CHAPTER 4

Numerical Ranges, Matrix Norms, and Special Operations

Introduction: This chapter is devoted to a few basic topics on matrices. We first study the numerical range and radius of a square matrix and matrix norms. We then introduce three important special matrix operations: the Kronecker product, the Hadamard product, and compound matrices.

4.1 Numerical Range and Radius

Let $A$ be an $n \times n$ complex matrix. For $x = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$, as usual, $\|x\| = (\sum_{i=1}^{n} |x_i|^2)^{1/2}$ is the norm of $x$ (Section 1.4). The numerical range, also known as the field of values, of $A$ is defined by

$$W(A) = \{x^*Ax : \|x\| = 1, \ x \in \mathbb{C}^n\}.$$

For example, if

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

then $W(A)$ is the closed interval $[0, 1]$, and if

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

then $W(A)$ is the closed elliptical disc with foci at $(0, 0)$ and $(1, 0)$, minor axis 1, and major axis $\sqrt{2}$ (Problem 9).
One of the celebrated and fundamental results on numerical range is the Toeplitz–Hausdorff convexity theorem.

Theorem 4.1 (Toeplitz–Hausdorff) The numerical range of a square matrix is a convex compact subset of the complex plane.

Proof. For convexity, if $W(A)$ is a singleton, there is nothing to show. Suppose $W(A)$ has more than one point. We prove that the line segment joining any two distinct points in $W(A)$ lies in $W(A)$; that is, if $u, v \in W(A)$, then $tu + (1 - t)v \in W(A)$ for all $t \in [0, 1]$.

For any complex numbers $\alpha$ and $\beta$, it is easy to verify that $W(\alpha I + \beta A) = \{\alpha + \beta z : z \in W(A)\}$.

Intuitively the convexity of $W(A)$ does not change under shifting, scaling, and rotation. Thus, we may assume that the two points to be considered are 0 and 1, and show that $[0, 1] \subseteq W(A)$. Write

$$A = H + iK,$$

where

$$H = \frac{1}{2}(A + A^*) \quad \text{and} \quad K = \frac{1}{2i}(A - A^*)$$

are Hermitian matrices. Let $x$ and $y$ be unit vectors in $\mathbb{C}^n$ such that

$$x^*Ax = 0, \quad y^*Ay = 1.$$

It follows that $x$ and $y$ are linearly independent and that

$$x^*Hx = x^*Kx = y^*Ky = 0, \quad y^*Hy = 1.$$

We may further assume that $x^*Ky$ has real part zero; otherwise, one may replace $x$ with $cx$, $c \in \mathbb{C}$, and $|c| = 1$, so that $cx^*Ky$ is 0 or a pure complex number without changing the value of $x^*Ax$.

Note that $tx + (1 - t)y \neq 0$, $t \in [0, 1]$. Define for $t \in [0, 1]$

$$z(t) = \frac{1}{\|tx + (1 - t)y\|^2}(tx + (1 - t)y).$$

Then $z(t)$ is a unit vector. It is easy to compute that for all $t \in [0, 1]$

$$z(t)^*Kz(t) = 0.$$
The convexity of $W(A)$ then follows, for

$$\{z(t)^{*}Az(t) : 0 \leq t \leq 1\} = [0, 1].$$

The compactness of $W(A)$, meaning the boundary is contained in $W(A)$, is seen by noting that $W(A)$ is the range of the continuous function $x \mapsto x^{*}Ax$ on the compact set $\{x \in \mathbb{C}^n : \|x\| = 1\}$. (A continuous function maps a compact set to a compact set.)

When considering the smallest disc centered at the origin that covers the numerical range, we associate with $W(A)$ a number

$$w(A) = \sup\{|z| : z \in W(A)\} = \sup_{\|x\|=1} |x^{*}Ax|$$

and call it the numerical radius of $A \in \mathbb{M}_n$. Note that the “sup” can be attained by some $z \in W(A)$. It is immediate that for any $x \in \mathbb{C}^n$

$$|x^{*}Ax| \leq w(A)\|x\|^2. \quad (4.1)$$

We now make comparisons of the numerical radius $w(A)$ to the largest eigenvalue $\rho(A)$ in absolute value, or the spectral radius, i.e.,

$$\rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\},$$

and to the largest singular value $\sigma_{\max}(A)$, also called the spectral norm. It is easy to see (Problem 7) that

$$\sigma_{\max}(A) = \sup_{\|x\|=1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

and that for every $x \in \mathbb{C}^n$

$$\|Ax\| \leq \sigma_{\max}(A)\|x\|.$$

**Theorem 4.2** Let $A$ be a square complex matrix. Then

$$\rho(A) \leq w(A) \leq \sigma_{\max}(A) \leq 2w(A).$$

**Proof.** Let $\lambda$ be the eigenvalue of $A$ such that $\rho(A) = |\lambda|$, and let $u$ be a unit eigenvector corresponding to $\lambda$. Then

$$\rho(A) = |\lambda u^{*}u| = |u^{*}Au| \leq w(A).$$
The second inequality follows from the Cauchy–Schwarz inequality
\[ |x^*Ax| = |(Ax, x)| \leq \|Ax\|\|x\|. \]

