CHAPTER 2

A REVIEW OF SOME BASIC CONCEPTS AND RESULTS FROM THEORETICAL LINEAR ALGEBRA

Topics covered

- Basic Concepts for Theoretical Linear Algebra

2.1 INTRODUCTION

Although a first course in linear algebra is a prerequisite for this book, for the sake of completeness, we establish some notations and quickly review the basic definitions and concepts on matrices and vectors in this chapter. Fundamental results on **vector and matrix norms** are described in some details. These results will be used frequently in the later chapters of the book. *The students can review material of this chapter, as needed.*

2.2 ORTHOGONALITY OF VECTORS AND SUBSPACES

Let $u = (u_1, u_2, \ldots, u_n)^{\mathrm{T}}$ and $v = (v_1, v_2, \ldots, v_n)^{\mathrm{T}}$ be two n-dimensional column vectors. The angle θ between two nonzero vectors u and v is given by

$$\cos(\theta) = \frac{u^*v}{\|u\|\|v\|},$$

where $u^*v = \sum_{i=1}^{n} \bar{u}_i v_i$, is the *inner product* of the vectors u and v. The vectors u and v are **orthogonal** if $\theta = 90°$, that is, if $u^*v = 0$. The symbol $\perp$ is used to denote orthogonality. The set of vectors $\{x_1, x_2, \ldots, x_k\}$ in $\mathbb{C}^n$ are **mutually**

orthogonal if $x_i^* x_j = 0$ for $i \neq j$, and **orthonormal** if $x_i^* x_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta function; that is, $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$, and "$*$" denotes complex conjugate transpose.

Let S be a nonempty subset of $\mathbb{C}^n$. Then S is called a **subspace** of $\mathbb{C}^n$ if $s_1, s_2 \in S$ implies $c_1 s_1 + c_2 s_2 \in S$, where c_1 and c_2 are arbitrary scalars. That is, S is a subspace if any linear combination of two vectors in S is also in S.

For every subspace there is a unique smallest positive integer r such that every vector in the subspace can be expressed as a linear combination of at most r vectors in the subspace; r is called the **dimension** of the subspace and is denoted by $\dim[S]$.

Any set of r linearly independent vectors from S of $\dim[S] = r$ forms a **basis** of the subspace.

The **orthogonal complement** of a subspace S is defined by $S^\perp = \{y \in \mathbb{C}^n \mid y^* x = 0 \text{ for all } x \in S\}$.

The set of vectors $\{v_1, v_2, \ldots, v_n\}$ form an **orthonormal basis** of a subspace S if these vectors form a basis of S and are orthonormal.

Two subspaces S_1 and S_2 of $\mathbb{C}^n$ are said to be **orthogonal** if $s_1^* s_2 = 0$ for every $s_1 \in S_1$ and every $s_2 \in S_2$. Two orthogonal subspaces S_1 and S_2 will be denoted by $S_1 \perp S_2$.

2.3 MATRICES

In this section, we state some fundamental concepts and results involving the *eigenvalues and eigenvectors: rank, range, nulspaces*, and the *inverse of a matrix.*

2.3.1 The Characteristic Polynomial, the Eigenvalues, and the Eigenvectors of a Matrix

Let A be an $n \times n$ matrix. Then the polynomial $p_A(\lambda) = \det(\lambda I - A)$ is called the **characteristic polynomial**. The zeros of the characteristic polynomial are called the **eigenvalues** of A. This is equivalent to the following: $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if there exists a nonzero vector x such that $Ax = \lambda x$.

The vector x is called a **right eigenvector** (or just an **eigenvector**) of A. A nonzero vector y is called a **left eigenvector** if $y^* A = \lambda y^*$ for some $\lambda \in \mathbb{C}$.

If an eigenvalue of A is repeated s times, then it is called a multiple eigenvalue of multiplicity s. If $s = 1$, then the eigenvalue is a **simple eigenvalue**.

