MATH 434/534

Theoretical Assignment 1 Solution

Chapter 2
No. 2.0
Answer “True” or “False” to the following. Give reasons for your answers.

-Solution-

(a) The eigenvalues of an upper triangular matrix $T$ are its diagonal entries.
   True. See Theorem 2.15-3.

(b) The eigenvalues of a real symmetric matrix are real.
   True. See Corollary 2.18.

(c) A matrix is nonsingular if and only if all its eigenvalues are nonzero.
   True. See Theorem 2.7-2.

(d) The eigenvalues of an orthogonal matrix are all equal to 1.
   False. Consider the following matrix
   \[
   O = \begin{pmatrix}
   1 & 0 \\
   0 & -1
   \end{pmatrix},
   \]
   Thus, $O^T O = O O^T = I$. So, $O$ is an orthogonal matrix. However, the eigenvalues of $O$ are $\lambda = \pm 1$. It can be shown that all eigenvalues of an orthogonal matrix have modulus one.

(e) An orthogonal matrix is not necessarily invertible.
   False, because if $O$ is an orthogonal matrix, then from the definition $O^T = O^{-1}$.

(f) A real symmetric or a complex Hermitian matrix can be always transformed into a diagonal matrix by similarity transformation.
   True. See Theorem 2.17 and Theorem 2.20.

(g) Two similar matrices have the same eigenvalues.
   True. See the property of similar matrices. See page 14.

(h) If two matrices have the same eigenvalues, they must be similar.
   False. Consider the following two matrices:
   \[
   A = \begin{pmatrix}
   1 & 0 \\
   0 & 1
   \end{pmatrix},
   \]
   and
   \[
   B = \begin{pmatrix}
   1 & 1 \\
   0 & 1
   \end{pmatrix}.
   \]
They have the same eigenvalues, but they are not similar.

(i) The product of two upper (lower) triangular matrices does not need to be an upper (lower) triangular matrix.

False. By Theorem 2.15, the product of two upper (lower) triangular matrices is an upper (lower) triangular matrix.

(j) $\|I\| = 1$ for any norm.

False. For the identity matrix $I$, $\|I\|_1 = \|I\|_2 = \|I\|_\infty = 1$. However, $\|I\|_F = \sqrt{n}$.

(k) The length of a vector is preserved by an orthogonal multiplication (multiplication of a vector or a matrix by an orthogonal matrix).

True. Let $O$ be an orthogonal matrix, that is, $O^T O = O O^T = I$. Then $\|Ox\|^2 = (Ox)^T (Ox) = x^T O^T O x = x^T x = \|x\|^2$. This implies $\|Ox\| = \|x\|$.

(l) If $\|A\| < 1$, then $I - A$ is nonsingular.

From Problem No. 2.10, we have $|\lambda| \leq \|A\|$ for every eigenvalues of $\lambda$ of $A$. Then we have

$$|\lambda| \leq \|A\| < 1.$$ 

Thus, none of quantities of $1 - \lambda_i$ for $i = 1, 2, \ldots, n$ is zero. This shows that $I - A$ is nonsingular.

(m) If $\|A\|_2 = 1$, then $A$ must be orthogonal.

False. Consider the following matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

Then we have

$$A^T = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

$$A^T A = A A^T = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$$

So, the eigenvalues of $A^T A$ are $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{4}$. This implies $\|A\|_2 = 1$. However, $A^T A = A A^T \neq I$. Therefore, $A$ is not an orthogonal matrix.

(n) The product of two orthogonal (unitary) matrices is an orthogonal (unitary) matrix.

True. Let $U_1$ and $U_2$ be unitary matrices. Then $U_i^* U_i = U_i U_i^* = I$ for $i = 1, 2$. 

$$(U_1 U_2)^* (U_1 U_2) = U_2^* U_1^* U_1 U_2 = I$$

and

$$(U_1 U_2) (U_1 U_2)^* = U_1 U_2 U_2^* U_1^* = I.$$ 

Therefore, the product of two unitary matrices is a unitary matrix. Similarly, it is shown that the product of two orthogonal matrices is an orthogonal matrix.

