MATH 434/534

Theoretical Assignment 7 Solution

Chapter 7

(7.1)

(i)

\[
H = I - 2 \frac{uu^T}{u^Tu}
\]

\[
Hu = (I - 2 \frac{uu^T}{u^Tu})u
= u - 2 \frac{uu^Tu}{u^Tu} = u - 2u = -u
\]

(ii)

\[
Hv = (I - 2 \frac{uu^T}{u^Tu})v = v - 2 \frac{uu^Tv}{u^Tu} = v - 0 = v
\]

because \( v^Tu = 0 \).

implies that \( u^Tv = 0 \).

7.2

(ii)

\[
Hv = (I - 2 \frac{uu^T}{u^Tu})v = v - 2 \frac{uu^Tv}{u^Tu} = v - 0 = v
\]

because \( v^Tu = 0 \).

implies that \( u^Tv = 0 \).

7.19

For an \( m \times n \) matrix \( A \), prove the following results using the SVD of \( A \):

(i) \( \text{rank}(A^T A) = \text{rank}(AA^T) = \text{rank}(A) = \text{rank}(A^T) \).

(ii) \( A^T A \) and \( AA^T \) have the same nonzero eigenvalues.

(iii) If the eigenvectors \( u_1 \) and \( u_2 \) of \( A^T A \) are orthogonal, then \( Au_1 \) and \( Au_2 \) are orthogonal.

-Solution-

Without the loss of generality, we will assume that \( m \geq n \) for all problems of No. 7.19.

Singular Value Decomposition

For a matrix \( A \in \mathbb{R}^{m \times n} \), there exist orthogonal matrices \( U \in \mathbb{R}^{m \times m} \) and \( V \in \mathbb{R}^{n \times n} \) such that \( A = U \Sigma V^T \) where \( \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m \times n} \) and \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0 \).

(i)

From the SVD of \( A \), there are orthogonal matrices \( U \) and \( V \) such that \( A = U \Sigma V^T \). Then we have

\[
A^T = V \Sigma^T U^T.
\]

However, we know \( \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m \times n} \). This implies \( \Sigma^T = \text{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{n \times m} \).

Also, we know that \( \text{rank}(A) = \text{The number of nonzero singular values} \).

Since the singular values of \( A \) and \( A^T \) are \( \sigma_1, \sigma_2, \ldots, \sigma_n \), the number of nonzero singular values of \( A \) is equal to the number of nonzero singular values of \( A^T \). Therefore, \( \text{rank}(A) = \text{rank}(A^T) \).

Now consider \( AA^T \) and \( A^T A \). Since \( A = U \Sigma V^T \) and \( A^T = V \Sigma^T U^T \), we have

\[
AA^T = (U \Sigma V^T)(V \Sigma^T U^T) = U(\Sigma \Sigma^T) U^T \quad \cdots \quad (1)
\]

\[
A^T A = (V \Sigma^T U^T)(U \Sigma V^T) = V(\Sigma^T \Sigma) V^T \quad \cdots \quad (2)
\]
Looking at $\Sigma^T \Sigma$ and $\Sigma \Sigma^T$, we can see

$$
\Sigma \Sigma^T = \begin{pmatrix}
\sigma_1 & 0 & \cdots & 0 \\
0 & \sigma_2 & 0 & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
0 & 0 & \cdots & \sigma_n \\
\vdots & \vdots & \ddots & \ddots \\
\vdots & \vdots & \ddots & \ddots \\
0 & 0 & \cdots & 0
\end{pmatrix}_{m \times n}
\begin{pmatrix}
\sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & \sigma_2 & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & \cdots & \sigma_n & 0 & \cdots & 0
\end{pmatrix}_{n \times m}
$$

Similarly, we can obtain

$$
\Sigma^T \Sigma = \begin{pmatrix}
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \sigma_2^2 & 0 & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
0 & 0 & \cdots & \sigma_n^2 \\
\vdots & \vdots & \ddots & \ddots \\
\vdots & \vdots & \ddots & \ddots \\
0 & 0 & \cdots & 0
\end{pmatrix}_{m \times m}
$$

From (1) and (2), the singular values of $AA^T$ and $A^TA$ are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Therefore, the number of nonzero singular values of $AA^T$ is equal to the number of nonzero singular values of $A^TA$, that is, $\text{rank}(AA^T) = \text{rank}(A^TA)$.

