

פתרון מטלה 4 – חישוביות וקוגניציה, 6119

1 בדצמבר 2025



שאלה הינה

סעיף א'

שי וקטור $0 \neq u$ ונרצה לתאר את u על ידי בסיס אורתונורמלי $\{v^l\}_{l=1}^N$. נסמן $u = \sum_{l=1}^N a_l v^l$ ובודוק אילו טענות נכונות במקרה זה.
פתרון: שינוי הכוון של u יגרור שינוי מקדמים a_l (תשובות ג'), באופן דומה גם כפל בסקלר של u יגרור גם שינוי מקדמים (תשובות ד'). אם $\sum_{l=1}^N a_l = 1$ או נובע שגם $\sum_{l=1}^N a_l = \langle u, v^l \rangle$ (תשובות ט').

סעיף ב'

תהי מטריצת קורלציה C בעלת שלושה ערכים עצמיים $\lambda_1 = 2\lambda_2 = 3\lambda_3$. נחשב את השונות שתתקבל מהטלה של דוגמה על הווקטור העצמי הגדל ביותר.
פתרון: ראיינו כי הטלה זו מниבה את ערך השונות הגדל ביותר, וכן נוכל להסיק ללא חישוב שהתשובה היא ד'.

שאלה 1

היא מטריצת $C = \mathbb{E}(xx^t) : \mathbb{R}^N \rightarrow \mathbb{R}$ נורון לינארי המקיים $w^t x = y$ כאשר w וקטור משקלות. נניח ש- $0 = \mathbb{E}(x)$ וכן נניח ש- y מטרת הרשת היא למקסם על שונות הפלט תחת אילוץ על w .

סעיף א'

נחשב את $\text{Var}(y) = 1 \|w\|$ כאשר w וקטור עצמי של C . נבין מה הקשר בין התוצאות לבין פתרון אופטימלי ל-PCA.

$$\text{פתרון לפי הגדרת השונות} (\text{Var}(y)) = \mathbb{E}((y - \mathbb{E}(y))^2)$$

$$\text{Var}(y) = \mathbb{E}(y^2) - 2(\mathbb{E}(y))^2 + (\mathbb{E}(y))^2 = \mathbb{E}(y^2) - (\mathbb{E}(y))^2$$

מעבר לחישוב ערכים אליה,

$$\mathbb{E}(y) = \mathbb{E}(w^t x) = w^t \mathbb{E}(x) = 0$$

ישירות מהנתון $0 = \mathbb{E}(x)$, וכן,

$$\mathbb{E}(y^2) = \mathbb{E}((w^t x)^t w^t x) \stackrel{(1)}{=} \mathbb{E}(w^t x x^t w) = w^t C w = w^t \lambda w = \lambda$$

כאשר λ הערך העצמי של הווקטור w ב- C וכאש המעבר (1) הוכח בכיתה. או מצאנו ש- $\lambda = \text{Var}(y)$ בדיק. ככל שהשונות גדולה יותר בקבוצה סגורה ונוכל להסיק שアイידנו פחות אינפורמציה, שכן נצפה שהadol בבחירה w עם ערך עצמי גדול ביותר מאשר תניב את איבוד המידע הקטן ביותר, וקיבלנו את בדיק אופן הפעולה של מזעור השגיאה ב-PCA.

סעיף ב'

נניח ש- $2 = N$ וכן $x = (x_1, x_2)$ בלתי-תלויים.

i

נחשב את מטריצת הקורלציה.

פתרון נבחין כי $\text{Cov}(x_1, x_2) = 0$ ולכן עבור $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ נתון כי

$$C = \mathbb{E}(xx^t) = \mathbb{E} \begin{pmatrix} x_1 x_1 & x_1 x_2 \\ x_2 x_1 & x_2 x_2 \end{pmatrix} = \begin{pmatrix} \mathbb{E}(x_1^2) & \mathbb{E}(x_1)\mathbb{E}(x_2) \\ \mathbb{E}(x_1)\mathbb{E}(x_2) & \mathbb{E}(x_2^2) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