Definition 2.3.1. *If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the n eigenvalues of A, then* $\max |\lambda_i|$, *$i = 1, \ldots, n$ is called the* **spectral radius** *of A. It is denoted by $\rho(A)$.*

Invariant Subspaces

A subspace S of $\mathbb{C}^n$ is called the invariant subspace or A-**invariant** if $Ax \in S$ for every $x \in S$.

Clearly, an eigenvector x of A defines a one-dimensional invariant subspace.

An A-invariant subspace $S \subseteq \mathbb{C}^n$ is called a **stable invariant subspace** if the eigenvectors in S correspond to the eigenvalues of A with negative real parts.

The Cayley–Hamilton Theorem

The **Cayley–Hamilton theorem** states that the characteristic polynomial of A is an annihilating polynomial of A. That is, if $p_A(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n I$, then $p_A(A) = A^n + a_1 A^{n-1} + \cdots + a_n I = 0$.

Definition 2.3.2. *An $n \times n$ matrix A having fewer than n linearly independent eigenvectors is called a* **defective matrix**.

Example 2.3.1. The matrix

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

is defective. It has only one eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

2.3.2 Range and Nullspaces

For every $m \times n$ matrix A, there are two important associated subspaces: the **range** of A, denoted by $R(A)$, and the **null space** of A, denoted by $N(A)$, defined as follows:

$$R(A) = \{b \mid b = Ax \text{ for some } x\},$$
$$N(A) = \{x \mid Ax = 0\}.$$

The dimension of $N(A)$ is called the **nullity** of A and is denoted by $\mathbf{null}(A)$.

2.3.3 Rank of a Matrix

Let A be an $m \times n$ matrix. Then the subspace spanned by the row vectors of A is called the **row space** of A. The subspace spanned by the columns of A is called the **column space** of A. The range of A, $R(A)$, is the same as the column space of A.

The **rank** of a matrix A is the dimension of the column space of A. It is denoted by $\text{rank}(A)$.

An $m \times n$ matrix is said to have **full column rank if its columns are linearly independent**. The **full row rank** is similarly defined. A matrix A is said to have **full rank** if it has either full row rank or full column rank. If A does not have full rank, it is called **rank deficient**.

The best way to find the rank of a matrix in a computational setting is via the *singular value decomposition* (SVD) of a matrix (see **Chapter 4**).

2.3.4 The Inverse of a Matrix

An $n \times n$ matrix A is said to be invertible if there exists an $n \times n$ matrix B such that $AB = BA = I$. The inverse of A is denoted by A^{-1}. An invertible matrix A is often called **nonsingular**.

An interesting property of the inverse of the product of two invertible matrices is: $(AB)^{-1} = B^{-1}A^{-1}$.

Theorem 2.3.1. *For an $n \times n$ matrix A, the following are equivalent:*

- *A is nonsingular.*
- $\det(A)$ *is nonzero.*
- $\operatorname{rank}(A) = n$.
- $N(A) = \{0\}$.
- *A^{-1} exists.*
- *A has linearly independent rows and columns.*
- *The eigenvalues of A are nonzero.*
- *For all x, $Ax = 0$ implies that $x = 0$.*
- *The system $Ax = b$ has a unique solution.*

2.3.5 The Generalized Inverse of a Matrix

Let A^* be the complex conjugate transpose of A; that is, $A^* = \left(\bar{A}\right)^{\mathrm{T}}$.

The (Moore–Penrose) **generalized inverse** of a matrix A, denoted by $A^{\dagger}$, is a unique matrix satisfying the following properties: (i) $AA^{\dagger}A = A$, (ii) $A^{\dagger}AA^{\dagger} = A^{\dagger}$, (iii) $(AA^{\dagger})^* = AA^{\dagger}$, and (iv) $(A^{\dagger}A)^* = A^{\dagger}A$.

Note: If A is square and invertible, then $A^{\dagger} = A^{-1}$.

