תרגיל בית 4:

<u>מגישים:</u>

-אלחנדרו מוסקוס 332336908

-שירז בנייטוב 313592958

שאלה 1.1:

סעיף 1.

$$\Sigma_y = V^T \Sigma_x V$$

$$\Sigma_y = E\left(yy^T\right) = E\left(\left(u_i^T x + a_i\right)\left(u_i^T x + a_i\right)^T\right) = E\left[\begin{array}{c}\left(u_i^T x - u_i^T \mu_x + a_i + u_i^T \mu_x\right)\cdot \\ \left(u_i^T x - u_i^T \mu_x + a_i + u_i^T \mu_x\right)^T\end{array}\right] =$$

$$= E\left[\left(u_i^T\left(x-\mu_x\right) + a_i + u_i^T\mu_x\right)\left(u_i^T\left(x-\mu_x\right) + a_i + u_i^T\mu_x\right)^T\right] =$$

$$= E\left[\left(u_i^T\left(x-\mu_x\right) + \left(a_i + u_i^T\mu_x\right)\right)\left(\left(x-\mu_x\right)^T u_i + \left(a_i + \mu_x u_i\right)\right)\right] =$$

$$= E\left[\left(\begin{array}{l} u_i^T\left(x-\mu_x\right)\left(x-\mu_x\right)^T u_i + u_i^T\left(x-\mu_x\right)\left(a_i + \mu_x u_i\right) \\ +\left(a_i + \mu_x u_i\right)\left(x-\mu_x\right)^T u_i + \left(a_i + u_i^T\mu_x\right)\left(a_i + \mu_x u_i\right)\end{array}\right)\right] =$$

$$= \left[\begin{array}{l} u_i^T E\left[\left(x-\mu_x\right)\left(x-\mu_x\right)^T\right] u_i + E\left[u_i^T\left(x-\mu_x\right)\left(a_i^T + \mu_x u_i\right)\right] + \\ +E\left[\left(a_i + \mu_x u_i\right)\left(x-\mu_x\right)^T\right] u_i + \left(a_i + u_i^T\mu_x\right)\left(a_i^T + \mu_x u_i\right)\end{array}\right] =$$

$$= \left[\begin{array}{l} u_i^T\Sigma_x u_i + u_i^T \underbrace{E\left[\left(x-\mu_x\right)\right]}_{=\,0}\left(a_i + \mu_x u_i\right) + \\ +\left(a_i + \mu_x u_i\right)\underbrace{E\left[\left(x-\mu_x\right)^T\right]}_{=\,0} u_i + \left(a_i + u_i^T\mu_x\right)\left(a_i^T + \mu_x u_i\right)\end{array}\right]\Bigg|_{a_i = -u_i^T\mu_x} =$$

$$= u_i^T \Sigma_x u_i$$

$$\Rightarrow V = \left[\begin{pmatrix} u_1 \\ \dots \\ u_D \end{pmatrix}_1, \begin{pmatrix} u_1 \\ \dots \\ u_D \end{pmatrix}_2, \dots, \begin{pmatrix} u_1 \\ \dots \\ u_D \end{pmatrix}_d\right]$$

סעיף 2.

$$a_i = -u_i^T \mu_x \Leftrightarrow$$

$$a_i + u_i^T \mu_x = 0 \Rightarrow$$

$$E\left[a_i + u_i^T \mu_x\right] = 0 \Rightarrow$$

$$E\left[y_i\right] = 0$$

סעיף 3.

$$Trace\left(\Sigma_y\right)=\sum_{i=1}^{d}\operatorname{var}\left(y_i\right)=\sum_{i=1}^{d}\lambda_i$$

כלומר סכום הערכים העצמיים שהתקבלו ממטריצת הקווריאנס של X.

סעיף 4.

