tioners have achieved remarkable progress in the development and use of effective iterative methods. Unfortunately, fewer elegant results have been discovered since the 1950s and 1960s. The field has moved in other directions. Methods have gained not only in efficiency but also in robustness and in generality. The traditional techniques, which required rather complicated procedures to determine optimal acceleration parameters, have yielded to the parameter-free conjugate gradient class of methods.

The primary aim of this book is to describe some of the best techniques available today, from both preconditioners and accelerators. One of the secondary aims of the book is to provide a good mix of theory and practice. It also addresses some of the current research issues, such as parallel implementations and robust preconditioners. The emphasis is on Krylov subspace methods, currently the most practical and common group of techniques used in applications. Although there is a tutorial chapter that covers the discretization of partial differential equations, the book is not biased toward any specific application area. Instead, the matrices are assumed to be general sparse and possibly irregularly structured.

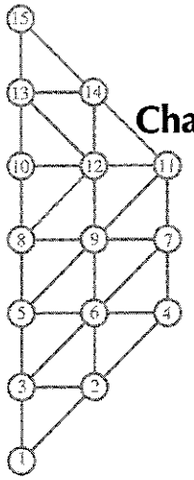
The book has been structured in four distinct parts. The first part, Chapters 1 to 4, presents the basic tools. The second part, Chapters 5 to 8, presents projection methods and Krylov subspace techniques. The third part, Chapters 9 and 10, discusses preconditioning. The fourth part, Chapters 11 to 13, discusses parallel implementations and parallel algorithms.

Acknowledgments

I am grateful to a number of colleagues who proofread or reviewed different versions of the manuscript. Among them are Randy Bramley (University of Indiana at Bloomington), Xiao-Chuan Cai (University of Colorado at Boulder), Tony Chan (University of California at Los Angeles), Jane Cullum (IBM, Yorktown Heights), Alan Edelman (Massachusetts Institute of Technology), Paul Fischer (Brown University), David Keyes (Old Dominion University), Beresford Parlett (University of California at Berkeley), and Shang-Hua Teng (University of Minnesota). Their numerous comments and corrections and their encouragement were a highly appreciated contribution. In particular, they helped improve the presentation considerably and prompted the addition of a number of topics missing from earlier versions.

This book evolved from several successive improvements of a set of lecture notes for the course "Iterative Methods for Linear Systems," which I taught at the University of Minnesota in the last few years. I apologize to those students who used the earlier error-laden and incomplete manuscripts. Their input and criticism contributed significantly to improving the manuscript. I also wish to thank those students at MIT (with Alan Edelman) and UCLA (with Tony Chan) who used this book in manuscript form and provided helpful feedback. My colleagues at the University of Minnesota, staff and faculty members, have helped in different ways. I wish to thank in particular Ahmed Sameh for his encouragement and for fostering a productive environment in the department. Finally, I am grateful to the National Science Foundation for its continued financial support of my research, part of which is represented in this work.

Yousef Saad



Chapter 1
Background in Linear Algebra

This chapter gives an overview of the relevant concepts in linear algebra that are useful in later chapters. It begins with a review of basic matrix theory and introduces the elementary notation used throughout the book. The convergence analysis of iterative methods requires a good level of knowledge in mathematical analysis and in linear algebra. Traditionally, many of the concepts presented specifically for these analyses have been geared toward matrices arising from the discretization of partial differential equations (PDEs) and basic relaxation-type methods. These concepts are now becoming less important because of the trend toward projection-type methods, which have more robust convergence properties and require different analysis tools. The material covered in this chapter will be helpful in establishing some theory for the algorithms and defining the notation used throughout the book.

1.1 Matrices

For the sake of generality, all vector spaces considered in this chapter are complex, unless otherwise stated. A complex $n \times m$ matrix A is an $n \times m$ array of complex numbers

$$a_{ij}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m.$$

The set of all $n \times m$ matrices is a complex vector space denoted by $\mathbb{C}^{n \times m}$. The main operations with matrices are the following:

- Addition: $C = A + B$, where A, B, and C are matrices of size $n \times m$ and
$$c_{ij} = a_{ij} + b_{ij}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, m.$$
- Multiplication by a scalar: $C = \alpha A$, where
$$c_{ij} = \alpha\, a_{ij}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, m.$$

- Multiplication by another matrix:

$$C = AB,$$

where $A \in \mathbb{C}^{n\times m}$, $B \in \mathbb{C}^{m\times p}$, $C \in \mathbb{C}^{n\times p}$, and

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Sometimes, a notation with column vectors and row vectors is used. The column vector a_{*j} is the vector consisting of the jth column of A:

$$a_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

Similarly, the notation a_{i*} will denote the ith row of the matrix A:

$$a_{i*} = (a_{i1}, a_{i2}, \ldots, a_{im}).$$

For example, the following could be written:

$$A = (a_{*1}, a_{*2}, \ldots, a_{*m})$$

or

$$A = \begin{pmatrix} a_{1*} \\ a_{2*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

The *transpose* of a matrix A in $\mathbb{C}^{n\times m}$ is a matrix C in $\mathbb{C}^{m\times n}$ whose elements are defined by $c_{ij} = a_{ji}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. It is denoted by A^T. It is often more relevant to use the *transpose conjugate* matrix denoted by A^H and defined by

$$A^H = \bar{A}^T = \overline{A^T},$$

in which the bar denotes the (element-wise) complex conjugation.

Matrices are strongly related to linear mappings between vector spaces of finite dimension. This is because they represent these mappings with respect to two given bases: one for the initial vector space and the other for the image vector space, or *range*, of A.

1.2 Square Matrices and Eigenvalues

A matrix is *square* if it has the same number of columns and rows, i.e., if $m = n$. An important square matrix is the identity matrix

$$I = \{\delta_{ij}\}_{i,j=1,\ldots,n},$$

where δ_{ij} is the Kronecker symbol. The identity matrix satisfies the equality $AI = IA = A$ for every matrix A of size n. The inverse of a matrix, when it exists, is a matrix C such that

$$CA = AC = I.$$

The inverse of A is denoted by A^{-1}.