ii

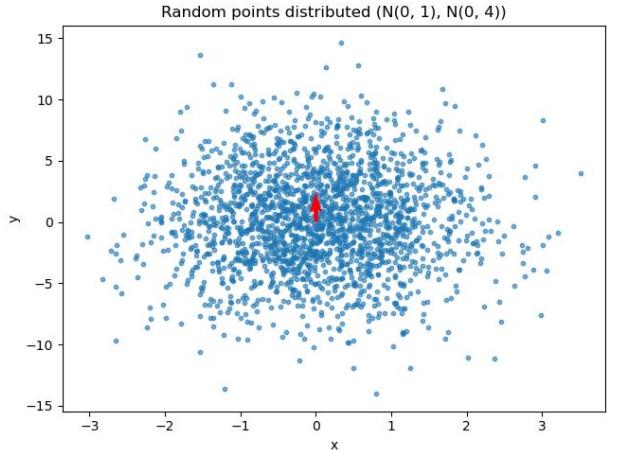
נמצא w אופטימלי תחת האילוץ $\|w\| = 1$.

פתרון קיבלנו ש- $C(0, 1)^t = 4(0, 1)^t$ מчисוב ישיר, ולכן $w = (0, 1)^t$ הוא האופטימלי.

iii

נشرط במשורר דוגמה להתפלגות אופיינית וכן נציג את הכיוון של w .

פתרון נشرط,



אנו יכולים לראות שהמרחק האופקי בין הנקודות הוא גדול מהמרחק האנכי של הנקודות, בכך מתקבל ההיגיון שמאחורי וקטור משקלות אשר תולא את כל ערכו בערך ה- y .

iv

נחשב את אחוז השונות המוסברת.
פתרון הגדרנו את השונות המוסברת על-ידי $5 \text{Var}(y) = w^t C w = 4$ וכן $\text{Var}(x^t x) = \text{trace}(C) = 5$ ולכן אחוז השונות המוסברת הוא .80.

v

נקבע ללא חישוב את וקטור המשקלות האופטימלי במרקם $x_1 \sim N(0, 1), x_2 \sim N(0, 2)$, $x_1 \sim N(0, 4), x_2 \sim N(0, 1)$ ו- $x_1 \sim N(0, 1), x_2 \sim N(0, 2)$ במרקם w בשל סימטריה במקרה שהפער בין הקודמים. במקרה השני נקבע וקטור זהה לווקטור w שהישבנו בסעיפים הקודמים, מאותה סיבה שקיבלנו את האחד שקיבלנו.

vi

נניח עתה ש- $x_1 \sim N(0, 1), x_2 \sim N(0, 1)$ בלתי-תלויים, נחשב את מטריצת הקורלציה ואת ערכיה העצמיים, נבין מה חריג במקרה זה ונחשב את וקטור המשקלות האופטימלי תחת האילוץ $w \in S(0, 1)$.
פתרון עתה נקבע,

$$C = \begin{pmatrix} \mathbb{E}(x_1^2) & 0 \\ 0 & \mathbb{E}(x_2^2) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

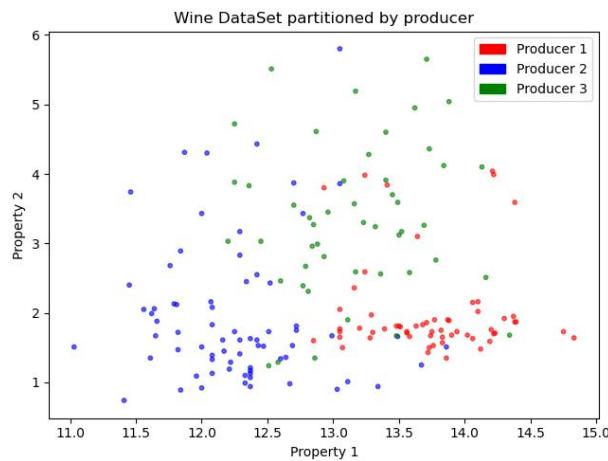
תווך שימוש במת-סעיף. הפעם קיבלנו מטריצה בה כל הערכיהם העצמיים הם 1 ואלכסונית, וכך כל וקטור $w \in S(0, 1)$ יקבל $Cw = 1 \cdot w$.
כלומר נוכל לבחור כל וקטור.

שאלה 2

לאורך התרגיל נעבדו עם dataset שמציג מידע על סוגי יינות המופוצרים על ידי שלושה יצרני ענבים באזורי מסויים באיטליה.

