Chapter 1

Vector and Matrix Analysis

The purpose of this chapter is to summarise the fundamental theoretical results from linear algebra to which we will frequently refer, and to provide some basic theoretical tools which we will use in our analysis. We study vector and matrix norms, inner-products, the eigenvalue problem, orthogonal projections and a variety of special matrices which arise frequently in computational linear algebra.

1.1 Vector Norms and Inner Products

Definition 1.1. A vector norm on $\mathbb{C}^n$ is a mapping $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ satisfying

a) $\|x\| \geq 0$ for all $x \in \mathbb{C}^n$ and $\|x\| = 0$ iff $x = 0$,

b) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{C}, x \in \mathbb{C}^n$, and

c) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathbb{C}^n$.

Remark. The definition of a norm on $\mathbb{R}^n$ is identical, but with $\mathbb{C}^n$ replaced by $\mathbb{R}^n$ and $\mathbb{C}$ replaced by $\mathbb{R}$.

Examples.

- the $p$-norm for $1 \leq p < \infty$:

  $\|x\|_p = \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p} \forall x \in \mathbb{C}^n$;

- for $p = 2$ we get the Euclidean norm:

  $\|x\|_2 = \sqrt{\sum_{j=1}^{n} |x_j|^2} \forall x \in \mathbb{C}^n$;

- for $p = 1$ we get

  $\|x\|_1 = \sum_{j=1}^{n} |x_j| \forall x \in \mathbb{C}^n$;

- Infinity norm: $\|x\|_{\infty} = \max_{1 \leq j \leq n} |x_j|$.

Theorem 1.2. All norms on $\mathbb{C}^n$ are equivalent: for each pair of norms $\|\cdot\|_a$ and $\|\cdot\|_b$ on $\mathbb{C}^n$ there are constants $0 < c_1 \leq c_2 < \infty$ with

$\ c_1 \|x\|_a \leq \|x\|_b \leq c_2 \|x\|_a \ \forall x \in \mathbb{C}^n.$
Remark. and multiplying the result by \( \lambda \) we get
\[
\text{Denition 1.3. If on the other hand } x \in S = \{ x \in \mathbb{C}^n \mid \| x \|_2 = 1 \}. \text{ Since } \| \cdot \|_a \text{ is non-zero on all of } S \text{ we can define } f: S \to \mathbb{R} \text{ by } f(x) = \| x \|_b/\| x \|_a. \text{ Because the function } f \text{ is continuous and the set } S \text{ is compact there are } x_1, x_2 \in S \text{ with } f(x_1) \leq f(x) \leq f(x_2) \text{ for all } x \in S. \text{ Setting } c_1 = f(x_1) > 0 \text{ and } c_2 = f(x_2) \text{ completes the proof.} \]

Remarks.
1. The same result holds for norms on \( \mathbb{R}^n \). The proof transfers to this situation without change.

2. We remark that, if \( A \in \mathbb{C}^{n \times n} \) is an invertible matrix and \( \| \cdot \| \) a norm on \( \mathbb{C}^n \) then \( \| \cdot \|_A := \| A \cdot \| \) is also a norm.

\text{Definition 1.3. An inner-product on } \mathbb{C}^n \text{ is a mapping } ⟨·, ·⟩: \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} \text{ satisfying:}

- a) \( ⟨x, x⟩ \in \mathbb{R}^+ \) for all \( x \in \mathbb{C}^n \) and \( ⟨x, x⟩ = 0 \) iff \( x = 0 \);
- b) \( ⟨x, y⟩ = ⟨y, x⟩ \) for all \( x, y \in \mathbb{C}^n \);
- c) \( ⟨x, αy⟩ = α⟨x, y⟩ \) for all \( α \in \mathbb{C}, x, y \in \mathbb{C}^n \);
- d) \( ⟨x, y + z⟩ = ⟨x, y⟩ + ⟨x, z⟩ \) for all \( x, y, z \in \mathbb{C}^n \);

\text{Remark. Conditions c) and d) above state that } ⟨·, ·⟩ \text{ is linear in the second component. Using the rules for inner products we get}
\[
⟨x + y, z⟩ = ⟨x, z⟩ + ⟨y, z⟩ \quad \text{for all } x, y, z \in \mathbb{C}^n
\]

and
\[
⟨αx, y⟩ = α⟨x, y⟩ \quad \text{for all } α \in \mathbb{C}, x, y \in \mathbb{C}^n.
\]
The inner product is said to be anti-linear in the first component.

\text{Example. The standard inner product on } \mathbb{C}^n \text{ is given by}
\[
⟨x, y⟩ = \sum_{j=1}^{n} x_j y_j \quad \forall x, y \in \mathbb{C}^n. \tag{1.1}
\]

\text{Definition 1.4. Two vectors } x, y \text{ are orthogonal with respect to an inner product } ⟨·, ·⟩ \text{ iff } ⟨x, y⟩ = 0.

\text{Lemma 1.5 (Cauchy-Schwarz inequality). Let } ⟨·, ·⟩: \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} \text{ be an inner product. Then}
\[
\left| ⟨x, y⟩ \right|^2 \leq ⟨x, x⟩ \cdot ⟨y, y⟩ \tag{1.2}
\]

for every } x, y \in \mathbb{C}^n \text{ and equality holds if and only if } x \text{ and } y \text{ are linearly dependent.

Proof. For every } λ \in \mathbb{C} \text{ we have}
\[
0 \leq ⟨x - λy, x - λy⟩ = ⟨x, x⟩ - λ⟨y, x⟩ - λ⟨x, y⟩ + λ^2⟨y, y⟩.
\tag{1.3}
\]

For } λ = ⟨y, x⟩/⟨y, y⟩ \text{ this becomes}
\[
0 \leq ⟨x, x⟩ - \frac{⟨x, y⟩⟨y, x⟩}{⟨y, y⟩} - \frac{⟨y, x⟩⟨x, y⟩}{⟨y, y⟩} + \frac{⟨y, x⟩⟨x, y⟩}{⟨y, y⟩} = ⟨x, x⟩ - \left| \frac{⟨x, y⟩}{⟨y, y⟩} \right|^2
\]

and multiplying the result by } ⟨y, y⟩ \text{ gives } (1.2).