The *determinant* of a matrix may be defined in several ways. For simplicity, the following recursive definition is used here. The determinant of a 1×1 matrix (a) is defined as the scalar a. Then the determinant of an $n \times n$ matrix is given by

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j}),$$

where A_{1j} is an $(n-1) \times (n-1)$ matrix obtained by deleting the first row and the jth column of A. A matrix is said to be *singular* when $\det(A) = 0$ and *nonsingular* otherwise. We have the following simple properties:

- $\det(AB) = \det(A)\det(B)$.
- $\det(A^T) = \det(A)$.
- $\det(\alpha A) = \alpha^n \det(A)$.
- $\det(\bar{A}) = \overline{\det(A)}$.
- $\det(I) = 1$.

From the above definition of determinants it can be shown by induction that the function that maps a given complex value λ to the value $p_A(\lambda) = \det(A - \lambda I)$ is a polynomial of degree n; see Exercise 8. This is known as the *characteristic polynomial* of the matrix A.

Definition 1.1. *A complex scalar λ is called an eigenvalue of the square matrix A if a nonzero vector u of $\mathbb{C}^n$ exists such that $Au = \lambda u$. The vector u is called an eigenvector of A associated with λ. The set of all the eigenvalues of A is called the spectrum of A and is denoted by $\sigma(A)$.*

A scalar λ is an eigenvalue of A *if and only if* (iff hereafter) $\det(A - \lambda I) \equiv p_A(\lambda) = 0$. That is true iff λ is a root of the characteristic polynomial. In particular, there are at most n distinct eigenvalues.

It is clear that a matrix is singular iff it admits zero as an eigenvalue. A well-known result in linear algebra is stated in the following proposition.

Proposition 1.2. *A matrix A is nonsingular iff it admits an inverse.*

Thus, the determinant of a matrix determines whether or not the matrix admits an inverse.

The maximum modulus of the eigenvalues is called the *spectral radius* and is denoted by $\rho(A)$:

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|.$$

The *trace* of a matrix is equal to the sum of all its diagonal elements:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

It can be easily shown that the trace of A is also equal to the sum of the eigenvalues of A counted with their multiplicities as roots of the characteristic polynomial.

Proposition 1.3. *If λ is an eigenvalue of A, then $\bar{\lambda}$ is an eigenvalue of A^H. An eigenvector v of A^H associated with the eigenvalue $\bar{\lambda}$ is called a left eigenvector of A.*

When a distinction is necessary, an eigenvector of A is often called a right eigenvector. Therefore, the eigenvalue λ as well as the right and left eigenvectors u and v satisfy the relations

$$Au = \lambda u, \quad v^H A = \lambda v^H$$

or, equivalently,

$$u^H A^H = \bar{\lambda} u^H, \quad A^H v = \bar{\lambda} v.$$

1.3 Types of Matrices

The choice of a method for solving linear systems will often depend on the structure of the matrix A. One of the most important properties of matrices is symmetry, because of its impact on the eigenstructure of A. A number of other classes of matrices also have particular eigenstructures. The most important ones are listed below:

- *Symmetric matrices:* $A^T = A$.
- *Hermitian matrices:* $A^H = A$.
- *Skew-symmetric matrices:* $A^T = -A$.
- *Skew-Hermitian matrices:* $A^H = -A$.
- *Normal matrices:* $A^H A = AA^H$.
- *Nonnegative matrices:* $a_{ij} \geq 0$, $i, j = 1, \ldots, n$ (similar definition for nonpositive, positive, and negative matrices).
- *Unitary matrices:* $Q^H Q = I$.

It is worth noting that a unitary matrix Q is a matrix whose inverse is its transpose conjugate Q^H, since

$$Q^H Q = I \quad \rightarrow \quad Q^{-1} = Q^H. \tag{1.1}$$

A matrix Q such that $Q^H Q$ is diagonal is often called orthogonal.

Some matrices have particular structures that are often convenient for computational purposes. The following list, though incomplete, gives an idea of these special matrices, which play an important role in numerical analysis and scientific computing applications.

- *Diagonal matrices:* $a_{ij} = 0$ for $j \neq i$. Notation:
$$A = \text{diag}(a_{11}, a_{22}, \ldots, a_{nn}).$$
- *Upper triangular matrices:* $a_{ij} = 0$ for $i > j$.
- *Lower triangular matrices:* $a_{ij} = 0$ for $i < j$.

- *Upper bidiagonal matrices:* $a_{ij} = 0$ for $j \neq i$ or $j \neq i + 1$.
- *Lower bidiagonal matrices:* $a_{ij} = 0$ for $j \neq i$ or $j \neq i - 1$.
- *Tridiagonal matrices:* $a_{ij} = 0$ for any pair i, j such that $|j - i| > 1$. Notation:
$$A = \text{tridiag}(a_{i,i-1}, a_{ii}, a_{i,i+1}).$$
- *Banded matrices:* $a_{ij} \neq 0$ only if $i - m_l \leq j \leq i + m_u$, where m_l and m_u are two nonnegative integers. The number $m_l + m_u + 1$ is called the bandwidth of A.
- *Upper Hessenberg matrices:* $a_{ij} = 0$ for any pair i, j such that $i > j + 1$. Lower Hessenberg matrices can be defined similarly.
- *Outer product matrices:* $A = uv^H$, where both u and v are vectors.
- *Permutation matrices:* the columns of A are a permutation of the columns of the identity matrix.
- *Block diagonal matrices:* generalizes the diagonal matrix by replacing each diagonal entry with a matrix. Notation:
$$A = \text{diag}(A_{11}, A_{22}, \ldots, A_{nn}).$$
- *Block tridiagonal matrices:* generalizes the tridiagonal matrix by replacing each nonzero entry with a square matrix. Notation:
$$A = \text{tridiag}(A_{i,i-1}, A_{ii}, A_{i,i+1}).$$

The above properties emphasize structure, i.e., the positions of the nonzero elements with respect to the zeros. Also, they assume that there are many zero elements or that the matrix is of low rank. This is in contrast with the classifications listed earlier, such as symmetry and normality.