We next show that \( \sigma_{\text{max}}(A) \leq 2w(A) \). It can be verified that
\[
4(Ax, y) = \left( A(x + y), x + y \right) - \left( A(x - y), x - y \right) \\
+ i \left( A(x + iy), x + iy \right) - i \left( A(x - iy), x - iy \right).
\]

Using (4.1), it follows that
\[
4|(Ax, y)| \leq w(A)(\|x + y\|^2 + \|x - y\|^2) \\
+ \|x + iy\|^2 + \|x - iy\|^2) = 4w(A)(\|x\|^2 + \|y\|^2).
\]

Thus, for any unit \( x \) and \( y \) in \( \mathbb{C}^n \), we have
\[
|(Ax, y)| \leq 2w(A).
\]

The inequality follows immediately from Problem 7. \( \blacksquare \)

**Theorem 4.3** Let \( A \in M_n \). Then \( \lim_{k \to \infty} A^k = 0 \) if and only if \( \rho(A) < 1 \); that is, all the eigenvalues of \( A \) have moduli less than 1.

**Proof.** Let \( A = P^{-1}TP \) be a Jordan decomposition of \( A \), where \( P \) is invertible and \( T \) is a direct sum of Jordan blocks with the eigenvalues \( \lambda_1, \ldots, \lambda_n \) of \( A \) on the main diagonal. Then \( A^k = P^{-1}T^kP \) and \( \rho(A^k) = (\rho(A))^k \). Thus, if \( A^k \) tends to zero, so does \( T^k \). It follows that \( \lambda^k \to 0 \) as \( k \to 0 \) for every eigenvalue of \( A \). Therefore \( \rho(A) < 1 \). Conversely, suppose \( \rho(A) < 1 \). We show that \( A^k \to 0 \) as \( k \to \infty \). It suffices to show that \( J^k \to 0 \) as \( k \to \infty \) for each Jordan block \( J \).

Suppose \( J \) is an \( m \times m \) Jordan block:
\[
J = \begin{pmatrix}
\lambda & 1 & \ldots & 0 \\
0 & \lambda & \ddots & \\
\vdots & \ddots & \ddots & 1 \\
0 & \ldots & 0 & \lambda
\end{pmatrix}.
\]
Upon computation, we have

\[ J^k = \begin{pmatrix}
\lambda^k & (k)\lambda^{k-1} & \cdots & (k)\lambda^{k-m+1} \\
0 & \lambda^k & \ddots & \vdots \\
\vdots & \ddots & \ddots & (k)\lambda^{k-1} \\
0 & \ldots & 0 & \lambda^k
\end{pmatrix}. \]

Recall from calculus that for any constants \( l \) and \( \lambda < 1 \)

\[ \lim_{k \to \infty} \left( \frac{k}{l} \right) \lambda^k = 0. \]

It follows that \( J^k \), thus \( A^k \), converges to 0 as \( k \to \infty \). ■

Problems

1. Find a nonzero matrix \( A \) so that \( \rho(A) = 0 \).

2. Find the eigenvalues, singular values, numerical radius, spectral radius, spectral norm, and numerical range for each of the following:

\[ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \]

3. Let \( A \) be an \( n \)-square complex matrix. Show that the numerical radius, spectral radius, spectral norm, and numerical range are unitarily invariant. That is, for instance, \( w(U^*AU) = w(A) \) for any \( n \)-square unitary matrix \( U \).

4. Show that the diagonal entries and the eigenvalues of a square matrix are contained in the numerical range of the matrix.

5. Let \( A \in \mathbb{M}_n \). Show that \( \frac{1}{n} \text{tr} A \) is contained in \( W(A) \). Conclude that for any nonsingular \( P \in \mathbb{M}_n \), \( W(P^{-1}AP - PAP^{-1}) \) contains 0.

6. Let \( A \) be a square complex matrix. Show that \( \frac{\|Ax\|}{\|x\|} \) is constant for all \( x \neq 0 \) if and only if all the singular values of \( A \) are identical.

7. Let \( A \) be a complex matrix. Show that

\[ \sigma_{\text{max}}(A) = \sqrt{\rho(A^*A)} = \sup_{\|x\| = 1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| = \|y\| = 1} |(Ax, y)|. \]
8. Show that for any square matrices $A$ and $B$ of the same size,
\[
\sigma_{\text{max}}(AB) \leq \sigma_{\text{max}}(A)\sigma_{\text{max}}(B),
\]
and
\[
\sigma_{\text{max}}(A + B) \leq \sigma_{\text{max}}(A) + \sigma_{\text{max}}(B).
\]

9. Show that the numerical range of \((0 0)
\begin{pmatrix}
0 & 0 \\
1 & 1
\end{pmatrix}
\) is a closed elliptical disc.

10. Take
\[
A = \begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]
and let $B = A^2$. Show that $w(A) < 1$. Find $w(B)$ and $w(AB)$.

