Study Questions 1 Numerical Linear Algebra I Fall 2002

Problem 1

This problem contains some simple facts that you might find useful.

(a) Show that $(AB)^T = B^T A^T$ for matrices $A$ and $B$ for which the multiplication is well-defined.

(b) Let $A ∈ \mathbb{R}^{n×n}$ and $B ∈ \mathbb{R}^{n×n}$ be nonsingular matrices. Show $(AB)^{-1} = B^{-1} A^{-1}$.

Solution (a):

We are given $A$ and $B$ with dimensions such that the matrix product $AB$ is well-defined.

We define the following matrices:

\[
\begin{align*}
C &= AB \\
M^{(1)} &= C^T \\
M^{(2)} &= B^T \\
M^{(3)} &= A^T \\
M^{(4)} &= M^{(2)} M^{(3)}
\end{align*}
\]

We assume the following notation:

\[
\begin{align*}
α_{ij} &= e_i^T Ae_j \\
β_{ij} &= e_i^T Be_j \\
γ_{ij} &= e_i^T Ce_j \\
μ^{(1)}_{ij} &= e_i^T M^{(1)} e_j \\
μ^{(2)}_{ij} &= e_i^T M^{(2)} e_j \\
μ^{(3)}_{ij} &= e_i^T M^{(3)} e_j \\
μ^{(4)}_{ij} &= e_i^T M^{(4)} e_j
\end{align*}
\]

We then have

\[
\begin{align*}
μ^{(2)}_{ij} &= β_{ji} \\
μ^{(3)}_{ij} &= α_{ji} \\
γ_{ij} &= Σ_kα_{ik} β_{kj} \\
μ^{(1)}_{ij} &= γ_{ji}
\end{align*}
\]

We must show that $M^{(1)} = M^{(4)}$. To see this note:

\[
μ^{(4)}_{ij} = Σ_k μ^{(2)}_{ik} μ^{(3)}_{kj} \\
= Σ_k β_{ki} α_{jk} \\
= Σ_k α_{jk} β_{ki} \\
= γ_{ji} \\
= μ^{(1)}_{ij}
\]
Solution (b):

By assumption, the $n \times n$ nonsingular matrices $A^{-1}$ and $B^{-1}$ exist, are unique and $AA^{-1} = BB^{-1} = A^{-1}A = B^{-1}B = I$. Let $M = B^{-1}A^{-1}$.

We have

$$
(AB)M = ABB^{-1}A^{-1} \\
= A(BB^{-1})A^{-1} \\
= AA^{-1} \\
= I
$$

$$
M(AB) = B^{-1}A^{-1}AB \\
= B^{-1}(A^{-1}A)B \\
= B^{-1}IB \\
= B^{-1}B \\
= I
$$

To see that $M$ is unique, suppose there is another matrix $Q \neq M$ such that $ABQ = I$. We have

$$
Q = IQ = (MAB)Q = M(ABQ) = M
$$

Now suppose that $Q \neq M$ such that $QAB = I$. We have

$$
Q = QI = Q(ABM) = (QAB)M = M
$$

(Strictly speaking you need only prove one of these to show uniqueness.)

Problem 2

2(a)

Suppose you have two vectors $x, y \in \mathbb{R}^n$ that are linearly independent but not orthogonal. Is it possible to create two vectors $u$ and $v$, via linear combinations of $x$ and $y$, so that the $\text{span}[u, v] = \text{span}[x, y]$, $u$ and $v$ are orthogonal, and $\|u\|_2 = \|v\|_2 = 1$?

Hint: Let $u = x/\|x\|_2$ and determine how to define $v$ from $u$ and $y$.

2(b)

Let $A \in \mathbb{R}^{n \times 2}$ have as its two columns $x$ and $y$ from 1(b), i.e., $Ae_1 = x$ and $Ae_2 = y$. Similarly let $B \in \mathbb{R}^{n \times 2}$ be such that $Be_1 = u$ and $Be_2 = v$. Show that $A$ and $B$ are related to each other via a nonsingular $2 \times 2$ matrix $M$, i.e., $AM = B$.