No. 2.10
Prove that for a subordinate matrix norm $\|\cdot\|$, $|\lambda| \leq \|A\|$ for every eigenvalues of $\lambda$ of $A$.

-Solution-
Let $\lambda$ be an eigenvalue of $A$ and $x$ be an eigenvector of $A$ corresponding to $\lambda$. Then we have $Ax = \lambda x$. By the property of a subordinate matrix norm, we have

$$\|Ax\| \leq \|A\| \|x\|$$

However, since $Ax = \lambda x$, we have $\|Ax\| = \|\lambda x\| = |\lambda| \|x\|$. Therefore, we have

$$|\lambda| \|x\| \leq \|A\| \|x\|$$

$$\Rightarrow |\lambda| \leq \|A\|$$

**No. 2.22**

If $x$ and $y$ are two vectors, then prove that

(a) $\left|x^T y\right| \leq \|x\|_2 \|y\|_2$

(b) $\|xy^T\|_2 \leq \|x\|_2 \|y\|_2$

---

**Solution**

(a)

If $x$ and $y$ are linearly dependent, that is, $y = cx$ where $c$ is constant. Then we have

$$\left|x^T y\right| = \left|x^T (cx)\right|$$

$$= |c| \left|x^T x\right|$$

$$= |c| \|x\|_2^2$$

$$= \|x\|_2 \|cx\|_2 = \|x\|_2 \|y\|_2$$

Suppose that $x$ and $y$ are linearly independent. Let $\epsilon$ be any number. Now consider the following inequality.

$$0 \leq \|x - \epsilon y\|_2^2 = (x - \epsilon y)^T (x - \epsilon y)$$

$$= (x^T x) - (\epsilon x^T y) - (\epsilon y^T x) + (\epsilon^2 y^T y)$$

$$= \|x\|_2^2 - 2\epsilon (x^T y) + \epsilon^2 \|y\|_2^2$$

$$= \|y\|_2^2 \epsilon^2 - 2(x^T y)\epsilon + \|x\|_2^2$$

Now the right hand side of the above inequality is a quadratic equation of $\epsilon$. Since the equation must be nonnegative, it is necessary that the discriminant $D \leq 0$. Therefore

$$0 \geq D = 4(x^T y)^2 - 4\|y\|_2^2 \|x\|_2^2 = 4 \left(\left|x^T y\right|^2 - \|y\|_2^2 \|x\|_2^2\right)$$

Therefore, we have $\left|x^T y\right|^2 - \|y\|_2^2 \|x\|_2^2 \leq 0$. This implies

$$\left|x^T y\right| \leq \|x\|_2 \|y\|_2$$

Therefore, $\left|x^T y\right| \leq \|x\|_2 \|y\|_2$.

Or using $x^T y = \|x\| \|y\| \cos \theta$ where $\theta$ is the angle between two vectors $x$ and $y$. Since $|\cos \theta| \leq 1$, we have

$$\left|x^T y\right| = \|x\| \|y\| |\cos \theta|$$

$$= \|x\|_2 \|y\|_2 |\cos \theta| \leq \|x\|_2 \|y\|_2$$
because \( \| \cdot \| = \| \cdot \|_2 \) on vector norms. Therefore, \( |x^T y| \leq \|x\|_2 \|y\|_2 \).

(b) Let \( x \) and \( y \) be \( n \times 1 \) vectors. Then \( xy^T \) is an \( n \times n \) matrix such that each entry of \( xy^T \) is expressed as \( x_i y_j \), that is,

\[
xy^T = \begin{pmatrix}
x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\
x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\
\vdots & \vdots & \ddots & \vdots \\
x_n y_1 & x_n y_2 & \cdots & x_n y_n
\end{pmatrix}
\]