2.3.6 Similar Matrices

Two matrices A and B are called **similar** if there exists a nonsingular matrix T such that

$$T^{-1}AT = B.$$

An important property of similar matrices: **Two similar matrices have the same eigenvalues**. However, two matrices having the same eigenvalues need not be similar.

2.3.7 Orthogonal Projection

Let S be a subspace of $\mathbb{C}^n$. Then an $n \times n$ matrix P having the properties: (i) $R(P) = S$, (ii) $P^* = P$ (P is **Hermitian**), (iii) $P^2 = P$ (P is **idempotent**) is called the **orthogonal projection** onto S or simply the **projection matrix**. We denote the orthogonal projection P onto S by P_S. **The orthogonal projection onto a subspace is unique**.

Let $V = (v_1, \ldots, v_k)$, where $\{v_1, \ldots, v_k\}$ is an orthonormal basis for a subspace S. Then,

$$P_S = VV^*$$

is the unique orthogonal projection onto S. **Note that V is not unique, but P_S is.**

A Relationship Between P_S and $P_{S^\perp}$

If P_S is the orthogonal projection onto S, then $I - P_S$, where I is the identity matrix of the same order as P_S, is the orthogonal projection onto $S^\perp$. It is denoted by $P_S^\perp$.

The Orthogonal Projection onto $R(A)$

It can be shown that if A is $m \times n$ $(m \geq n)$ and has **full rank**, then the orthogonal projection P_A onto $R(A)$ is given by:

$$P_A = A(A^*A)^{-1}A^*.$$

2.4 SOME SPECIAL MATRICES

2.4.1 Diagonal and Triangular Matrices

An $m \times n$ matrix $A = (a_{ij})$ is a **diagonal matrix** if $a_{ij} = 0$ for $i \neq j$. We write $A = \mathrm{diag}(a_{11}, \ldots, a_{ss})$, where $s = \min(m, n)$. An $n \times n$ matrix A is a **block diagonal matrix** if it is a diagonal matrix whose each diagonal entry is a square matrix. It is written as:

$$A = \mathrm{diag}(A_{11}, \ldots, A_{kk}),$$

where each A_{ii} is a square matrix. The sum of the orders of A_{ii}, $i = 1, \ldots, k$ is n.

An $m \times n$ matrix $A = (a_{ij})$ is an **upper triangular** matrix if $a_{ij} = 0$ for $i > j$.

The transpose of an upper triangular matrix is **lower triangular**; that is, $A = (a_{ij})$ is lower triangular if $a_{ij} = 0$ for $i < j$.

2.4.2 Unitary (Orthogonal) Matrix

A complex square matrix U is **unitary** if $UU^* = U^*U = I$, where $U^* = (\overline{U})^{\mathrm{T}}$. A real square matrix O is **orthogonal** if $OO^{\mathrm{T}} = O^{\mathrm{T}}O = I$. If U is an $n \times k$ matrix such that $U^*U = I_k$, then U is said to be **orthonormal**.

Orthogonal matrices play a very important role in numerical matrix computations.

The following important properties of orthogonal (unitary) matrices are attractive for numerical computations: (i) The inverse of an orthogonal (unitary) matrix O is just its transpose (conjugate transpose), (ii) The product of two orthogonal (unitary) matrices is an orthogonal (unitary) matrix, (iii) The 2-norm and the Frobenius norm are invariant under multiplication by an orthogonal (unitary) matrix **(See Section 2.6)**, and (iv) The error in multiplying a matrix by an orthogonal matrix is not magnified by the process of numerical matrix multiplication **(See Chapter 3)**.

2.4.3 Permutation Matrix

A nonzero square matrix P is called a **permutation matrix** if there is exactly one nonzero entry in each row and column which is 1 and the rest are all zero.

Effects of Premultiplication and Postmultiplication by a permutation matrix

When a matrix A is premultiplied by a permutation matrix P, the effect is a permutation of the rows of A. Similarly, if A is postmultiplied by a permutation matrix, the effect is a permutation of the columns of A.