1.4 Vector Inner Products and Norms

An inner product on a (complex) vector space $\mathbb{X}$ is any mapping s from $\mathbb{X} \times \mathbb{X}$ into $\mathbb{C}$,

$$x \in \mathbb{X}, y \in \mathbb{X} \quad \rightarrow \quad s(x, y) \in \mathbb{C},$$

that satisfies the following conditions:

1. $s(x, y)$ is linear with respect to x; i.e.,
$$s(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 s(x_1, y) + \lambda_2 s(x_2, y) \quad \forall\, x_1, x_2 \in \mathbb{X}, \forall\, \lambda_1, \lambda_2 \in \mathbb{C}.$$
2. $s(x, y)$ is *Hermitian*; i.e.,
$$s(y, x) = \overline{s(x, y)} \quad \forall\, x, y \in \mathbb{X}.$$
3. $s(x, y)$ is *positive definite*; i.e.,
$$s(x, x) > 0 \quad \forall\, x \neq 0.$$

Note that (2) implies that $s(x, x)$ is real and, therefore, (3) adds the constraint that $s(x, x)$ must also be positive for any nonzero x. For any x and y,

$$s(x, 0) = s(x, 0 \cdot y) = 0 \cdot s(x, y) = 0.$$

Similarly, $s(0, y) = 0$ for any y. Hence, $s(0, y) = s(x, 0) = 0$ for any x and y. In particular condition (3) can be rewritten as

$$s(x, x) \geq 0 \quad \text{and} \quad s(x, x) = 0 \quad \text{iff} \quad x = 0,$$

as can be readily shown. A useful relation satisfied by any inner product is the so-called Cauchy–Schwarz inequality

$$|s(x, y)|^2 \leq s(x, x)\, s(y, y). \tag{1.2}$$

The proof of this inequality begins by expanding $s(x - \lambda y, x - \lambda y)$ using the properties of s:

$$s(x - \lambda y, x - \lambda y) = s(x, x) - \bar{\lambda} s(x, y) - \lambda s(y, x) + |\lambda|^2 s(y, y).$$

If $y = 0$ then the inequality is trivially satisfied. Assume that $y \neq 0$ and take $\lambda = s(x, y)/s(y, y)$. Then, from the above equality, $s(x - \lambda y, x - \lambda y) \geq 0$ shows that

$$\begin{aligned} 0 \leq s(x - \lambda y, x - \lambda y) &= s(x, x) - 2\frac{|s(x, y)|^2}{s(y, y)} + \frac{|s(x, y)|^2}{s(y, y)} \\ &= s(x, x) - \frac{|s(x, y)|^2}{s(y, y)}, \end{aligned}$$

which yields the result (1.2).

In the particular case of the vector space $\mathbb{X} = \mathbb{C}^n$, a *canonical* inner product is the *Euclidean inner product*. The Euclidean inner product of two vectors $x = (x_i)_{i=1,\ldots,n}$ and $y = (y_i)_{i=1,\ldots,n}$ of $\mathbb{C}^n$ is defined by

$$(x, y) = \sum_{i=1}^{n} x_i \bar{y}_i, \tag{1.3}$$

which is often rewritten in matrix notation as

$$(x, y) = y^H x. \tag{1.4}$$

It is easy to verify that this mapping does indeed satisfy the three conditions required for inner products listed above. A fundamental property of the Euclidean inner product in matrix computations is the simple relation

$$(Ax, y) = (x, A^H y) \quad \forall\, x, y \in \mathbb{C}^n. \tag{1.5}$$

The proof of this is straightforward. The *adjoint* of A *with respect to an arbitrary inner product* is a matrix B such that $(Ax, y) = (x, By)$ for all pairs of vectors x and y. A matrix is *self-adjoint*, or Hermitian, with respect to this inner product if it is equal to its adjoint. The following proposition is a consequence of equality (1.5).

Proposition 1.4. *Unitary matrices preserve the Euclidean inner product; i.e.,*

$$(Qx, Qy) = (x, y)$$

for any unitary matrix Q and any vectors x and y.

Proof. Indeed, $(Qx, Qy) = (x, Q^H Qy) = (x, y)$. □

A vector norm on a vector space $\mathbb{X}$ is a real-valued function $x \rightarrow \|x\|$ on $\mathbb{X}$ that satisfies the following three conditions:

1. $\|x\| \geq 0 \quad \forall\, x \in \mathbb{X}$ and $\|x\| = 0$ iff $x = 0$.
2. $\|\alpha x\| = |\alpha| \|x\| \quad \forall\, x \in \mathbb{X} \quad \forall\, \alpha \in \mathbb{C}$.
3. $\|x + y\| \leq \|x\| + \|y\| \quad \forall\, x, y \in \mathbb{X}$.

For the particular case when $\mathbb{X} = \mathbb{C}^n$, we can associate with the inner product (1.3) the *Euclidean norm* of a complex vector defined by

$$\|x\|_2 = (x, x)^{1/2}.$$

It follows from Proposition 1.4 that a unitary matrix preserves the Euclidean norm metric; i.e.,

$$\|Qx\|_2 = \|x\|_2 \;\forall\, x.$$

The linear transformation associated with a unitary matrix Q is therefore an *isometry*.

The most commonly used vector norms in numerical linear algebra are special cases of the Hölder norms

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}. \tag{1.6}$$

Note that the limit of $\|x\|_p$ when p tends to infinity exists and is equal to the maximum modulus of the x_i's. This defines a norm denoted by $\|\cdot\|_\infty$. The cases $p = 1$, $p = 2$, and $p = \infty$ lead to the most important norms in practice:

$$\begin{aligned} \|x\|_1 &= |x_1| + |x_2| + \cdots + |x_n|, \\ \|x\|_2 &= \left[|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 \right]^{1/2}, \\ \|x\|_\infty &= \max_{i=1,\ldots,n} |x_i|. \end{aligned}$$

The Cauchy–Schwarz inequality of (1.2) becomes

$$|(x, y)| \leq \|x\|_2 \|y\|_2.$$

1.5 Matrix Norms

For a general matrix A in $\mathbb{C}^{n \times m}$, we define the following special set of norms:

$$\|A\|_{pq} = \max_{x \in \mathbb{C}^m,\, x \neq 0} \frac{\|Ax\|_p}{\|x\|_q}. \tag{1.7}$$

The norm $\|\cdot\|_{pq}$ is *induced* by the two norms $\|\cdot\|_p$ and $\|\cdot\|_q$. These norms satisfy the usual properties of norms; i.e.,

$$\|A\| \geq 0 \quad \forall\, A \in \mathbb{C}^{n\times m} \quad \text{and} \quad \|A\| = 0 \quad \text{iff} \quad A = 0, \tag{1.8}$$

$$\|\alpha A\| = |\alpha|\|A\| \quad \forall\, A \in \mathbb{C}^{n\times m}, \quad \forall\, \alpha \in \mathbb{C}, \tag{1.9}$$

$$\|A + B\| \leq \|A\| + \|B\| \quad \forall\, A, B \in \mathbb{C}^{n\times m}. \tag{1.10}$$

A norm that satisfies the above three properties is nothing but a *vector norm* applied to the matrix considered as a vector consisting of the m columns stacked into a vector of size nm.