The inequality holds for either \( x = 0 \) or \( y = 0 \). If \( x \neq 0 \) and \( y \neq 0 \), then \( x^T x \neq 0 \) and \( y^T y \neq 0 \). Therefore we have

\[
\|xy^T\|_2 \leq \frac{1}{\sqrt{n}} \|xy^T\|_F
\]

\[
= \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i y_j|^2 \right)
\]

\[
= \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} |x_i|^2 \sum_{j=1}^{n} |y_j|^2 \right)
\]

\[
= \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} |x_i|^2 \sqrt{\sum_{j=1}^{n} |y_j|^2} \right)
\]

\[
= \frac{1}{\sqrt{n}} \|x\|_2 \|y\|_2 \leq \|x\|_2 \|y\|_2
\]

because \( n \) is an integer. Therefore, \( \|xy^T\|_2 \leq \|x\|_2 \|y\|_2 \).

No. 2.23
Let \( x \) and \( y \) be two orthogonal vectors. Then prove that \( \|x + y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 \)

-Solution-
Since \( x \) and \( y \) are orthogonal vectors, \( x^T y = 0 \). Now consider \( \|x + y\|_2^2 \).

\[
\|x + y\|_2^2 = (x + y)^T(x + y) = (x^T x) + 2(x^T y) + (y^T y) = \|x\|_2^2 + \|y\|_2^2
\]

Note that \( x^T y = y^T x \) is scalar.

No. 2.28
Prove that (i) \( \|I\|_2 = 1 \), and (ii) \( \|I\|_F = \sqrt{n} \).

-Solution-(i)
Since \( \|A\|_2 = \sqrt{\text{maximum eigenvalue of } A^T A} \), we have \( \|I\|_2 = \sqrt{\text{maximum eigenvalue of } I^T I} \). However, \( I^T I = I \) and the eigenvalues of the identity matrix \( I \) are 1’s. Therefore, \( \|I\|_2 = 1 \).

(ii)
First consider the meaning of \( \sum_{j=1}^{n} \sum_{i=1}^{n} |a_{ij}|^2 \), that is, taking \( \cdot \)^2 on each entry of a matrix \( A \) and adding up these results. So, if you take \( \cdot \)^2 on each entry of the \( n \) by \( n \) matrix \( I \) and
add these up, the sum must be \( n \). Since \( \|I\|_F = \left[ \sum_{j=1}^{n} \sum_{i=1}^{n} |I_{ij}|^2 \right]^{\frac{1}{2}} \) where \( I_{ij} \) is the \( ij \) entry of \( I \). You have

\[
\|I\|_F = \left[ \sum_{j=1}^{n} \sum_{i=1}^{n} |I_{ij}|^2 \right]^{\frac{1}{2}} = (n)^{\frac{1}{2}} = \sqrt{n}
\]

**No. 2.29**
Prove that if \( Q \) and \( P \) are orthogonal matrices, then (a) \( \|QAP\|_F = \|A\|_F \), and (b) \( \|QAP\|_2 = \|A\|_2 \)

- **Solution**-
(a)
Since \( P \) and \( Q \) are orthogonal matrices, \( Q^TQ = QQ^T = I \). By a property of the Frobenius norm, we have \( \|A\|_F^2 = \text{trace}(A^TA) \). Now consider \( \|QAP\|_F^2 \).

\[
\|QAP\|_F^2 = \text{trace}((QAP)^T(QAP)) = \text{trace}(P^TQA^TQAP) = \text{trace}(P^TAP)^T(QAP)
\]

Since the trace of a matrix is just the sum of its eigenvalues, and two similar matrices have the same eigenvalues, \( \text{trace}(P^TAP) = \text{trace}(A^TA) \). Thus,

\[
\|QAP\|_F^2 = \text{trace}(A^TA) = \|A\|_F^2.
\]

Therefore, \( \|QAP\|_F = \|A\|_F \).

