CS-663 Assignment 4 Q3

Soham Naha (193079003)
Akshay Bajpai (193079002)
Mohit Agarwala (19307R004)

November 6, 2020

3

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).

- **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?**
- **If $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{P}$ with eigenvalue λ, show that $\boldsymbol{A}\boldsymbol{u}$ 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 $\boldsymbol{v}$?**
- **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 $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i\boldsymbol{v}_i$.**
- **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]**

(a) Given $P = A^TA$ $n*n$ and $Q = AA^T \in \mathbb{R}^{m*m}$, and $y \in \mathbb{R}^{n*1}$, we have (given the correct dimension of real-valued y):

$$y^TPy = y^TA^TAy$$
$$= (Ay)^T(Ay)$$

Now let $Ay = w \in \mathbb{R}^{n*1}$. Hence, we have:

$$(Ay)^T(Ay) = w^Tw = ||w||^2 \geq= 0$$

$$y^TPy \geq 0, \forall y$$

Similarly, for a real-valued $m*1$ vector z, we have (given the correct dimension of real-valued z):

$$z^TQy = z^TAA^Tz$$
$$= (A^Tz)^T(A^Tz)$$

Now let $A^Tz = u \in \mathbb{R}^{m*1}$. Hence, we have:

$$(A^Tz)^T(A^Tz) = u^Tu = ||u||^2 \geq 0$$

$$z^TQz \geq 0, \forall z$$

Proof of non-negative eigenvalues of P and Q: For P , let the non-zero eigen vector be v with corresponding eigenvalue λ:

$$Pv = \lambda v$$

[1]This follows because $\boldsymbol{P}$ and $\boldsymbol{Q}$ are symmetric matrices. Consider $\boldsymbol{P}\boldsymbol{u}_1 = \lambda_1\boldsymbol{u}_1$ and $\boldsymbol{P}\boldsymbol{u}_2 = \lambda_2\boldsymbol{u}_2$. Then $\boldsymbol{u}_2^T\boldsymbol{P}\boldsymbol{u}_1 = \lambda_1\boldsymbol{u}_2^T\boldsymbol{u}_1$. But $\boldsymbol{u}_2^T\boldsymbol{P}\boldsymbol{u}_1$ also equal to $(\boldsymbol{P}^T\boldsymbol{u}_2)^T\boldsymbol{u}_1 = (\boldsymbol{P}\boldsymbol{u}_2)^T\boldsymbol{u}_1 = (\lambda_2\boldsymbol{u}_2)^T\boldsymbol{u}_1 = \lambda_2\boldsymbol{u}_2^T\boldsymbol{u}_1$. Hence $\lambda_2\boldsymbol{u}_2^T\boldsymbol{u}_1 = \lambda_1\boldsymbol{u}_2^T\boldsymbol{u}_1$. Since $\lambda_2 \neq \lambda_1$, we must have $\boldsymbol{u}_2^T\boldsymbol{u}_1 = 0$.

Pre-multiplying both sides by v^T , we get,

$$v^T Pv = \lambda v^T v = \lambda ||v||^2$$

$$\implies \lambda = \frac{v^T Pv}{||v||^2}$$

Now, $||v||^2$ is always positive (by-definition). Also, $v^T Pv \geq 0$ (proved earlier). Hence, value of λ is always non-negative. Since we made no assumptions for v or λ , the condition holds for all eigenvalues of P . Similarly, for Q , let:

$$Qv = \lambda v$$

Pre-multiplying both sides by v^T , we get,

$$v^T Qv = \lambda v^T v = \lambda ||v||^2$$

$$\implies \lambda = \frac{v^T Qv}{||v||^2}$$

Now, $||v||^2$ is always positive (by-definition). Also, $v^T Qv \geq 0$ (proved earlier). Hence, value of λ is always non-negative. Since we made no assumptions for v or λ , the condition holds for all eigenvalues of Q .

