Chapter 2
Linear Algebra

The purpose of this chapter is to briefly review the notations, definitions, and results of linear algebra that are used in the rest of the book. There is no claim of completeness as the reader is assumed to be familiar with the basic concepts of college linear algebra such as vector spaces, linear mappings, matrix decompositions, etc. More systematic exposition and additional material can be found in the books by Demmel [1], Laub [2], or Golub and Van Loan [3].

2.1 Vectors and Matrices

In this book, we work with *n-dimensional arithmetic vectors* $\mathbf{v} \in \mathbb{R}^n$, where $\mathbb{R}$ denotes the set of real numbers. The only exception is Sect. 2.7, where vectors with complex entries are considered. We denote the ith component of an arithmetic vector $\mathbf{v} \in \mathbb{R}^n$ by $[\mathbf{v}]_i$. Thus $[\mathbf{v}]_i = v_i$ if $\mathbf{v} = [v_i]$ is defined by its components v_i. All the arithmetic vectors are considered by default to be column vectors. The relations between vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are defined componentwise. Thus $\mathbf{u} \leq \mathbf{v}$ is equivalent to $[\mathbf{u}]_i \leq [\mathbf{v}]_i$, $i = 1, \ldots, n$. We sometimes call the elements of $\mathbb{R}^n$ *points* to indicate that the concepts of length and direction are not important.

The vector analog of $0 \in \mathbb{R}$ is the *zero vector* $\mathbf{o}_n \in \mathbb{R}^n$ with all the entries equal to zero. When the dimension can be deduced from the context, possibly using the assumption that all the expressions in our book are well defined, we often drop the subscript and write simply $\mathbf{o}$.

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, the set

$$\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k, \ \alpha_i \in \mathbb{R}\}$$

is a vector space called the *linear span* of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. For example

$$\mathrm{Span}\{\mathbf{s}_1, \ldots, \mathbf{s}_n\} = \mathbb{R}^n, \quad [\mathbf{s}_i]_j = \delta_{ij}, \quad i, j = 1, \ldots, n,$$

Z. Dostál et al., *Scalable Algorithms for Contact Problems*, Advances in Mechanics and Mathematics 36, DOI 10.1007/978-1-4939-6834-3_2

where δ_{ij} denotes the *Kronecker symbol* defined by $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$, is the *standard basis* of $\mathbb{R}^n$.

We sometimes use the componentwise extensions of scalar functions to vectors. Thus, if $\mathbf{v} \in \mathbb{R}^n$, then $\mathbf{v}^+$ and $\mathbf{v}^-$ are the vectors the ith components of which are $\max\{[\mathbf{v}]_i, 0\}$ and $\min\{[\mathbf{v}]_i, 0\}$, respectively.

If $\mathcal{I}$ is a nonempty subset of $\{1, \ldots, n\}$ and $\mathbf{v} \in \mathbb{R}^n$, then we denote by $[\mathbf{v}]_\mathcal{I}$ or simply $\mathbf{v}_\mathcal{I}$ the subvector of $\mathbf{v}$ with components $[\mathbf{v}]_i$, $i \in \mathcal{I}$. Thus if $\mathcal{I}$ has m elements, then $\mathbf{v}_\mathcal{I} \in \mathbb{R}^m$, so we can refer to the components of $\mathbf{v}_\mathcal{I}$ either by the *global indices* $i \in \mathcal{I}$ or by the *local indices* $j \in \{1, \ldots, m\}$. We usually rely on the reader's judgment to recognize the appropriate type of indexing.

Similarly to the related convention for vectors, the (i, j)th component of a matrix $\mathsf{A} \in \mathbb{R}^{m \times n}$ is denoted by $[\mathsf{A}]_{ij}$, so that $[\mathsf{A}]_{ij} = a_{ij}$ for $\mathsf{A} = [a_{ij}]$ which is defined by its entries a_{ij}. A matrix $\mathsf{A} \in \mathbb{R}^{m \times n}$ is called an (m, n)-*matrix*.

The matrix analog of 0 is the *zero matrix* $\mathsf{O}_{mn} \in \mathbb{R}^{m \times n}$ with all the entries equal to zero. When the dimension is clear from the context, we often drop the subscripts and write simply O.

The matrix counterpart of $1 \in \mathbb{R}$ in $\mathbb{R}^{n \times n}$ is the *identity matrix* $\mathsf{I}_n = [\delta_{ij}]$ of the order n. When the dimension may be deduced from the context, we often drop the subscripts and write simply I. Thus, we can write

$$\mathsf{A} = \mathsf{I}\mathsf{A} = \mathsf{A}\mathsf{I}$$

for any matrix A, having in mind that the order of I on the left may be different from that on the right.

A matrix A is *positive definite* if $\mathbf{x}^T \mathsf{A}\mathbf{x} > 0$ for any $\mathbf{x} \neq \mathbf{o}$, *positive semidefinite* if $\mathbf{x}^T \mathsf{A}\mathbf{x} \geq 0$ for any $\mathbf{x}$, and *indefinite* if neither A nor $-\mathsf{A}$ is positive definite or semidefinite. We are especially interested in *symmetric positive definite (SPD)* or *symmetric positive semidefinite (SPS)* matrices.

If $\mathsf{A} \in \mathbb{R}^{m \times n}$, $\mathcal{I} \subseteq \{1, \ldots, m\}$, and $\mathcal{J} \subseteq \{1, \ldots, n\}$, $\mathcal{I}$ and $\mathcal{J}$ nonempty, we denote by $\mathsf{A}_{\mathcal{I}\mathcal{J}}$ the submatrix of A with the components $[\mathsf{A}]_{ij}$, $i \in \mathcal{I}$, $j \in \mathcal{J}$. The local indexing of the entries of $\mathsf{A}_{\mathcal{I}\mathcal{J}}$ is used whenever it is convenient in a similar way as the local indexing of subvectors which was introduced in Sect. 2.1. The full set of indices may be replaced by * so that $\mathsf{A} = \mathsf{A}_{**}$ and $\mathsf{A}_{\mathcal{I}*}$ denotes the submatrix of A with the row indices belonging to $\mathcal{I}$. Occasionally we simplify $\mathsf{A}_\mathcal{I} = \mathsf{A}_{\mathcal{I}*}$.

Sometimes it is useful to rearrange the matrix operations into manipulations with submatrices of given matrices called blocks. A *block matrix* $\mathsf{A} \in \mathbb{R}^{m \times n}$ is defined by its blocks $\mathsf{A}_{ij} = \mathsf{A}_{\mathcal{I}_i \mathcal{J}_j}$, where $\mathcal{I}_i$ and $\mathcal{J}_j$ denote nonempty contiguous sets of indices decomposing $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$, respectively. We can use the block structure to implement matrix operations only when the block structure of the involved matrices matches.

The matrices in our applications are often *sparse* in the sense that they have a small number of nonzero entries distributed in a pattern which can be exploited to the efficient implementation of matrix operations or to the reduction of storage requirements.