If equality holds in } (1.2) \text{ then } x - λy \text{ in } (1.3) \text{ must be } 0 \text{ and thus } x \text{ and } y \text{ are linearly dependent. If on the other hand } x \text{ and } y \text{ are linearly dependent, say } x = αy, \text{ then } λ = ⟨y, αy⟩/⟨y, y⟩ = α \text{ and } x - λy = 0 \text{ giving equality in } (1.3) \text{ and thus in } (1.2). \qed
Lemma 1.6. Let \( \langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C} \) be an inner product. Then \( \| \cdot \| : \mathbb{C}^n \rightarrow \mathbb{R} \) defined by
\[
\| x \| = \sqrt{\langle x, x \rangle} \quad \forall x \in \mathbb{C}^n
\]
is a vector norm.

Proof. a) Since \( \langle \cdot, \cdot \rangle \) is an inner product we have \( \langle x, x \rangle \geq 0 \) for all \( x \in \mathbb{C}^n \), i.e. \( \sqrt{\langle x, x \rangle} \) is real and positive. Also we get
\[
\| x \| = 0 \iff \langle x, x \rangle = 0 \iff x = 0.
\]
b) We have
\[
\| \alpha x \| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha \overline{\alpha} \langle x, x \rangle} = |\alpha| \cdot \| x \|.
\]
c) Using the Cauchy-Schwarz inequality
\[
\big| \langle x, y \rangle \big| \leq \| x \| \| y \| \quad \forall x, y \in \mathbb{C}^n
\]
from Lemma 1.5 we get
\[
\| x + y \|^2 = \langle x + y, x + y \rangle
\]
\[
= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle
\]
\[
\leq \| x \|^2 + 2\langle x, y \rangle + \| y \|^2
\]
\[
\leq \| x \|^2 + 2\| x \| \| y \| + \| y \|^2
\]
\[
= (\| x \| + \| y \|)^2 \quad \forall x, y \in \mathbb{C}^n.
\]
This completes the proof. \( \square \)

Remark. The angle between two vectors \( x \) and \( y \) is the unique value \( \varphi \in [0, \pi] \) with
\[
\cos(\varphi) \| x \| \| y \| = \langle x, y \rangle.
\]

When considering the Euclidean norm and inner product on \( \mathbb{R}^n \), this definition of angle coincides with the usual, geometric meaning of angles. In any case, two vectors are orthogonal, if and only if they have angle \( \pi/2 \).

We write matrices \( A \in \mathbb{C}^{m \times n} \) as
\[
A = \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{pmatrix};
\]
we write \( (A)_{ij} = a_{ij} \) for the \( ij^{th} \) entry of \( A \).

Definition 1.7. Given \( A \in \mathbb{C}^{m \times n} \) we define the adjoint \( A^* \in \mathbb{C}^{n \times m} \) by \( (A^*)_{ij} = \overline{a_{ji}} \). (For \( A \in \mathbb{R}^{m \times n} \) we write \( A^T \) instead of \( A^* \).)

By identifying the space \( \mathbb{C}^n \) of vectors with the space \( \mathbb{C}^{n \times 1} \) of \( n \times 1 \)-matrices, we can take the adjoint of a vector. Then we can write the standard inner product as
\[
\langle x, y \rangle = x^* y.
\]
Thus, the standard inner product satisfies
\[
\langle Ax, y \rangle = (Ax)^* y = x^* A^* y = \langle x, A^* y \rangle
\]
for all \( x \in \mathbb{C}^n \), \( y \in \mathbb{C}^m \) and all \( A \in \mathbb{C}^{m \times n} \). Unless otherwise specified, we will use \( \langle \cdot, \cdot \rangle \) to denote the standard inner product (1.1) and \( \| \cdot \|_2 \) to denote the corresponding Euclidean norm.
The following families of special matrices will be central in what follows:

**Definition 1.8.**

1. \( Q \in \mathbb{C}^{m \times n} \) is unitary if \( Q^*Q = I \). (If \( Q \) is real then \( Q^TQ = I \) and we say \( Q \) is orthogonal.)

2. \( A \in \mathbb{C}^{n \times n} \) is Hermitian if \( A^* = A \). (If \( A \) is real, we say \( A \) is symmetric.)

3. A Hermitian matrix \( A \in \mathbb{C}^{n \times n} \) is positive-definite (resp. positive semi-definite) if \( x^*Ax = \langle Ax, x \rangle > 0 \) (resp. \( \geq 0 \)) for all \( x \in \mathbb{C}^n \setminus \{0\} \). In this text, whenever we use the terminology positive-definite or positive semi-definite we are necessarily referring to Hermitian matrices.

**Remarks.** Unitary matrices have the following properties:

- A matrix \( Q \) is unitary, if and only if the columns of \( Q \) are orthonormal with respect to the standard inner-product. In particular unitary matrices cannot have more columns than rows.

- If \( Q \) is a square matrix, \( Q^{-1} = Q^* \) and thus \( QQ^* = I \).

- A square matrix \( Q \) is unitary, if and only if \( Q^* \) is unitary.

The standard inner product and norm are invariant under multiplication by a unitary matrix:

**Theorem 1.9.** Let \( \langle \cdot, \cdot \rangle \) denote the standard inner product. Then for any unitary \( Q \in \mathbb{C}^{m \times n} \) and any \( x, y \in \mathbb{C}^n \) we have \( \langle Qx, Qy \rangle = \langle x, y \rangle \) and \( \|Qx\|_2 = \|x\|_2 \).

**Proof.** The first claim follows from \( \langle Qx, Qy \rangle = \langle x, Q^*Qy \rangle = \langle x, y \rangle \) and using the relation \( \|x\|_2 = \langle x, x \rangle \) gives the second claim.

Other inner products with appropriate properties can give rise to other norms; for example, for matrices \( A \) which are Hermitian and positive-definite,

\[
\langle x, y \rangle_A = \langle x, Ay \rangle
\]

(1.4)

is an inner product and

\[
\|x\|_A = \sqrt{\langle x, x \rangle_A}.
\]

(1.5)

defines a norm (see Exercise 1-2).