Some Important Properties of Permutation Matrices

- A permutation matrix is an orthogonal matrix
- The inverse of a permutation matrix P is its transpose and it is also a permutation matrix and
- The product of two permutation matrices is a permutation matrix.

2.4.4 Hessenberg (Almost Triangular) Matrix

A square matrix A is **upper Hessenberg** if $a_{ij} = 0$ for $i > j + 1$. The transpose of an upper Hessenberg matrix is a lower Hessenberg matrix, that is, a square matrix $A = (a_{ij})$ is a **lower Hessenberg matrix** if $a_{ij} = 0$ for $j > i + 1$. A square matrix A that is both upper and lower Hessenberg is **tridiagonal**.

$$\underset{\text{Lower Hessenberg}}{\begin{pmatrix} * & * & & 0 \\ \vdots & & \ddots & \\ * & \vdots & & * \\ * & * & \cdots & * \end{pmatrix}} \qquad \underset{\text{Upper Hessenberg}}{\begin{pmatrix} * & \cdots & * & * \\ * & \cdots & * & * \\ & \ddots & \vdots & \vdots \\ 0 & & * & * \end{pmatrix}}$$

An upper Hessenberg matrix $A = (a_{ij})$ is **unreduced** if $a_{i,i-1} \neq 0$ for $i = 2, 3, \ldots, n$.

Similarly, a lower Hessenberg matrix $A = (a_{ij})$ is **unreduced** if $a_{i,i+1} \neq 0$ for $i = 1, 2, \ldots, n-1$.

2.4.5 Companion Matrix

An unreduced upper Hessenberg matrix of the form

$$C = \begin{pmatrix} 0 & 0 & \cdots & \cdots & c_1 \\ 1 & 0 & \cdots & \cdots & c_2 \\ 0 & 1 & \cdots & \cdots & \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & c_n \end{pmatrix}$$

is called an **upper companion matrix**. The transpose of an upper companion matrix is a **lower companion matrix**.

The **characteristic polynomial** of the companion matrix C is:

$$\det(\lambda I - C) = \det(\lambda I - C^{\mathrm{T}}) = \lambda^n - c_n\lambda^{n-1} - c_{n-1}\lambda^{n-2} - \cdots - c_2\lambda - c_1.$$

2.4.6 Nonderogatory Matrix

A matrix A is **nonderogatory** if and only if it is similar to a companion matrix of its characteristic polynomial. That is, A is a nonderogatory matrix if and only if there exists a nonsingular matrix T such that $T^{-1}AT$ is a companion matrix.

Remark

- An unreduced Hessenberg matrix is nonderogatory, but the converse is not true.

2.4.7 The Jordan Canonical Form of a Matrix

For an $n \times n$ complex matrix A, there exists a nonsingular matrix T such that

$$T^{-1}AT = J = \mathrm{diag}(J_1, \ldots, J_k),$$

where

$$J_i = \begin{pmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}$$

is $m_i \times m_i$ and $m_1 + \cdots + m_k = n$.

The matrices J_i are called **Jordan matrices** or **Jordan blocks** and J is called the **Jordan Canonical Form (JCF) of** A. For each $j = 1, 2, \ldots, k$, λ_j is the eigenvalue of A with multiplicity m_j. **The same eigenvalue can appear in more than one block**.

Note: The matrix A is **nonderogatory** if its JCF has only one Jordan block associated with each distinct eigenvalue.

If $T = (t_1, t_2, \ldots, t_{m_1}; t_{m_1+1}, \ldots, t_{m_2}; \ldots, t_n)$.

Then $t_1, \ldots, t_{m_1}$ must satisfy

$$At_1 = \lambda_1 t_1$$

and $At_{i+1} = \lambda_1 t_{i+1} + t_i$, $i = 1, 2, \ldots, m_1 - 1$.

Similarly, relations hold for the other vectors in T. The vectors t_i are called the **generalized eigenvectors** or **principal vectors** of A.

2.4.8 Positive Definite Matrix

A real symmetric matrix A is **positive definite** (positive semidefinite) if $x^{\mathrm{T}}Ax > 0$ (≥ 0) for every nonzero vector x.