2.2 Matrices and Mappings

Each matrix $\mathsf{A} \in \mathbb{R}^{m\times n}$ defines the mapping which assigns to each $\mathbf{x} \in \mathbb{R}^n$ the vector $\mathsf{A}\mathbf{x} \in \mathbb{R}^m$. Two important subspaces associated with this mapping are its *range* or *image space* $\mathrm{Im}\mathsf{A}$ and its *kernel* or *null space* $\mathrm{Ker}\mathsf{A}$; they are defined by

$$\mathrm{Im}\mathsf{A} = \{\mathsf{A}\mathbf{x} : \ \mathbf{x} \in \mathbb{R}^n\} \quad \text{and} \quad \mathrm{Ker}\mathsf{A} = \{\mathbf{x} \in \mathbb{R}^n : \ \mathsf{A}\mathbf{x} = \mathbf{o}\}.$$

The range of A is the span of its columns. The *rank* and the *defect* of a matrix are defined as the dimension of its image and kernel, respectively.

If f is a mapping defined on $\mathscr{D} \subseteq \mathbb{R}^n$ and $\Omega \subseteq \mathscr{D}$, then $f|\Omega$ denotes the *restriction* of f to Ω, that is, the mapping defined on Ω which assigns to each $\mathbf{x} \in \Omega$ the value $f(\mathbf{x})$. If $\mathsf{A} \in \mathbb{R}^{m\times n}$ and V is a subspace of $\mathbb{R}^n$, we define $\mathsf{A}|V$ as a restriction of the mapping associated with A to V. The restriction $\mathsf{A}|V$ is said to be positive definite if $\mathbf{x}^T\mathsf{A}\mathbf{x} > 0$ for $\mathbf{x} \in V$, $\mathbf{x} \neq \mathbf{o}$, and positive semidefinite if $\mathbf{x}^T\mathsf{A}\mathbf{x} \geq 0$ for $\mathbf{x} \in V$.

The mapping associated with A is *injective* if $\mathsf{A}\mathbf{x} = \mathsf{A}\mathbf{y}$ implies $\mathbf{x} = \mathbf{y}$. It is easy to check that the mapping associated with A is injective if and only if $\mathrm{Ker}\mathsf{A} = \{\mathbf{o}\}$. If $m = n$, then A is injective if and only if $\mathrm{Im}\mathsf{A} = \mathbb{R}^n$.

A subspace $V \subseteq \mathbb{R}^n$ which satisfies

$$\mathsf{A}V = \{\mathsf{A}\mathbf{x} : \ \mathbf{x} \in V\} \subseteq V$$

is an *invariant subspace* of A. Obviously

$$\mathsf{A}(\mathrm{Im}\mathsf{A}) \subseteq \mathrm{Im}\mathsf{A},$$

so that $\mathrm{Im}\mathsf{A}$ is an invariant subspace of A.

A *projector* is a square matrix P that satisfies

$$\mathsf{P}^2 = \mathsf{P}.$$

A vector $\mathbf{x} \in \mathrm{Im}\mathsf{P}$ if and only if there is $\mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{x} = \mathsf{P}\mathbf{y}$, so that

$$\mathsf{P}\mathbf{x} = \mathsf{P}(\mathsf{P}\mathbf{y}) = \mathsf{P}\mathbf{y} = \mathbf{x}.$$

If P is a projector, then also $\mathsf{Q} = \mathsf{I} - \mathsf{P}$ and P^T are projectors as

$$(\mathsf{I} - \mathsf{P})^2 = \mathsf{I} - 2\mathsf{P} + \mathsf{P}^2 = \mathsf{I} - \mathsf{P} \quad \text{and} \quad \left(\mathsf{P}^T\right)^2 = \left(\mathsf{P}^2\right)^T = \mathsf{P}^T.$$

Since for any $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x} = \mathsf{P}\mathbf{x} + (\mathsf{I} - \mathsf{P})\mathbf{x},$$

it simply follows that $\mathrm{Im}\mathsf{Q} = \mathrm{Ker}\mathsf{P}$,

$$\mathbb{R}^n = \mathrm{Im}\mathsf{P} + \mathrm{Ker}\mathsf{P}, \quad \text{and} \quad \mathrm{Ker}\mathsf{P} \cap \mathrm{Im}\mathsf{P} = \{\mathbf{o}\}.$$

We say that P is a projector onto $U = \mathrm{Im}\mathsf{P}$ along $V = \mathrm{Ker}\mathsf{P}$ and Q is a complementary projector onto V along U. The above relations may be rewritten as

$$\mathrm{Im}\mathsf{P} \oplus \mathrm{Ker}\mathsf{P} = \mathbb{R}^n. \tag{2.1}$$

Let $(\pi(1), \ldots, \pi(n))$ be a permutation of numbers $1, \ldots, n$. Then, the mapping which assigns to each $\mathbf{v} = [v_i] \in \mathbb{R}^n$ a vector $[v_{\pi(1)}, \ldots, v_{\pi(n)}]^T$ is associated with the *permutation matrix*

$$\mathsf{P} = [\mathbf{s}_{\pi(1)}, \ldots, \mathbf{s}_{\pi(n)}],$$

where $\mathbf{s}_i$ denotes the ith column of the identity matrix I_n. If P is a permutation matrix, then

$$\mathsf{P}\mathsf{P}^T = \mathsf{P}^T\mathsf{P} = \mathsf{I}.$$

Notice that if B is a matrix obtained from a matrix A by reordering of the rows of A, then there is a permutation matrix P such that $\mathsf{B} = \mathsf{P}\mathsf{A}$. Similarly, if B is a matrix obtained from A by reordering of the columns of A, then there is a permutation matrix P such that $\mathsf{B} = \mathsf{A}\mathsf{P}$.

2.3 Inverse and Generalized Inverse

If A is a square full rank matrix, then there is the unique *inverse matrix* A^{-1} such that

$$\mathsf{A}\mathsf{A}^{-1} = \mathsf{A}^{-1}\mathsf{A} = \mathsf{I}. \tag{2.2}$$

The mapping associated with A^{-1} is inverse to that associated with A.

If A^{-1} exists, we say that A is *nonsingular*. A square matrix is *singular* if its inverse matrix does not exist. If P is a permutation matrix, then P is nonsingular and

$$\mathsf{P}^{-1} = \mathsf{P}^T.$$

If A is a nonsingular matrix, then $\mathsf{A}^{-1}\mathbf{b}$ is the unique solution of $\mathsf{A}\mathbf{x} = \mathbf{b}$.

If A is nonsingular, then we can transpose (2.2) to get

$$(\mathsf{A}^{-1})^T\mathsf{A}^T = \mathsf{A}^T(\mathsf{A}^{-1})^T = \mathsf{I},$$

so that

$$(\mathsf{A}^T)^{-1} = (\mathsf{A}^{-1})^T. \tag{2.3}$$

It follows that if A is symmetric, then A^{-1} is symmetric.