Also, from the facts that the singular values of $A$ are $\sigma_1, \sigma_2, \ldots, \sigma_n$ and that the singular values of $AA^T$ are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, the number of nonzero singular values of $A$ is equal to the number of nonzero singular values of $AA^T$. Therefore, $\text{rank}(A) = \text{rank}(AA^T)$.

Finally, we showed that

$$
\text{rank}(A^TA) = \text{rank}(AA^T) = \text{rank}(A) = \text{rank}(A^T).
$$

(ii)

By the Singular Value Decomposition, we have

$$
A = U \Sigma V^T \quad \text{where} \quad \Sigma_{m \times n} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)
$$

$$
A^T = V \Sigma^T U^T \quad \text{where} \quad \Sigma_{n \times m} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)
$$

By the definition of singular values, for a real matrix $A$, the square roots of the eigenvalues of $A^TA$ are called the singular values of $A$, that is, the eigenvalues of $A^TA$ are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Since $AA^T = (A^T)^T A^T$ and the singular values of $A^T$ are $\sigma_1, \sigma_2, \ldots, \sigma_n$, the eigenvalues of $AA^T = (A^T)^T A$ are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Therefore, $A^TA$ and $AA^T$ have the same nonzero eigenvalues.
Suppose that the eigenvectors \( u_1 \) and \( u_2 \) of \( A^T A \) are orthogonal. Then we have

\[
egin{align*}
(A^T A)u_1 & = \lambda_1 u_1 \quad \cdots (1) \\
(A^T A)u_2 & = \lambda_2 u_2 \quad \cdots (2) \\
u_1^T u_2 & = u_2^T u_1 = 0 \quad \cdots (3)
\end{align*}
\]

Then consider \((Au_1)^T(Au_2)\) and \((Au_2)^T(Au_1)\). Using (2) and (3), we have

\[
(Au_1)^T(Au_2) = u_1^T A^T A u_2 = u_1^T (\lambda_2 u_2) = \lambda_2 u_1^T u_2 = 0
\]

Similarly, using (1) and (2), we have

\[
(Au_2)^T(Au_1) = u_2^T A^T A u_1 = u_2^T (\lambda_1 u_1) = \lambda_1 u_2^T u_1 = 0
\]

Therefore, \(Au_1\) and \(Au_2\) are orthogonal.

**7.21**

Let \(U\) have orthogonal columns. Then using SVD, prove that

(i) \(\|AU\|_2 = \|A\|_2\),

(ii) \(\|AU\|_F = \|A\|_F\).

---

**Solution**

Again without the loss of generality, we will assume that \(m \geq n\) for all problems of No. 7.21. Since \(U\) has orthogonal columns, \(U = [u_1 u_2 \cdots u_n]_{m \times n}\), where \(u_i\) is the \(i\)th column vector expressed as \(u_i = (u_{1i}, u_{2i}, \ldots, u_{mi})^T\) and satisfying

\[
u_i^T u_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}
\]

This implies

\[
U^T U = \begin{pmatrix}
    u_1^T \\
    u_2^T \\
    \vdots \\
    u_n^T
\end{pmatrix}
    \begin{pmatrix}
    u_1 & u_2 & \cdots & u_n^T
\end{pmatrix}_{m \times n} = I_{n \times n}
\]

Therefore, \(U\) is an orthogonal matrix.

From the Singular Value Decomposition, for a matrix \(A\), there exist orthogonal matrices \(\hat{U}\) and \(\hat{V}\) such that \(A = \hat{U} \Sigma \hat{V}^T\), where \(\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)\).

Now multiplying \(A = \hat{U} \Sigma \hat{V}^T\) by \(U\), we have

\[
AU = \hat{U} \Sigma \hat{V}^T U
\]

Since the product of orthogonal matrices is an orthogonal matrix, letting \(W^T = \hat{V}^T U\), we have

\[
AU = \hat{U} \Sigma W^T \quad \cdots (\ast)
\]

(i) Since \(\|A\|_2 = \sigma_{max}\), from (\ast) we have \(\|AU\|_2 = \sigma_{max}\). Therefore, \(\|AU\|_2 = \|A\|_2\).
(ii) Since $\|A\|_F = (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)^{\frac{1}{2}}$, again from (*), we have

$$\|AU\|_F = (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)^{\frac{1}{2}}.$$ 

Therefore, we showed $\|A\|_F = \|AU\|_F$.

