DIP Assignment 5: Question 4

October 22, 2015

1 Question

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

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

1.1 Solution

Substituting for P in $\mathbf{y}^t\mathbf{P}\mathbf{y}$, we get

$$\mathbf{y}^t\mathbf{A}^T\mathbf{A}\mathbf{y}$$

$$= (\mathbf{A}\mathbf{y})^t\mathbf{A}\mathbf{y}$$

$$= \mathbf{b}^t\mathbf{b}$$

where $b = \mathbf{A}\mathbf{y}$. Now, dimension of b is m x 1. So, $\mathbf{b}^t\mathbf{b}$ is a scalar which is nothing but inner product of b with itself, i.e. $< b, b >$.
Thus, $\mathbf{y}^t\mathbf{P}\mathbf{y}$ is sum of square of elements of $\mathbf{A}\mathbf{y}$, which is ought to be non-negative. So,

$$\mathbf{y}^t\mathbf{P}\mathbf{y} \geq 0$$

On similar lines, $\mathbf{z}^t\mathbf{Q}\mathbf{z}$ can be expanded as

$$\mathbf{z}^t\mathbf{A}\mathbf{A}^T\mathbf{z}$$

$$= (\mathbf{A}^T\mathbf{z})^t\mathbf{A}^T\mathbf{z}$$

$$= \mathbf{c}^t\mathbf{c}$$

where $c = \mathbf{A}^T\mathbf{z}$. Now, dimension of c is n x 1. So, $\mathbf{c}^t\mathbf{c}$ is a scalar.
Thus, $\mathbf{z}^t\mathbf{Q}\mathbf{z}$ is sum of square of elements of $\mathbf{A}^T\mathbf{z}$, which is ought to be non-negative. So,

$$\mathbf{z}^t\mathbf{Q}\mathbf{z} \geq 0$$

Say, $\mathbf{u}$ is an eigenvector of $\mathbf{P}$ with eigenvalue λ, then,

$$\mathbf{P}\mathbf{u} = \lambda\mathbf{u}$$

Pre-multiplying both sides by $\mathbf{u}^T$, we have

$$\mathbf{u}^T\mathbf{P}\mathbf{u} = \mathbf{u}^T\lambda\mathbf{u}$$

$$\mathbf{u}^T\mathbf{P}\mathbf{u} = \lambda\mathbf{u}^T\mathbf{u}$$

Now, from the result proved above, LHS is always non-negative. $\mathbf{u}^T\mathbf{u}$ is a scalar which is sum of square of elements of u, and so is non-negative. Thus,

$$\lambda = \frac{\mathbf{u}^T\mathbf{P}\mathbf{u}}{\mathbf{u}^T\mathbf{u}}$$

λ is a product of two non-negative quantities and so is non-negative.
Following the similar arguments for eigenvalues and eigenvectors of $\mathbf{Q}$ and using the result for Q proved above, we can prove that eigenvalues of $\mathbf{Q}$ are always non-negative.

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

1.2 Solution

We have,

$$\mathbf{P}\mathbf{u} = \lambda\mathbf{u}$$

$$\mathbf{A}^T\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$$

Pre-multiplying both sides by $\mathbf{A}$, we get

$$\mathbf{A}\mathbf{A}^T\mathbf{A}\mathbf{u} = \mathbf{A}\lambda\mathbf{u}$$

replacing $\mathbf{A}\mathbf{A}^T$ by Q

$$\mathbf{Q}\mathbf{A}\mathbf{u} = \lambda\mathbf{A}\mathbf{u}$$

where, dimension of $\mathbf{Au}$ is m x 1. Thus, $\mathbf{Au}$ is an eigenvector of $\mathbf{Q}$ with corresponding eigenvalue as λ

On similar lines,

$$\mathbf{Qv} = \mu\mathbf{v}$$

$$\mathbf{AA}^T\mathbf{v} = \mu\mathbf{v}$$

Pre-multiplying both sides by $\mathbf{A}^T$, we get

$$\mathbf{A}^T\mathbf{AA}^T\mathbf{v} = \mathbf{A}^T\mu\mathbf{v}$$

replacing $\mathbf{A}^T\mathbf{A}$ by P

$$\mathbf{PA}^T\mathbf{v} = \mu\mathbf{A}^T\mathbf{v}$$

where, dimension of $\mathbf{A}^T\mathbf{v}$ is n x 1. Thus, $\mathbf{A}^T\mathbf{v}$ is an eigenvector of $\mathbf{P}$ with corresponding eigenvalue as μ.