Hint: Deduce $M$ from your answer to (a) and prove that it is nonsingular.

Solution (a): The initial direction used is arbitrary in general and therefore we can choose

$$
u = \frac{x}{\|x\|_2}
$$

i.e. a scaled version of $x$. We have

$$
u^T u = \frac{x^T x}{\|x\|_2}
$$

$$
= 1
$$
In order to preserve the span we must take \( v \) as a linear combination of \( x \) and \( y \) (or \( u \)). Consider

\[
\hat{v} = y - (u^T y)u
\]

We can verify orthogonality

\[
\begin{align*}
    u^T \hat{v} &= u^T (y - (u^T y)u) \\
    &= u^T y - (u^T y)u^T u \\
    &= u^T y - u^T u \\
    &= 0
\end{align*}
\]

Normalizing \( \hat{v} \) to give \( v = \frac{\hat{v}}{\|v\|_2} \) yields the final vector.

**Solution (b):** To determine the matrix \( M \) define it in terms of three matrices \( M_i, i = 1, 2, 3 \) where each represent a single step of the algorithm in (b).

\[
\begin{bmatrix}
    x & y \\
    u & y
\end{bmatrix} = A \\
\begin{bmatrix}
    x & y
\end{bmatrix} \begin{bmatrix}
    \frac{1}{\|x\|_2} & 0 \\
    0 & 1
\end{bmatrix} = AM_1 \\
\begin{bmatrix}
    u & \hat{v}
\end{bmatrix} \begin{bmatrix}
    1 & -u^T y \\
    0 & 1
\end{bmatrix} = AM_1 M_2 \\
\begin{bmatrix}
    u & v
\end{bmatrix} \begin{bmatrix}
    1 & 0 \\
    0 & \frac{1}{\|v\|_2}
\end{bmatrix} = AM_1 M_2 M_3
\]

The assumption of linear independence guarantees that \( \|x\|_2 \neq 0 \) and \( \|\hat{v}\|_2 \neq 0 \).

The matrices \( M_1 \) and \( M_3 \) are therefore nonsingular since they are diagonal with nonzero diagonal elements. \( M_2 \) is also nonsingular since it is upper triangular with nonzero diagonal elements. (Recall in the notes we show that nonzero diagonal elements for lower triangular matrices implies nonsingularity. The same argument applies to upper triangular matrices.) The product of nonsingular matrices must be nonsingular (by Problem 0).

Finally we have the form of \( M \) must be upper triangular since it is the product of upper triangular matrices. (Diagonal matrices are trivially upper triangular.)

**Problem 3**

Consider the definition \( \|A\| = \max_{i,j} |a_{i,j}| \) where \( e_i^T Ae_j = a_{i,j} \). Does this satisfy all of the required properties of a matrix norm?

**Solution:**

Recall that for the complex numbers the absolute value, \(|\alpha|\) function is a norm and therefore satisfies all three required properties:

\[
\begin{align*}
    \alpha \neq 0 & \Rightarrow |\alpha| > 0 \\
    |\alpha \beta| &= |\alpha| |\beta| \\
    |\alpha + \beta| &\leq |\alpha| + |\beta|
\end{align*}
\]
So we can check each of the properties for the proposed matrix norm.

The proposed norm is definite since \( A \neq 0 \rightarrow \exists p, q \) such that \( \alpha_{pq} \neq 0 \) and therefore
\[
\| A \| = \max_{ij} |\alpha_{ij}| \geq |\alpha_{pq}| > 0
\]

The proposed norm is homogeneous. We have
\[
\| \beta A \| = \max_{ij} |\beta \alpha_{ij}|
\]
\[
= \max_{ij} |\beta| |\alpha_{ij}|
\]
\[
= |\beta| \max_{ij} |\alpha_{ij}|
\]
\[
= \beta \| A \|
\]

The proposed norm satisfies the triangle inequality.
\[
\| A + B \| = \max_{ij} |\alpha_{ij} + \beta_{ij}|
\]
\[
\leq \max_{ij} (|\alpha_{ij}| + |\beta_{ij}|)
\]
\[
\leq \max_{ij} (|\alpha_{ij}|) + \max_{ij} (|\beta_{ij}|)
\]
\[
= \| A \| + \| B \|
\]

**Problem 4**

Suppose \( Q \in \mathbb{R}^{n \times n} \) is an orthogonal matrix and \( x \in \mathbb{R}^n \).

