CS663 Assignment 4 Question 3 Report

Shreya Laddha, Archishman Biswas, Shreyan Jabade, Rishabh Arya

November 7, 2020

Q3. 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 y with appropriate number of elements, we have $y^t P y \geq 0$. Similarly show that $z^t Q z \geq 0$ for a vector z with appropriate number of elements. Why are the eigenvalues of P and Q non-negative?

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

(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 $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i \boldsymbol{v}_i$

(d) It can be shown that $\boldsymbol{u}_i^T\boldsymbol{u}_j = 0$ for $\boldsymbol{i} \neq \boldsymbol{j}$ and likewise $\boldsymbol{v}_i^T\boldsymbol{v}_j = 0$ for $\boldsymbol{i} \neq \boldsymbol{j}$ for correspondingly distinct eigenvalues [1]] Now, define $\boldsymbol{U} = [\boldsymbol{v}_1 | \boldsymbol{v}_2 | \boldsymbol{v}_3 | \ldots | \boldsymbol{v}_m]$ and $\boldsymbol{V} = [\boldsymbol{u}_1 | \boldsymbol{u}_2 | \boldsymbol{u}_3 | \ldots | \boldsymbol{u}_m]$. Now show that $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{V}^T$ where $\boldsymbol{\Gamma}$ is a diagonal matrix containing the non-negative values $\gamma_1, \gamma_2, \ldots, \gamma_m$. 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.

Ans.

(a). For any vector $\boldsymbol{y}$, we have $\boldsymbol{y}^t\boldsymbol{P}\boldsymbol{y} = \boldsymbol{y}^t\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{y})^t(\boldsymbol{A}\boldsymbol{y}) = \|\boldsymbol{A}\boldsymbol{y}\|_2^2 \geq 0$

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

Let u be an eigenvector of P with eigenvalue λ. $Pu = \lambda u$. Thus, $\boldsymbol{u}^t\boldsymbol{P}\boldsymbol{u} = \lambda\boldsymbol{u}^t\boldsymbol{u} = \lambda$ and we know that $\boldsymbol{u}^t\boldsymbol{P}\boldsymbol{u} \geq 0$. Therefore, eigenvalues of P are non-negative.

Let v be an eigenvector of Q with eigenvalue μ. $Qv = \mu v$. Thus, $\boldsymbol{v}^t\boldsymbol{Q}\boldsymbol{v} = \mu\boldsymbol{v}^t\boldsymbol{v} = \mu$ and we know that $\boldsymbol{v}^t\boldsymbol{Q}\boldsymbol{v} \geq 0$ for any vector $\boldsymbol{v}$. Thus, eigenvalues of Q are non-negative.

Hence, P&Q are positive semi-definite matrices.

(b). $Pu = \lambda u \Rightarrow A^TAu = \lambda u$.

We get $\left(AA^T\right)(Au) = \lambda(Au) \Rightarrow Q(Au) = \lambda(Au) \Rightarrow Au$ is an eigenvector of Q with eigenvalue λ. Also, $Qv = \mu v \Rightarrow AA^Tv = \mu v$. We can write $\left(\boldsymbol{A}^T\boldsymbol{A}\right)\left(\boldsymbol{A}^T\boldsymbol{v}\right) = \mu\left(\boldsymbol{A}^T\boldsymbol{v}\right) \Rightarrow \boldsymbol{P}\left(\boldsymbol{A}^T\boldsymbol{v}\right) = \mu\left(\boldsymbol{A}^T\boldsymbol{v}\right)$ which means A^Tv is an eigenvector of P with eigenvalue μ.

Thus, u has n elements whereas v has m elements.

(c). μ_i is the eigenvalue of Q corresponding to the eigenvector $v_i \Rightarrow Qv_i = \mu_i v_i$

$$u_i \triangleq \frac{A^T v_i}{\left\|A^T v_i\right\|_2}.$$

Multiply by A, we get $Au_i = \frac{AA^T v_i}{\left\|A^T v_i\right\|_2} = \frac{Q v_i}{\left\|A^T v_i\right\|_2} = \frac{\mu_i v_i}{\left\|A^T v_i\right\|_2}$
$\Rightarrow Au_i = \gamma_i v_i,$ where $\gamma_i = \frac{\mu_i}{\left\|A^T v_i\right\|_2}$.
We know that $\mu_i \geq 0$ (from part (a)) and $\left\|A^T v_i\right\|_2 \geq 0$ (magnitude of vector).
Thus, γ_i is non-negative.

(d). We have, $U_{m\times m} = [v_1\,|v_2|\,v_3|\ldots|v_m]$ and $V_{n\times n} = [u_1\,|u_2|\,u_3|\ldots|u_n]$. From part (c), we know that $Au_i = \gamma_i v_i$. Consider,

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

We can write the above statement as $\boldsymbol{A}_{m\times n}\boldsymbol{V}_{n\times n} = \boldsymbol{U}_{m\times m}\boldsymbol{\Gamma}_{m\times n}$ where Γ is a diagonal matrix containing γ_i 's as diagonal elements. Γ has at most $\boldsymbol{m}$ non-zero values. Also, from the arguments of part (c), all elements of Γ are non-negative.
The v_i 's are eigenvectors of $\boldsymbol{Q}$as given. So, they are orthogonal to each other and of unit magnitude. For $\boldsymbol{i} \neq \boldsymbol{j}$

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

Thus, $\boldsymbol{u}_i$'s are also orthogonal to each other & are of unit magnitude.
This makes both U&V orthonormal matrices. We can write $AVV^T = U\Gamma V^T \Rightarrow A = U\Gamma V^T$ This represents the Singular Value Decomposition for matrix $\boldsymbol{A}_{m\times n}$.