11. Show that the numerical range of a normal matrix is the convex hull of its eigenvalues. That is, if $A \in \mathbb{M}_n$ is a normal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, then
\[
W(A) = \{t_1\lambda_1 + \cdots + t_n\lambda_n : t_1 + \cdots + t_n = 1, \text{ each } t_i \geq 0\}.
\]

12. Show that $W(A)$ is a polygon inscribed in the unit circle if $A$ is unitary, and that $W(A) \subseteq \mathbb{R}$ if $A$ is Hermitian. What can be said about $W(A)$ if $A$ is positive semidefinite?

13. Show that $w(A) = \rho(A) = \sigma_{\text{max}}(A)$ if $A$ is normal. Discuss the converse by considering
\[
A = \text{diag}(1, i, -1, -i) \oplus \begin{pmatrix}
0 & 1 \\
0 & 0
\end{pmatrix}.
\]

14. Prove or disprove that for any $n$-square complex matrices $A$ and $B$
\begin{enumerate}
\item $\rho(AB) \leq \rho(A)\rho(B)$.
\item $w(AB) \leq w(A)w(B)$.
\item $\sigma_{\text{max}}(AB) \leq \sigma_{\text{max}}(A)\sigma_{\text{max}}(B)$.
\end{enumerate}

15. Let $A$ be a square matrix. Show that for every positive integer $k$
\[
w(A^k) \leq (w(A))^k.
\]
Is it true in general that
\[
w(A^{k+m}) \leq w(A^k)w(A^m)?
\]
A matrix may be assigned numerical items in various ways. In addition to determinant, trace, eigenvalues, singular values, numerical radius, and spectral radius, matrix norm is another important one.

Recall from Section 1.4 of Chapter 1 that vectors can be measured by their norms. If $V$ is an inner product space, then the norm of a vector $v$ in $V$ is $\|v\| = \sqrt{(v, v)}$. The norm $\| \cdot \|$ on $V$ satisfies

i. $\|v\| \geq 0$ with equality if and only if $v = 0$,

ii. $\|cv\| \leq |c|\|v\|$ for all scalars $c$ and vectors $v$, and

iii. $\|u + v\| \leq \|u\| + \|v\|$ for all vectors $u, v$.

Like the norms for vectors being introduced to measure the magnitudes of vectors, norms for matrices are used to measure the “sizes” of matrices. We call a matrix function $\| \cdot \| : \mathbb{M}_n \rightarrow \mathbb{R}$ a matrix norm if for all $A, B \in \mathbb{M}_n$ and $c \in \mathbb{C}$, the following conditions are satisfied:

1. $\|A\| \geq 0$ with equality if and only if $A = 0$,

2. $\|cA\| \leq |c|\|A\|$,

3. $\|A + B\| \leq \|A\| + \|B\|$, and

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

We call $\| \cdot \|$ for matrices satisfying (1)–(3) a matrix-vector norm. In this book by a matrix norm we mean that all conditions (1)–(4) are met. Such a matrix norm is sometimes referred to as a multiplicative matrix norm. We use the notation $\| \cdot \|$ for both vector norm and matrix norm. Generally speaking, this won’t cause confusion as one can easily tell from what is being studied.

If a matrix is considered as a linear operator on an inner product space $V$, a matrix operator norm $\| \cdot \|_{\text{op}}$ can be induced as follows:

$$\|A\|_{\text{op}} = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| = 1} \|Ax\|.$$
From the previous section (Problems 7 and 8, Section 4.1), we see that the spectral norm is a matrix (operator) norm on $\mathbb{M}_n$ induced by the ordinary inner product on $\mathbb{C}^n$.

Matrices can be viewed as vectors in the matrix space $\mathbb{M}_n$ equipped with the inner product $(A, B) = \text{tr}(B^*A)$. Matrices as vectors under the inner product have vector norms. One may check that this vector norm for matrices is also a (multiplicative) matrix norm.

Two observations on matrix norms follow: first, a matrix norm $\| \cdot \| : A \mapsto \| A \|$ is a continuous function on the matrix space $\mathbb{M}_n$ (Problem 2), and second, $\rho(\cdot) \leq \| \cdot \|$ for any matrix norm $\| \cdot \|$. Reason: If $Ax = \lambda x$, where $x \neq 0$ and $\rho(A) = |\lambda|$, then, by letting $X$ be the $n \times n$ matrix with all columns equal to the eigenvector $x$,

$$\rho(A)\|X\| = \|\lambda X\| = \|AX\| \leq \|A\|\|X\|.$$ 

Nevertheless, the numerical radius $\rho(\cdot)$ is not a matrix norm. (Why?) The following result reveals a relation between the two.