(b)
Since \( P \) and \( Q \) are orthogonal matrices, \( Q^TQ = QQ^T = I \). By the definition of \( \|QAP\|_2 \), we have \( \|QAP\|_2 = \sqrt{\text{max eigenvalue of } (QAP)^T(QAP)} \).

\[
(QAP)^T(QAP) = P^T(QA)^T(QAP) = P^TQA^TQAP = P^TAP
\]

Since two similar matrices have the same eigenvalues, we have

\[
\|QAP\|_2^2 = \sqrt{\text{max eigenvalue of } (P^T(A^TA)P)} = \sqrt{\text{max eigenvalue of } (A^TA)} = \|A\|_2^2.
\]

Therefore, \( \|QAP\|_2 = \sqrt{\text{max}(\lambda_1, \lambda_2, \ldots, \lambda_n)} = \|A\|_2 \)

**No. 2.33**
Prove that (i) \( \|A^T\|_2 = \|A\|_2 \), and (ii) \( \|A^TA\|_2 = \|A\|_2^2 \)

- **Solution**-
(i)

\[
\|A^T\|_2 = \sqrt{\rho(A^TA)} = \sqrt{\rho(A^TA)} = \|A\|_2.
\]

(ii)
Preliminaries
Let $\lambda$ be the eigenvalues of a matrix $A$. Then we have $Ax = \lambda x$ where $x$ is an eigenvector corresponding to $\lambda$. Multiplying both sides of $Ax = \lambda x$ by $A$, we have

$$A Ax = A \lambda x = \lambda Ax = \lambda(\lambda x) = \lambda^2 x$$

Therefore, $A^2 x = \lambda^2 x$, that is, the eigenvalues of $A^2$ are the square of the eigenvalues of $A$. Also $(A^T A)^T (A^T A) = (A^T A)(A^T A) = (A^T A)^2$. Now consider $\|A\|_2^2$.

$$\|A\|_2^2 = \sqrt{\text{max eigenvalue of } (A^T A)}$$

$$= \sqrt{\text{max eigenvalue of } ((A^T A)^2)}$$

$$= \text{max eigenvalues of } (A^T A) = \|A\|_2^2$$

because the eigenvalues of $(A^T A)^2$ are the square of the eigenvalues of $(A^T A)$.

**No. 2.34**

Let $A = (a_1, \ldots, a_N)$, where $a_j$ is the $j^{th}$ column of $A$. Then prove that $\|A\|_F^2 = \sum_{i=1}^n \|a_i\|_2^2$.

-Solution-

By the definition of $\|A\|_F$, we have $\|A\|_F^2 = \sum_{j=1}^n \sum_{i=1}^n |a_{ij}|^2$. First considering $\sum_{i=1}^n |a_{ij}|^2$.

For a fixed $j$, we have

$$\sum_{i=1}^n |a_{ij}|^2 = |a_{1j}|^2 + |a_{2j}|^2 + \cdots + |a_{mj}|^2$$

$$= \left( \sqrt{|a_{1j}|^2 + |a_{2j}|^2 + \cdots + |a_{mj}|^2} \right)^2$$

$$= \|a_j\|_2^2$$

where $a_j$ is the $j^{th}$ column of $A$. Therefore we have

$$\|A\| = \sum_{j=1}^n \|a_j\|_2^2 = \sum_{i=1}^n \|a_i\|_2^2$$

**No. 2.35**

Prove that if $A$ and $B$ are two matrices compatible for matrix multiplication, then

(a) $\|AB\|_F \leq \|A\|_F \|B\|_F$

(b) $\|AB\|_F \leq \|A\|_2 \|B\|_F$

-Solution-

(a) Let $A$ be an $m \times n$ matrix and $B$ be an $n \times l$ matrix. Then $AB$ is an $m \times l$ matrix and $AB = \sum_{k=1}^l a_{ik} B_{kj}$.

$$\|AB\|_F^2 = \sum_{j=1}^l \sum_{i=1}^m \sum_{k=1}^n |a_{ik} b_{kj}|$$

$$\leq \sum_{j=1}^l \sum_{i=1}^m \left( \sum_{k=1}^n |a_{ik}|^2 \right) \left( \sum_{k=1}^n |b_{kj}|^2 \right)$$

$$= \sum_{i=1}^m \sum_{k=1}^n |a_{ik}|^2 \sum_{j=1}^l \sum_{k=1}^n |b_{kj}|^2$$

$$= \|A\|_F^2 \|B\|_F^2$$
Therefore, \( \|AB\|_F \leq \|A\|_F \|B\|_F \).