If $\mathsf{A} \in \mathbb{R}^{n\times n}$ is positive definite, then also A^{-1} is positive definite, as any vector $\mathbf{x} \neq \mathbf{o}$ can be expressed as $\mathbf{x} = \mathsf{A}\mathbf{y}$, $\mathbf{y} \neq \mathbf{o}$, and

$$\mathbf{x}^T\mathsf{A}^{-1}\mathbf{x} = (\mathsf{A}\mathbf{y})^T\mathsf{A}^{-1}\mathsf{A}\mathbf{y} = \mathbf{y}^T\mathsf{A}^T\mathbf{y} = \mathbf{y}^T\mathsf{A}\mathbf{y} > 0.$$

If A and B are nonsingular matrices, then it is easy to check that also AB is nonsingular and

$$(\mathsf{AB})^{-1} = \mathsf{B}^{-1}\mathsf{A}^{-1}.$$

If

$$\mathsf{H} = \begin{bmatrix} \mathsf{H}_{\mathscr{I}\mathscr{I}} & \mathsf{H}_{\mathscr{I}\mathscr{J}} \\ \mathsf{H}_{\mathscr{J}\mathscr{I}} & \mathsf{H}_{\mathscr{J}\mathscr{J}} \end{bmatrix} = \begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & \mathsf{C} \end{bmatrix}$$

is an SPD block matrix, then we can directly evaluate

$$\mathsf{H}^{-1} = \begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & \mathsf{C} \end{bmatrix}^{-1} = \begin{bmatrix} \mathsf{A}^{-1} - \mathsf{A}^{-1}\mathsf{B}^T\mathsf{S}^{-1}\mathsf{B}\mathsf{A}^{-1}, & -\mathsf{A}^{-1}\mathsf{B}^T\mathsf{S}^{-1} \\ -\mathsf{S}^{-1}\mathsf{B}\mathsf{A}^{-1}, & \mathsf{S}^{-1} \end{bmatrix}, \tag{2.4}$$

where $\mathsf{S} = \mathsf{C} - \mathsf{B}\mathsf{A}^{-1}\mathsf{B}^T$ denotes the *Schur complement* of H with respect to A. Thus

$$\left[\mathsf{A}^{-1}\right]_{\mathscr{J}\mathscr{J}} = \mathsf{S}^{-1}. \tag{2.5}$$

If $\mathsf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathrm{Im}\mathsf{A}$, then we can express a solution of the system of linear equations $\mathsf{A}\mathbf{x} = \mathbf{b}$ by means of a *left generalized inverse matrix* $\mathsf{A}^+ \in \mathbb{R}^{n\times m}$ which satisfies $\mathsf{A}\mathsf{A}^+\mathsf{A} = \mathsf{A}$. Indeed, if $\mathbf{b} \in \mathrm{Im}\mathsf{A}$, then there is $\mathbf{y}$ such that $\mathbf{b} = \mathsf{A}\mathbf{y}$ and $\overline{\mathbf{x}} = \mathsf{A}^+\mathbf{b}$ satisfies

$$\mathsf{A}\overline{\mathbf{x}} = \mathsf{A}\mathsf{A}^+\mathbf{b} = \mathsf{A}\mathsf{A}^+\mathsf{A}\mathbf{y} = \mathsf{A}\mathbf{y} = \mathbf{b}.$$

Thus A^+ acts on the range of A like the inverse matrix. If A is a nonsingular square matrix, then obviously

$$\mathsf{A}^+ = \mathsf{A}^{-1}.$$

Moreover, if $\mathsf{A} \in \mathbb{R}^{n\times n}$ and $\mathsf{S} \in \mathbb{R}^{n\times p}$ are such that $\mathsf{A}\mathsf{S} = \mathsf{O}$, then $(\mathsf{A}^+) + \mathsf{S}\mathsf{S}^T$ is also a left generalized inverse as

$$\mathsf{A}\left(\mathsf{A}^+ + \mathsf{S}\mathsf{S}^T\right)\mathsf{A} = \mathsf{A}\mathsf{A}^+\mathsf{A} + \mathsf{A}\mathsf{S}\mathsf{S}^T\mathsf{A} = \mathsf{A}.$$

If A is a symmetric singular matrix, then there is a permutation matrix P such that

$$\mathsf{A} = \mathsf{P}^T\begin{bmatrix} \mathsf{B} & \mathsf{C}^T \\ \mathsf{C} & \mathsf{C}\mathsf{B}^{-1}\mathsf{C}^T \end{bmatrix}\mathsf{P},$$

where B is a nonsingular matrix the dimension of which is equal to the rank of A. It may be verified directly that the matrix

$$\mathsf{A}^{\#} = \mathsf{P}^T\begin{bmatrix} \mathsf{B}^{-1} & \mathsf{O}^T \\ \mathsf{O} & \mathsf{O} \end{bmatrix}\mathsf{P} \tag{2.6}$$

is a left generalized inverse of A. If A is SPS, then $\mathsf{A}^{\#}$ is also SPS. Notice that if $\mathsf{A}\mathsf{S} = \mathsf{O}$, then $\mathsf{A}^+ = \mathsf{A}^{\#} + \mathsf{S}\mathsf{S}^T$ is also an SPS generalized inverse.

2.4 Direct Methods for Solving Linear Equations

The inverse matrix is a useful tool for theoretical developments, but not for computations. It is often much more efficient to implement the multiplication of a vector by the inverse matrix by solving the related system of linear equations. We recall here briefly the *direct methods*, which reduce solving of the original system of linear equations to solving of a system or systems of equations with triangular matrices.

A matrix $\mathsf{L} = [l_{ij}]$ is *lower triangular* if $l_{ij} = 0$ for $i < j$. It is easy to solve a system $\mathsf{L}\mathbf{x} = \mathbf{b}$ with the nonsingular lower triangular matrix $\mathsf{L} \in \mathbb{R}^n$. As there is only one unknown in the first equation, we can find it and then substitute it into the remaining equations to obtain a system with the same structure, but with only $n - 1$ remaining unknowns. Repeating the procedure, we can find all the components of $\mathbf{x}$.

A similar procedure, but starting from the last equation, can be applied to a system with the nonsingular *upper triangular matrix* $\mathsf{U} = [u_{ij}]$ with $u_{ij} = 0$ for $i > j$.

The solution costs of a system with triangular matrices is proportional to the number of its nonzero entries. In particular, the solution of a system of linear equations with a *diagonal matrix* $\mathsf{D} = [d_{ij}]$, $d_{ij} = 0$ for $i \neq j$, reduces to the solution of a sequence of linear equations with one unknown.

If we are to solve the system of linear equations with a nonsingular matrix, we can use systematically *equivalent transformations* that do not change the solution in order to modify the original system to that with an upper triangular matrix. It is well known that the solutions of a system of linear equations are the same as the solutions of a system of linear equations obtained from the original system by interchanging two equations, replacing an equation by its nonzero multiple, or adding a multiple of one equation to another equation. The *Gauss elimination* for the solution of a system of linear equations with a nonsingular matrix thus consists of two steps: the *forward reduction*, which exploits equivalent transformations to reduce the original system to the system with an upper triangular matrix, and the *backward substitution*, which solves the resulting system with the upper triangular matrix.