**Theorem 4.4** Let $\| \cdot \|$ be a matrix norm. Then for every $A \in \mathbb{M}_n$

$$\rho(A) = \lim_{k \to \infty} \| A^k \|^{1/k}.$$ 

**Proof.** The eigenvalues of $A^k$ are the $k$th powers of those of $A$. Since spectral radius is dominated by norm, for every positive integer $k$,

$$(\rho(A))^k = \rho(A^k) \leq \| A^k \| \quad \text{or} \quad \rho(A) \leq \| A^k \|^{1/k}.$$ 

On the other hand, for any $\epsilon > 0$, let

$$A_\epsilon = \frac{1}{\rho(A) + \epsilon} A.$$ 

Then $\rho(A_\epsilon) < 1$. By Theorem 4.3, $A_\epsilon^k$ tends to 0 as $k \to \infty$. Thus, because the norm is a continuous function, for $k$ large enough,

$$\| A_\epsilon^k \| < 1 \quad \text{or} \quad \| A^k \| \leq (\rho(A) + \epsilon)^k.$$ 

Therefore

$$\| A^k \|^{1/k} \leq \rho(A) + \epsilon.$$
In summary, for any \( \epsilon > 0 \) and \( k \) large enough
\[
\rho(A) \leq \|A^k\|^{1/k} \leq \rho(A) + \epsilon.
\]
The conclusion follows immediately by letting \( \epsilon \) approach 0.

Now we turn our attention to an important class of matrix norms: unitarily invariant norms. We say a matrix (vector) norm \( \|\cdot\| \) on \( \mathbb{M}_n \) is unitarily invariant if for any \( A \in \mathbb{M}_n \) and for all unitary \( U, V \in \mathbb{M}_n \)
\[
\|UAV\| = \|A\|.
\]
The spectral norm \( \sigma_{\text{max}} : A \mapsto \sigma_{\text{max}}(A) \) is a matrix norm and it is unitarily invariant because \( \sigma_{\text{max}}(UAV) = \sigma_{\text{max}}(A) \). The Frobenius norm (also known as the Euclidean norm or Hilbert–Schmidt norm) is the matrix norm induced by the inner product \( (A, B) = \text{tr}(B^*A) \) on the matrix space \( \mathbb{M}_n \)
\[
\|A\|_F = \left( \text{tr}(A^*A) \right)^{1/2} = \left( \sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2}.
\]
(4.2)

With \( \sigma_i(A) \) denoting the singular values of \( A \), we see that
\[
\|A\|_F = \left( \sum_{i=1}^n \sigma_i^2(A) \right)^{1/2}.
\]
Thus \( \|A\|_F \) is uniquely determined by the singular values of \( A \). Consequently, the Frobenius norm is a unitarily invariant matrix norm.

Actually, the spectral norm and the Frobenius norm belong to two larger families of unitarily invariant norms: the Ky Fan \( k \)-norms and the Schatten \( p \)-norms, which we study more in Chapter 10.

**Ky Fan \( k \)-norm:** Let \( k \leq n \) be a positive integer. Define
\[
\|A\|_{(k)} = \sum_{i=1}^k \sigma_i(A), \quad A \in \mathbb{M}_n.
\]

**Schatten \( p \)-norm:** Let \( p \geq 1 \) be a real number. Define
\[
\|A\|_p = \left( \sum_{i=1}^n \sigma_i^p(A) \right)^{1/p}, \quad A \in \mathbb{M}_n.
\]
It is readily seen that the spectral norm is the Ky Fan norm when \( k = 1 \), it also equals the limit of \( \| A \|_p \) as \( p \to \infty \), i.e., \( \| A \|_\infty \), whereas the Frobenius norm is the Schatten 2-norm, i.e., \( \| A \|_F = \| A \|_2 \).

**Problems**

1. Let \( \| \cdot \| \) be a vector norm on \( \mathbb{C}^n \) (or \( \mathbb{R}^n \)). Define
   \[
   \| x \|_D = \max \{ |(x, y)| : y \in \mathbb{C}^n, \| y \| = 1 \}.
   \]
   Show that \( \| \cdot \|_D \) is a vector norm on \( \mathbb{C}^n \) (known as the *dual norm*).

2. Show that for any matrix norm \( \| \cdot \| \) on \( M_n \) and \( A = (a_{ij}), B \in M_n \)
   
   \[ |\| A \| - \| B \| | \leq \| A - B \| \quad \text{and} \quad \| A \| \leq \sum_{i,j} |a_{ij}| \| E_{ij} \|, \]
   
   where \( E_{ij} \) is the matrix with \((i, j)\)-entry 1 and elsewhere 0 for all \( i, j \).

3. Let \( A, B \in M_n \). Show that
   
   \[ \| A + B \|_F \leq \| A \|_F + \| B \|_F \quad \text{and} \quad \| AB \|_F \leq \| A \|_F \| B \|_F. \]

4. Let \( A \in M_n \). Show that for any matrix norm \( \| \cdot \| \) and integer \( k \geq 1 \),
   
   \[ \| A^k \| \leq \| A \|_F^k \quad \text{and} \quad \| A^k \|^{-1} \| I \| \leq \| A^{-1} \|_F^k \] if \( A \) is invertible.

5. Let \( A \in M_n \) be given. If there exists a matrix norm \( \| \cdot \| \) such that \( \| A \| < 1 \), show that \( A^k \to 0 \) as \( k \to 0 \).