7.22 Let $U\Sigma V^T$ be the SVD of $A$. Then prove that

$$\left\| U^T AV \right\|_F^2 = \sum_{i=1}^{p} \sigma_i^2,$$

where $\sigma_i$ are the singular values of $A$.

-Solution-

From the Singular Value Decomposition, for an $m \times n$ matrix $A$, there exists $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that $A = U\Sigma V^T$ where $\Sigma = diag(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m \times n}$ and $p = \min(m, n)$. Since $U$ and $V$ are orthogonal matrices, we have

$$A = U\Sigma V^T \Rightarrow U^T AV = \Sigma.$$

Now taking the Frobenius norm on $U^T AV = \Sigma$, we have

$$\left\| U^T AV \right\|_F^2 = \|\Sigma\|_F^2 \quad \cdots \quad (1)$$

By the definition of the Frobenius norm, $\|A\|_F = \left[ \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2 \right]^{\frac{1}{2}}$, since $\Sigma = diag(\sigma_1, \ldots, \sigma_p)$, we have

$$\|\Sigma\|_F^2 = \sum_{i=1}^{p} \sigma_i^2 \quad \cdots \quad (2)$$

From (1) and (2), we showed that $\left\| U^T AV \right\|_F^2 = \sum_{i=1}^{p} \sigma_i^2$.

7.23 Let $A$ be an $m \times n$ matrix.

(a) Using the SVD of $A$, prove that

(i) $\left\| A^T A \right\|_2 = \|A\|_2^2$;

(ii) $\text{Cond}_2(A^T A) = (\text{Cond}_2(A))^2$;

(iii) $\text{Cond}_2(A) = \text{Cond}_2(U^T AV)$, where $U$ and $V$ are orthogonal.

(b) Let $\text{rank}(A_{m\times n}) = n$, and let $B_{m\times r}$ be a matrix obtained by deleting $(n - r)$ columns from $A$. Then prove that $\text{Cond}_2(B) \leq \text{Cond}_2(A)$.

-Solution-

(a)

Using the Singular Value Decomposition, we have an orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that $A = U\Sigma V^T$. Without the loss of generality, we assume $m \geq n$.

(i)
By the Singular Value Decomposition, we have $A = U \Sigma V^T$. Therefore, $\|A\|_2 = \sigma_{max} = \sigma_1$. Also, we have $A^T = V \Sigma^T U^T$. Considering $A^T A$, we have

$$A^T A = U \Sigma^T V V^T \Sigma m \times n = U \Sigma \Sigma^T U^T$$

Since $\Sigma \Sigma^T = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$, $\|A^T A\|_2 = \sigma_{max}^2 = \sigma_1^2$. Therefore, $\|A\|_2^2 = \|A^T A\|_2^2$.

(ii)

$$\text{Cond}_2(A^T A) = \frac{\sigma_{max}}{\sigma_{min}}$$

(iii)

By the Singular Value Decomposition, we have

$$A = U \Sigma V^T \Rightarrow U^T AV = \Sigma$$

Since $\text{Cond}_2(\Sigma) = \frac{\sigma_{max}}{\sigma_{min}}$, we have

$$\text{Cond}_2(U^T AV) = \text{Cond}_2(\Sigma) = \frac{\sigma_{max}}{\sigma_{min}} \cdot \cdots \quad (1)$$

Also, $\text{Cond}_2(A) = \frac{\sigma_{max}}{\sigma_{min}}$. Therefore, we have $\text{Cond}_2(A) = \text{Cond}_2(U^T AV)$.

(b)

Let

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m \times n}$$

Then we have

$$B_{m \times r} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{pmatrix}_{m \times r}$$

Since $\text{rank}(A) = n$, the number of nonzero singular values of $A$ is $n$. This tell us that $\sigma_{min} = \sigma_n \leq \sigma_r$. Therefore, we have

$$\text{Cond}_2(A) = \frac{\sigma_{max}}{\sigma_{min}} = \frac{\sigma_1}{\sigma_n} \geq \frac{\sigma_1}{\sigma_r} = \text{Cond}_2(B)$$