(b) u is an eigenvector of P with eigenvalue λ. Now,

$$A \in \mathbb{R}^{m*n} \implies A^T A = P \in \mathbb{R}^{n*n} \implies u \in \mathbb{R}^n and$$

$$Pu = \lambda u$$

Pre-multiplying both sides with A, we get

$$APu = \lambda Au$$

Using $P = A^T A$, we simplify this to,

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

Simplify $AA^T = Q$, and $Au = v \in R^m$ to get,

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

Hence, Au is an eigenvector of Q with eigenvalue λ.
Similarly to prove for Q, we assume an eigenvector v with eigenvalue μ. Now,

$$A \in \mathbb{R}^{m*n} \implies AA^T = Q \in \mathbb{R}^{m*m} \implies v \in \mathbb{R}^m and$$

$$Qv = \mu v$$

Pre-multiplying both sides with A , we get

$$A^T Qv = \mu A^T v$$

Using $Q = AA^T$, we simplify this to,

$$A^T AA^T v = (A^T A)A^T v = \mu A^T v$$

Simplify $A^T A = P$, and $A^T v = w \in R^n$ to get,
Hence, $A^T v$ is an eigenvector of P with eigenvalue μ.

(c) We are given that v_i is an eigenvector of Q (with eigenvector α_i). Which gives us the first equation

$$Qv_i = AA^T v_i = \alpha_i v_i$$

.

Now, given

$$u_i = \frac{A^T v_i}{||A^T v_i||_2}$$

We have

$$Au_i = \frac{AA^T v_i}{||A^T v_i||_2}$$

$$\implies Au_i = \frac{Qv_i}{||A^Tv_i||_2} = \frac{\alpha_i v_i}{||A^Tv_i||_2} = \left(\frac{\alpha_i}{||A^Tv_i||_2}\right) v_i$$

Substitute $\frac{\alpha_i}{||A^Tv_i||_2} = \gamma_i$ to get the required equation.

$$Au_i = \gamma_i v_i$$

Now, α_i is non-negative since it's an eigenvalue of Q (proved earlier), and $||A^Tv_i||_2$ is positive by definition, the value of γ_i i is non-negative. There exists a non-negative γ_i for which the given equation holds.

(d) We know as a result that

$$u_i^T u_j = 0 for i \neq j$$

and for

$$u_i^T u_i = \frac{(A^Tv_i)^T(A^Tv_i)}{||A^Tv_i||^2} = 1$$

So, the matrix $V = [u_1|u_2|...|u_n]$ is orthonormal because $VV^T = V^TV = I_n$ Similarly, for $U = [v_1|v_2|...|v_m], UU^T = U^TU = I_m$ because $v_i^Tv_j = 0$ for $i \neq j$ and $v_i^Tv_j = 1$ (because v_i and v_j are eigen-vectors of P , and which can be assumed to be of unit length for consistency. Given these results, the value of U^TAV

$$= U^TAV = U^TA[u_1|u_2|...|u_n] = \begin{bmatrix} v_1^T \\ v_2^T \\ .. \\ v_m^T \end{bmatrix} A[u_1|u_2|...|u_n]$$

From the previous results, we have $Au_i = \gamma_i v_i$

$$= \begin{bmatrix} v_1^T \\ v_2^T \\ .. \\ v_m^T \end{bmatrix} [\gamma_1 v_1|\gamma_2 v_2|...|\gamma_n v_n]$$

$$= [\Gamma_{ij}^T]_m * n$$

Where

$$\Gamma_i j = v_i^T \gamma_j v_j = \begin{cases} 0 & \text{if } i \neq j \\ \gamma_j & \text{otherwise} \end{cases}$$

Hence, $U^TAV = \Gamma$ where T is a diagonal matrix, with i^th diagonal element $= \gamma_i$

$$\implies (UU^T)A(VV^T) = U\Gamma V^T \implies A = U\Gamma V^T$$

This holds since U and V are orthonormal. Hence, proved.