6. Let \( \| \cdot \| \) be a matrix norm on \( M_n \). Show that for any invertible matrix \( P \in M_n \), \( \| \cdot \|_P : M_n \to \mathbb{R} \) defined by \( \| A \|_P = \| P^{-1} A P \| \) for all matrices \( A \in M_n \) is also a matrix norm.

7. Let \( A = (a_{ij}) \in M_n \) and define \( \| A \|_{\infty} = \max_{1 \leq i, j \leq n} |a_{ij}|. \) Show that \( \| \cdot \|_\infty \) is a matrix-vector norm, but not a multiplicative matrix norm.

8. Show that \( \| \cdot \|_{\infty}, \| \cdot \|_1, \) and \( \| \cdot \|_\infty \) are matrix norms on \( M_n \), where
   
   \[ \| A \|_{\infty} = n \| A \|_\infty, \quad \| A \|_1 = \sum_{1 \leq i, j \leq n} |a_{ij}|, \quad \| A \|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|. \]
   
   Are these (multiplicative) matrix norms unitarily invariant?
4.3 The Kronecker and Hadamard Products

Matrices can be multiplied in different ways. The Kronecker product and Hadamard product, defined below, used in many fields, are almost as important as the ordinary product. Another basic matrix operation is “compounding” matrices, which is evidently a useful tool in deriving matrix inequalities. This section introduces the three concepts and presents their properties.

The Kronecker product, also known as tensor product or direct product, of two matrices $A$ and $B$ of sizes $m \times n$ and $s \times t$, respectively, is defined to be the $(ms) \times (nt)$ matrix

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

In other words, the Kronecker product $A \otimes B$ is an $(ms) \times (nt)$ matrix, partitioned into $mn$ blocks with the $(i, j)$ block the $s \times t$ matrix $a_{ij}B$. Note that $A$ and $B$ can have any different sizes.

The Hadamard product, or the Schur product, of two matrices $A$ and $B$ of the same size is defined to be the entrywise product

$$A \circ B = (a_{ij}b_{ij}).$$

In particular, for $u = (u_1, u_2, \ldots, u_n)$, $v = (v_1, v_2, \ldots, v_n) \in \mathbb{C}^n$,

$$u \otimes v = (u_1v_1, \ldots, u_1v_n, u_2v_1, \ldots, u_2v_n, \ldots, u_nv_1, \ldots, u_nv_n)$$

and

$$u \circ v = (u_1v_1, u_2v_2, \ldots, u_nv_n).$$

Note that $A \otimes B \neq B \otimes A$ in general and $A \circ B = B \circ A$.

We take, for example, $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

$$A \otimes B = \begin{pmatrix} a & b & 2a & 2b \\ c & d & 2c & 2d \\ 3a & 3b & 4a & 4b \\ 3c & 3d & 4c & 4d \end{pmatrix}, \quad A \circ B = \begin{pmatrix} a & 2b \\ 3c & 4d \end{pmatrix}.$$
The Kronecker product has the following basic properties, each of which can be verified by definition and direct computations.

**Theorem 4.5** Let \( A, B, C \) be matrices of appropriate sizes. Then

1. \((kA) \otimes B = A \otimes (kB) = k(A \otimes B)\), where \( k \) is a scalar.
2. \((A + B) \otimes C = A \otimes C + B \otimes C\).
3. \(A \otimes (B + C) = A \otimes B + A \otimes C\).
4. \((A \otimes B) \otimes C = A \otimes (B \otimes C)\).
5. \(A \otimes B = 0\) if and only if \( A = 0 \) or \( B = 0 \).
6. \((A \otimes B)^T = A^T \otimes B^T\). If \( A \) and \( B \) are symmetric, so is \( A \otimes B\).
7. \((A \otimes B)^* = A^* \otimes B^*\). If \( A \) and \( B \) are Hermitian, so is \( A \otimes B\).

**Theorem 4.6** Let \( A, B, C \) be matrices of appropriate sizes. Then

1. \((A \otimes B)(C \otimes D) = (AC) \otimes (BD)\).
2. \((A \otimes B)^{-1} = A^{-1} \otimes B^{-1}\) if \( A \) and \( B \) are invertible.
3. \(A \otimes B\) is unitary if \( A \) and \( B \) are unitary.
4. \(A \otimes B\) is normal if \( A \) and \( B \) are normal.

**Proof.** For (1), let \( A \) have \( n \) columns. Then \( C \) has \( n \) rows as indicated in the product \( AC \) on the right-hand side of (1). We write \( A \otimes B = (a_{ij}B) \), \( C \otimes D = (c_{ij}D) \). Then the \((i, j)\) block of \((A \otimes B)(C \otimes D)\) is

\[
\sum_{t=1}^{n} a_{it}Bc_{tj}D = \sum_{t=1}^{n} a_{it}c_{tj}BD.
\]

But this is the \((i, j)\)-entry of \( AC \) times \( BD \), which is the \((i, j)\) block of \((AC) \otimes (BD)\). (1) follows. The rest are immediate from (1).

