Assignment 4 Question 3: CS 663, Digital Image Processing

Tanmay Randhavane, 110050010
Nilesh Kulkarni, 110050007
Deepali Adlakha, 11d170020

October 20, 2014

Consider a matrix $\mathbf{A}$ of size $m \times n$. Define $\mathbf{P} = \mathbf{A}^T\mathbf{A}$ and $\mathbf{Q} = \mathbf{A}\mathbf{A}^T$.

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?
   $\mathbf{A}$ is of size $m \times n$ and $\mathbf{P} = \mathbf{A}^T\mathbf{A}$
   $\therefore$ $\mathbf{P}$ is of size $n \times n$
   Let $\mathbf{y}$ be any vector of size $n \times 1$,
   Consider,

$$\begin{aligned}
\mathbf{y}^t\mathbf{P}\mathbf{y} &= \mathbf{y}^t\mathbf{A}^T\mathbf{A}\mathbf{y} \\
&= (\mathbf{A}\mathbf{y})^T\mathbf{A}\mathbf{y} \\
&= \Sigma(\mathbf{A}\mathbf{y})_i^2 \\
&\geq 0
\end{aligned}$$

   $\mathbf{Q} = \mathbf{A}\mathbf{A}^T$
   $\therefore$ $\mathbf{Q}$ is of size $m \times m$
   Let $\mathbf{z}$ be any vector of size $m \times 1$,
   Consider,

$$\begin{aligned}
\mathbf{z}^t\mathbf{Q}\mathbf{z} &= \mathbf{z}^t\mathbf{A}\mathbf{A}^T\mathbf{z} \\
&= (\mathbf{A}^T\mathbf{z})^T\mathbf{A}^T\mathbf{z} \\
&= \Sigma(\mathbf{A}^T\mathbf{z})_i^2 \\
&\geq 0
\end{aligned}$$

Let λ be any eigenvalue of $\mathbf{P}$, then

$$\begin{aligned}
\mathbf{Py} &= \lambda\mathbf{y} \\
\mathbf{y}^t\mathbf{Py} &= \lambda\mathbf{y}^t\mathbf{y} \\
\therefore\ \lambda\mathbf{y}^t\mathbf{y} &\geq 0\,(from\,previous\,results) \\
But,\ \mathbf{y}^t\mathbf{y} &\geq 0\,(length\,of\,\mathbf{y}) \\
\therefore\ \lambda &\geq 0
\end{aligned}$$

Similary, let λ be any eigenvalue of $\mathbf{P}$, then

$$\begin{aligned}
\mathbf{Qy} &= \lambda\mathbf{y} \\
\mathbf{y}^t\mathbf{Qy} &= \lambda\mathbf{y}^t\mathbf{y} \\
\therefore\ \lambda\mathbf{y}^t\mathbf{y} &\geq 0\,(from\,previous\,results) \\
But,\ \mathbf{y}^t\mathbf{y} &\geq 0\,(length\,of\,\mathbf{y}) \\
\therefore\ \lambda &\geq 0
\end{aligned}$$

2. If $\mathbf{u}$ is an eigenvector of $\mathbf{P}$ with eigenvalue λ, show that $\mathbf{Au}$ 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 $\mathbf{v}$?
Let $\mathbf{u}$ be a eigenvector of $\mathbf{P}$ with eigenvalue λ, then

$$\begin{aligned}
\mathbf{Pu} &= \lambda\mathbf{u} \\
\mathbf{APu} &= \lambda\mathbf{Au} \\
\mathbf{AA}^T\mathbf{Au} &= \lambda\mathbf{Au} \\
\mathbf{QAu} &= \lambda\mathbf{Au}
\end{aligned}$$

Therefore, $\mathbf{Au}$ is an eigenvector of $\mathbf{Q}$ with eigenvalue λ.

Similarly,
Let $\mathbf{v}$ be a eigenvector of $\mathbf{Q}$ with eigenvalue μ, then

$$\begin{aligned}
\mathbf{Qv} &= \mu\mathbf{v} \\
\mathbf{A}^T\mathbf{Qv} &= \mu\mathbf{A}^T\mathbf{v} \\
\mathbf{A}^T\mathbf{AA}^T\mathbf{v} &= \mu\mathbf{A}^T\mathbf{v} \\
\mathbf{PA}^T\mathbf{v} &= \mu\mathbf{A}^T\mathbf{v}
\end{aligned}$$

Therefore, $\mathbf{A}^T\mathbf{v}$ is an eigenvector of $\mathbf{P}$ with eigenvalue μ.

