\sum(x_1+0.6)^2 = \frac{1.4^2+0.1^2+0.4^2+1.1^2\ 0.6^2}{5} = 0.74$$

$$J_{x_1} = \frac{(0.4+0.6)^2}{0.34+0.74} = \frac{1}{1.08}$$

ובאופן דומה עבור מאפיין x_2:

$$\hat{\mu}_1 = 0.8 \qquad \hat{\mu}_2 = -1$$

$$\hat{\sigma}_1^2 = 0.66 \qquad \hat{\sigma}_2^2 = 0.5$$

$$J_{x_2} = \frac{(0.8+1)^2}{0.66+0.5} = \frac{3.24}{1.16} > J_{x_1}$$

מדד ההפרדה גדול יותר עבור מאפיין x_2, ולכן אם נצטרך לבחור באחד משני המאפיינים, נבחר ב-x_2.

ט) העתקת MDA משני מאפיינים למאפיין יחיד נותנת את ההפרדה הטובה ביותר :

$$J(\underline{w}) = \frac{|m_1 - m_2|^2}{\sigma_1^2 + \sigma_2^2} = \frac{\underline{w}^T S_B \underline{w}}{\underline{w}^T S_w \underline{w}}$$

וההעתקה $\underline{w}$ היא הוקטור העצמי של המטריצה $S_w^{-1} S_B$ כאשר

$$S_w = \Sigma_1 + \Sigma_2 \qquad S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

התוחלת של וקטורי המאפיינים בשתי הקבוצות משוערכת לפי

$$\underline{\mu}_i = \frac{1}{5} \sum_{x \in c_i} \underline{x} \qquad i = 1, 2$$

ומקבל

$$\underline{\mu}_1^T = (0.4, 0.8) \qquad \underline{\mu}_2^T = (-0.6, -1)$$

$$S_B = \begin{pmatrix} 1.0 \\ 1.8 \end{pmatrix} (1.0 \quad 1.8) = \begin{pmatrix} 1.0 & 1.8 \\ 1.8 & 1.0 \end{pmatrix}$$

ב. שערוך מטריצת הקווריאנס הוא לפי

$$\hat{\Sigma}_i = \frac{1}{N_i} \sum_{x \in c_i} (\underline{x} - \underline{\mu}_i)(\underline{x} - \underline{\mu}_i)^T \qquad i = 1, 2$$

$$\hat{\Sigma} = (\hat{\Sigma}_1 + \hat{\Sigma}_2) = \frac{1}{5} \left[\sum_{x \in c_1} (\underline{x} - \underline{\mu}_1)(\underline{x} - \underline{\mu}_1)^T + \sum_{x \in c_2} (\underline{x} - \underline{\mu}_2)(\underline{x} - \underline{\mu}_2)^T \right]$$

$$\hat{\Sigma} = \begin{pmatrix} 0.54 & -0.35 \\ -0.35 & 0.68 \end{pmatrix} *2$$

$$S_W = \hat{\Sigma} = \begin{pmatrix} 1.08 & -0.7 \\ -0.7 & 1.36 \end{pmatrix} \quad S_W^{-1} = \begin{pmatrix} 1.3895 & 0.7152 \\ 0.7152 & 1.1034 \end{pmatrix}$$

$$S_W^{-1} S_B = \begin{pmatrix} 2.6767 & 3.2162 \\ 2.7013 & 2.3907 \end{pmatrix}$$

$$\overline{W} = S_W^{-1}(\overline{\mu}_1 - \overline{\mu}_2) = \begin{pmatrix} 1.3895 & 0.7152 \\ 0.7152 & 1.1034 \end{pmatrix} \begin{pmatrix} 1.0 \\ 1.8 \end{pmatrix} = \begin{pmatrix} 2.6767 \\ 2.7013 \end{pmatrix}$$

וההעתקה היא: $y = \bar{w}^T \bar{x} = 2.6767x_1 + 2.7013x_2$

ערכי המאפיין החדש y בקבוצה 1 הם:

$$y = \{1.3629 \quad 5.4026 \quad 5.3780 \quad 2.6890 \quad 1.3260\}$$

ערכי המאפיין החדש y בקבוצה 2 הם:

$$y = \{-5.3534 \quad -2.6890 \quad -6.7287 \quad -1.3629 \quad -5.4026\}$$

ניתן לראות שהמאפיין היחיד y הוא בעל ערכים שונים באופן בולט בשתי הקבוצות ללא חפיפה ביניהן