$\mathbf{u}$ is an eigenvector of n x n matrix $\mathbf{P}$ and so has $\mathbf{n}$ elements.

$\mathbf{v}$ is an eigenvector of m x m matrix $\mathbf{Q}$ and so has $\mathbf{m}$ elements.

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

1.3 Solution

$$\mathbf{u}_i = \frac{\mathbf{A}^T\mathbf{v}_i}{\|\mathbf{A}^T\mathbf{v}_i\|}$$

So,

$$\mathbf{Au}_i = \mathbf{A}\frac{\mathbf{A}^T\mathbf{v}_i}{\|\mathbf{A}^T\mathbf{v}_i\|}$$

replacing $\mathbf{AA}^T$ by $\mathbf{Q}$

$$\mathbf{Au}_i = \mathbf{Q}\frac{\mathbf{v}_i}{\|\mathbf{A}^T\mathbf{v}_i\|}$$

Now, as $\mathbf{v}_i$ is an eigenvector of $\mathbf{Q}$, let λ_i be the corresponding eigenvector. We now have,

$$\mathbf{Au}_i = \frac{\lambda_i}{\|\mathbf{A}^T\mathbf{v}_i\|}\mathbf{v}_i$$

Now, λ_i, being eigenvalue of $\mathbf{Q}$ is non-negative as proved in part 1; and so is $\|\mathbf{A}^T\mathbf{v}_i\|$. So, replacing $\frac{\lambda_i}{\|\mathbf{A}^T\mathbf{v}_i\|}$ by some non-negative and real $\gamma_\mathbf{i}$, we have

$$\mathbf{A}\mathbf{u}_i = \gamma_i\mathbf{v}_i$$

4. It can be shown that $\mathbf{u}_i^T\mathbf{u}_j = 0$ for $i \neq j$ and likewise $\mathbf{v}_i^T\mathbf{v}_j = 0$ for $i \neq j$ for correspondingly distinct eigenvalues. Now, define $\mathbf{U} = [\mathbf{v}_1|\mathbf{v}_2|\mathbf{v}_3|...|\mathbf{v}_m]$ and $\mathbf{V} = [\mathbf{u}_1|\mathbf{u}_2|\mathbf{u}_3|...|\mathbf{u}_n]$. Now show that $\mathbf{A} = \mathbf{U}\boldsymbol{\Gamma}\mathbf{V}^T$ where $\boldsymbol{\Gamma}$ is a diagonal matrix containing the non-negative values $\gamma_1, \gamma_2, ..., \gamma_n$. With this, you have just established the existence of the singular value decomposition of any matrix $\mathbf{A}$. 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]

1.4 Solution

Let us first consider $\mathbf{U}\Gamma$ with the matrices as defined in the question. We have

$$\mathbf{U}\Gamma = [\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_m]\begin{pmatrix}\gamma_1 & 0 & \cdots & 0\\ 0 & \gamma_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \gamma_m\end{pmatrix}$$

which can also be written as

$$\mathbf{U}\Gamma = [\gamma_1\mathbf{v}_1|\gamma_2\mathbf{v}_2|\cdots|\gamma_m\mathbf{v}_m]$$

As proved in the previous part, replacing $\gamma_i\mathbf{v}_i$ by $\mathbf{A}\mathbf{u}_i$, we have

$$\mathbf{U}\Gamma = [\mathbf{A}\mathbf{u}_1|\mathbf{A}\mathbf{u}_2|\cdots|\mathbf{A}\mathbf{u}_n]$$

$$\mathbf{U}\Gamma = \mathbf{A}[\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_n]$$

Substituting this in $\mathbf{U}\Gamma\mathbf{V}^T$,

$$\mathbf{U}\Gamma\mathbf{V}^T = \mathbf{A}[\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_n][\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_n]^T$$

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

Now $\mathbf{V}\mathbf{V}^T$ will be $\mathbf{I}_n$ because orthogonality between all the eigenvectors and assuming that the eigenvectors are normalized, all eigenvectors will be orthonormal. Basically using,

$$\mathbf{u}_i^T\mathbf{u}_j = 0 \text{ for } i \neq j$$

and

$$\mathbf{u}_i^T \mathbf{u}_j = 1 \text{ for } i = j$$

So, all diagonal elements will be 1 and rest all the elements will be zero. Thus,

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

Thus proved.