The most important cases are again those associated with $p, q = 1, 2, \infty$. The case $q = p$ is of particular interest and the associated norm $\|\cdot\|_{pq}$ is simply denoted by $\|\cdot\|_p$ and called a *p-norm*. A fundamental property of a p-norm is that

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

an immediate consequence of the definition (1.7). Matrix norms that satisfy the above property are sometimes called *consistent*. Often a norm satisfying the properties (1.8)–(1.10) that is consistent is called a *matrix norm*. A result of consistency is that, for any square matrix A,

$$\|A^k\|_p \leq \|A\|_p^k.$$

In particular the matrix A^k converges to zero if *any* of its p-norms is less than 1.

The Frobenius norm of a matrix is defined by

$$\|A\|_F = \left(\sum_{j=1}^{m} \sum_{i=1}^{n} |a_{ij}|^2 \right)^{1/2}. \tag{1.11}$$

This can be viewed as the 2-norm of the column (or row) vector in $\mathbb{C}^{n^2}$ consisting of all the columns (resp., rows) of A listed from 1 to m (resp., 1 to n). It can be shown that this norm is also consistent, in spite of the fact that it is not induced by a pair of vector norms; i.e., it is not derived from a formula of the form (1.7); see Exercise 5. However, it does not satisfy some of the other properties of the p-norms. For example, the Frobenius norm of the identity matrix is not equal to one. To avoid these difficulties, *we will only use the term matrix norm for a norm that is induced by two norms, as in the definition* (1.7). Thus, we will not consider the Frobenius norm to be a proper matrix norm, according to our conventions, even though it is consistent.

The following equalities satisfied by the matrix norms defined above lead to alternative definitions that are often easier to work with:

$$\|A\|_1 = \max_{j=1,\ldots,m} \sum_{i=1}^{n} |a_{ij}|, \tag{1.12}$$

$$\|A\|_\infty = \max_{i=1,\ldots,n} \sum_{j=1}^{m} |a_{ij}|, \tag{1.13}$$

$$\|A\|_2 = \left[\rho(A^H A)\right]^{1/2} = \left[\rho(AA^H)\right]^{1/2}, \tag{1.14}$$

$$\|A\|_F = \left[\mathrm{tr}(A^H A)\right]^{1/2} = \left[\mathrm{tr}(AA^H)\right]^{1/2}. \tag{1.15}$$

As will be shown later, the eigenvalues of $A^H A$ are nonnegative. Their square roots are called *singular values* of A and are denoted by σ_i, $i = 1, \ldots, m$. Thus, the relation (1.14) states that $\|A\|_2$ is equal to σ_1, the largest singular value of A.

Example 1.1. From the relation (1.14), it is clear that the spectral radius $\rho(A)$ is equal to the 2-norm of a matrix when the matrix is Hermitian. However, it is not a matrix norm in general. For example, the first property of norms is not satisfied, since, for

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

we have $\rho(A) = 0$ while $A \neq 0$. Also, the triangle inequality is not satisfied for the pair A and $B = A^T$, where A is defined above. Indeed,

$$\rho(A + B) = 1 \quad \text{while} \quad \rho(A) + \rho(B) = 0.$$

1.6 Subspaces, Range, and Kernel

A subspace of $\mathbb{C}^n$ is a subset of $\mathbb{C}^n$ that is also a complex vector space. The set of all linear combinations of a set of vectors $G = \{a_1, a_2, \ldots, a_q\}$ of $\mathbb{C}^n$ is a vector subspace called the linear span of G:

$$\begin{aligned} \mathrm{span}\{G\} &= \mathrm{span}\{a_1, a_2, \ldots, a_q\} \\ &= \left\{ z \in \mathbb{C}^n \;\middle|\; z = \sum_{i=1}^{q} \alpha_i a_i,\ \{\alpha_i\}_{i=1,\ldots,q} \in \mathbb{C}^q \right\}. \end{aligned}$$

If the a_i's are linearly independent, then each vector of span$\{G\}$ admits a unique expression as a linear combination of the a_i's. The set G is then called a *basis* of the subspace span$\{G\}$.

Given two vector subspaces S_1 and S_2, their *sum* S is a subspace defined as the set of all vectors that are equal to the sum of a vector of S_1 and a vector of S_2. The intersection of two subspaces is also a subspace. If the intersection of S_1 and S_2 is reduced to $\{0\}$, then the sum of S_1 and S_2 is called their direct sum and is denoted by $S = S_1 \bigoplus S_2$. When S is equal to $\mathbb{C}^n$, then every vector x of $\mathbb{C}^n$ can be written in a unique way as the sum of an element x_1 of S_1 and an element x_2 of S_2. The transformation P that maps x into x_1 is a linear transformation that is *idempotent*, i.e., such that $P^2 = P$. It is called a *projector* onto S_1 along S_2.

Two important subspaces that are associated with a matrix A of $\mathbb{C}^{n\times m}$ are its *range*, defined by

$$\mathrm{Ran}(A) = \{Ax \mid x \in \mathbb{C}^m\}, \tag{1.16}$$

and its *kernel* or *null space*

$$\mathrm{Ker}(A) = \{x \in \mathbb{C}^m \mid Ax = 0\}.$$

The range of A is clearly equal to the linear *span* of its columns. The *rank* of a matrix is equal to the dimension of the range of A, i.e., to the number of linearly independent columns. This *column rank* is equal to the *row rank*, the number of linearly independent

rows of A. A matrix in $\mathbb{C}^{n\times m}$ is of *full rank* when its rank is equal to the smallest of m and n. A fundamental result of linear algebra is stated by the following relation:

$$\mathbb{C}^n = \text{Ran}(A) \oplus \text{Ker}(A^T). \tag{1.17}$$

The same result applied to the transpose of A yields $\mathbb{C}^m = \text{Ran}(A^T) \oplus \text{Ker}(A)$.