### 1.2 Eigenvalues and Eigenvectors

**Definition 1.10.** Given a matrix \( A \in \mathbb{C}^{n \times n} \), a vector \( x \in \mathbb{C}^n \) is an eigenvector and \( \lambda \in \mathbb{C} \) is an eigenvalue (also called a right eigenvalue) of \( A \) if

\[
Ax = \lambda x \quad \text{and} \quad x \neq 0.
\]

(1.6)

When \( x \) is an eigenvector of \( A \), then for every \( \alpha \neq 0 \) the vector \( \alpha x \) is an eigenvector for the same eigenvalue, since both sides of (1.6) are linear in \( x \). Sometimes it is convenient to normalise \( x \) by choosing \( \|x\|_2 = 1 \). Then the eigenvalue problem is to find \( (x, \lambda) \in \mathbb{C}^n \times \mathbb{C} \) satisfying

\[
Ax = \lambda x \quad \text{and} \quad \|x\|_2 = 1.
\]

**Definition 1.11.** Given a matrix \( A \in \mathbb{C}^{n \times n} \) we define the characteristic polynomial of \( A \) as

\[
\rho_A(z) := \det(A - zI).
\]

**Theorem 1.12.** A value \( \lambda \in \mathbb{C} \) is an eigenvalue of the matrix \( A \), if and only if \( \rho_A(\lambda) = 0 \).

**Proof.** \( \lambda \) is an eigenvalue of \( A \), if and only if there is an \( x \neq 0 \) with \( (A - \lambda I)x = 0 \). This is equivalent to the condition that \( A - \lambda I \) is singular which in turn is equivalent to \( \det(A - \lambda I) = 0 \).
Since \( \rho_A \) is a polynomial of degree \( n \), there will be \( n \) (some possibly repeated) eigenvalues, denoted by \( \lambda_1, \ldots, \lambda_n \) and determined by \( \rho_A(\lambda) = 0 \).

**Definition 1.13.** An eigenvalue \( \lambda \) has **algebraic multiplicity** \( q \) if \( q \) is the largest integer such that \( (z - \lambda)^q \) is a factor of the characteristic polynomial \( \rho_A(z) \). The **geometric multiplicity**, \( r \), is the dimension of the null space of \( A - \lambda I \). An eigenvalue is **simple** if \( q = r = 1 \).

If \( \lambda \) is an eigenvalue of \( A \in \mathbb{C}^{n \times n} \) then \( \det(A - \lambda I) = 0 \) which implies that \( \det(A^* - \bar{\lambda} I) = 0 \) and so \( (A^* - \bar{\lambda} I) \) has non-trivial null space. Thus there is a vector \( y \) (the eigenvector of \( A^* \) corresponding to the eigenvalue \( \bar{\lambda} \)) with \( y^* A = \bar{\lambda} y^* \) and \( y \neq 0 \).

**Definition 1.14.** A vector \( y \in \mathbb{C}^n \) with
\[
y^* A = \lambda y^* \quad \text{and} \quad y \neq 0
\]
is known as a **left eigenvector** of \( A \in \mathbb{C}^{n \times n} \) corresponding to the eigenvalue \( \lambda \).

Note that, even though the corresponding eigenvalues are the same, the right and left eigenvectors of a matrix are usually different.

**Definition 1.15.** Matrices \( A, B \in \mathbb{C}^{n \times n} \) are **similar**, if \( B = S^{-1} A S \) with \( S \in \mathbb{C}^{n \times n} \) invertible. The matrix \( S \) is a **similarity transform**.

**Remarks.** If a matrix \( A \in \mathbb{C}^{n \times n} \) has \( n \) linearly independent eigenvectors \( x_i \) and we arrange them as columns of the matrix \( X \), then \( X \) is invertible. If we let \( \Lambda \) denote a diagonal matrix with eigenvalues of \( A \) on the diagonal, then we may write
\[
AX = X\Lambda.
\]
By invertibility of \( X \) we have
\[
A = X\Lambda X^{-1}. \tag{1.7}
\]

Thus \( \Lambda \) is a similarity transform of \( A \). It reveals the eigenstructure of \( A \) and is hence very useful in many situations. However, in general a matrix does not have \( n \) linearly independent eigenvalues and hence generalizations of this factorization are important. Two which will arise in the next chapter are:

- **Jordan Canonical Form:** \( A = SJS^{-1} \) (see Theorem 2.7)
- **Schur Factorization:** \( A = QTQ^* \) (see Theorem 2.2)

These are both similarity transformations which reveal the eigenvalues of \( A \) on the diagonals of \( J \) and \( T \), respectively. The Jordan Canonical Form is not stable to perturbations, but the Schur Factorization is. Hence Schur Factorization will form the basis of good algorithms while the Jordan Canonical Form is useful for more theoretical purposes, such as defining the matrix exponential \( e^{At} \).

**Theorem 1.16 (Similarity and Eigenvalues).** If \( B \) is similar to \( A \), then \( B \) and \( A \) have the same eigenvalues with the same algebraic and geometric multiplicities.

**Proof.** Exercise 2-4.

**Lemma 1.17.** For a simple eigenvalue \( \mu \),
\[
\dim(\ker(A - \mu I)^2) = 1.
\]

**Proof.** (Sketch) We prove this in the case where \( A \) has \( n \) linearly independent eigenvalues \( x_i \) all of which correspond to simple eigenvalues \( \lambda_i \). Then \( A \) may be factorized as in (1.7) with \( \Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \). Hence
\[
(A - \mu I)^2 = X\Omega X^{-1},
\]
where \( \Omega = \text{diag}\{(\lambda_1 - \mu)^2, (\lambda_2 - \mu)^2, \ldots, (\lambda_n - \mu)^2\} \).

Without loss of generality let \( \mu = \lambda_1 \), noting that \( \lambda_j \neq \mu \) by simplicity. \( \ker(\Omega) \) is one dimensional, spanned by \( e_1 \). Hence \( \ker(A - \mu I)^2 \) is one dimensional by Theorem 1.16.