To perform the Hadamard product, matrices need to have the same size. In the case of square matrices, an interesting and important observation is that the Hadamard product \( A \circ B \) is contained in the Kronecker product \( A \otimes B \) as a principal submatrix.
Theorem 4.7 Let $A, B \in \mathbb{M}_n$. Then the Hadamard product $A \circ B$ is a principal submatrix of the Kronecker product $A \otimes B$ lying on the intersections of rows and columns $1, n + 2, 2n + 3, \ldots, n^2$.

Proof. Let $e_i$ be, as usual, the column vector of $n$ components with the $i$th position 1 and 0 elsewhere, $i = 1, 2, \ldots, n$, and let

$$E = (e_1 \otimes e_1, \ldots, e_n \otimes e_n).$$

Then for every pair of $i$ and $j$, we have by computation

$$a_{ij}b_{ij} = (e_i^T A e_j) \otimes (e_i^T B e_j) = (e_i \otimes e_i)^T (A \otimes B) (e_j \otimes e_j),$$

which equals the $(i, j)$-entry of the matrix $E^T (A \otimes B) E$. Thus,

$$E^T (A \otimes B) E = A \circ B.$$ 

This says that $A \circ B$ is the principal submatrix of $A \otimes B$ lying on the intersections of rows and columns $1, n + 2, 2n + 3, \ldots, n^2$.

The following theorem, relating the eigenvalues of the Kronecker product to those of individual matrices, presents in its proof a common method of decomposing a Kronecker product.

Theorem 4.8 Let $A$ and $B$ be $m$-square and $n$-square complex matrices with eigenvalues $\lambda_i$ and $\mu_j$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, respectively. Then the eigenvalues of $A \otimes B$ are

$$\lambda_i \mu_j, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n,$$

and the eigenvalues of $A \otimes I_n + I_m \otimes B$ are

$$\lambda_i + \mu_j, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

Proof. By the Schur decomposition (Theorem 3.3), let $U$ and $V$ be unitary matrices of sizes $m$ and $n$, respectively, such that

$$U^* AU = T_1 \quad \text{and} \quad V^* BV = T_2,$$

where $T_1$ and $T_2$ are upper-triangular matrices with diagonal entries $\lambda_i$ and $\mu_j$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, respectively. Then

$$T_1 \otimes T_2 = (U^* AU) \otimes (V^* BV) = (U^* \otimes V^*)(A \otimes B)(U \otimes V).$$
Note that $U \otimes V$ is unitary. Thus $A \otimes B$ is unitarily similar to $T_1 \otimes T_2$. The eigenvalues of the latter matrix are $\lambda_i \mu_j$.

For the second part, let $W = U \otimes V$. Then

$$W^*(A \otimes I_n)W = T_1 \otimes I_n = \begin{pmatrix} \lambda_1 I_n & * \\ \vdots & \ddots \\ 0 & \lambda_m I_n \end{pmatrix}$$

and

$$W^*(I_m \otimes B)W = I_m \otimes T_2 = \begin{pmatrix} T_2 & 0 \\ \vdots & \ddots \\ 0 & T_2 \end{pmatrix}.$$ 

Thus

$$W^*(A \otimes I_n + I_m \otimes B)W = T_1 \otimes I_n + I_m \otimes T_2$$

is an upper-triangular matrix with eigenvalues $\lambda_i + \mu_j$.  

**Problems**

1. Compute $A \otimes B$ and $B \otimes A$ for

$$A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \sqrt{2} \\ \pi & 2 \\ -1 & 7 \end{pmatrix}.$$ 

2. Let $J_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Compute $I_n \otimes J_2$ and $J_2 \otimes I_n$.

3. Let $A$, $B$, $C$, and $D$ be complex matrices. Show that

   (a) $(A \otimes B)^k = A^k \otimes B^k$.
   (b) $\text{tr}(A \otimes B) = \text{tr} A \text{tr} B$.
   (c) $\text{rank}(A \otimes B) = \text{rank}(A) \text{rank}(B)$.
   (d) $\det(A \otimes B) = (\det A)^n (\det B)^m$, if $A \in M_m$ and $B \in M_n$.
   (e) If $A \otimes B = C \otimes D \neq 0$, where $A$ and $C$ are of the same size, then $A = aC$ and $B = bD$ with $ab = 1$, and vice versa.

4. Let $A$ and $B$ be $m$- and $n$-square matrices, respectively. Show that

$$(A \otimes I_n)(I_m \otimes B) = A \otimes B = (I_m \otimes B)(A \otimes I_n).$$
5. Let $A \in \mathbb{M}_n$ have characteristic polynomial $p$. Show that
\[
\det(A \otimes I + I \otimes A) = (-1)^n \det(p(-A)).
\]

6. Let $A, B \in \mathbb{M}_n$. Show that for some permutation matrix $P \in \mathbb{M}_{n^2}$
\[
P^{-1}(A \otimes B)P = B \otimes A.
\]

7. Let $x, y, u, v \in \mathbb{C}^n$. With $(x, y) = y^*x$ and $\|x\|^2 = x^*x$, show that
\[
(x, y)(u, v) = (x \otimes u, y \otimes v).
\]
Derive
\[
\|x \otimes y\| = \|x\| \|y\|.
\]

8. Let $A, B \in \mathbb{M}_n$. Show that $A \circ I_n = \text{diag}(a_{11}, \ldots, a_{nn})$ and that
\[
D_1(A \circ B)D_2 = (D_1AD_2) \circ B = A \circ (D_1BD_2)
\]
for any $n$-square diagonal matrices $D_1$ and $D_2$.