A subspace S is said to be *invariant* under a (square) matrix A whenever $AS \subset S$. In particular, for any eigenvalue λ of A the subspace $\text{Ker}(A - \lambda I)$ is invariant under A. The subspace $\text{Ker}(A - \lambda I)$ is called the eigenspace associated with λ and consists of all the eigenvectors of A associated with λ, in addition to the zero vector.

1.7 Orthogonal Vectors and Subspaces

A set of vectors $G = \{a_1, a_2, \ldots, a_r\}$ is said to be *orthogonal* if

$$(a_i, a_j) = 0 \quad \text{when} \quad i \neq j.$$

It is *orthonormal* if, in addition, every vector of G has a 2-norm equal to unity. A vector that is orthogonal to all the vectors of a subspace S is said to be orthogonal to this subspace. The set of all the vectors that are orthogonal to S is a vector subspace called the *orthogonal complement* of S and denoted by $S^\perp$. The space $\mathbb{C}^n$ is the direct sum of S and its orthogonal complement. Thus, any vector x can be written in a unique fashion as the sum of a vector in S and a vector in $S^\perp$. The operator that maps x into its component in the subspace S is the *orthogonal projector* onto S.

Every subspace admits an orthonormal basis that is obtained by taking any basis and *orthonormalizing* it. The orthonormalization can be achieved by an algorithm known as the Gram–Schmidt process, which we now describe.

Given a set of linearly independent vectors $\{x_1, x_2, \ldots, x_r\}$, first normalize the vector x_1, which means divide it by its 2-norm, to obtain the scaled vector q_1 of norm unity. Then x_2 is orthogonalized against the vector q_1 by subtracting from x_2 a multiple of q_1 to make the resulting vector orthogonal to q_1; i.e.,

$$x_2 \leftarrow x_2 - (x_2, q_1)q_1.$$

The resulting vector is again normalized to yield the second vector q_2. The ith step of the Gram–Schmidt process consists of orthogonalizing the vector x_i against all previous vectors q_j.

ALGORITHM 1.1. Gram–Schmidt

1. *Compute* $r_{11} := \|x_1\|_2$. *If* $r_{11} = 0$ *Stop, else compute* $q_1 := x_1/r_{11}$
2. *For* $j = 2, \ldots, r$, *Do*
3. *Compute* $r_{ij} := (x_j, q_i)$ *for* $i = 1, 2, \ldots, j-1$
4. $\hat{q} := x_j - \sum_{i=1}^{j-1} r_{ij} q_i$
5. $r_{jj} := \|\hat{q}\|_2$
6. *If* $r_{jj} = 0$ *then Stop, else* $q_j := \hat{q}/r_{jj}$
7. *EndDo*

It is easy to prove that the above algorithm will not break down; i.e., all r steps will be completed iff the set of vectors $x_1, x_2, \ldots, x_r$ is linearly independent. From lines 4 and 5, it is clear that at every step of the algorithm the following relation holds:

$$x_j = \sum_{i=1}^{j} r_{ij} q_i.$$

If $X = [x_1, x_2, \ldots, x_r]$, $Q = [q_1, q_2, \ldots, q_r]$, and R denotes the $r \times r$ upper triangular matrix whose nonzero elements are the r_{ij}'s defined in the algorithm, then the above relation can be written as

$$X = QR. \tag{1.18}$$

This is called the QR decomposition of the $n \times r$ matrix X. From what was said above, the QR decomposition of a matrix exists whenever the column vectors of X form a linearly independent set of vectors.

The above algorithm is the standard Gram–Schmidt process. There are alternative formulations of the algorithm that have better numerical properties. The best known of these is the modified Gram–Schmidt (MGS) algorithm.

ALGORITHM 1.2. MGS

1. *Define* $r_{11} := \|x_1\|_2$. *If* $r_{11} = 0$ *Stop, else* $q_1 := x_1/r_{11}$
2. *For* $j = 2, \ldots, r$, *Do*
3. *Define* $\hat{q} := x_j$
4. *For* $i = 1, \ldots, j-1$, *Do*
5. $r_{ij} := (\hat{q}, q_i)$
6. $\hat{q} := \hat{q} - r_{ij} q_i$
7. *EndDo*
8. *Compute* $r_{jj} := \|\hat{q}\|_2$
9. *If* $r_{jj} = 0$ *then Stop, else* $q_j := \hat{q}/r_{jj}$
10. *EndDo*

Yet another alternative for orthogonalizing a sequence of vectors is the Householder algorithm. This technique uses Householder *reflectors*, i.e., matrices of the form

$$P = I - 2ww^T, \tag{1.19}$$

in which w is a vector of 2-norm unity. Geometrically, the vector Px represents a mirror image of x with respect to the hyperplane $\text{span}\{w\}^\perp$.

To describe the Householder orthogonalization process, the problem can be formulated as that of finding a QR factorization of a given $n \times m$ matrix X. For any vector x, the vector w for the Householder transformation (1.19) is selected in such a way that

$$Px = \alpha e_1,$$

where α is a scalar. Writing $(I - 2ww^T)x = \alpha e_1$ yields

$$2w^T x \; w = x - \alpha e_1. \tag{1.20}$$

This shows that the desired w is a multiple of the vector $x - \alpha e_1$:

$$w = \pm \frac{x - \alpha e_1}{\|x - \alpha e_1\|_2}.$$

For (1.20) to be satisfied, we must impose the condition

$$2(x - \alpha e_1)^T x = \|x - \alpha e_1\|_2^2,$$

which gives $2(\|x\|_1^2 - \alpha\xi_1) = \|x\|_2^2 - 2\alpha\xi_1 + \alpha^2$, where $\xi_1 \equiv e_1^T x$ is the first component of the vector x. Therefore, it is necessary that

$$\alpha = \pm\|x\|_2.$$

In order to avoid the resulting vector w being small, it is customary to take

$$\alpha = -\text{sign}(\xi_1)\|x\|_2,$$

which yields

$$w = \frac{x + \text{sign}(\xi_1)\|x\|_2 e_1}{\|x + \text{sign}(\xi_1)\|x\|_2 e_1\|_2}. \tag{1.21}$$