The general case can be established by use of the Jordan form (see Theorem 2.7), using the fact that the Jordan block corresponding to a simple eigenvalue is diagonal. \( \square \)

**Theorem 1.18 (Eigenvalue Multiplicities).** For any eigenvalue of \( A \in \mathbb{R}^{n \times n} \) the algebraic and geometric multiplicities \( q \) and \( r \) respectively satisfy

\[
1 \leq r \leq q.
\]

**Proof.** Let \( \mu \) be the eigenvalue. Since \( (A - \mu I) \) is non-invertible, its null-space \( \mathcal{U} \) has dimension \( r \geq 1 \). Let \( V \in \mathbb{C}^{n \times r} \) have \( r \) columns comprising an orthonormal basis for \( \mathcal{U} \); then extend \( V \) to \( V \in \mathbb{C}^{n \times n} \) by adding orthonormal columns so that \( V \) is unitary. Now

\[
B = V^*AV = \begin{pmatrix} \mu I & C \\ 0 & D \end{pmatrix},
\]

where \( I \) is the \( r \times r \) identity, \( C \) is \( r \times (n-r) \) and \( D \) is \( (n-r) \times (n-r) \), and \( B \) and \( A \) are similar. Then

\[
\det(B - zI) = \det(\mu I - zI) \det(D - zI) = (\mu - z)^r \det(D - zI).
\]

Thus \( B \) has algebraic multiplicity \( \geq r \) for \( B \) and hence \( A \) has algebraic multiplicity \( \geq r \), by Theorem 1.16. \( \square \)

**Definition 1.19.** The spectral radius of a matrix \( A \in \mathbb{C}^{n \times n} \) is defined by

\[
\rho(A) = \max\{|\lambda| \mid \lambda \text{ is an eigenvalue of } A\}.
\]

By considering the eigenvectors of a matrix, we can define an important class of matrices

**Definition 1.20.** A matrix \( A \in \mathbb{C}^{n \times n} \) is normal iff it has \( n \) orthogonal eigenvectors.

The importance of this concept lies in the fact, that normal matrices can always be diagonalised: if \( Q \in \mathbb{C}^{n \times n} \) is a matrix where the columns form an orthonormal system of eigenvectors and \( \Lambda \in \mathbb{C}^{n \times n} \) is the diagonal matrix with the corresponding eigenvalues on the diagonal, then we have

\[
A = Q\Lambda Q^*.
\]

Using this relation, we see that every normal matrix satisfies \( A^*A = QA^*AQ^* = Q\Lambda\Lambda^*Q^* = AA^* \). In Theorem 2.3 we will see that the condition \( A^*A = AA^* \) is actually equivalent to \( A \) being normal in the sense of the definition above. Sometimes this alternative condition is used to define when a matrix is normal. As a consequence of this equivalence, every Hermitian matrix is also normal.

Let now \( A \) be Hermitian and positive definite. Then \( Ax = \lambda x \) implies

\[
\lambda(x, x) = \langle x, \lambda x \rangle = \langle x, Ax \rangle = \langle Ax, x \rangle = \langle \lambda x, x \rangle = \bar{\lambda} \langle x, x \rangle.
\]

Thus, all eigenvalues of \( A \) are real and we can arrange them in increasing order \( \lambda_1 \leq \cdots \leq \lambda_n = \lambda_{\text{max}} \). The following lemma uses \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \) to estimate the values of the norm \( \| \cdot \|_A \) from (1.5) by \( \| \cdot \| \).

**Lemma 1.21.** Let \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \) be the smallest and largest eigenvalues of a Hermitian, positive definite matrix \( A \in \mathbb{C}^{n \times n} \). Then

\[
\lambda_{\text{min}}\|x\|^2 \leq \|x\|_A^2 \leq \lambda_{\text{max}}\|x\|^2 \quad \forall x \in \mathbb{C}^n.
\]
Proof. Let \( \varphi_1, \ldots, \varphi_n \) be an orthonormal system of eigenvectors of \( A \) with corresponding eigenvalues \( \lambda_{\min} = \lambda_1 \leq \cdots \leq \lambda_n = \lambda_{\max} \). By writing \( x \) as \( x = \sum_{i=1}^{n} \xi_i \varphi_i \), we get

\[
\|x\|^2 = \sum_{i=1}^{n} |\xi_i|^2
\]

and

\[
\|x\|^2_A = \sum_{i=1}^{n} \lambda_i |\xi_i|^2.
\]

This gives the upper bound

\[
\|x\|^2 \leq \sum_{i=1}^{n} \lambda_{\max} |\xi_i|^2 \leq \lambda_{\max} \|x\|^2
\]

and similarly we get the lower bound. \( \square \)

1.3 Dual Spaces

Let \( \langle \cdot, \cdot \rangle \) denote the standard inner-product.

Definition 1.22. Given a norm \( \| \cdot \| \) on \( \mathbb{C}^n \), the pair \( (\mathbb{C}^n, \| \cdot \|) \) is a Banach space \( B \). The Banach space \( B' \), the dual of \( B \), is the pair \( (\mathbb{C}^n, \| \cdot \|_{B'}) \), where

\[
\|x\|_{B'} = \max_{\|y\| = 1} \langle x, y \rangle.
\]

See Exercise 1-5 to deduce that the preceding definition satisfies the norm axioms.

Theorem 1.23. The spaces \( (\mathbb{C}^n, \| \cdot \|_1) \) and \( (\mathbb{C}^n, \| \cdot \|_{\infty}) \) are the duals of one another.

Proof. Firstly, we must show

\[
\|x\|_1 = \max_{\|y\| = 1} |\langle x, y \rangle|.
\]

This is clearly true for \( x = 0 \) and so we consider only \( x \neq 0 \). Now

\[
|\langle x, y \rangle| \leq \max_{i} |y_i| \sum_{j=1}^{n} |x_j| = \|y\|_{\infty} \|x\|_1,
\]

and therefore

\[
\max_{\|y\|_{\infty} = 1} |\langle x, y \rangle| \leq \|x\|_1.
\]

We need to show that this upper-bound is achieved. If \( y_j = x_j/|x_j| \) (with the convention that this is 0 when \( x_j \) is 0) then \( \|y\|_{\infty} = 1 \) (since \( x \neq 0 \)) and

\[
\langle x, y \rangle = \sum_{j=1}^{n} |x_j|^2/|x_j| = \sum_{j=1}^{n} |x_j| = \|x\|_1.
\]

Hence \( \max_{\|y\|_{\infty} = 1} |\langle x, y \rangle| = \|x\|_1 \).