9. Let $A, B,$ and $C$ be square matrices. Show that
\[
(A \oplus B) \otimes C = (A \otimes C) \oplus (B \otimes C).
\]
But it need not be true that
\[
(A \otimes B) \oplus C = (A \oplus C) \otimes (B \oplus C).
\]

10. Consider the vector space $\mathbb{M}_2$, $2 \times 2$ complex matrices, over $\mathbb{C}$.
(a) What is the dimension of $\mathbb{M}_2$?
(b) Find a basis for $\mathbb{M}_2$.
(c) For $A, B \in \mathbb{M}_2$, define
\[
\mathcal{L}(X) = AXB, \quad X \in \mathbb{M}_2.
\]
Show that $\mathcal{L}$ is a linear transformation on $\mathbb{M}_2$.
(d) Show that if $\lambda$ and $\mu$ are eigenvalues of $A$ and $B$, respectively, then $\lambda\mu$ is an eigenvalue of $\mathcal{L}$.

11. Let $A$ and $B$ be square matrices (of possibly different sizes). Show that
\[
e^{A \otimes I} = e^A \otimes I, \quad e^{I \otimes B} = I \otimes e^B, \quad e^{A \oplus B} = e^A \otimes e^B.
\]
4.4 Compound Matrices

We now turn our attention to compound matrices. Roughly speaking, the Kronecker product and the Hadamard product are operations on two (or more) matrices. Unlike the Kronecker and Hadamard products, “compounding” matrix is a matrix operation on a single matrix that arranges in certain order all minors of a given size from the given matrix. A rigorous definition is given as follows.

Let $A$ be an $m \times n$ matrix, $\alpha = \{i_1, \ldots, i_s\}$, and $\beta = \{j_1, \ldots, j_t\}$, $1 \leq i_1 < \cdots < i_s \leq m$, $1 \leq j_1 < \cdots < j_t \leq n$. Denote by $A[i_1, \ldots, i_s, j_1, \ldots, j_t]$, or simply $A[\alpha, \beta]$, the submatrix of $A$ consisting of the entries in rows $i_1, \ldots, i_s$ and columns $j_1, \ldots, j_t$.

Given a positive integer $k \leq \min\{m, n\}$, there are $\binom{m}{k} \times \binom{n}{k}$ possible minors (numbers) that we can get from the $m \times n$ matrix $A$. We now form a matrix, denoted by $A^{(k)}$ and called the $k$th compound matrix of $A$, of size $\binom{m}{k} \times \binom{n}{k}$ by ordering these numbers lexicographically; that is, the $(1, 1)$ position of $A^{(k)}$ is det $A[1, \ldots, k|1, \ldots, k]$, the $(1, 2)$ position of $A^{(k)}$ is det $A[1, \ldots, k|1, \ldots, k-1, k+1]$, ..., whereas the $(2, 1)$ position is det $A[1, \ldots, k-1, k+1|1, \ldots, k]$, ..., and so on. For convenience, we say that the minor det $A[\alpha, \beta]$ is in the $(\alpha, \beta)$ position of the compound matrix and denote it by $A^{(k)}_{\alpha, \beta}$. Clearly, $A^{(1)} = A$ and $A^{(n)} = \det A$ if $A$ is an $n \times n$ matrix.

As an example, let $m = n = 3$, $k = 2$, and take

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$ 

Then

$$A^{(2)} = \begin{pmatrix} \det A[1, 2|1, 2] & \det A[1, 2|1, 3] & \det A[1, 2|2, 3] \\ \det A[1, 3|1, 2] & \det A[1, 3|1, 3] & \det A[1, 3|2, 3] \\ \det A[2, 3|1, 2] & \det A[2, 3|1, 3] & \det A[2, 3|2, 3] \end{pmatrix}$$

$$= \begin{pmatrix} \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} \\ \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} & \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} & \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -3 & -6 & -3 \\ -6 & -12 & -6 \\ -3 & -6 & -3 \end{pmatrix}.$$
If $A$ is an $n$-square matrix, then the main diagonal entries of $A^{(k)}$ are $\det A[\alpha|\alpha]$, i.e., the principal minors of $A$. For an $n$-square upper-triangular matrix $A$, $\det A[\alpha|\beta] = 0$ if $\alpha$ is after $\beta$ in lexicographic order. This leads to the result that if $A$ is upper (lower)-triangular, then so is $A^{(k)}$. As a consequence, if $A$ is diagonal, then so is $A^{(k)}$.