Given an $n \times m$ matrix, its first column can be transformed to a multiple of the column e_1 by premultiplying it by a Householder matrix P_1:

$$X_1 \equiv P_1 X, \qquad X_1 e_1 = \alpha e_1.$$

Assume, inductively, that the matrix X has been transformed in $k-1$ successive steps into the partially upper triangular form

$$X_k \equiv P_{k-1} \cdots P_1 X_1 = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \cdots & \cdots & \cdots & x_{1m} \\ & x_{22} & x_{23} & \cdots & \cdots & \cdots & x_{2m} \\ & & x_{33} & \cdots & \cdots & \cdots & x_{3m} \\ & & & \ddots & \cdots & \cdots & \vdots \\ & & & & x_{kk} & \cdots & \vdots \\ & & & & x_{k+1,k} & \cdots & x_{k+1,m} \\ & & & & \vdots & \vdots & \vdots \\ & & & & x_{n,k} & \cdots & x_{n,m} \end{pmatrix}.$$

This matrix is upper triangular up to column number $k-1$. To advance by one step, it must be transformed into one that is upper triangular up to the kth column, leaving the previous columns in the same form. To leave the first $k-1$ columns unchanged, select a w-vector that has zeros in positions 1 through $k-1$. So the next Householder reflector matrix is defined as

$$P_k = I - 2w_k w_k^T, \tag{1.22}$$

in which the vector w_k is defined as

$$w_k = \frac{z}{\|z\|_2}, \tag{1.23}$$

where the components of the vector z are given by

$$z_i = \begin{cases} 0 & \text{if } i < k, \\ \beta + x_{ii} & \text{if } i = k, \\ x_{ik} & \text{if } i > k, \end{cases} \tag{1.24}$$

with

$$\beta = \text{sign}(x_{kk}) \times \left(\sum_{i=k}^{n} x_{ik}^2 \right)^{1/2}. \tag{1.25}$$

We note in passing that the premultiplication of a matrix X by a Householder transform requires only a rank-one update since

$$(I - 2ww^T)X = X - wv^T, \quad \text{where} \quad v = 2X^T w.$$

Therefore, the Householder matrices need not, and should not, be explicitly formed. In addition, the vectors w need not be explicitly scaled.

Assume now that $m-1$ Householder transforms have been applied to a certain matrix X of dimension $n \times m$ to reduce it into the upper triangular form

$$X_m \equiv P_{m-1} P_{m-2} \cdots P_1 X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1m} \\ & x_{22} & x_{23} & \cdots & x_{2m} \\ & & x_{33} & \cdots & x_{3m} \\ & & & \ddots & \vdots \\ & & & & x_{m,m} \\ & & & & 0 \\ & & & & \vdots \\ & & & & \vdots \end{pmatrix}. \tag{1.26}$$

Recall that our initial goal was to obtain a QR factorization of X. We now wish to recover the Q- and R-matrices from the P_k's and the above matrix. If we denote by P the product of the P_i on the left side of (1.26), then (1.26) becomes

$$PX = \begin{pmatrix} R \\ O \end{pmatrix}, \tag{1.27}$$

in which R is an $m \times m$ upper triangular matrix and O is an $(n-m) \times m$ zero block. Since P is unitary, its inverse is equal to its transpose and, as a result,

$$X = P^T \begin{pmatrix} R \\ O \end{pmatrix} = P_1 P_2 \cdots P_{m-1} \begin{pmatrix} R \\ O \end{pmatrix}.$$

If E_m is the matrix of size $n \times m$ that consists of the first m columns of the identity matrix, then the above equality translates into

$$X = P^T E_m R.$$

The matrix $Q = P^T E_m$ represents the first m columns of P^T. Since

$$Q^T Q = E_m^T P P^T E_m = I,$$

Q and R are the matrices sought. In summary,

$$X = QR,$$

in which R is the triangular matrix obtained from the Householder reduction of X (see (1.26) and (1.27)) and

$$Qe_j = P_1 P_2 \cdots P_{m-1} e_j.$$

ALGORITHM 1.3. Householder Orthogonalization

1. *Define* $X = [x_1, \ldots, x_m]$
2. *For* $k = 1, \ldots, m$, *Do*
3. *If* $k > 1$ *compute* $r_k := P_{k-1} P_{k-2} \cdots P_1 x_k$
4. *Compute* w_k *using* (1.23), (1.24), (1.25)
5. *Compute* $r_k := P_k r_k$ *with* $P_k = I - 2 w_k w_k^T$
6. *Compute* $q_k = P_1 P_2 \cdots P_k e_k$
7. *EndDo*

Note that line 6 can be omitted since the q_i's are not needed in the execution of the next steps. It must be executed only when the matrix Q is needed at the completion of the algorithm. Also, the operation in line 5 consists only of zeroing the components $k+1, \ldots, n$ and updating the kth component of r_k. In practice, a work vector can be used for r_k, and its nonzero components after this step can be saved into an upper triangular matrix. Since the components 1 through k of the vector w_k are zero, the upper triangular matrix R can be saved in those zero locations that would otherwise be unused.

1.8 Canonical Forms of Matrices

This section discusses the reduction of square matrices into matrices that have simpler forms, such as diagonal, bidiagonal, or triangular. Reduction means a transformation that preserves the eigenvalues of a matrix.

Definition 1.5. *Two matrices A and B are said to be similar if there is a nonsingular matrix X such that*

$$A = XBX^{-1}.$$

The mapping $B \rightarrow A$ is called a similarity transformation.

It is clear that *similarity* is an equivalence relation. Similarity transformations preserve the eigenvalues of matrices. An eigenvector u_B of B is transformed into the eigenvector $u_A = X u_B$ of A. In effect, a similarity transformation amounts to representing the matrix B in a different basis.

We now introduce some terminology.

1. An eigenvalue λ of A has *algebraic multiplicity* μ if it is a root of multiplicity μ of the characteristic polynomial.
2. If an eigenvalue is of algebraic multiplicity one, it is said to be *simple*. A nonsimple eigenvalue is *multiple*.

3. The *geometric multiplicity* γ of an eigenvalue λ of A is the maximum number of independent eigenvectors associated with it. In other words, the geometric multiplicity γ is the dimension of the eigenspace $\mathrm{Ker}(A - \lambda I)$.
4. A matrix is *derogatory* if the geometric multiplicity of at least one of its eigenvalues is larger than one.
5. An eigenvalue is *semisimple* if its algebraic multiplicity is equal to its geometric multiplicity. An eigenvalue that is not semisimple is called *defective*.

