Assignment 4

Question 5, 6

Kumar Ashutosh - 16D070043
Nisha Brahmankar - 16D070019
Adarsh Kumar - 160110071

5. Consider a set of N vectors $\mathcal{X} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_N}\}$ each in $\mathbb{R}^d$, with average vector $\bar{\boldsymbol{x}}$. We have seen in class that the direction $\boldsymbol{e}$ such that $\sum_{i=1}^{N} \|\boldsymbol{x_i} - \bar{\boldsymbol{x}} - (\boldsymbol{e} \cdot (\boldsymbol{x_i} - \bar{\boldsymbol{x}}))\boldsymbol{e}\|^2$ is minimized, is obtained by maximizing $\boldsymbol{e}^t \boldsymbol{C} \boldsymbol{e}$, where $\boldsymbol{C}$ is the covariance matrix of the vectors in $\mathcal{X}$. This vector $\boldsymbol{e}$ is the eigenvector of matrix $\boldsymbol{C}$ with the highest eigenvalue. Prove that the direction $\boldsymbol{f}$ perpendicular to $\boldsymbol{e}$ for which $\boldsymbol{f}^t \boldsymbol{C} \boldsymbol{f}$ is maximized, is the eigenvector of $\boldsymbol{C}$ with the second highest eigenvalue. For simplicity, assume that all non-zero eigenvalues of $\boldsymbol{C}$ are distinct and that rank$(\boldsymbol{C}) > 2$. [10 points]

We have $\quad J(e) = \sum_{i=1}^{N} \| x_i - \bar{x} - (e\cdot(x_i - \bar{x}))e \|^2$

We obtain e s.t. $e^t C e$ is maximized

Additional constraints

$$f \perp e \quad \text{and} \quad \|f\|^2 = 1$$

Forming Lagrangian,

$$\frac{\partial}{\partial f}\left[f^T S f - \lambda(f^T f - 1) - \mu(e^T f - 0)\right] = 0$$

$$\Rightarrow \quad 2Sf - 2\lambda f - \mu e = 0$$

pre-multiplying by e^T

$$2e^T S f - 2\lambda \underbrace{e^T f}_{=0 \,\because\, f \perp e} - \mu \underbrace{e^T e}_{=1} = 0$$

Also, $\quad e^T S = (S^T e)^T = (Se)^T = \lambda e^T \quad \because S^T = S$

$$\Rightarrow \quad 2\lambda e^T f - 0 - \mu = 0$$

$$\Rightarrow \quad \boxed{\mu = 0} \qquad \Rightarrow \boxed{Sf = \lambda f}$$

Thus f is eigenvector (Let corr. eigenvalue be α)

putting back

$$J(f) = \alpha^2.$$

Clearly, we can't take highest eigenvector since it will not be $\perp e$.

Hence

$$\boxed{\text{f is the eigenvector corresponding to second largest eigenvector}}$$

6. Consider a matrix $\boldsymbol{A}$ of size $m \times n, m \leq n$. Define $\boldsymbol{P} = \boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{Q} = \boldsymbol{A}\boldsymbol{A}^T$. (Note: all matrices, vectors and scalars involved in this question are real-valued).

   (a) Prove that for any vector $\boldsymbol{y}$ with appropriate number of elements, we have $\boldsymbol{y}^t\boldsymbol{P}\boldsymbol{y} \geq 0$. Similarly show that $\boldsymbol{z}^t\boldsymbol{Q}\boldsymbol{z} \geq 0$ for a vector $\boldsymbol{z}$ with appropriate number of elements. Why are the eigenvalues of $\boldsymbol{P}$ and $\boldsymbol{Q}$ non-negative?

We have, $P = A^T A$, $Q = AA^T$

Now, $y^t P y = y^t A^T A y$

$= (y^t A^T)(y^t A^T)^T$ $\quad \because (AB)^T = B^T A^T$

$= \| y^t A^T \|^2$

≥ 0 [Norm of a vector]

Similarly, $z^t Q z = z^t A A^T z$

$= (z^t A)(z^t A)^T$ $\quad \because (AB)^T = B^T A^T$

$= \| z^t A \|^2$

≥ 0

Now, let λ be an eigenvalue of P and v be corr. eigenvector

$\Rightarrow Pv = \lambda v$

$\Rightarrow v^t P v = \lambda v^t v$

$\Rightarrow v^t P v = \lambda \|v\|^2$

We proved above that $v^t P v \geq 0 \; \forall \, v$

$\Rightarrow \lambda \|v\|^2 \geq 0 \Rightarrow \boxed{\lambda \geq 0}$

Similarly, μ be an eigenvalue of Q and w be eigenvector

$Qw = \mu w$

$\Rightarrow w^t Q w = w^t \mu w$

$\Rightarrow w^t Q w = \mu \|w\|^2$

$\Rightarrow \mu \|w\|^2 \geq 0 \Rightarrow \boxed{\mu \geq 0}$

(b) If $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{P}$ with eigenvalue λ, show that $\boldsymbol{Au}$ is an eigenvector of $\boldsymbol{Q}$ with eigenvalue λ. If $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{Q}$ with eigenvalue μ, show that $\boldsymbol{A}^T\boldsymbol{v}$ is an eigenvector of $\boldsymbol{P}$ with eigenvalue μ. What will be the number of elements in $\boldsymbol{u}$ and $\boldsymbol{v}$?