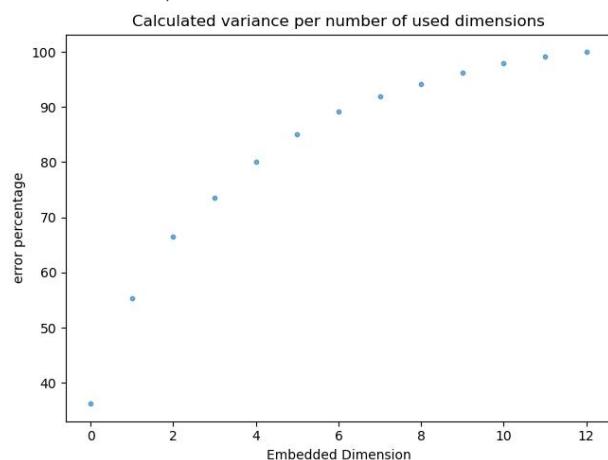
סעיף ב'

מציג את הווקטורים על מישור על-ידי ההעתקה $T(\bar{x}) = (x_1, x_2)$



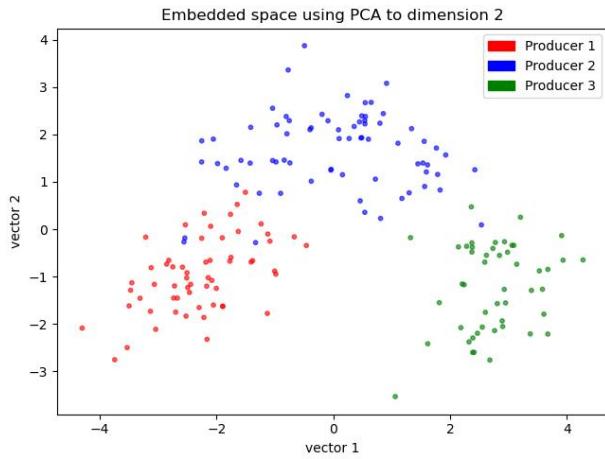
סעיף ג'

מציג גרף של אחוז השונות המצטברת, בציר x יוצג מספר הרכיבים שאנו מטילים עליהם (בהתאם להגדרת PCA) ובציר y מציג את אחוז השונות.



סעיף ד'

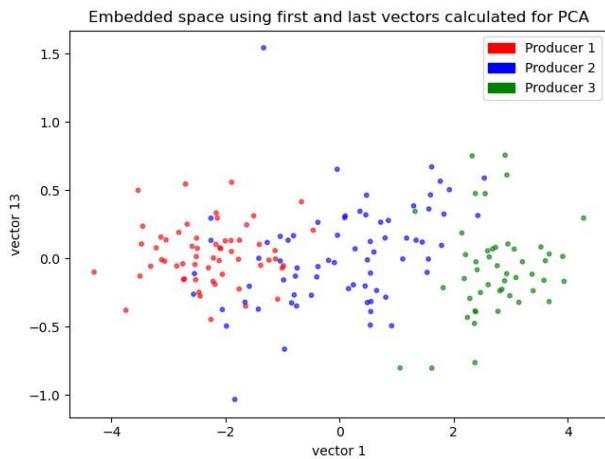
נשרטט את הנקודות כאשר הן מותלות על-ידי שני הרכיבים הראשונים שהישבנו עם PCA עבור הווקטורים.



נבחן כי גרפ זה נראה מהותית שונה מהגרף בסעיף ב', זאת שכן נקודות בצלבים שונים מופיעות באופן יותר צפוף ומרוחק מאשר הנקודות. ככלומר ישائزו מתאם שהצגה זו יוצרת בין יצרן ותכונת היין שלו, תכונות ששימוש בשתי תכונות היין הראשונות (לפי סדר) לא מצליח לתפוס.

סעיף ז'

הפעם נשרטט גרף שמשתמש ברכיב הראשון והאחרון שנחשב עם PCA ונשווה את התוצאות הקודומות,



בזמן שהרכיב הראשון מצליח ליזור בידול ויוזלוי בין הנקודות המיצוגות יין של יצרנים שונים, הרכיב השני משפיע באופן זניח על בידול התוצאות ולמעשה בלתי ניתן לראותה.