Alternatively, we can use suitable matrix factorizations. For example, it is well known that any SPD matrix A can be decomposed into the product

$$\mathsf{A} = \mathsf{L}\mathsf{L}^T, \tag{2.7}$$

where L is a nonsingular lower triangular matrix with positive diagonal entries. Having the decomposition, we can evaluate $\mathbf{z} = \mathsf{A}^{-1}\mathbf{x}$ by solving the systems

$$\mathsf{L}\mathbf{y} = \mathbf{x} \quad \text{and} \quad \mathsf{L}^T\mathbf{z} = \mathbf{y}.$$

The factorization-based solvers may be especially useful when we are to solve several systems of equations with the same coefficients but different right-hand sides.

The method of evaluation of the factor L is known as the *Cholesky factorization*. The Cholesky factor L can be computed column by column. Suppose that

$$\mathsf{A} = \begin{bmatrix} a_{11} & \mathbf{a}_1^T \\ \mathbf{a}_1 & \mathsf{A}_{22} \end{bmatrix} \quad \text{and} \quad \mathsf{L} = \begin{bmatrix} l_{11} & \mathbf{o} \\ \mathbf{l}_1 & \mathsf{L}_{22} \end{bmatrix}.$$

Substituting for A and L into (2.7) and comparing the corresponding terms immediately reveals that

$$l_{11} = \sqrt{a_{11}}, \qquad \mathbf{l}_1 = l_{11}^{-1}\mathbf{a}_1, \qquad \mathsf{L}_{22}\mathsf{L}_{22}^T = \mathsf{A}_{22} - \mathbf{l}_1\mathbf{l}_1^T. \tag{2.8}$$

This gives us the first column of L, and the remaining factor L_{22} is simply the Cholesky factor of the Schur complement $\mathsf{A}_{22} - \mathbf{l}_1\mathbf{l}_1^T$ which is known to be positive definite, so we can find its first column by the above procedure. The algorithm can be implemented to exploit a sparsity pattern of A, e.g., when $\mathsf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ is a *band matrix* with $a_{ij} = 0$ for $|i - j| > b$, $b \ll n$.

If $\mathsf{A} \in \mathbb{R}^{n\times n}$ is only positive semidefinite, it can happen that $a_{11} = 0$. Then

$$0 \le \mathbf{x}^T\mathsf{A}\mathbf{x} = \mathbf{y}^T\mathsf{A}_{22}\mathbf{y} + 2x_1\mathbf{a}_1^T\mathbf{y}$$

for any vector $\mathbf{x} = \left[x_1, \mathbf{y}^T\right]^T$. The inequality implies that $\mathbf{a}_1 = \mathbf{o}$, as otherwise we could take $\mathbf{y} = -\mathbf{a}_1$ and large x_1 to get

$$\mathbf{y}^T\mathsf{A}_{22}\mathbf{y} + 2x_1\mathbf{a}_1^T\mathbf{y} = \mathbf{a}_1^T\mathsf{A}_{22}\mathbf{a}_1 - 2x_1\|\mathbf{a}_1\|^2 < 0.$$

Thus for A symmetric positive semidefinite and $a_{11} = 0$, (2.8) reduces to

$$l_{11} = 0, \qquad \mathbf{l}_1 = \mathbf{o}, \qquad \mathsf{L}_{22}\mathsf{L}_{22}^T = \mathsf{A}_{22}. \tag{2.9}$$

This simple modification assumes exact arithmetics. In the computer arithmetics, the decision whether a_{11} is to be treated as zero depends on some small $\varepsilon > 0$. Alternatively, it is possible to exploit some additional information. For example, any orthonormal basis of the kernel of a matrix can be used to identify the zero rows (and columns) of a Cholesky factor by means of the following lemma.

Lemma 2.1 *Let $\mathsf{A} \in \mathbb{R}^{n\times n}$ denote an SPS matrix the kernel of which is spanned by the full column rank matrix $\mathsf{R} \in \mathbb{R}^{n\times d}$ with orthonormal columns. Let*

$$\mathscr{I} = \{i_1, \ldots, i_d\}, \quad 1 \le i_1 < i_2 < \cdots < i_d \le n,$$

denote a set of indices, and let $\mathscr{J} = \mathscr{N} - \mathscr{I}$, $\mathscr{N} = \{1, 2, \ldots, n\}$.

Then

$$\lambda_{\min}(\mathsf{A}_{\mathscr{J}\mathscr{J}}) \ge \overline{\lambda}_{\min}(\mathsf{A})\,\sigma_{\min}^4(\mathsf{R}_{\mathscr{J}*}), \tag{2.10}$$

where $\overline{\lambda}_{\min}(\mathsf{A})$ and $\sigma_{\min}(\mathsf{R}_{\mathscr{J}*})$ denote the least nonzero eigenvalue of A and the least singular value of $\mathsf{R}_{\mathscr{J}*}$, respectively.

Proof See Dostál et al. [4]. □

If we can identify a nonsingular block $\mathsf{A}_{\mathcal{J}\mathcal{J}}$ of A, then we can reduce the problems related to the manipulation with singular matrices to those with smaller matrices using the decomposition

$$\mathsf{A} = \begin{bmatrix} \mathsf{A}_{\mathcal{J}\mathcal{J}} & \mathsf{A}_{\mathcal{J}\mathcal{I}} \\ \mathsf{A}_{\mathcal{I}\mathcal{J}} & \mathsf{A}_{\mathcal{I}\mathcal{I}} \end{bmatrix} = \begin{bmatrix} \mathsf{L}_{\mathcal{J}\mathcal{J}} & \mathsf{O} \\ \mathsf{L}_{\mathcal{I}\mathcal{J}} & \mathsf{I} \end{bmatrix} \begin{bmatrix} \mathsf{L}^T_{\mathcal{J}\mathcal{J}} & \mathsf{L}^T_{\mathcal{I}\mathcal{J}} \\ \mathsf{O} & \mathsf{S} \end{bmatrix}, \tag{2.11}$$

where $\mathsf{L}_{\mathcal{J}\mathcal{J}} \in \mathbb{R}^{r\times r}$ is a lower factor of the Cholesky decomposition of $\mathsf{A}_{\mathcal{J}\mathcal{J}}$, $\mathsf{L}_{\mathcal{I}\mathcal{J}} \in \mathbb{R}^{s\times r}$, $r = n - s$, $\mathsf{L}_{\mathcal{I}\mathcal{J}} = \mathsf{A}_{\mathcal{I}\mathcal{J}}\mathsf{L}^{-T}_{\mathcal{J}\mathcal{J}}$, and $\mathsf{S} \in \mathbb{R}^{s\times s}$ is the *Schur complement matrix of the block* $\mathsf{A}_{\mathcal{I}\mathcal{I}}$ defined by

$$\mathsf{S} = \mathsf{A}_{\mathcal{I}\mathcal{I}} - \mathsf{A}_{\mathcal{I}\mathcal{J}}\mathsf{A}^{-1}_{\mathcal{J}\mathcal{J}}\mathsf{A}_{\mathcal{J}\mathcal{I}}.$$

The decomposition (2.11) is a useful tool for the effective construction of a generalized inverse or for the effective evaluation of the multiplication of a vector by a generalized inverse.