אנו מניחים שכל הערכים העצמיים $\{\lambda_i\}$ של מטריצת הקווריאנס של X הם שונים כלומר הוקטורים העצמיים אורתוגונליים $\{u_i\}$ ולכן בלתי תלויים ונוכל לבצע לכסון עבור מטריצת הקוואריאנס Σ_x

סעיף 5.

$$u_i=\lambda_i u_i$$

$$\Rightarrow$$

$$u_i=\lambda_i u_i \Sigma_x^{-1}$$

שאלה 1.2:

- עבור geometric PCA המטרה הייתה לחפש ייצוג במימד נמוך יותר $\{y_1,...,y_n\} \in R^d$ כך שיתקיים $x_j = Uy_j + \varepsilon_j \ , \quad j = 1,...,\mathrm{n}$

ולכן רצינו שההפרש בין הייצוג למקור יהיה מינימלי והשתמשנו ב-MSE כדי לתאר את הבעיה

$$\min_{U,\{y_j\}} \sum_{j=1}^{n} \left\| x_j - Uy_j \right\|^2 \ , \ s.t\, U^T U = I_d$$

לעומת זאת עבור statistical PCA אנו מחפשים את הכיוון u_1 שעבור הטלה בכיוון שלו נקבל את השונות המקסימלית כלומר $u_1^* = \arg\max_{u_1} \mathrm{E}\left(u_1^T x\right)^2 \ , \quad s.t.\ \ u_1^T u_1 = 1$

- עבור statistical PCA לא קיים השימוש בMSE מכיוון שבו אנו רוצים למצוא את הכיוון הטוב ביותר להטלה מבחינת שונות, ולא מנסים למצוא קירוב הכי טוב למדידות.

שאלה 2:

סעיף 1.1.

$$k(x,y)=\left(x^T y\right)^n=\begin{pmatrix} {x_1}^n \\ {x_2}^n \\ \dots \\ {x_D}^n \end{pmatrix}^T \cdot \begin{pmatrix} {y_1}^n \\ {y_2}^n \\ \dots \\ {y_D}^n \end{pmatrix}=\phi(x)^T \cdot \phi(y)$$

$$\Rightarrow$$

$$\phi(x)=\begin{pmatrix} {x_1}^n \\ {x_2}^n \\ \dots \\ {x_D}^n \end{pmatrix}=\overline{x}^n$$

$$\Rightarrow$$

$$R^D \rightarrow R^M$$

when dimension of feature space M is D

סעיף 1.2.

$$\phi(x)=\frac{x^n}{\sqrt{n!}}\exp\left[-\frac{x^2}{2\sigma^2}\right], \quad m=0,1,2,\dots$$

$$\Rightarrow$$

$$k(x,y)=<\phi(x),\phi(y)>=\sum_{n=0}^{\infty}\left(\frac{x^n}{\sqrt{n!}}\exp\left[-\frac{x^2}{2\sigma^2}\right]\right)\left(\frac{y^n}{\sqrt{n!}}\exp\left[-\frac{y^2}{2\sigma^2}\right]\right)=$$

$$=\exp\left[-\frac{1}{2\sigma^2}\left(\|x\|^2+\|y\|^2\right)\right]\cdot\sum_{n=0}^{\infty}\frac{(xy)^n}{n!}=\exp\left[-\frac{1}{2\sigma^2}\left(\|x\|^2+\|y\|^2\right)\right]\exp[xy]=$$

$$=\exp\left[-\frac{1}{2\sigma^2}\left(\|x\|^2-2\sigma^2 xy+\|y\|^2\right)\right]=\exp\left[-\frac{1}{2\sigma^2}\left(\|x\|^2-2\sigma^2 x^T y+\|y\|^2\right)\right]=$$

$$=\exp\left[-\frac{1}{2\sigma^2}\left(\|x-y\|^2\right)\right]$$

$$\Rightarrow$$

$$R^D \rightarrow R^M$$

when dimension of feature space M is ∞

סעיף 2.

