

# נושאים באנליזה סטטיסטית מרובת משתנים: תרגיל מס' 1

תאריך הגשה: 23.5.2022

## שאלה 1: קירוב אורתונורמלי של וקטור אקראי

נתון וקטור אקראי  $\mathbf{x} \in \mathbb{R}^p$  המקיים את המשוואה הבאה:

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{w}$$

כאשר  $\mathbf{A} \in \mathbb{R}^{p \times q}$ ,  $p > q$ , היא מטריצה דטרמיניסטית עם עמודות בלתי-תלויות לינאריות,  $\mathbf{s} \in \mathbb{R}^q$  הוא וקטור

אקראי עם תוחלת  $\mathbf{0}$  ומטריצת קווריאנס  $\mathbf{R}_s$  הפיכה ו- $\mathbf{w} \in \mathbb{R}^p$  הוא וקטור אקראי בת"ס ב- $\mathbf{s}$  עם תוחלת  $\mathbf{0}$

וקווריאנס  $\mathbf{R}_w = \sigma_w^2 \mathbf{I}_p$  (מסמנת מטריצת יחידה  $p \times p$ ). נסמן ב- $\mathbf{R}_x$  את מטריצת הקווריאנס של  $\mathbf{x}$  וב-

$\mathcal{R}(\mathbf{A}) \triangleq \{\mathbf{y} = \mathbf{A}\mathbf{b} : \mathbf{b} \in \mathbb{R}^q\}$  את מרחב העמודות של המטריצה  $\mathbf{A}$ . במקרה המדובר, ידוע כי  $q$

מהוקטורים העצמיים של  $\mathbf{R}_x$  שנסמנם ב- $\{\mathbf{v}_k\}_{k=1}^q$  פורשים את  $\mathcal{R}(\mathbf{A})$ . נגדיר את הוקטור האקראי

$\hat{\mathbf{x}} \triangleq \sum_{k=1}^q \beta_k \mathbf{u}_k$  כאשר  $\{\mathbf{u}_k\}_{k=1}^q$  הוא סט כלשהוא של וקטורי בסיס אורתונורמליים הפורש תת-מרחב של  $\mathbb{R}^p$

$$\text{ו-} \beta_k \triangleq \mathbf{u}_k^T \mathbf{x}. \text{ בנוסף, נגדיר את היחסים } r_1 \triangleq \frac{E[\|\hat{\mathbf{x}}\|^2]}{E[\|\mathbf{x}\|^2]} \text{ ו-} r_2 \triangleq \frac{E[\|\mathbf{A}\mathbf{s}\|^2]}{E[\|\mathbf{w}\|^2]}.$$

א. מצאו ביטוי מפורש עבור  $r_1$  בתלות ב- $\mathbf{R}_s$ ,  $\mathbf{A}$ , ו- $\{\mathbf{u}_k\}_{k=1}^q$  ו- $\sigma_w^2$ .

ב. מצאו את  $\{\mathbf{u}_k\}_{k=1}^q$  שעבורם  $r_1$  מקבל ערך מקסימאלי.

ג. מצאו ביטוי מפורש בתלות ב- $\mathbf{R}_s$ ,  $\mathbf{A}$  ו- $\sigma_w^2$  עבור  $r_2$ .

ד. מצאו ביטוי מפורש בתלות ב- $r_2$  עבור הערך המקסימאלי של  $r_1$ .

## שאלה 2: דחיסה ושחזור של תמונה באמצעות KLT (PCA)

הורידו מאתר Moodle את התמונה תחת שם הקובץ Castle.JPG.

א. המירו את התמונה ל-GrayScale באמצעות פקודת "rgm2gray.m" MATLAB.

ב. בצעו דחיסה ושחזור של התמונה באמצעות הגרסא האמפירית של KLT פעם עבור יחס דחיסה  $\rho_1 = 0.95$  ופעם נוספת עבור יחס דחיסה  $\rho_2 = 0.65$ . כתבו במפורט את שלבי הדחיסה והשחזור והציגו את התמונות המשוחזרות.

ג. השוו בין איכות השחזור עבור  $\rho_1$  ו- $\rho_2$ .

ד. חשבו את מספר האיברים (המרכיבים את מקדמי ההתמרה ואת הוקטורים העצמיים הנבחרים של מטריצת הקווריאנס האמפירית) שיש לשמור בזיכרון על מנת לשחזר את התמונה בכל אחד מהמקרים והשוו למספר הפיקסלים של התמונה המקורית.

## שאלה 3: חישוב מקדמי קורלציה וכיוונים קנוניים

יהיו  $\mathbf{x} \in \mathbb{R}^p$  ו- $\mathbf{y} \in \mathbb{R}^q$  וקטורים אקראיים עם מטריצות קווריאנס וקרוס-קווריאנס:

$$\boxed{\Sigma_{\mathbf{x}} = (1-g)\mathbf{I}_p + g\mathbf{e}_1\mathbf{e}_1^T} \quad \boxed{\Sigma_{\mathbf{y}} = (1-h)\mathbf{I}_q + h\mathbf{e}_2\mathbf{e}_2^T} \quad \boxed{\Sigma_{\mathbf{xy}} = s\mathbf{e}_1\mathbf{e}_2^T}$$

בהתאמה, כאשר  $1 > g > (1-p)^{-1}$ ,  $1 > h > (1-q)^{-1}$ ,  $\|\mathbf{e}_1\| = \|\mathbf{e}_2\| = 1$  ו- $s$  הוא קבוע.

א. חשבו את מקדמי הקורלציה ואת כיווני הקורלציה הקנוניים עבור הבעיה הנתונה.

ב. מהו תחום הערכים שהקבוע  $s$  יכול לקבל?

#### שאלה 4: ICA תחת מקורות לא-סטציונאריים

נתון מודל הערבוב הבא:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad t \in \mathbb{R}$$

כאשר  $\mathbf{x}(t) \in \mathbb{R}^p$  הוא וקטור תצפיות בנקודת זמן  $t$ ,  $\{\mathbf{s}(t) \triangleq [s_1(t), \dots, s_p(t)]^T\}$  הוא תהליך אקראי חבוי

לא-סטציונארי עם קווריאנס משתנה בזמן ו- $\mathbf{A} \in \mathbb{R}^{p \times p}$  היא מטריצת ערבוב לא-יוניטרית הפיכה לא ידועה. מניחים כי בכל נקודת זמן  $t$  המקורות  $s_1(t), \dots, s_p(t)$  הם בת"ס.

א. הוכיחו כי מטריצת הערבוב  $\mathbf{A}$  מקיימת:

$$\Sigma_{\mathbf{x}}(t_1) \Sigma_{\mathbf{x}}^{-1}(t_2) \mathbf{A} = \mathbf{A} \Sigma_{\mathbf{s}}(t_1) \Sigma_{\mathbf{s}}^{-1}(t_2)$$

ב. תחת אילו תנאים ניתן לזהות ביחידות את מטריצת הערבוב  $\mathbf{A}$  (עד כדי סימן ופרמוטציה של עמודות) באמצעות פיתרון המשוואה לע"ל.

ג. השתמשו בתוצאה של סעיפים א' ו-ב' על מנת לתכנן אלגוריתם ICA להפרדת מקורות סטציונאריים למקוטעין באינטרוולי זמן של  $T$  שניות.