Similarly, a complex Hermitian matrix A is positive definite (positive semidefinite) if $x^*Ax > 0$ (≥ 0) for every nonzero complex vector x.

A commonly used notation for a symmetric positive definite (positive semidefinite) matrix is $A > 0\,(\geq 0)$.

Unless otherwise mentioned, a real symmetric or a complex Hermitian positive definite matrix will be referred to as a **positive definite matrix**.

A symmetric positive definite matrix A admits the **Cholesky factorization** $A = HH^{\mathrm{T}}$, where H is a lower triangular matrix with positive diagonal entries. **The most numerically efficient and stable way to check if a real symmetric matrix is positive definite is to compute its Cholesky factorization and see if the diagonal entries of the Cholesky factor are all positive**. See Chapter 3 (Section 3.4.2) for details.

2.4.9 Block Matrices

A matrix whose each entry is a matrix is called a **block matrix**. A **block diagonal matrix** is a diagonal matrix whose each entry is a matrix. A **block triangular matrix** is similarly defined.

The JCF is an example of a **block diagonal matrix**.

Suppose A is partitioned in the form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

then A is nonsingular if and only if $A_S = A_{22} - A_{21}A_{11}^{-1}A_{12}$, called the **Schur-Complement** of A, is nonsingular (assuming that A_{11} is nonsingular) and in this case, the inverse of A is given by:

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}A_S^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_S^{-1} \\ -A_S^{-1}A_{21}A_{11}^{-1} & A_S^{-1} \end{pmatrix}.$$

2.5 VECTOR AND MATRIX NORMS

2.5.1 Vector Norms

Let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

be an n-vector in $\mathbb{C}^n$. Then, a vector norm, denoted by the symbol $\|x\|$, is a real-valued **continuous** function of the components $x_1, x_2, \ldots, x_n$ of x, satisfying the following properties:

1. $\|x\| > 0$ for every nonzero x. $\|x\| = 0$ if and only if x is the zero vector.
2. $\|\alpha x\| = |\alpha|\|x\|$ for all x in $\mathbb{C}^n$ and for all scalars α.
3. $\|x + y\| \leq \|x\| + \|y\|$ for all x and y in $\mathbb{C}^n$.

The last property is known as the **Triangle Inequality**.

Note: $\|-x\| = \|x\|$ and $\big|\, \|x\| - \|y\| \,\big| \leq \|x - y\|$. It is simple to verify that the following are vector norms.

Some Commonly Used Vector Norms

1. $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$ (sum norm or 1-norm)
2. $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ (Euclidean norm or 2-norm)
3. $\|x\|_\infty = \max_i |x_i|$ (maximum or ∞-norm)

The above three are special case of the ***p*-norm** or **Hölder norm** defined by $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$ for any $p \geq 1$.

Unless otherwise stated, by $\|x\|$ we will mean $\|x\|_2$.

Example 2.5.1. Let $x = (1, 1, -2)^{\mathrm{T}}$. Then $\|x\|_1 = 4$, $\|x\|_2 = \sqrt{1^2 + 1^2 + (-2)^2} = \sqrt{6}$, and $\|x\|_\infty = 2$.

An important property of the Hölder norm is the **Hölder inequality**

$$|x^* y| \leq \|x\|_p \, \|y\|_q, \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

A special case of the Hölder inequality is the **Cauchy–Schwartz** inequality: $|x^* y| \leq \|x\|_2 \|y\|_2$.

Equivalence Property of the Vector norms

All vector norms are **equivalent** in the sense that there exist positive constants α and β such that $\alpha\|x\|_\mu \leq \|x\|_\nu \leq \beta\|x\|_\mu$, for all x, where μ and ν specify the nature of norms.