Often, $\lambda_1, \lambda_2, \ldots, \lambda_p$ $(p \leq n)$ are used to denote the *distinct* eigenvalues of A. It is easy to show that the characteristic polynomials of two similar matrices are identical; see Exercise 9. Therefore, the eigenvalues of two similar matrices are equal and so are their algebraic multiplicities. Moreover, if v is an eigenvector of B, then Xv is an eigenvector of A and, conversely, if y is an eigenvector of A, then $X^{-1}y$ is an eigenvector of B. As a result, the number of independent eigenvectors associated with a given eigenvalue is the same for two similar matrices; i.e., their geometric multiplicity is also the same.

1.8.1 Reduction to the Diagonal Form

The simplest form into which a matrix can be reduced is undoubtedly the diagonal form. Unfortunately, this reduction is not always possible. A matrix that can be reduced to the diagonal form is called *diagonalizable*. The following theorem characterizes such matrices.

Theorem 1.6. *A matrix of dimension n is diagonalizable iff it has n linearly independent eigenvectors.*

Proof. A matrix A is diagonalizable iff there exists a nonsingular matrix X and a diagonal matrix D such that $A = XDX^{-1}$ or, equivalently, $AX = XD$, where D is a diagonal matrix. This is equivalent to saying that n linearly independent vectors exist—the n column vectors of X—such that $Ax_i = d_i x_i$. Each of these column vectors is an eigenvector of A. □

A matrix that is diagonalizable has only semisimple eigenvalues. Conversely, if all the eigenvalues of a matrix A are semisimple, then A has n eigenvectors. It can be easily shown that these eigenvectors are linearly independent; see Exercise 2. As a result, we have the following proposition.

Proposition 1.7. *A matrix is diagonalizable iff all its eigenvalues are semisimple.*

Since every simple eigenvalue is semisimple, an immediate corollary of the above result is as follows: When A has n distinct eigenvalues, then it is diagonalizable.

1.8.2 The Jordan Canonical Form

From the theoretical viewpoint, one of the most important canonical forms of matrices is the well-known Jordan form. A full development of the steps leading to the Jordan form is beyond the scope of this book. Only the main theorem is stated. Details, including the

proof, can be found in standard books of linear algebra such as [164]. In the following, m_i refers to the algebraic multiplicity of the individual eigenvalue λ_i and l_i is the *index* of the eigenvalue, i.e., the smallest integer for which $\mathrm{Ker}(A-\lambda_i I)^{l_i+1} = \mathrm{Ker}(A-\lambda_i I)^{l_i}$.

Theorem 1.8. *Any matrix A can be reduced to a block diagonal matrix consisting of p diagonal blocks, each associated with a distinct eigenvalue λ_i. Each of these diagonal blocks has itself a block diagonal structure consisting of γ_i sub-blocks, where γ_i is the geometric multiplicity of the eigenvalue λ_i. Each of the sub-blocks, referred to as a Jordan block, is an upper bidiagonal matrix of size not exceeding $l_i \leq m_i$, with the constant λ_i on the diagonal and the constant one on the superdiagonal.*

The ith diagonal block, $i = 1, \ldots, p$, is known as the ith Jordan submatrix (sometimes "Jordan box"). The Jordan submatrix number i starts in column $j_i \equiv m_1 + m_2 + \cdots + m_{i-1} + 1$. Thus,

$$X^{-1}AX = J = \begin{pmatrix} J_1 & & & & \\ & J_2 & & & \\ & & \ddots & & \\ & & & J_i & \\ & & & & \ddots \\ & & & & & J_p \end{pmatrix},$$

where each J_i is associated with λ_i and is of size m_i, the algebraic multiplicity of λ_i. It has itself the following structure:

$$J_i = \begin{pmatrix} J_{i1} & & & \\ & J_{i2} & & \\ & & \ddots & \\ & & & J_{i\gamma_i} \end{pmatrix}, \text{ with } J_{ik} = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{pmatrix}.$$

Each of the blocks J_{ik} corresponds to a different eigenvector associated with the eigenvalue λ_i. Its size l_i is the index of λ_i.

1.8.3 The Schur Canonical Form

Here, it will be shown that any matrix is unitarily similar to an upper triangular matrix. The only result needed to prove the following theorem is that any vector having a 2-norm can be completed by $n-1$ additional vectors to form an orthonormal basis of $\mathbb{C}^n$.

Theorem 1.9. *For any square matrix A, there exists a unitary matrix Q such that*

$$Q^H A Q = R$$

is upper triangular.

Proof. The proof is by induction over the dimension n. The result is trivial for $n = 1$. Assume that it is true for $n-1$ and consider any matrix A of size n. The matrix admits

at least one eigenvector u that is associated with an eigenvalue λ. Also assume without loss of generality that $\|u\|_2 = 1$. First, complete the set consisting of the vector u into an orthonormal set; i.e., find an $n \times (n-1)$ matrix V such that the $n \times n$ matrix $U = [u, V]$ is unitary. Then $AU = [\lambda u, AV]$ and, hence,

$$U^H A U = \begin{bmatrix} u^H \\ V^H \end{bmatrix} [\lambda u, AV] = \begin{pmatrix} \lambda & u^H A V \\ 0 & V^H A V \end{pmatrix}. \tag{1.28}$$

Now use the induction hypothesis for the $(n-1) \times (n-1)$ matrix $B = V^H A V$: There exists an $(n-1) \times (n-1)$ unitary matrix Q_1 such that $Q_1^H B Q_1 = R_1$ is upper triangular. Define the $n \times n$ matrix

$$\hat{Q}_1 = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$$

and multiply both members of (1.28) by $\hat{Q}_1^H$ from the left and $\hat{Q}_1$ from the right. The resulting matrix is clearly upper triangular, which shows that the result is true for A, with $Q = \hat{Q}_1 U$, which is a unitary $n \times n$ matrix. □

A simpler proof that uses the Jordan canonical form and the QR decomposition is the subject of Exercise 7. Since the matrix R is triangular and similar to A, its diagonal elements are equal to the eigenvalues of A ordered in a certain manner. In fact, it is easy to extend the proof of the theorem to show that this factorization can be obtained with *any order* of the eigenvalues. Despite its simplicity, the above theorem has far-reaching consequences, some of which will be examined in the next section.