$\mathbf{P}$ is of size $n \times n$
$\therefore$ $\mathbf{u}$ has n elements.
While, $\mathbf{Q}$ is of size $m \times m$
$\therefore$ $\mathbf{v}$ has 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 $\mathbf{A}\mathbf{u}_i = \gamma_i\mathbf{v}_i$.
   $\mathbf{v}_i$ is an eigenvector of $\mathbf{Q}$ with eigenvalue λ.
   Therefore,

$$\begin{aligned} \mathbf{Q}\mathbf{v}_i &= \lambda\mathbf{v}_i \\ \mathbf{A}\mathbf{A}^T\mathbf{v}_i &= \lambda\mathbf{v}_i \\ \mathbf{A}\frac{\mathbf{A}^T\mathbf{v}_i}{\|\mathbf{A}^T\mathbf{v}_i\|} &= \frac{\lambda}{\|\mathbf{A}^T\mathbf{v}_i\|}\mathbf{v}_i \\ \mathbf{A}\mathbf{u}_i &= \gamma_i\mathbf{v}_i \\ where,\ \gamma_i = \frac{\lambda}{\|\mathbf{A}^T\mathbf{v}_i\|} \end{aligned}$$

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.[1]. 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]
   Let,
   $\mathbf{v}_i$ be an eigenvector of $\mathbf{Q}$.
   From, result of part (b) and (c), $\mathbf{u}_i = \dfrac{\mathbf{A}^T\mathbf{v}_i}{\sigma_i}$ is an eigenvector of $\mathbf{P}$
   where, $\sigma_i = \|\mathbf{A}^T\mathbf{v}_i\|$
   Consider,

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

$$
\begin{aligned}
\mathbf{u}_i^T \mathbf{A}^T \mathbf{v}_j &= \frac{(\mathbf{A}^T \mathbf{v}_i)^T \mathbf{A}^T \mathbf{v}_j}{\sigma_i} \\
&= \frac{\mathbf{v}_i^T \mathbf{A} \mathbf{A}^T \mathbf{v}_j}{\sigma_i} \\
&= \frac{\mathbf{v}_i^T \mathbf{Q} \mathbf{v}_j}{\sigma_i} \\
&= \frac{\mathbf{v}_i^T \mathbf{Q} \mathbf{v}_j}{\sigma_i} \\
&= \frac{\mu_i \mathbf{v}_i^T \mathbf{v}_j}{\sigma_i}
\end{aligned}
$$

Therefore, $\mathbf{u}_i^T \mathbf{A}^T \mathbf{v}_j = 0$ if $i \neq j$ and γ_i for $i = j$
Now, define $\mathbf{U} = [\mathbf{v}_1 | \mathbf{v}_2 | \mathbf{v}_3 | ... | \mathbf{v}_m]$ and $\mathbf{V} = [\mathbf{u}_1 | \mathbf{u}_2 | \mathbf{u}_3 | ... | \mathbf{u}_n]$
Therefore,

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

where Σ is a $n \times m$ matrix with $\Sigma_i i = \gamma_i$ and $\Sigma_i j = 0$ for $i \neq j$
Taking transpose,

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

Since, $\mathbf{u}_i^T \mathbf{u}_j = 0$ for $i \neq j$ and likewise $\mathbf{v}_i^T \mathbf{v}_j = 0$ for $i \neq j$,
$\mathbf{U}$ and $\mathbf{V}$ are orthonormal. Therefore,

$$
\begin{aligned}
\mathbf{U} \mathbf{U}^T \mathbf{A} \mathbf{V} \mathbf{V}^T &= \mathbf{U} \mathbf{\Sigma}^T \mathbf{V}^T \\
\mathbf{A} &= \mathbf{U} \mathbf{\Sigma}^T \mathbf{V}^T \\
\mathbf{A} &= \mathbf{U} \mathbf{\Gamma} \mathbf{V}^T
\end{aligned}
$$