לאחר הוספת מדידה

$$new\ sample\ x^{(N+1)} \in R^D$$

$$\left(when\ \ x^{(N+1)} = \begin{pmatrix} x_1^{(N+1)} \\ x_2^{(N+1)} \\ \dots \\ x_D^{(N+1)} \end{pmatrix} \right)$$

נרצה למצוא את הנקודה המתאימה $y^{(N+1)} \in R^d$.

- עבור PCA נוכל להשתמש באותה מטריצה $U^T \in R^{d\times D}$ שנמצאה עבור פתרון הבעיה

$$\min_{U,\{y_j\}} \sum_{j=1}^{n} \left\| x_j - Uy_j \right\|^2 \ \ ,\ \ s.t\, U^T U = I_d$$

ולכן נקבל

$$y^{(N+1)} = U^T{}_{[d\times D]} \cdot x^{(N+1)}{}_{[D\times 1]}$$

- עבור KPCA , קודם פתרנו בשימוש ב-

$$y_i = \omega_i^T \left[\tilde{\kappa}(x_1, x), \tilde{\kappa}(x_2, x), ..., \tilde{\kappa}(x_N, x) \right]^T$$

$$\tilde{k}(x,\mathrm{y}) = k(x,y) - \frac{1}{N}\sum_{j=1}^{N} k(x, x_j) - \frac{1}{N}\sum_{i=1}^{N} k(x_i, x) + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} k(x_i, y_j)$$

$$\tilde{K} = \Phi^T \Phi = \left[\tilde{k}_{ij} \right]$$

$$when\ \omega_i\ are\ the\ eigenvectors\ of\ \ \tilde{K} -$$

$$\tilde{K}\omega_i = \lambda_i \omega_i \ \ ,\ \ \left\| \omega_i \right\| = \lambda_i^{-1}$$

ולכן כעת בתוספת מדידה, $y_i = \omega_i^T \left[\tilde{\kappa}(x_1, x), \tilde{\kappa}(x_2, x), ..., \tilde{\kappa}(x_N, x), \tilde{\kappa}(x_{N+1}, x) \right]^T$

כאשר גם כל אחד מהאיברים השתנה עקב ההגדרה של $\tilde{k}(x,\mathrm{y})$, אך אם נניח שמדובר ב- $N >> 1$ כך שתוספת של מדידה אחת לסכומים היא זניחה אז נוכל לומר שאפשר להשתמש בחלק מהחישוב הקודם

$$y_i = \omega_i^T \left[\underbrace{\boxed{\tilde{\kappa}(x_1, x), \tilde{\kappa}(x_2, x), ..., \tilde{\kappa}(x_N, x)}}_{from\ previous\ calculation} , \tilde{\kappa}(x_{N+1}, x) \right]^T$$

ובנוגע ל- ω

$$\tilde{K}\omega_i = \lambda_i \omega_i$$

$$\tilde{K} = \left[\tilde{k}_{ij} \right] \ ,\ \ i, j \in \{1, 2, ..., \mathrm{N}, N+1\}$$

כלומר נצטרך רק להוסיף חישוב עבור המדידה החדשה $i = N+1$ שנוספה ובכל שאר החישוב שנעשה עבור $i, j \in \{1, 2, ..., N\}$ נוכל להשתמש שוב.

סעיף 3.

עבור KPCA חישבנו את מטריצת הקרנל הממורכזת $\tilde{K}=\Phi^T\Phi$. אם נבחר גרעין לינארי כלומר

$$k(x,y)=x^T y=\phi(x)^T\phi(y)$$
$$\phi(x)=x$$

אז הגרעין מתנוון ל- $\tilde{K}=XX^T$

ולכן בשלב חיפוש הוקטורים העצמיים הבעיה מתנוונת לאותה הבעיה של PCA רגיל (כאשר חיפשנו הטלה בכיוון שתמקסם את השונות)-

$$\tilde{K}\omega_i=\lambda_i\omega_i$$
$$\Rightarrow$$
$$P_x u_i=\eta_i u_i$$