Lemma 2.2 *Let* A *denote an SPS block matrix as in (2.11), let* $\mathbf{e} \in \mathrm{Ker}\mathsf{A}$*, and let* S *denote the Schur complement matrix of the block* $\mathsf{A}_{\mathcal{I}\mathcal{I}}$*. Then* $\mathbf{e}_{\mathcal{J}} \in \mathrm{Ker}\mathsf{S}$ *and if* S^+ *denote any generalized inverse of* S*, then the matrix* A^+ *defined by*

$$\mathsf{A}^+ = \begin{bmatrix} \mathsf{L}^{-T}_{\mathcal{J}\mathcal{J}} & -\mathsf{L}^{-T}_{\mathcal{J}\mathcal{J}}\mathsf{L}^T_{\mathcal{I}\mathcal{J}}\mathsf{S}^+ \\ \mathsf{O} & \mathsf{S}^+ \end{bmatrix} \begin{bmatrix} \mathsf{L}^{-1}_{\mathcal{J}\mathcal{J}} & \mathsf{O} \\ -\mathsf{L}_{\mathcal{I}\mathcal{J}}\mathsf{L}^{-1}_{\mathcal{J}\mathcal{J}} & \mathsf{I} \end{bmatrix} \tag{2.12}$$

is a generalized inverse of A *which satisfies*

$$[\mathsf{A}^+]_{\mathcal{I}\mathcal{I}} = \mathsf{S}^+. \tag{2.13}$$

Proof If $\mathsf{A}\mathbf{e} = \mathbf{o}$, then

$$\mathsf{A}_{\mathcal{J}\mathcal{J}}\mathbf{e}_{\mathcal{J}} + \mathsf{A}_{\mathcal{J}\mathcal{I}}\mathbf{e}_{\mathcal{I}} = \mathbf{o}, \qquad \mathsf{A}_{\mathcal{I}\mathcal{J}}\mathbf{e}_{\mathcal{J}} + \mathsf{A}_{\mathcal{I}\mathcal{I}}\mathbf{e}_{\mathcal{I}} = \mathbf{o},$$

and

$$\begin{aligned} \mathsf{S}\mathbf{e}_{\mathcal{I}} &= (\mathsf{A}_{\mathcal{I}\mathcal{I}} - \mathsf{A}_{\mathcal{I}\mathcal{J}}\mathsf{A}^{-1}_{\mathcal{J}\mathcal{J}}\mathsf{A}_{\mathcal{J}\mathcal{I}})\mathbf{e}_{\mathcal{I}} \\ &= \mathsf{A}_{\mathcal{I}\mathcal{I}}\mathbf{e}_{\mathcal{I}} - \mathsf{A}_{\mathcal{I}\mathcal{J}}\mathsf{A}^{-1}_{\mathcal{J}\mathcal{J}}(-\mathsf{A}_{\mathcal{J}\mathcal{J}}\mathbf{e}_{\mathcal{J}}) = \mathbf{o}, \end{aligned}$$

i.e., $\mathbf{e}_{\mathcal{J}} \in \mathrm{Ker}\mathsf{S}$. The rest can be verified directly. □

Notice that

$$\mathsf{S}^{+} = \left[\mathsf{A}^{+}\right]_{\mathscr{J}\,\mathscr{J}}. \tag{2.14}$$

If A is a full rank matrix, then we get

$$\mathsf{S}^{-1} = \left[\mathsf{A}^{-1}\right]_{\mathscr{J}\,\mathscr{J}}, \tag{2.15}$$

which agrees with (2.5).

2.5 Norms

General concepts of size and distance in a vector space are expressed by norms. A *norm* on $\mathbb{R}^n$ is a function which assigns to each $\mathbf{x} \in \mathbb{R}^n$ a number $\mathbb{I}\mathbf{x}\mathbb{I} \in \mathbb{R}$ in such a way that for any vectors $\mathbf{x},\ \mathbf{y} \in \mathbb{R}^n$ and any scalar $\alpha \in \mathbb{R}$, the following three conditions are satisfied:

(i) $\mathbb{I}\mathbf{x}\mathbb{I} \geq 0$, and $\mathbb{I}\mathbf{x}\mathbb{I} = 0$ if and only if $\mathbf{x} = \mathbf{o}$.
(ii) $\mathbb{I}\mathbf{x} + \mathbf{y}\mathbb{I} \leq \mathbb{I}\mathbf{x}\mathbb{I} + \mathbb{I}\mathbf{y}\mathbb{I}$.
(iii) $\mathbb{I}\alpha\mathbf{x}\mathbb{I} = |\alpha|\ \mathbb{I}\mathbf{x}\mathbb{I}$.

It is easy to check that the functions

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}_1^2 + \cdots + \mathbf{x}_n^2} \quad \text{and} \quad \|\mathbf{x}\|_\infty = \max\{|\mathbf{x}_1|, \ldots, |\mathbf{x}_n|\}$$

are norms. They are called ℓ_2 (Euclidean) and ℓ_∞ norms, respectively.

Given a norm defined on the domain and the range of a matrix A, we can define the *induced norm* $\mathbb{I}\mathsf{A}\mathbb{I}$ of A by

$$\mathbb{I}\mathsf{A}\mathbb{I} = \sup_{\mathbb{I}\mathbf{x}\mathbb{I}=1} \mathbb{I}\mathsf{A}\mathbf{x}\mathbb{I} = \sup_{\mathbf{x}\neq\mathbf{o}} \frac{\mathbb{I}\mathsf{A}\mathbf{x}\mathbb{I}}{\mathbb{I}\mathbf{x}\mathbb{I}}.$$

If $\mathsf{B} \neq \mathsf{O}$, then

$$\mathbb{I}\mathsf{A}\mathsf{B}\mathbb{I} = \sup_{\mathbf{x}\neq\mathbf{o}} \frac{\mathbb{I}\mathsf{A}\mathsf{B}\mathbf{x}\mathbb{I}}{\mathbb{I}\mathbf{x}\mathbb{I}} = \sup_{\mathsf{B}\mathbf{x}\neq\mathbf{o}} \frac{\mathbb{I}\mathsf{A}\mathsf{B}\mathbf{x}\mathbb{I}}{\mathbb{I}\mathsf{B}\mathbf{x}\mathbb{I}} \frac{\mathbb{I}\mathsf{B}\mathbf{x}\mathbb{I}}{\mathbb{I}\mathbf{x}\mathbb{I}} \leq \sup_{\substack{\mathbf{y}\in\mathrm{Im}\mathsf{B},\\ \mathbf{y}\neq\mathbf{o}}} \frac{\mathbb{I}\mathsf{A}\mathbf{y}\mathbb{I}}{\mathbb{I}\mathbf{y}\mathbb{I}} \sup_{\mathbf{x}\neq\mathbf{o}} \frac{\mathbb{I}\mathsf{B}\mathbf{x}\mathbb{I}}{\mathbb{I}\mathbf{x}\mathbb{I}}.$$

