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$(\boldsymbol{C}) > 2$.

Answer:
Let $J(f) = f^tCf$, we need to maxmise $J(f)$ with constraints $f^te = 0$ (f is orthogonal to e) and we already know that f is a unit vector hence $f^tf = 1$
So, using lagrange's therorm,
Let $\widetilde{J}(f) = f^tCf - \lambda_1(f^tf - 1) - \lambda_2(e^tf)$, by lagranges property, for J to be maximim under the given constraints $\widetilde{J}$'s derivative wrt to f should be 0,

$$\begin{aligned} \frac{d(\widetilde{J}(f))}{df} &= \frac{d}{df}(f^tCf - \lambda_1(f^tf - 1) - \lambda_2(e^tf)) \\ &= 2Cf - 2\lambda_1 f - \lambda_2 e \end{aligned}$$

Hence $2Cf - 2\lambda_1 f - \lambda_2 e = 0$
Taking right vector multiplation with e^t on both sides

$$\begin{aligned} \implies & e^t(2Cf - 2\lambda_1 f - \lambda_2 e) = 0 \\ & 2e^tCf - 2\lambda_1 e^tf - \lambda_2 e^te = 0 \\ & 2\lambda_e e^tf - 2\lambda_1 e^tf - \lambda_2 e^te = 0 \text{ (Where } \lambda_e \text{ is the eigen value corresponding to the eigen vector } e) \\ & 0 - 0 - 2\lambda_2 = 0 \\ & \lambda_2 = 0 \end{aligned}$$

Hence f is a eigen vector of C, now as f^tCf has to maximised and $f \neq e$, f is the eigen vector cooresponding to the second largest eigen value

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?
Answer:

$$\begin{aligned} y^tPy &= y^tA^tAy \\ &= (Ay)^t(Ay) \because \; Ay \text{ is a column vector, and } v^tv \geq 0 \text{ (also called norm) for every column vector } v \\ &\geq 0 \end{aligned}$$

$$\begin{aligned} z^tQz &= z^tAA^tz \\ &= (A^tz)^t(A^tz) \because \; A^tz \text{ is a column vector, and } v^tv \geq 0 \text{ (also called norm) for every column vector } v \\ &\geq 0 \end{aligned}$$

Consider any eigenvalue of P, let it be e
We know $Pe = \lambda e \implies e^tPe = \lambda e^te \implies e^tPe = \lambda$, since as LHS is always non negative RHS must be non negative, hence λ is non-negative.
Hence P has only non negative eigen values
Similarly Q also have non-negative eigen values

(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 $\boldsymbol{v}$?

Answer:

$$\begin{aligned} & Pu = \lambda u \\ \implies & A^t A u = \lambda u \\ \implies & A A^t A u = \lambda A u \\ \implies & Q(Au) = \lambda (Au) \end{aligned}$$

Hence Au is a eigen vector of Q with eigen value λ

$$\begin{aligned} & Qv = \lambda v \\ \implies & A A^t u = \lambda v \\ \implies & A^t A A^v = \lambda A^t v \\ \implies & P(A^t v) = \lambda (A^t v) \end{aligned}$$

Hence $A^t v$ is a eigen vector of P with eigen value λ
Number of elements in u are n and in v are m

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

Answer:

$$\begin{aligned} Au_i &= A\frac{A^T v_i}{\|A^T v_i\|_2} \\ &= \frac{AA^T v_i}{\|A^T v_i\|_2} \\ &= \frac{Qv_i}{\|A^T v_i\|_2} \\ &= \frac{\lambda_{v_i} v_i}{\|A^T v_i\|_2} \end{aligned}$$

Where λ_{v_i} is the eigen value corresponding to eigen vector v_i
Hence there exits $\gamma_i = \dfrac{\lambda_{v_i}}{\|A^T v_i\|_2}$ and $\gamma_i \geq 0$ (as $\lambda_{v_i} \geq 0$) with $Au = \gamma_i v_i$

(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.. 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 $\boldsymbol{A} = \boldsymbol{U\Gamma V}^T$ 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.

Answer:

Let $K = U^T U$

$$K(i,j) = \begin{cases} 1 & i \neq j \text{ as } v_i^t v_i = 0 \\ 0 & i = j \text{ as } v_i^t v_j = 1 \end{cases}$$

Hence $K = I$, so U orthogonal matix, similarly V is also a orthogonal matrix(as v_i are eigenvectors, hence $v_i^t v_j = 0$ for $i \neq j$ and $v_i^t v_i = 1$

So, $UU^T = I$ and $VV^T = I$

$$\begin{aligned} A &= IAI \\ &= UU^T AVV^T \\ &= U(U^T AV)V^T = U\Theta V^T \end{aligned}$$

Now,

$$\Theta(i,j) = U^T AV = \begin{cases} \gamma_i & i = j \text{ as } v_i^t A u_i = v_i^t \gamma_i v_i = \gamma_i \\ 0 & i \neq j \text{ as } v_i^t A u_j = v_i^t \gamma_i v_j = 0 \end{cases}$$

So $\Theta = \Gamma$

Hence $A = U\Gamma V^T$