(b) Let \( A \) be an \( m \times n \) matrix and \( B \) be an \( n \times l \) matrix. Let \( B = (b_1, \ldots, b_l) \) where \( b_1, \ldots, b_l \) are columns of \( B \). Now consider \( \|AB\|_F^2 \).

\[
\|AB\|_F^2 = \|A(b_1, \ldots, b_l)\|_F^2 = (Ab_1)^T Ab_1 + \cdots + (Ab_l)^T Ab_l = \|Ab_1\|_2^2 + \cdots + \|Ab_l\|_2^2 \leq A\|_2^2 (\|b_1\|_2^2 + \cdots + \|b_l\|_2^2)
\]

because of the property of a subordinate matrix norm, \( \|Ax\|_p \leq \|A\|_p \|x\|_p \). Since \( \|B\|_F^2 = \sum_{i=1}^n \|b_i\|_2^2 \), we have

\[
\|AB\|_F^2 \leq \|A\|_2^2 \sum_{i=1}^n \|b_i\|_2^2 = \|A\|_2^2 \|B\|_F^2.
\]

Therefore, \( \|AB\|_F \leq \|A\|_2 \|B\|_F \).

Note
By the compatibility of \( F \)-norm, \( \|Ax\|_2 \leq \|A\|_F \|x\|_2 \) (see Page 20), similarly part (a) can be shown.

No. 2.38
Let \( \text{trace}(A) = \sum_{i=1}^n a_{ii} \). Then prove the following
(a) \( \text{trace}(AB) = \text{trace}(BA) \),
(b) \( \text{trace}(AA^*) = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \), where \( A = (a_{ij}) \) is \( m \times n \).
(c) \( \text{trace}(A+B) = \text{trace}(A) + \text{trace}(B) \)

Consider \( AB \) and \( BA \)

\[
\text{trace}(AB) = \sum_{i=1}^n \sum_{j=1}^m (a_{ij}b_{ji})
\]

\[
\text{trace}(BA) = \sum_{j=1}^m \sum_{i=1}^n (b_{ij}a_{ji}) = \sum_{i=1}^n \sum_{j=1}^m (a_{ij}b_{ji}) = \text{trace}(AB)
\]

(b) \( \text{trace}(AA^*) = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \), where \( A = (a_{ij}) \) is \( m \times n \).

Let \( a_{ij} \) be an entry of \( A \). Since \( A^* = (\bar{A})^T \) where \( \bar{A} \) is the complex conjugate of \( A \) and \( A \) is an \( m \times n \) matrix, \( A^* \) is an \( n \times m \) matrix and each entry of \( A^* \) is \( \bar{a}_{ji} \). Similarly, we have

\[
\text{trace}(AA^*) = \sum_{i=1}^m \sum_{j=1}^n (a_{ij}\bar{a}_{ji}) = \sum_{i=1}^m \sum_{j=1}^n (|a_{ij}|^2)
\]

because \( zz^* = |z|^2 \) where \( z \) is a complex number.

(c) \( \text{trace}(A+B) = \text{trace}(A) + \text{trace}(B) \)

To carry out \( A + B \), the size of \( A \) must be the same as the size of \( B \). Suppose the size of these matrices is \( m \times n \).

\[
\text{trace}(A + B) = (a_{11} + b_{22}) + (a_{22} + b_{22}) + \cdots + (a_{mm} + b_{mm})
\]

\[
= (a_{11} + a_{22} + \cdots + a_{mm}) + (b_{11} + b_{22} + \cdots + b_{mm})
\]

\[
= \text{trace}(A) + \text{trace}(B)
\]
(d) $\text{trace}(TAT^{-1}) = \text{trace}(A)$

Using (a), we have

$$\text{trace}(TAT^{-1}) = \text{trace}((TA)T^{-1}) = \text{trace}(T^{-1}(TA)) = \text{trace}(A)$$