It follows easily that the induced norm is *submultiplicative*, i.e.,

$$\mathbb{I}\mathsf{A}\mathsf{B}\mathbb{I} \leq \mathbb{I}\mathsf{A}|\mathrm{Im}\mathsf{B}\mathbb{I}\ \mathbb{I}\mathsf{B}\mathbb{I} \leq \mathbb{I}\mathsf{A}\mathbb{I}\ \mathbb{I}\mathsf{B}\mathbb{I}. \tag{2.16}$$

If $\mathsf{A} = [a_{ij}] \in \mathbb{R}^{m\times n}$ and $\mathbf{x} = [x_i] \in \mathbb{R}^n$, then

$$\|\mathsf{A}\mathbf{x}\|_\infty = \max_{i=1,\dots,m} |\sum_{j=1}^{n} a_{ij}x_j| \le \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}||x_j| \le \|\mathbf{x}\|_\infty \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}|,$$

that is, $\|\mathsf{A}\|_\infty \le \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}|$. Since the last inequality turns into the equality for a vector $\mathbf{x}$ with suitably chosen entries $x_i \in \{1, -1\}$, we have

$$\|\mathsf{A}\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}|. \tag{2.17}$$

2.6 Scalar Products

General concepts of length and angle in a vector space are introduced by means of a *scalar product*; it is the mapping which assigns to each couple $\mathbf{x},\ \mathbf{y} \in \mathbb{R}^n$ a number $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}$ in such a way that for any vectors $\mathbf{x},\ \mathbf{y},\ \mathbf{z} \in \mathbb{R}^n$ and any scalar $\alpha \in \mathbb{R}$, the following four conditions are satisfied:

(i) $(\mathbf{x}, \mathbf{y} + \mathbf{z}) = (\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{z})$.
(ii) $(\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y})$.
(iii) $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$.
(iv) $(\mathbf{x}, \mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{o}$.

The scalar product is an SPD form, see also Chap. 4.

We often use the *Euclidean scalar product* or the *Euclidean inner product* which assigns to each couple of vectors $\mathbf{x},\ \mathbf{y} \in \mathbb{R}^n$ a number defined by

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y}.$$

In more complicated expressions, we often denote the Euclidean scalar product in $\mathbb{R}^3$ by dot, so that

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}.$$

If A is an SPD matrix, then we can define the more general A-*scalar product* on $\mathbb{R}^n$ by

$$(\mathbf{x}, \mathbf{y})_\mathsf{A} = \mathbf{x}^T\mathsf{A}\mathbf{y}.$$

We denote for any $\mathbf{x} \in \mathbb{R}^n$ its *Euclidean norm* and A-*norm* by

$$\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2}, \quad \|\mathbf{x}\|_\mathsf{A} = (\mathbf{x}, \mathbf{x})_\mathsf{A}^{1/2}.$$

It is easy to see that any norm induced by a scalar product satisfies the properties (i) and (iii) of the norm. The property (ii) follows from the *Cauchy–Schwarz inequality*

$$(\mathbf{x}, \mathbf{y})^2 \leq \|\mathbf{x}\|^2 \, \|\mathbf{y}\|^2, \tag{2.18}$$

which is valid for any $\mathbf{x}, \ \mathbf{y} \in \mathbb{R}^n$ and any scalar product. The bound is tight in the sense that the inequality becomes the equality when $\mathbf{x}, \ \mathbf{y}$ are dependent.

A pair of vectors $\mathbf{x}$ and $\mathbf{y}$ is *orthogonal* (with respect to a given scalar product) if

$$(\mathbf{x}, \mathbf{y}) = 0.$$

The vectors $\mathbf{x}$ and $\mathbf{y}$ that are orthogonal in A-scalar product are called A-*conjugate* or briefly *conjugate*.

Two *sets of vectors* $\mathscr{E}$ *and* $\mathscr{F}$ *are orthogonal* (also stated "$\mathscr{E}$ orthogonal to $\mathscr{F}$") if any $\mathbf{x} \in \mathscr{E}$ is orthogonal to any $\mathbf{y} \in \mathscr{F}$. The set $\mathscr{E}^\perp$ of all the vectors of $\mathbb{R}^n$ that are orthogonal to $\mathscr{E} \subseteq \mathbb{R}^n$ is a vector space called an *orthogonal complement* of $\mathscr{E}$. If $\mathscr{E} \subseteq \mathbb{R}^n$, then

$$\mathbb{R}^n = \operatorname{Span} \mathscr{E} \oplus \mathscr{E}^\perp.$$

A *set of vectors* $\mathscr{E}$ *is orthogonal* if its elements are pairwise orthogonal, i.e., any $\mathbf{x} \in \mathscr{E}$ is orthogonal to any $\mathbf{y} \in \mathscr{E}$, $\mathbf{y} \neq \mathbf{x}$. A set of vectors $\mathscr{E}$ is *orthonormal* if it is orthogonal and $(\mathbf{x}, \mathbf{x}) = 1$ for any $\mathbf{x} \in \mathscr{E}$.

A square matrix U *is orthogonal* if $\mathsf{U}^T\mathsf{U} = \mathsf{I}$, that is, $\mathsf{U}^{-1} = \mathsf{U}^T$. Multiplication by an orthogonal matrix U preserves both the angles between any two vectors and the Euclidean norm of any vector as

$$(\mathsf{U}\mathbf{x})^T \mathsf{U}\mathbf{y} = \mathbf{x}^T \mathsf{U}^T \mathsf{U}\mathbf{y} = \mathbf{x}^T \mathbf{y}.$$

A matrix $\mathsf{P} \in \mathbb{R}^{n \times n}$ is an *orthogonal projector* if P is a projector, i.e., $\mathsf{P}^2 = \mathsf{P}$, and $\mathrm{Im}\mathsf{P}$ is orthogonal to $\mathrm{Ker}\mathsf{P}$. The latter condition can be rewritten equivalently as

$$\mathsf{P}^T(\mathsf{I} - \mathsf{P}) = \mathsf{O}.$$

It simply follows that

$$\mathsf{P}^T = \mathsf{P}^T\mathsf{P} = \mathsf{P},$$

so that orthogonal projectors are symmetric matrices and symmetric projectors are orthogonal projectors. If P is an orthogonal projector, then $\mathsf{I} - \mathsf{P}$ is also an orthogonal projector as

$$(\mathsf{I} - \mathsf{P})^2 = I - 2\mathsf{P} + \mathsf{P}^2 = \mathsf{I} - \mathsf{P} \quad \text{and} \quad (\mathsf{I} - \mathsf{P})^T\mathsf{P} = (\mathsf{I} - \mathsf{P})\mathsf{P} = \mathsf{O}.$$

If $U \subseteq \mathbb{R}^n$ is the subspace spanned by the columns of a full column rank matrix $\mathsf{U} \in \mathbb{R}^{m \times n}$, then

$$\mathsf{P} = \mathsf{U}(\mathsf{U}^T\mathsf{U})^{-1}\mathsf{U}^T$$

is an orthogonal projector as

$$\mathsf{P}^2 = \mathsf{U}(\mathsf{U}^T\mathsf{U})^{-1}\mathsf{U}^T\mathsf{U}(\mathsf{U}^T\mathsf{U})^{-1}\mathsf{U}^T = \mathsf{P} \quad \text{and} \quad \mathsf{P}^T = \mathsf{P}.$$