we have $\quad Pu = \lambda u \qquad P = A^TA\ ,\ Q = AA^T$

$$A^TAu = \lambda u$$

pre-multiplying by A on both sides

$$AA^T(Au) = \lambda(Au)$$

$$Q(Au) = \lambda(Au)$$

Hence Au is an eigenvector of Q with eigenvalue λ

Similarly, $\quad Qv = \mu v$

$$AA^Tv = \mu v$$

pre-multiplying by A^T on both sides

$$A^TA(A^Tv) = \mu(A^Tv)$$

$$P(A^Tv) = \mu(A^Tv)$$

Hence A^Tv is an eigenvector of P with eigenvalue μ

Clearly, $\quad u \rightarrow n \times 1$

$$v \rightarrow m \times 1$$

(c) If $\boldsymbol{v}_i$ is an eigenvector of $\boldsymbol{Q}$ and we define $\boldsymbol{u}_i \triangleq \frac{\boldsymbol{A}^T \boldsymbol{v}_i}{\|\boldsymbol{A}^T \boldsymbol{v}_i\|_2}$. Then prove that there will exist some real, non-negative γ_i such that $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i \boldsymbol{v}_i$.

We have

$$Qv_i = \lambda_i v_i \quad \Rightarrow \quad AA^T v_i = \lambda_i v_i \quad - (1)$$

$$u_i = \frac{A^T v_i}{\|A^T v_i\|^2} \quad \Rightarrow \quad Au_i = \frac{AA^T v_i}{\|A^T v_i\|^2}$$

$$\Rightarrow \quad Au_i = \frac{\lambda_i v_i}{\|A^T v_i\|^2}$$

$$\therefore \boxed{Au_i = \gamma_i v_i}$$

where $\gamma_i = \frac{\lambda_i}{\|A^T v_i\|^2}$

To show that $\gamma_i \geq 0$

need to only show $\lambda_i \geq 0$

From $-(1)$

$$AA^T v_i = \lambda_i v_i$$

$$v_i^T AA^T v_i = \lambda_i v_i^T v_i$$

$$\|v_i^T A\|^2 = \lambda_i \|v_i\|^2$$

$$\Rightarrow \lambda_i = \frac{\|v_i^T A\|^2}{\|v_i\|^2} \geq 0$$

$$\therefore \boxed{\gamma_i \geq 0}$$

non-negative γ_i such that $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i \boldsymbol{v}_i$.

(d) It can be shown that $\boldsymbol{u}_i^T \boldsymbol{u}_j = 0$ for $i \neq j$ and likewise $\boldsymbol{v}_i^T \boldsymbol{v}_j = 0$ for $i \neq j$ for correspondingly distinct eigenvalues.[1]. Now, define $\boldsymbol{U} = [\boldsymbol{v}_1|\boldsymbol{v}_2|\boldsymbol{v}_3|...|\boldsymbol{v}_m]$ and $\boldsymbol{V} = [\boldsymbol{u}_1|\boldsymbol{u}_2|\boldsymbol{u}_3|...|\boldsymbol{u}_m]$. Now show that $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{V}^T$ where $\boldsymbol{\Gamma}$ is a diagonal matrix containing the non-negative values $\gamma_1, \gamma_2, ..., \gamma_m$. With this, you have just established the existence of the singular value decomposition of any matrix $\boldsymbol{A}$. This is a key result in linear algebra and it is widely used in image processing, computer vision, computer graphics, statistics, machine learning, numerical analysis, natural language processing and data mining. [5 + 5 + 5 + 5 = 20 points]

We have, $\quad U = [v_1 | v_2 | v_3 | \ldots | v_m]$

so that

$$U\Gamma = [v_1 | v_2 | v_3 | \ldots | v_m] \begin{bmatrix} \gamma_1 & & & \\ & \gamma_2 & & \\ & & \ddots & \\ & & & \gamma_m \end{bmatrix}$$

$$= [\gamma_1 v_1 | \gamma_2 v_2 | \gamma_3 v_3 | \ldots | \gamma_m v_m]$$

$$= [Au_1 | Au_2 | Au_3 | \ldots | Au_m]$$

$$= A [u_1 | u_2 | u_3 | \ldots | u_m]$$

$$= AV \quad \text{— (1)}$$

CLAIM $\quad V$ is orthonormal, ie. $VV^T = I$

PROOF

$$V = [u_1 | u_2 | u_3 | \ldots | u_m]$$

Observe that $\quad u_i = \dfrac{A^T v_i}{\|A^T v_i\|^2} \Rightarrow u_i^T u_j = \dfrac{v_i^T A A^T v_j}{\|A^T v_i\|^2 \, \|A^T v_j\|^2}$

$$= \delta_{ij}$$

Hence $\quad VV^T = I$

From — (1)

$$U\Gamma V^T = AVV^T$$

$$\boxed{U\Gamma V^T = A}$$