The goal of this section is to show that the compound matrix of the product of matrices is the product of their compound matrices. For this purpose, we need to borrow a well-known result on determinant expansion, the Binet–Cauchy formula. (A good reference on this formula and its proof is Lancaster and Tismenetsky’s book, The Theory of Matrices, 1985, pp. 36–42.)

**Theorem 4.9 (Binet–Cauchy formula)** Let $C = AB$, where $A$ is $m \times n$ and $B$ is $n \times m$, $m \leq n$, and let $\alpha = \{1, 2, \ldots, m\}$. Then

$$\det C = \sum_{\beta} \det A[\alpha|\beta] \det B[\beta|\alpha],$$

where $\beta$ runs over all sequences $\{j_1, \ldots, j_m\}$, $1 \leq j_1 < \cdots < j_m \leq n$.

The following theorem is of fundamental importance for compound matrices, whose corollary plays a pivotal role in deriving matrix inequalities involving eigenvalue and singular value products.

**Theorem 4.10** Let $A$ be an $m \times p$ matrix and $B$ be a $p \times n$ matrix. If $k$ is a positive integer, $k \leq \min\{m, p, n\}$, then $(AB)^{(k)} = A^{(k)}B^{(k)}$.

**Proof.** For $\alpha = \{i_1, \ldots, i_k\}$ and $\beta = \{j_1, \ldots, j_k\}$, $1 \leq i_1 < \cdots < i_k \leq m$, $1 \leq j_1 < \cdots < j_k \leq m$, we compute the entry in the $(\alpha, \beta)$ position (in lexicographic order) of $(AB)^{(k)}$ by the Binet–Cauchy determinant expansion formula and get

$$(AB)^{(k)}_{\alpha, \beta} = \det ((AB)[\alpha|\beta]) = \sum_{\gamma} \det A[\alpha|\gamma] \det B[\gamma|\beta] = (A^{(k)}B^{(k)})_{\alpha, \beta},$$

where $\gamma$ runs over all possible sequences $1 \leq \gamma_1 < \cdots < \gamma_k \leq p$. □

If $A$ is Hermitian, it is readily seen that $A^{(k)}$ is Hermitian too.
Corollary 4.1 Let $A \in M_n$ be a positive semidefinite matrix with eigenvalues $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$. Then the largest eigenvalue of $A^{(k)}$ is the product of the first $k$ largest eigenvalues of $A$; that is,

$$\lambda_{\text{max}}(A^{(k)}) = \prod_{i=1}^{k} \lambda_i(A).$$

Problems

1. Find $A^{(2)}$, where

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & -1 \end{pmatrix}.$$  

2. Show that $A^*[\alpha|\beta] = (A[\beta|\alpha])^*$.  

3. Show that $I_{n}^{(k)} = I_{\binom{n}{k}}$, where $I_l$ is the $l \times l$ identity matrix.  

4. Show that $(A^{(k)})^* = (A^*)^{(k)}$, $(A^{(k)})^T = (A^T)^{(k)}$.  

5. Show that $(A^{(k)})^{-1} = (A^{-1})^{(k)}$ if $A$ is nonsingular.  

6. Show that $\det A^{(k)} = (\det A)^{\binom{n-1}{k-1}}$ when $A$ is $n$-square.  

7. If $\text{rank}(A) = r$, show that $\text{rank}(A^{(k)}) = \binom{r}{k}$ or 0 if $r < k$.  

8. Show that if $A$ is unitary, symmetric, positive (semi-)definite, Hermitian, or normal, then so is $A^{(k)}$, respectively.  

9. If $A = \text{diag}(a_1, \ldots, a_n)$, show that $A^{(k)}$ is an $\binom{n}{k} \times \binom{n}{k}$ diagonal matrix with diagonal entries $a_{i_1} \cdots a_{i_k}$, $1 \leq i_1 < \cdots < i_k \leq n$.  

10. If $A \in M_n$ has eigenvalues $\lambda_1, \ldots, \lambda_n$, show that $A^{(k)}$ has eigenvalues $\lambda_{i_1} \cdots \lambda_{i_k}$, $1 \leq i_1 < \cdots < i_k \leq n$.  

11. If $A \in M_n$ has singular values $\sigma_1, \ldots, \sigma_n$, show that $A^{(k)}$ has singular values $\sigma_{i_1} \cdots \sigma_{i_k}$, $1 \leq i_1 < \cdots < i_k \leq n$.  

12. If $A \in M_n$ has eigenvalues $\lambda_1, \ldots, \lambda_n$, show that $\text{tr}(A^{(k)})$ equals $\sum \gamma \lambda_{i_1} \cdots \lambda_{i_k}$, denoted by $s_k(\lambda_1, \ldots, \lambda_n)$ and called $k$th elementary symmetric function, where $\gamma$ is any sequence $1 \leq i_1 < \cdots < i_k \leq n$.  

13. If $A$ is a positive semidefinite matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, show that the smallest eigenvalue of $A^{(k)}$ is $\lambda_{\text{min}}(A^{(k)}) = \prod_{i=1}^{k} \lambda_{n-i+1}$.  

.  

.  

.  

.  

⊙