Secondly, it remains to show that

\[
\|x\|_{\infty} = \max_{\|y\|_1 = 1} |\langle x, y \rangle|.
\]

We have

\[
|\langle x, y \rangle| \leq \|y\|_1 \|x\|_{\infty}
\]

\[
\Rightarrow \max_{\|y\|_1 = 1} |\langle x, y \rangle| \leq \|x\|_{\infty}.
\]

If \( x = 0 \) we have equality; if not then, for some \( k \) such that \( |x_k| = \|x\|_{\infty} > 0 \), choose \( y_j = \delta_{jk} x_k/|x_k| \). Then

\[
\langle x, y \rangle = |x_k| = \|x\|_{\infty} \quad \text{and} \quad \|y\|_1 = 1.
\]

Thus \( \max_{\|y\|_1 = 1} |\langle x, y \rangle| = \|x\|_{\infty} \). \( \square \)
Theorem 1.24. If \( p, q \in (1, \infty) \) with \( p^{-1} + q^{-1} = 1 \) then the Banach spaces \( (\mathbb{C}^n, \| \cdot \|_p) \) and \( (\mathbb{C}^n, \| \cdot \|_q) \) are the duals of one another.

Proof. See Exercise 1-6. \( \square \)

1.4 Matrix Norms

Since we can consider the space \( \mathbb{C}^{m \times n} \) of all \( m \times n \)-matrices to be a copy of the \( m \times n \)-dimensional vector space \( \mathbb{C}^{mn} \), we can use all vector norms on \( \mathbb{C}^{mn} \) as vector norms on the matrices \( \mathbb{C}^{m \times n} \).

Examples of vector norms on the space of matrices include

- maximum norm: \( \| A \|_{\text{max}} = \max_{i,j} |a_{ij}| \)
- Frobenius norm: \( \| A \|_F = \left( \sum_{i,j=1}^{m,n} |a_{ij}|^2 \right)^{\frac{1}{2}} \)
- operator norm \( \mathbb{C}^n \to \mathbb{C}^m \): if \( A \in \mathbb{C}^{m \times n} \),
  \[
  \| A \|_{(\hat{m},\hat{n})} = \max_{\| x \|_{\hat{n}} = 1} \| Ax \|_{\hat{m}}
  \]
  where \( \| \cdot \|_{\hat{m}} \) is a norm on \( \mathbb{C}^m \), and \( \| \cdot \|_{\hat{n}} \) is a norm on \( \mathbb{C}^n \). Note that, for any operator norm,
  \[
  \| A \|_{(\hat{m},\hat{n})} = \max_{\| x \|_{\hat{n}} \leq 1} \| Ax \|_{\hat{m}} = \max_{\| x \|_{\hat{n}} = 1} \| Ax \|_{\hat{m}} = \max_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\| Ax \|_{\hat{m}}}{\| x \|_{\hat{n}}}
  \]

Sometimes it is helpful to consider special vector norms on a space of matrices, which are compatible with the matrix-matrix multiplication.

Definition 1.25. A matrix norm on \( \mathbb{C}^{n \times n} \) is a mapping \( \| \cdot \| : \mathbb{C}^{n \times n} \to \mathbb{R} \) with

a) \( \| A \| \geq 0 \) for all \( A \in \mathbb{C}^{n \times n} \) and \( \| A \| = 0 \) iff \( A = 0 \),
b) \( \| \alpha A \| = |\alpha| \| A \| \) for all \( \alpha \in \mathbb{C} \), \( A \in \mathbb{C}^{n \times n} \),
c) \( \| A + B \| \leq \| A \| + \| B \| \) for all \( A, B \in \mathbb{C}^{n \times n} \),
d) \( \| AB \| \leq \| A \| \| B \| \) for all \( A, B \in \mathbb{C}^{n \times n} \).

Remark. Conditions a), b) and c) state that \( \| \cdot \| \) is a vector norm on the vector space \( \mathbb{C}^{n \times n} \). Condition d) only makes sense for matrices, since general vectors spaces are not equipped with a product.

Examples of matrix norms include

- \( p \)-operator norm \( \mathbb{C}^n \to \mathbb{C}^n \): if \( A \in \mathbb{C}^{n \times n} \),
  \[
  \| A \|_p = \max_{\| x \|_p = 1} \| Ax \|_p, \quad 1 \leq p \leq \infty
  \]

  The vector operator norm from \( \mathbb{C}^n \) into \( \mathbb{C}^m \) reduces to the \( p \)-operator norm if \( n = m \) and the \( p \)-norm is chosen in the range and image spaces.

Definition 1.26. Given a vector norm \( \| \cdot \|_v \) on \( \mathbb{C}^n \) we define the induced norm \( \| \cdot \|_m \) on \( \mathbb{C}^{n \times n} \) by

\[
\| A \|_m = \max_{x \neq 0} \frac{\| Ax \|_v}{\| x \|_v}
\]

for all \( A \in \mathbb{C}^{n \times n} \).

We now show that the induced norm is indeed a norm.
Theorem 1.27. The induced norm $\| \cdot \|_m$ of a vector norm $\| \cdot \|_v$ is a matrix norm with

$$\| I \|_m = 1$$

and

$$\| Ax \|_v \leq \| A \|_m \| x \|_v$$

for all $A \in \mathbb{C}^{n \times n}$ and $x \in \mathbb{C}^n$.

Proof. a) $\| A \|_m \in \mathbb{R}$ and $\| A \|_m \geq 0$ for all $A \in \mathbb{C}^{n \times n}$ is obvious from the definition. Also from the definition we get

$$\| A \|_m = 0 \iff \| Ax \|_v = 0 \quad \forall x \neq 0 \iff A x = 0 \quad \forall x \neq 0 \iff A = 0.$$

b) For $\alpha \in \mathbb{C}$ and $A \in \mathbb{C}^{n \times n}$ we get

$$\| \alpha A \|_m = \max_{x \neq 0} \frac{\| \alpha Ax \|_v}{\| x \|_v} = \max_{x \neq 0} \frac{\| \alpha \| \| Ax \|_v}{\| x \|_v} = |\alpha| \| A \|_m.$$

c) For $A, B \in \mathbb{C}^{n \times n}$ we get

$$\| A + B \|_m = \max_{x \neq 0} \frac{\| (A + B)x \|_v}{\| x \|_v} \leq \max_{x \neq 0} \frac{\| Ax \|_v + \| Bx \|_v}{\| x \|_v} \leq \max_{x \neq 0} \frac{\| Ax \|_v}{\| x \|_v} + \max_{x \neq 0} \frac{\| Bx \|_v}{\| x \|_v} = \| A \|_m + \| B \|_m.$$

