CS 663 : Digital Image Processing : Assignment 4

Siddharth Saha [170100025], Tezan Sahu [170100035]

Due Date :- 16rd October, 2019

Question 6

Question:
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).

1. 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?

2. 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 $\boldsymbol{v}$?

3. 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 $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i\boldsymbol{v}_i$.

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

Solution:

1. Clearly, $\boldsymbol{y}^t\boldsymbol{P}\boldsymbol{y} = \boldsymbol{y}^t\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{y})^t(\boldsymbol{A}\boldsymbol{y}) = ||\boldsymbol{A}\boldsymbol{y}||_2^2 \geq 0$

   Similarly, $\boldsymbol{z}^t\boldsymbol{Q}\boldsymbol{z} = \boldsymbol{z}^t\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{z} = (\boldsymbol{A}^T\boldsymbol{z})^t(\boldsymbol{A}^T\boldsymbol{z}) = ||\boldsymbol{A}^T\boldsymbol{z}||_2^2 \geq 0$

   Let $\boldsymbol{u}$ be an eigenvector of $\boldsymbol{P}$ with eigenvalue λ. $\boldsymbol{P}\boldsymbol{u} = \lambda\boldsymbol{u}$.

   Thus, $\boldsymbol{u}^t\boldsymbol{P}\boldsymbol{u} = \lambda\boldsymbol{u}^t\boldsymbol{u} = \lambda$ and we showed above that $\boldsymbol{y}^t\boldsymbol{P}\boldsymbol{y} \geq 0$ for any vector $\boldsymbol{u}$ of appropriate dimensions. Therefore, eigenvalues of $\boldsymbol{P}$ are non-negative.

   Similarly, let $\boldsymbol{v}$ be an eigenvector of $\boldsymbol{Q}$ with eigenvalue μ. $\boldsymbol{Q}\boldsymbol{v} = \mu\boldsymbol{v}$.

   Thus, $\boldsymbol{v}^t\boldsymbol{Q}\boldsymbol{v} = \mu\boldsymbol{v}^t\boldsymbol{v} = \mu$ and we showed above that $\boldsymbol{v}^t\boldsymbol{Q}\boldsymbol{v} \geq 0$ for any vector $\boldsymbol{v}$ of appropriate dimensions. Therefore, eigenvalues of $\boldsymbol{Q}$ are non-negative.

   Thus, $\boldsymbol{P}$ & $\boldsymbol{Q}$ are positive semi-definite matrices.

2. Given, $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{P}$ with eigenvalue $\lambda \Rightarrow \boldsymbol{P}\boldsymbol{u} = \lambda\boldsymbol{u} \Rightarrow \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{u}$.
   Pre-multiplying by $\boldsymbol{A}$, we get $(\boldsymbol{A}\boldsymbol{A}^T)(\boldsymbol{A}\boldsymbol{u}) = \lambda(\boldsymbol{A}\boldsymbol{u}) \Rightarrow \boldsymbol{Q}(\boldsymbol{A}\boldsymbol{u}) = \lambda(\boldsymbol{A}\boldsymbol{u})$
   Thus, $\boldsymbol{A}\boldsymbol{u}$ is an eigenvector of $\boldsymbol{Q}$ with eigenvalue λ.

   Also, given $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{Q}$ with eigenvalue $\mu \Rightarrow \boldsymbol{Q}\boldsymbol{v} = \mu\boldsymbol{v} \Rightarrow \boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v} = \mu\boldsymbol{v}$.
   Pre-multiplying by $\boldsymbol{A}^T$, we get $(\boldsymbol{A}^T\boldsymbol{A})(\boldsymbol{A}^T\boldsymbol{v}) = \mu(\boldsymbol{A}^T\boldsymbol{v}) \Rightarrow \boldsymbol{P}(\boldsymbol{A}^T\boldsymbol{v}) = \mu(\boldsymbol{A}^T\boldsymbol{v})$
   Thus, $\boldsymbol{A}^T\boldsymbol{v}$ is an eigenvector of $\boldsymbol{P}$ with eigenvalue μ.