Since any vector $\mathbf{x} \in U$ may be written in the form $\mathbf{x} = \mathsf{U}\mathbf{y}$ and

$$\mathsf{P}\mathbf{x} = \mathsf{U}(\mathsf{U}^T\mathsf{U})^{-1}\mathsf{U}^T\mathsf{U}\mathbf{y} = \mathsf{U}\mathbf{y} = \mathbf{x},$$

it follows that

$$U = \mathrm{Im}\mathsf{P}.$$

Observe that $\mathsf{U}^T\mathsf{U}$ is nonsingular; since $\mathsf{U}^T\mathsf{U}\mathbf{x} = \mathbf{o}$ implies

$$\|\mathsf{U}\mathbf{x}\|^2 = \mathbf{x}^T(\mathsf{U}^T\mathsf{U}\mathbf{x}) = 0,$$

it follows that $\mathbf{x} = \mathbf{o}$ by the assumption on the full column rank of U.

2.7 Eigenvalues and Eigenvectors

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ denote a square matrix with complex entries. If a vector $\mathbf{e} \in \mathbb{C}^n$ and a scalar $\lambda \in \mathbb{C}$ satisfy

$$\mathsf{A}\mathbf{e} = \lambda\mathbf{e}, \tag{2.19}$$

then $\mathbf{e}$ is said to be an *eigenvector* of A associated with an *eigenvalue* λ. A vector $\mathbf{e}$ is an eigenvector of A if and only if $\mathrm{Span}\{\mathbf{e}\}$ is an invariant subspace of A; the restriction $\mathsf{A}|\mathrm{Span}\{\mathbf{e}\}$ reduces to the multiplication by λ. If $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ are eigenvectors of a symmetric matrix A, then it is easy to check that $\mathrm{Span}\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ and $\mathrm{Span}\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}^\perp$ are invariant subspaces.

The set of all eigenvalues of A is called the *spectrum* of A; we denote it by $\sigma(\mathsf{A})$. Obviously, $\lambda \in \sigma(\mathsf{A})$ if and only if $\mathsf{A} - \lambda\mathsf{I}$ is singular, and $0 \in \sigma(\mathsf{A})$ if and only if A is singular. If $\lambda \neq 0$, $\lambda \in \sigma(\mathsf{A})$, then we can multiply (2.19) by $\lambda^{-1}\mathsf{A}^{-1}$ to get $\mathsf{A}^{-1}\mathbf{e} = \lambda^{-1}\mathbf{e}$, so we can write

$$\sigma(\mathsf{A}^{-1}) = \sigma^{-1}(\mathsf{A}).$$

If $U \subseteq \mathbb{C}^n$ is an invariant subspace of $\mathsf{A} \in \mathbb{C}^{n\times n}$, then we denote by $\sigma(\mathsf{A}|U)$ the eigenvalues of A that correspond to the eigenvectors belonging to U.

Since it is well known that a matrix is singular if and only if its determinant is equal to zero, it follows that the eigenvalues of A are the roots of the *characteristic equation*

$$\det(\mathsf{A} - \lambda\mathsf{I}) = 0. \tag{2.20}$$

The *characteristic polynomial* $p_{\mathsf{A}}(\lambda) = \det(\mathsf{A} - \lambda\mathsf{I})$ is of the degree n. Thus there are at most n distinct eigenvalues and $\sigma(\mathsf{A})$ is not the empty set.

Even though it is in general difficult to evaluate the eigenvalues of a given matrix A, it is still possible to get nontrivial information about $\sigma(\mathsf{A})$ without heavy computations. Useful information about the location of eigenvalues can be obtained by *Gershgorin's theorem*, which guarantees that every eigenvalue of $\mathsf{A} = [a_{ij}] \in \mathbb{C}^{n\times n}$ is located in at least one of the n circular disks in the complex plane with the centers a_{ii} and radii $r_i = \sum_{j\neq i} |a_{ij}|$.

The *eigenvalues of a real symmetric matrix are real*. Since it is easy to check whether a matrix is symmetric, this gives us useful information about the location of eigenvalues.

Let $\mathsf{A} \in \mathbb{R}^{n\times n}$ denote a real symmetric matrix, let $\mathscr{I} = \{1, \ldots, n-1\}$, and let $\mathsf{A}^1 = \mathsf{A}_{\mathscr{I}\mathscr{I}}$. Let $\lambda_1 \geq \cdots \geq \lambda_n$ and $\lambda_1^1 \geq \cdots \geq \lambda_{n-1}^1$ denote the eigenvalues of A and A^1, respectively. Then by the *Cauchy interlacing theorem*

$$\lambda_1 \geq \lambda_1^1 \geq \lambda_2 \geq \lambda_2^1 \geq \cdots \geq \lambda_{n-1}^1 \geq \lambda_n. \tag{2.21}$$

2.8 Matrix Decompositions

If $\mathsf{A} \in \mathbb{R}^{n\times n}$ is a symmetric matrix, then it is possible to find n orthonormal eigenvectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ that form the basis of $\mathbb{R}^n$. Moreover, the corresponding eigenvalues are real. Denoting by $\mathsf{U} = [\mathbf{e}_1, \ldots, \mathbf{e}_n] \in \mathbb{R}^{n\times n}$ an orthogonal matrix the columns of which are the eigenvectors, we may write the *spectral decomposition* of A as

$$\mathsf{A} = \mathsf{U}\mathsf{D}\mathsf{U}^T, \tag{2.22}$$

where $\mathsf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^{n\times n}$ is the diagonal matrix the diagonal entries of which are the eigenvalues corresponding to the eigenvectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Reordering the columns of U, we can achieve that $\lambda_1 \geq \cdots \geq \lambda_n$.

The spectral decomposition reveals close relations between the properties of a symmetric matrix and its eigenvalues. Thus, a symmetric matrix is SPD if and only if all its eigenvalues are positive, and it is SPS if and only if they are nonnegative. The rank of a symmetric matrix is equal to the number of nonzero entries of D.