For the 2, 1, or ∞ norms, we can compute α and β easily and have the following inequalities:

Theorem 2.5.1. *Let x be in $\mathbb{C}^n$. Then*

1. $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$
2. $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$
3. $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$

2.5.2 Matrix Norms

Let A be an $m \times n$ matrix. Then, analogous to the vector norm, we define the matrix norm for $\|A\|$ in $\mathbb{C}^{m \times n}$ with the following properties:

1. $\|A\| \geq 0$; $\|A\| = 0$ only if A is the zero matrix
2. $\|\alpha A\| = |\alpha|\|A\|$ for any scalar α
3. $\|A + B\| \leq \|A\| + \|B\|$, where B is also an $m \times n$ matrix.

Subordinate Matrix Norms

Given a matrix A and a vector norm $\|\cdot\|_p$ on $\mathbb{C}^n$, a nonnegative number defined by:

$$\|A\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

satisfies all the properties of a matrix norm. This norm is called the matrix norm **subordinate** to (or induced by) the p-norm.

A very useful and frequently used property of a subordinate matrix norm $\| A \|_p$ **(we shall sometimes call it the p-norm of a matrix A)** is

$$\|Ax\|_p \leq \|A\|_p \|x\|_p.$$

Two important p-norms of an $m \times n$ matrix are: (i) $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|$ **(maximum column sum norm)** and (ii) $\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|$ **(maximum row sum norm)**.

The Frobenius Norm

An important matrix norm is the Frobenius norm:

$$\|A\|_{\mathrm{F}} = \left[\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right]^{1/2}.$$

A matrix norm $\|\cdot\|_{\mathrm{M}}$ and a vector norm $\|\cdot\|_{\mathrm{v}}$ are consistent if for all matrices A and vectors x, the following inequality holds:

$$\|Ax\|_{\mathrm{v}} \leq \|A\|_{\mathrm{M}} \, \|x\|_{\mathrm{v}}.$$

Consistency Property of the Matrix Norm

A matrix norm is consistent if, for any two matrices A and B compatible for matrix multiplication, the following property is satisfied:

$$\|AB\| \leq \|A\| \, \|B\|.$$

The Frobenius norm and all subordinate norms are consistent.

Notes

1. For the identity matrix I, $\|I\|_{\mathrm{F}} = \sqrt{n}$, whereas $\|I\|_1 = \|I\|_2 = \|I\|_\infty = 1$.
2. $\|A\|_{\mathrm{F}}^2 = \mathrm{trace}(A^*A)$, where trace (A) is defined as the sum of the diagonal entries of A, that is, if $A = (a_{ij})$, then trace $(A) = a_{11} + a_{22} + \cdots + a_{nn}$. The trace of A will, sometimes, be denoted by $\mathrm{Tr}(A)$ or $\mathrm{tr}(A)$.

Equivalence Property of Matrix Norms

As in the case of vector norms, the matrix norms are also related. There exist scalars α and β such that: $\alpha\|A\|_\mu \leq \|A\|_\nu \leq \beta\|A\|_\mu$. In particular, the following inequalities relating various matrix norms are true and are used very frequently in practice. We state the theorem without proof. For a proof, see Datta (1995, pp. 28–30).

Theorem 2.5.2. *Let A be $m \times n$. Then,*

1. $\frac{1}{\sqrt{n}}\|A\|_\infty \le \|A\|_2 \le \sqrt{m}\|A\|_\infty$.
2. $\|A\|_2 \le \|A\|_{\mathrm{F}} \le \sqrt{n}\|A\|_2$.
3. $\frac{1}{\sqrt{m}}\|A\|_1 \le \|A\|_2 \le \sqrt{n}\|A\|_1$.
4. $\|A\|_2 \le \sqrt{\|A\|_1\|A\|_\infty}$.

2.6 NORM INVARIANT PROPERTIES UNDER UNITARY MATRIX MULTIPLICATION

We conclude the chapter by listing some very useful norm properties of unitary matrices that are often used in practice.