Before we check condition d) from the definition of a matrix norm we verify

$$\| I \|_m = \max_{x \neq 0} \frac{\| Ix \|_v}{\| x \|_v} = \max_{x \neq 0} \frac{\| x \|_v}{\| x \|_v} = 1$$

and

$$\| A \|_m = \max_{y \neq 0} \frac{\| Ay \|_v}{\| y \|_v} \geq \frac{\| Ax \|_v}{\| x \|_v} \quad \forall x \in \mathbb{C}^n \setminus \{0\}$$

which gives

$$\| Ax \|_v \leq \| A \|_m \| x \|_v \quad \forall x \in \mathbb{C}^n.$$

d) Using this estimate we find

$$\| AB \|_m = \max_{x \neq 0} \frac{\| ABx \|_v}{\| x \|_v} \leq \max_{x \neq 0} \frac{\| A \|_m \| Bx \|_v}{\| x \|_v} = \| A \|_m \| B \|_m.$$

Remarks.

1. Usually one denotes the induced matrix norm with the same symbol as the corresponding vector norm. For the remainder of this text we will follow this convention.

2. As a consequence of theorem 1.27 we can see that not every matrix norm is an induced norm: If $\| \cdot \|_m$ is a matrix norm, then it is easy to check that $\| \cdot \|'_m = 2 \cdot \|_m$ is a matrix norm, too. But at most one of these two norms can equal 1 for the identity matrix, and thus the other one cannot be an induced matrix norm.
3. Recall that \(\|\cdot\|_A := \|A \cdot\|\) is a vector norm on \(\mathbb{C}^n\) whenever \(\|\cdot\|\) is, provided that \(A\) is invertible. The inequality from Theorem 1.27 gives the following upper and lower bounds for the norm \(\|\cdot\|\) in terms of the original norm:

\[
\frac{1}{\|A^{-1}\|} \|x\| \leq \|x\|_A \leq \|A\| \|x\|.
\]

**Theorem 1.28.** The matrix norm induced by the infinity norm is the maximum row sum:

\[
\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.
\]

**Proof.** For \(x \in \mathbb{C}^n\) we get

\[
\|Ax\|_\infty = \max_{1 \leq i \leq n} |(Ax)_i| = \max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}| \|x\|_\infty
\]

which gives

\[
\frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|
\]

for all \(x \in \mathbb{C}^n\) and thus \(\|A\|_\infty \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|\).

For the lower bound choose \(k \in \{1, 2, \ldots, n\}\) such that

\[
\max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}| = \sum_{j=1}^{n} |a_{kj}|
\]

and define \(x \in \mathbb{C}^n\) by \(x_j = \frac{a_{kj}}{|a_{kj}|}\) for all \(j = 1, \ldots, n\) (with the convention that this is 0 when \(a_{kj}\) is 0). Then we have \(\|x\|_\infty = 1\) and

\[
\|A\|_\infty \geq \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^{n} a_{ij} \frac{|a_{kj}|}{|a_{kj}|} = \sum_{j=1}^{n} |a_{kj}| = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.
\]

This is the required result. \(\Box\)

**Theorem 1.29.** Let \(A \in \mathbb{C}^{n \times n}\). Then \(\|A^*\|_1 = \|A\|_\infty\) and so

\[
\|A\|_1 = \|A^*\|_\infty = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|.
\]

This expression is known as the maximum column sum.
Proof. (Sketch)

\[ \|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty \]
\[ = \max_{\|x\|_\infty = 1} \left( \max_{\|y\|_1 = 1} |(Ax, y)| \right) \quad \text{(Dual definition)} \]
\[ = \max_{\|y\|_1 = 1} \left( \max_{\|x\|_\infty = 1} |(Ax, y)| \right) \quad \text{(needs careful justification)} \]
\[ = \max_{\|y\|_1 = 1} \left( \max_{\|x\|_\infty = 1} |(x, A^*y)| \right) \quad \text{(definition of } A^*) \]
\[ = \max_{\|y\|_1 = 1} \|A^*y\|_1 \quad \text{(Dual definition)} \]
\[ = \|A^*\|_1. \]

Since \((A^*)_{ij} = \overline{a_{ji}}\), the result follows. \( \Box \)

Remark. This result is readily extended to non-square matrices \( A \in \mathbb{C}^{m \times n} \).

Recall that the spectral radius of a matrix is defined as
\[ \rho(A) = \max \{ |\lambda| \mid \lambda \text{ is eigenvalue of } A \} . \]

Theorem 1.30. For any matrix norm \( \| \cdot \| \), any matrix \( A \in \mathbb{C}^{n \times n} \) and any \( k \in \mathbb{N} \) we have
\[ \rho(A)^k \leq \rho(A^k) \leq \|A^k\| \leq \|A\|^k . \]

Proof. Let \( B = A^k \). The first inequality is a consequence of the fact that, whenever \( x \) is an eigenvector of \( A \) with eigenvalue \( \lambda \), the vector \( x \) is also an eigenvector of \( B \), but with eigenvalue \( \lambda^k \). By definition of the spectral radius \( \rho(B) \) we can find an eigenvector \( x \) with \( Bx = \lambda x \) and \( \rho(B) = |\lambda| \). Let \( X \in \mathbb{C}^{n \times n} \) be the matrix where all \( n \) columns are equal to \( x \). Then we have \( BX = \lambda X \) and thus
\[ \|B\|\|X\| \geq \|BX\| = \|\lambda X\| = |\lambda|\|X\| = \rho(B)\|X\|. \]
Dividing by \( \|X\| \) gives \( \rho(B) \leq \|B\| \). The final inequality follows from property d) in the definition of a matrix norm. \( \Box \)

Theorem 1.31. If \( A \in \mathbb{C}^{n \times n} \) is normal, then
\[ \rho(A)^\ell = \|A^\ell\|_2 = \|A\|_2^\ell \quad \forall \ell \in \mathbb{N} . \]

Proof. Let \( x_1, \ldots, x_n \) be an orthonormal basis composed of eigenvectors of \( A \) with corresponding eigenvalues \( \lambda_1, \ldots, \lambda_n \). Without loss of generality we have \( \rho(A) = |\lambda_1| \).