It is important to note that, for any $k \leq n$, the subspace spanned by the first k columns of Q is invariant under A. Indeed, the relation $AQ = QR$ implies that, for $1 \leq j \leq k$, we have

$$Aq_j = \sum_{i=1}^{i=j} r_{ij} q_i.$$

If we let $Q_k = [q_1, q_2, \ldots, q_k]$ and if R_k is the principal leading submatrix of dimension k of R, the above relation can be rewritten as

$$AQ_k = Q_k R_k,$$

which is known as the partial Schur decomposition of A. The simplest case of this decomposition is when $k = 1$, in which case q_1 is an eigenvector. The vectors q_i are usually called Schur vectors. Schur vectors are not unique and depend, in particular, on the order chosen for the eigenvalues.

A slight variation of the Schur canonical form is the quasi-Schur form, also called the real Schur form. Here diagonal blocks of size 2×2 are allowed in the upper triangular matrix R. The reason for this is to avoid complex arithmetic when the original matrix is real. A 2×2 block is associated with each complex conjugate pair of eigenvalues of the matrix.

Example 1.2. Consider the 3×3 matrix

$$A = \begin{pmatrix} 1 & 10 & 0 \\ -1 & 3 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

The matrix A has the pair of complex conjugate eigenvalues

$$2.4069\ldots \pm i \times 3.2110\ldots$$

and the real eigenvalue $0.1863\ldots$. The standard (complex) Schur form is given by the pair of matrices

$$V = \begin{pmatrix} 0.3381 - 0.8462i & 0.3572 - 0.1071i & 0.1749 \\ 0.3193 - 0.0105i & -0.2263 - 0.6786i & -0.6214 \\ 0.1824 + 0.1852i & -0.2659 - 0.5277i & 0.7637 \end{pmatrix}$$

and

$$S = \begin{pmatrix} 2.4069 + 3.2110i & 4.6073 - 4.7030i & -2.3418 - 5.2330i \\ 0 & 2.4069 - 3.2110i & -2.0251 - 1.2016i \\ 0 & 0 & 0.1863 \end{pmatrix}.$$

It is possible to avoid complex arithmetic by using the quasi-Schur form, which consists of the pair of matrices

$$U = \begin{pmatrix} -0.9768 & 0.1236 & 0.1749 \\ -0.0121 & 0.7834 & -0.6214 \\ 0.2138 & 0.6091 & 0.7637 \end{pmatrix}$$

and

$$R = \begin{pmatrix} 1.3129 & -7.7033 & 6.0407 \\ 1.4938 & 3.5008 & -1.3870 \\ 0 & 0 & 0.1863 \end{pmatrix}.$$

We conclude this section by pointing out that the Schur and the quasi-Schur forms of a given matrix are in no way unique. In addition to the dependence on the ordering of the eigenvalues, any column of Q can be multiplied by a complex sign $e^{i\theta}$ and a new corresponding R can be found. For the quasi-Schur form, there are infinitely many ways to select the 2×2 blocks, corresponding to applying arbitrary rotations to the columns of Q associated with these blocks.

1.8.4 Application to Powers of Matrices

The analysis of many numerical techniques is based on understanding the behavior of the successive powers A^k of a given matrix A. In this regard, the following theorem plays a fundamental role in numerical linear algebra, more particularly in the analysis of iterative methods.

Theorem 1.10. *The sequence A^k, $k = 0, 1, \ldots$, converges to zero iff $\rho(A) < 1$.*

Proof. To prove the necessary condition, assume that $A^k \to 0$ and consider u_1 a unit eigenvector associated with an eigenvalue λ_1 of maximum modulus. We have

$$A^k u_1 = \lambda_1^k u_1,$$

which implies, by taking the 2-norms of both sides,

$$|\lambda_1^k| = \|A^k u_1\|_2 \to 0.$$

This shows that $\rho(A) = |\lambda_1| < 1$.

The Jordan canonical form must be used to show the sufficient condition. Assume that $\rho(A) < 1$. Start with the equality

$$A^k = X J^k X^{-1}.$$

To prove that A^k converges to zero, it is sufficient to show that J^k converges to zero. An important observation is that J^k preserves its block form. Therefore, it is sufficient to prove that each of the Jordan blocks converges to zero. Each block is of the form

$$J_i = \lambda_i I + E_i,$$

where E_i is a nilpotent matrix of index l_i; i.e., $E_i^{l_i} = 0$. Therefore, for $k \geq l_i$,

$$J_i^k = \sum_{j=0}^{l_i - 1} \frac{k!}{j!(k-j)!} \lambda_i^{k-j} E_i^j.$$

Using the triangle inequality for any norm and taking $k \geq l_i$ yields

$$\|J_i^k\| \leq \sum_{j=0}^{l_i - 1} \frac{k!}{j!(k-j)!} |\lambda_i|^{k-j} \|E_i^j\|.$$

Since $|\lambda_i| < 1$, each of the terms in this *finite* sum converges to zero as $k \to \infty$. Therefore, the matrix J_i^k converges to zero. □

An equally important result is stated in the following theorem.

Theorem 1.11. *The series*

$$\sum_{k=0}^{\infty} A^k$$

converges iff $\rho(A) < 1$. Under this condition, $I - A$ is nonsingular and the limit of the series is equal to $(I - A)^{-1}$.

Proof. The first part of the theorem is an immediate consequence of Theorem 1.10. Indeed, if the series converges, then $\|A^k\| \to 0$. By the previous theorem, this implies that $\rho(A) < 1$. To show that the converse is also true, use the equality

$$I - A^{k+1} = (I - A)(I + A + A^2 + \cdots + A^k)$$

and exploit the fact that, since $\rho(A) < 1$, then $I - A$ is nonsingular and, therefore,

$$(I - A)^{-1}(I - A^{k+1}) = I + A + A^2 + \cdots + A^k.$$

This shows that the series converges since the left-hand side will converge to $(I - A)^{-1}$. In addition, it shows the second part of the theorem. □

Another important consequence of the Jordan canonical form is a result that relates the spectral radius of a matrix to its matrix norm.

Theorem 1.12. *For any matrix norm $\|\cdot\|$, we have*

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

Proof. The proof is a direct application of the Jordan canonical form and is the subject of Exercise 10. □