Theorem 2.6.1. *Let U be an unitary matrix. Then,*

$$\|U\|_2 = 1.$$

Proof. $\| U \|_2 = \sqrt{\rho(U^*U)} = \sqrt{\rho(I)} = 1$. (Recall that $\rho(A)$ denotes the spectral radius of A.) ■

The next two theorems show that **2-norm and the Frobenius norm are invariant under multiplication by a unitary matrix**.

Theorem 2.6.2. *Let U be an unitary matrix and AU be defined. Then,*

1. $\|AU\|_2 = \|A\|_2$
2. $\|AU\|_{\mathrm{F}} = \|A\|_{\mathrm{F}}$

Proof.

1. $\|AU\|_2 = \sqrt{\rho(U^*A^*AU)} = \sqrt{\rho(A^*A)} = \|A\|_2$ (Note that $U^* = U^{-1}$, and two similar matrices have the same eigenvalues).
2. $\|AU\|_{\mathrm{F}} = \mathrm{trace}(U^*A^*AU) = \mathrm{trace}(A^*A) = \|A\|_{\mathrm{F}}^2$ (Note that the trace of a matrix remains invariant under similarity transformation).

Thus $\|AU\|_{\mathrm{F}} = \|A\|_{\mathrm{F}}$. ■

Similarly, if UA is defined, then we have

Theorem 2.6.3.

1. $||UA||_2 = ||A||_2$
2. $||UA||_{\mathrm{F}} = ||A||_{\mathrm{F}}$

Proof. The proof is similar to Theorem 2.6.2. ■

2.7 KRONECKER PRODUCT, KRONECKER SUM, AND VEC OPERATION

Let $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{r \times s}$, then the $mr \times ns$ matrix defined by:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}$$

is called the **Kronecker product** of A and B.

If A and B are invertible, then $A \otimes B$ is invertible and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

The Eigenvalues of the Kronecker Product and Kronecker Sum

Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $A \in \mathbb{C}^{n \times n}$, and $\mu_1, \dots, \mu_m$ be the eigenvalues of $B \in \mathbb{C}^{m \times m}$. Then it can be shown that the eigenvalues of $A \otimes B$ are the mn numbers $\lambda_i \mu_j, i = 1, \dots, n;\ j = 1, \dots, m$, and the eigenvalues of $A \oplus B$ are the mn numbers $\lambda_i + \mu_j, i = 1, \dots, n;\ j = 1, \dots, m$.

Vec Operation

Let $X \in \mathbb{C}^{m \times n}$ and $X = (x_{ij})$.

Then the vector obtained by stacking the columns of X in one vector is denoted by $\text{vec}(X)$:

$$\text{vec}(X) = (x_{11}, \dots, x_{m1}, x_{12}, \dots, x_{m2}, \dots, x_{1n}, \dots, x_{mn})^{\mathrm{T}}.$$

If $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$, then it can be shown that $\text{vec}(AX + XB) = ((I_n \otimes A) + (B^{\mathrm{T}} \otimes I_m))\text{vec}\, X$.

The Kronecker products and vec operations are useful in the study of the existence, uniqueness, sensitivity, and numerical solutions of the Lyapunov and Sylvester equations (**see Chapter 8**).

2.8 CHAPTER NOTES AND FURTHER READING

Most of the material in this chapter can be found in standard linear algebra text books. Some such books are cited below.

For further reading of material of Section 2.7, the readers are referred to Horn and Johnson (1985).

References

Anton H. and Rorres C. *Elementary Linear Algebra with Applications*, John Wiley, New York, 1987.

Dattta B.N. *Numerical Linear Algebra and Applications*, Brooks/Cole Publishing Company, Pacific Grove, 1995 (*Custom published by Brooks/Cole, 2003*).

Horn R.A. and Johnson C.R. *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

Lay D.C. *Linear Algebra and Its Applications*, Addison-Wesley, Reading, MA, 1994.

Leon S.J. *Linear Algebra with Applications*, 4th edn, Macmillan, New York, 1994.

Noble B. and Daniel J.W. *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

Strang G. *Linear Algebra and Its Applications*, 3rd edn, Academic Press, New York, 1988.