Let \( x \in \mathbb{C}^n \). Then we can write
\[ x = \sum_{j=1}^n \alpha_j x_j \]
and get
\[ \|x\|_2^2 = \sum_{j=1}^n |\alpha_j|^2 . \]
Similarly we find
\[ Ax = \sum_{j=1}^n \alpha_j \lambda_j x_j \quad \text{and} \quad \|Ax\|_2^2 = \sum_{j=1}^n |\alpha_j \lambda_j|^2 . \]
This shows
\[
\frac{\|Ax\|_2}{\|x\|_2} = \frac{(\sum_{j=1}^{n} |\alpha_j \lambda_j|^2)^{1/2}}{(\sum_{j=1}^{n} |\alpha_j|^2)^{1/2}} \\
\leq \frac{(\sum_{j=1}^{n} |\alpha_j|^2 |\lambda_j|^2)^{1/2}}{(\sum_{j=1}^{n} |\alpha_j|^2)^{1/2}} \\
= |\lambda_1| = \rho(A) \quad \forall x \in \mathbb{C}^n
\]
and consequently \(\|A\|_2 \leq \rho(A)\).

Using Theorem 1.30 we get
\[
\rho(A^\ell) \leq \|A\|_2^\ell \leq \|A\|_2^\ell \leq \rho(A)^\ell
\]
for all \(\ell \in \mathbb{N}\). This completes the proof. \(\square\)

Similar methods to those used in the proof of the previous result yield the following theorem.

**Theorem 1.32.** For all matrices \(A \in \mathbb{C}^{m \times n}\)
\[
\|A\|_2^2 = \rho(A^*A).
\]

**Proof.** See Exercise 1-9. \(\square\)

The matrix 2-norm has the special property that it is invariant under multiplication by a unitary matrix. This is the analog of Theorem 1.9 for vector norms.

**Theorem 1.33.** For all matrices \(A \in \mathbb{C}^{m \times n}\) and unitary matrices \(U \in \mathbb{C}^{m \times m}, V \in \mathbb{C}^{n \times n}\)
\[
\|UA\|_2 = \|A\|_2, \quad \|AV\|_2 = \|A\|_2.
\]

**Proof.** The first result follows from the previous theorem, after noting that \((UA)^*(UA) = A^*U^*UA = A^*A\). Because \((AV)^*(AV) = V^*(A^*A)V\) and because \(V^*(A^*A)V\) is a similarity transformation of \(A^*A\) the second result also follows from the previous theorem. \(\square\)

Let \(A, B \in \mathbb{C}^{m \times n}\). In the following it will be useful to employ the notation \(|A|\) to denote the matrix with entries \(|A|_{ij} = |a_{ij}|\) and the notation \(|A| \leq |B|\) as shorthand for \(|a_{ij}| \leq |b_{ij}|\) for all \(i,j\).

**Lemma 1.34.** If two matrices \(A, B \in \mathbb{C}^{m \times n}\) satisfy \(|A| \leq |B|\) then \(\|A\|_\infty \leq \|B\|_\infty\) and \(\|A\|_1 \leq \|B\|_1\). Furthermore \(\|AB\|_\infty \leq \|A\|_1 \|B\|_\infty\).

**Proof.** For the first two observations, it suffices to prove the first result since \(\|A\|_1 = \|A^*\|_\infty\) and \(|A| \leq |B|\) implies that \(|A^*| \leq |B^*|\). The first result is a direct consequence of the representation of the \(\infty\)-norm and 1-norm from theorems 1.28 and 1.29. To prove the last result note that
\[
(|AB|)_{ij} = \sum_k |A_{ik}B_{kj}|
\leq \sum_k |A_{ik}| |B_{kj}|
= (|A||B|)_{ij}.
\]
The first result completes the proof. \(\square\)

**Lemma 1.35.** Let \(A, B \in \mathbb{C}^{n \times n}\). Then
\[
\|A\|_\max \leq \|A\|_\infty \leq n\|A\|_\max, \quad \|AB\|_\max \leq \|A\|_\infty \|B\|_\max, \quad \|AB\|_\max \leq \|A\|_\max \|B\|_1.
\]
Definition 1.36. The outer product of two vectors \( a, b \in \mathbb{C}^n \) is the matrix \( a \otimes b \in \mathbb{C}^{n \times n} \) defined by\[ (a \otimes b)c = (b^* c)a = \langle b, c \rangle a \quad \forall c \in \mathbb{C}^n. \]
We sometimes write \( a \otimes b = ab^* \). The \( ij \)th entry of the outer product is \( (a \otimes b)_{ij} = a_i \bar{b}_j \).

Definition 1.37. Let \( S \) be a subspace of \( \mathbb{C}^n \). Then the orthogonal complement of \( S \) is defined by\[ S^\perp = \{ x \in \mathbb{C}^n \mid \langle x, y \rangle = 0 \quad \forall y \in S \}. \]
The orthogonal projection onto \( S \), \( P \), can be defined as follows: let \( \{ y_i \}_{i=1}^k \) be an orthonormal basis for \( S \), then\[ Px = \sum_{j=1}^k \langle y_j, x \rangle y_j = \left( \sum_{j=1}^k y_j y_j^* \right) x = \left( \sum_{j=1}^k y_j \otimes y_j \right) x. \]

Theorem 1.38. \( P \) is a projection, that is \( P^2 = P \). Furthermore, if \( P^\perp = I - P \), then \( P^\perp \) is the orthogonal projection onto \( S^\perp \).

Proof. Extend \( \{ y_i \}_{i=1}^k \) to a basis for \( \mathbb{C}^n \), denoted \( \{ y_i \}_{i=1}^n \), noting that \( S^\perp = \text{span} \{ y_{k+1}, \ldots, y_n \} \).
Any \( x \in \mathbb{C}^n \) can be written uniquely as\[ x = \sum_{j=1}^n \langle y_j, x \rangle y_j, \]
and so\[ Px = \sum_{j=1}^k \langle y_j, x \rangle y_j, \]
found by truncating to \( k \) terms. Clearly truncating again leaves the expression unchanged:\[ P^2 x = \sum_{j=1}^k \langle y_j, x \rangle y_j = Px, \quad \forall x \in \mathbb{C}^n. \]
Now \( (I - P)x = P^\perp x = \sum_{j=k+1}^n \langle y_j, x \rangle y_j \), proving the second result.