(a) Show that \( \| Qx \|_2 = \| x \|_2 \).

(b) Is \( \| QA \|_2 = \| A \|_2 \) where \( A \in \mathbb{R}^{n \times n} \)?

(c) Is \( \| QA \|_F = \| A \|_F \) where \( A \in \mathbb{R}^{n \times n} \)?

(d) Is \( \| AQ \|_2 = \| A \|_2 \) where \( A \in \mathbb{R}^{n \times n} \)?

(e) Is \( \| AQ \|_F = \| A \|_F \) where \( A \in \mathbb{R}^{n \times n} \)?

**Solution (a):** Given \( x \) and orthogonal \( Q \in \mathbb{R}^{n \times n} \), let \( y = Qx \). We then have
\[
\| x \|_2 \quad = \quad x^T x
\]
\[
\quad = \quad x^T I x
\]
\[
\quad = \quad x^T (Q^T Q)x
\]
\[
\quad = \quad (x^T Q^T)(Qx)
\]
\[
\quad = \quad (Qx)^T (Qx)
\]
\[
\quad = \quad y^T y
\]
\[
\quad = \quad \| y \|_2
\]

and therefore the 2-norm is invariant under an orthogonal transformation.
Solution (b): To prove $\|QA\|_2 = \|A\|_2$ we must exploit the result from (a). We have

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2$$
$$= \max_{\|x\|_2 = 1} \|QAx\|_2$$
$$= \max_{\|x\|_2 = 1} \|(QA)x\|_2$$
$$= \|QA\|_2$$

Solution (c): To show $\|QA\|_F = \|A\|_F$ we must also exploit the result from (a) and a simple observation about the relationship of the Frobenius matrix norm and the 2-norms of the columns of the matrix (we mentioned this in class).

$$\|A\|_F^2 = \Sigma_i \|Ae_i\|_2^2$$
$$= \Sigma_i \|Q(Ae_i)\|_2^2$$
$$= \Sigma_i \|(QA)e_i\|_2^2$$
$$= \|QA\|_F^2$$

Therefore taking the square root we have $\|QA\|_F = \|A\|_F$ as desired.

Solution (d): To show $\|AQ\|_2 = \|A\|_2$ where $A \in \mathbb{R}^{n \times n}$, and $Q \in \mathbb{R}^{n \times n}$ is orthogonal we note that we must have

$$Q^TQ = I$$
$$QQ^T = I$$
$$Q^T = Q^{-1}$$

and for any $x \in \mathbb{R}^n$, $\exists y$ (uniquely) such that $Qy = x$ and $\|x\|_2 = \|y\|_2$. Therefore it follows that

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2$$
$$= \max_{\|y\|_2 = 1} \|A(Qy)\|_2$$
$$= \max_{\|y\|_2 = 1} \|(AQ)y\|_2$$
$$= \|AQ\|_2$$

Solution (e):

To show $\|AQ\|_F = \|A\|_F$ where $A \in \mathbb{R}^{n \times n}$, and $Q \in \mathbb{R}^{n \times n}$ is orthogonal we exploit an observation relating the Frobenius norm of a matrix and its transpose.

We have

$$\|A\|_F^2 = \|A^T\|_F^2$$
$$= \Sigma_k \|A^T e_k\|_2^2$$
$$= \Sigma_k \|Q^T (A^T e_k)\|_2^2$$
$$= \Sigma_k \|(Q^T A^T)e_k\|_2^2$$
$$= \|Q^T A^T\|_F^2$$
$$= \|AQ\|_F^2$$

Therefore $\|AQ\|_F = \|A\|_F$ as desired.