   $\boldsymbol{u}$ has n elements whereas $\boldsymbol{v}$ has m elements.

3. Given, $\boldsymbol{Q}\boldsymbol{v}_i = \mu_i\boldsymbol{v}_i$ where μ_i is the eigenvalue of $\boldsymbol{Q}$ corresponding to the eigenvector $\boldsymbol{v}_i$.

$$\boldsymbol{u}_i \triangleq \frac{\boldsymbol{A}^T\boldsymbol{v}_i}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2}$$
$$\boldsymbol{A}\boldsymbol{u}_i = \frac{\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v}_i}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2} = \frac{\boldsymbol{Q}\boldsymbol{v}_i}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2} = \frac{\mu_i\boldsymbol{v}_i}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2}$$
$$\therefore \boldsymbol{A}\boldsymbol{u}_i = \gamma_i\boldsymbol{v}_i, \text{ where } \gamma_i = \frac{\mu_i}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2}$$

   Now, from part 1, we know that $\mu_i \geq 0$. Also, $\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2$ is the magnitude of vector $\boldsymbol{A}^T\boldsymbol{v}_i$, which cannot be negative. Thus, γ_i is non-negative.

4. From part (3), we know that $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i\boldsymbol{v}_i$.
   $\boldsymbol{A}$ is of size $m \times n$. We have defined $\boldsymbol{U}_{m\times m} = [\boldsymbol{v}_1|\boldsymbol{v}_2|\boldsymbol{v}_3|...|\boldsymbol{v}_m]$ and $\boldsymbol{V}_{n\times n} = [\boldsymbol{u}_1|\boldsymbol{u}_2|\boldsymbol{u}_3|...|\boldsymbol{u}_n]$
   Without loss of generality, we consider $m < n$.

$$\text{For } i \in \{1, 2, ..., m\},\ \boldsymbol{A}\boldsymbol{u}_i = \gamma_i\boldsymbol{v}_i$$
$$\text{For } i \in \{m+1, m+2, ..., n\},\ \boldsymbol{A}\boldsymbol{u}_i = 0$$

   The above statement could be written in matrix form as:

$$\boldsymbol{A}_{m\times n}\boldsymbol{V}_{n\times n} = \boldsymbol{U}_{m\times m}\boldsymbol{\Gamma}_{m\times n}$$

   where $\boldsymbol{\Gamma}$ is a diagonal matrix containing γ_i's as diagonal elements. Clearly, $\boldsymbol{\Gamma}$ has at most m non-zero values. Also, from the arguments of part (3), all elements of $\boldsymbol{\Gamma}$ are non-negative.

   $\boldsymbol{v}_i$'s are eigenvectors of $\boldsymbol{Q}$. So, they are orthogonal toeach other and are of unit magnitude. Now, $\boldsymbol{u}_i$'s are defined such that their magnitude is 1. Also for $i \neq j$,

$$\boldsymbol{u}_i^t\boldsymbol{u}_j = \frac{\boldsymbol{v}_i^t\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v}_i}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2\|\boldsymbol{A}^T\boldsymbol{v}_j\|_2} = \frac{\boldsymbol{v}_i^t\boldsymbol{Q}\boldsymbol{v}_j}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2\|\boldsymbol{A}^T\boldsymbol{v}_j\|_2} = \frac{\mu\boldsymbol{v}_i^t\boldsymbol{v}_i}{\|\boldsymbol{A}^T\boldsymbol{v}_i\|_2\|\boldsymbol{A}^T\boldsymbol{v}_j\|_2} = 0$$

   Thus, $\boldsymbol{u}_i$'s are also orthogonal to each other & are of unit magnitude.
   This makes both $\boldsymbol{U}$ & $\boldsymbol{V}$ orthonormal matrices.
   Pre-multiplying by $\boldsymbol{V}^T$, we get: $\boldsymbol{A}\boldsymbol{V}\boldsymbol{V}^T = \boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{V}^T \Rightarrow \boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{V}^T$
   This represents the Singular Value Decomposition (SVD) for matrix $\boldsymbol{A}_{m\times n}$.