1.5 Structured Matrices

Definition 1.39. A matrix \( A \in \mathbb{C}^{n \times n} \) is\[ \begin{array}{ll} \text{diagonal} & \text{if } a_{ij} = 0 \quad i \neq j \\ \text{(strictly) upper-triangular} & \text{if } a_{ij} = 0 \quad i > j \\
\text{(strictly) lower-triangular} & \text{if } a_{ij} = 0 \quad i < j \\
\text{upper Hessenberg} & \text{if } a_{ij} = 0 \quad i > j + 1 \\
\text{upper bidiagonal} & \text{if } a_{ij} = 0 \quad i > j \land i < j - 1 \\
\text{tridiagonal} & \text{if } a_{ij} = 0 \quad i > j + 1 \land i < j - 1 \end{array} \]

Definition 1.40. A matrix \( P \in \mathbb{R}^{n \times n} \) is called a permutation matrix if every row and every column contains \( n - 1 \) zeros and 1 one.
Remarks.

1. If $P$ is a permutation matrix, then we have

\[(P^TP)_{ij} = \sum_{k=1}^{n} p_{ki}p_{kj} = \delta_{ij}\]

and thus $P^TP = I$. This shows that permutation matrices are orthogonal.

2. If $\pi: \{1, \ldots, n\} \to \{1, \ldots, n\}$ is a permutation, then the matrix $P = (p_{ij})$ with

\[p_{ij} = \begin{cases} 1 & \text{if } j = \pi(i) \text{ and} \\ 0 & \text{else} \end{cases}\]

is a permutation matrix. Indeed every permutation matrix is of this form. In particular the identity matrix is a permutation matrix.

3. If $P$ is the permutation matrix corresponding to the permutation $\pi$, then $(P^{-1})_{ij} = 1$ if and only if $j = \pi^{-1}(i)$. Thus the permutation matrix $P^{-1}$ corresponds to the permutation $\pi^{-1}$.

4. We get

\[(PA)_{ij} = \sum_{k=1}^{n} p_{ik}a_{kj} = a_{\pi(i),j}\]

for all $i, j \in \{1, \ldots, n\}$. This shows that multiplying a permutation matrix from the left reorders the rows of $A$. Furthermore we have

\[(AP)_{ij} = \sum_{k=1}^{n} a_{ik}p_{kj} = a_{i,\pi^{-1}(j)}\]

and hence multiplying a permutation matrix from the right reorders the columns of $A$. 

5. If $P$ is a permutation matrix, then $P^T$ is also a permutation matrix.

Bibliography

Excellent treatments of matrix analysis may be found in [Bha97] and [HJ85]. More advanced treatment of the subject includes [Lax97]. Theorem 1.16 and Lemma 1.17 are proved in [Ner71]. The proof of Theorem 1.24 may be found in [Rob01]. Theorem 1.32 is proved in the solutions for instructors.

Exercises

Exercise 1-1. Show that the following relations hold for all $x \in \mathbb{C}^n$:

a) $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$,

b) $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$ and

c) $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$.

Exercise 1-2. Prove that, for Hermitian positive definite $A$, equations (1.4) and (1.5) define an inner-product and norm, respectively.

Exercise 1-3. For matrices in $\mathbb{R}^{m \times n}$ prove that

\[\|A\|_{\text{max}} \leq \|A\|_F \leq \sqrt{mn}\|A\|_{\text{max}}.\]
Exercise 1-4. Define an inner product $\langle \cdot, \cdot \rangle$ on matrices in $\mathbb{R}^{n \times n}$ such that
$$\|A\|_F^2 = \langle A, A \rangle.$$

Exercise 1-5. Prove that the norm $\| \cdot \|_{B'}$ appearing in Definition 1.22 is indeed a norm.


Exercise 1-7. Show that $\|A\|_{\text{max}} = \max_{i,j} |a_{ij}|$ for all $A \in \mathbb{C}^{n \times n}$ defines a vector norm on the space of $n \times n$-matrices, but not a matrix norm.

Exercise 1-8. Prove Lemma 1.35.

Exercise 1-9. Show that $\|A\|_2^2 = \rho(A^T A)$ for every matrix $A \in \mathbb{R}^{n \times n}$ (this is the real version of Theorem 1.32).

Exercise 1-10. For $A \in \mathbb{R}^{n \times n}$ recall the definition of $|A|$, namely $(|A|)_{ij} = |a_{ij}|$. Show that
$$\| |A| \| = \| A \|$$
holds in the Frobenius, infinity and 1-norms. Is the result true in the Euclidean norm? Justify your assertion.

Exercise 1-11. Let $\| \cdot \|$ be an operator norm. Prove that if $\|X\| < 1$, then
- $I - X$ is invertible,
- the series $\sum_{i=0}^{\infty} X^i$ converges, and
- $(I - X)^{-1} = \sum_{i=0}^{\infty} X^i$.

Moreover, prove that in the same norm
$$\|(I - X)^{-1}\| \leq (1 - \|X\|)^{-1}.$$

Exercise 1-12. Let $K$ be a matrix in $\mathbb{R}^{n \times n}$ with non-negative entries and let $f, g$ be two vectors in $\mathbb{R}^n$ with strictly positive entries which satisfy
$$(Kf)_i/g_i < \lambda, \ (K^T g)_i/f_i < \mu \quad \forall i \in \{1, \ldots, n\}.$$

Prove that $\|K\|_2^2 \leq \lambda \mu$.

Exercise 1-13. Let $A \in \mathbb{R}^{k \times l}, B \in \mathbb{R}^{l \times m}$ and $C \in \mathbb{R}^{m \times m}$. Here $k \geq l$. If $A$ and $C$ are orthogonal, that is if $C^T C = I$ (the identity on $\mathbb{R}^m$) and $A^T A = I$ (the identity on $\mathbb{R}^l$) then show that $\|ABC\|_2 = \|B\|_2$.

Exercise 1-14. Prove that the Frobenius norm of a matrix is unchanged by multiplication by unitary matrices. This is an analogue of Theorem 1.33.

Exercise 1-15. Show that for every vector norm $\| \cdot \|$ on $\mathbb{R}^{n \times n}$ there is a number $\lambda > 0$ such that
$$\|A\|_{\lambda} = \lambda \|A\| \quad A \in \mathbb{R}^{n \times n}$$
defines a matrix norm.