If A is symmetric, then we can use the spectral decomposition (2.22) to check that for any nonzero $\mathbf{x}$

$$\lambda_1 = \lambda_{\max} \geq \|\mathbf{x}\|^{-2}\mathbf{x}^T\mathsf{A}\mathbf{x} \geq \lambda_{\min} = \lambda_n. \tag{2.23}$$

Thus for any symmetric positive definite matrix A

$$\|\mathsf{A}\| = \lambda_{\max},\ \|\mathsf{A}^{-1}\| = \lambda_{\min}^{-1},\ \|\mathbf{x}\|_{\mathsf{A}} \leq \lambda_{\max}\|\mathbf{x}\|,\ \|\mathbf{x}\|_{\mathsf{A}^{-1}} \leq \lambda_{\min}^{-1}\|\mathbf{x}\|. \tag{2.24}$$

The *spectral condition number* $\kappa(\mathsf{A}) = \|\mathsf{A}\|\|\mathsf{A}^{-1}\|$, which is a measure of departure from the identity, can be expressed for real symmetric matrix by

$$\kappa(\mathsf{A}) = \lambda_{\max}/\lambda_{\min}.$$

If A is a real symmetric matrix and f is a real function defined on $\sigma(\mathsf{A})$, we can use the spectral decomposition to define the *scalar function* by

$$f(\mathsf{A}) = \mathsf{U} f(\mathsf{D})\mathsf{U}^T,$$

where $f(\mathsf{D}) = \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))$. It is easy to check that if a is the identity function on $\mathbb{R}$ defined by $a(x) = x$, then

$$a(\mathsf{A}) = \mathsf{A},$$

and if f and g are real functions defined on $\sigma(\mathsf{A})$, then

$$(f+g)(\mathsf{A}) = f(\mathsf{A}) + g(\mathsf{A}) \quad \text{and} \quad (f \cdot g)(\mathsf{A}) = f(\mathsf{A})g(\mathsf{A}).$$

Moreover, if $f(x) \geq 0$ for $x \in \sigma(\mathsf{A})$, then $f(\mathsf{A})$ is SPS, and if $f(x) > 0$ for $x \in \sigma(\mathsf{A})$, then $f(\mathsf{A})$ is SPD. For example, if A is SPD, then

$$\mathsf{A} = \mathsf{A}^{1/2}\mathsf{A}^{1/2}.$$

Obviously

$$\sigma(f(\mathsf{A})) = f(\sigma(\mathsf{A})), \tag{2.25}$$

and if $\mathbf{e}_i$ is an eigenvector corresponding to $\lambda_i \in \sigma(\mathsf{A})$, then it is also an eigenvector of $f(\mathsf{A})$ corresponding to $f(\lambda_i)$. It follows easily that for any SPS matrix

$$\operatorname{Im}\mathsf{A} = \operatorname{Im}\mathsf{A}^{1/2} \quad \text{and} \quad \operatorname{Ker}\mathsf{A} = \operatorname{Ker}\mathsf{A}^{1/2}. \tag{2.26}$$

A key to understanding nonsymmetric matrices is the *singular value decomposition* (SVD). If $\mathsf{B} \in \mathbb{R}^{m\times n}$, then SVD of B is given by

$$\mathsf{B} = \mathsf{U}\mathsf{S}\mathsf{V}^T, \tag{2.27}$$

where $\mathsf{U} \in \mathbb{R}^{m\times m}$ and $\mathsf{V} \in \mathbb{R}^{n\times n}$ are orthogonal, and $\mathsf{S} \in \mathbb{R}^{m\times n}$ is a diagonal matrix with nonnegative diagonal entries $\sigma_1 \geq \cdots \geq \sigma_{\min\{m,n\}} = \sigma_{\min}$ called *singular values* of B. If A is not a full rank matrix, then it is often more convenient to use the *reduced singular value decomposition* (RSVD)

$$\mathsf{B} = \widehat{\mathsf{U}}\widehat{\mathsf{S}}\widehat{\mathsf{V}}^T, \tag{2.28}$$

where $\widehat{\mathsf{U}} \in \mathbb{R}^{m\times r}$ and $\widehat{\mathsf{V}} \in \mathbb{R}^{n\times r}$ are matrices with orthonormal columns, $\widehat{\mathsf{S}} \in \mathbb{R}^{r\times r}$ is a nonsingular diagonal matrix with positive diagonal entries $\sigma_1 \geq \cdots \geq \sigma_r = \overline{\sigma}_{\min}$, and $r \leq \min\{m,n\}$ is the rank of B. The matrices $\widehat{\mathsf{U}}$ and $\widehat{\mathsf{V}}$ are formed by the first r columns of U and V. If $\mathbf{x} \in \mathbb{R}^m$, then

$$\mathsf{B}\mathbf{x} = \widehat{\mathsf{U}}\widehat{\mathsf{S}}\widehat{\mathsf{V}}^T\mathbf{x} = (\widehat{\mathsf{U}}\widehat{\mathsf{S}}\widehat{\mathsf{V}}^T)(\widehat{\mathsf{V}}\widehat{\mathsf{S}}\widehat{\mathsf{U}}^T)(\widehat{\mathsf{U}}\widehat{\mathsf{S}}^{-1}\mathsf{V}^T\mathbf{x}) = \mathsf{B}\mathsf{B}^T\mathbf{y},$$

so that

$$\mathrm{Im}\mathsf{B} = \mathrm{Im}\mathsf{B}\mathsf{B}^T. \tag{2.29}$$

If $\mathsf{B} = \widehat{\mathsf{U}}\widehat{\mathsf{S}}\widehat{\mathsf{V}}^T$ is RSVD, then

$$\mathrm{Im}\,\mathsf{B} = \mathrm{Im}\,\widehat{\mathsf{U}}, \quad \mathrm{Ker}\,\mathsf{B} = \left(\mathrm{Im}\,\widehat{\mathsf{V}}\right)^{\perp}.$$

It follows that

$$\mathrm{Im}\,\mathsf{B}^T = (\mathrm{Ker}\,\mathsf{B})^{\perp}. \tag{2.30}$$

The SVD reveals close relations between the properties of a matrix and its singular values. Thus, the rank of $\mathsf{B} \in \mathbb{R}^{m\times n}$ is equal to the number of its nonzero singular values,

$$\|\mathsf{B}\| = \|\mathsf{B}^T\| = \sigma_1, \tag{2.31}$$

and for any vector $\mathbf{x} \in \mathbb{R}^n$

$$\sigma_{\min}\|\mathbf{x}\| \le \|\mathsf{B}\mathbf{x}\| \le \|\mathsf{B}\|\|\mathbf{x}\|. \tag{2.32}$$

Let $\overline{\sigma}_{\min}$ denote the least nonzero singular value of $\mathsf{B} \in \mathbb{R}^{m\times n}$, let $\mathbf{x} \in \mathrm{Im}\mathsf{B}^T$, and consider the RSVD $\mathsf{B} = \widehat{\mathsf{U}}\widehat{\mathsf{S}}\widehat{\mathsf{V}}^T$ with $\widehat{\mathsf{U}} \in \mathbb{R}^{m\times r}$, $\widehat{\mathsf{V}} \in \mathbb{R}^{n\times r}$, and $\widehat{\mathsf{S}} \in \mathbb{R}^{r\times r}$. Then there is $\mathbf{y} \in \mathbb{R}^r$ such that $\mathbf{x} = \widehat{\mathsf{V}}\mathbf{y}$ and

$$\|\mathsf{B}\mathbf{x}\| = \|\widehat{\mathsf{U}}\widehat{\mathsf{S}}\widehat{\mathsf{V}}^T\widehat{\mathsf{V}}\mathbf{y}\| = \|\widehat{\mathsf{U}}\widehat{\mathsf{S}}\mathbf{y}\| = \|\widehat{\mathsf{S}}\mathbf{y}\| \ge \overline{\sigma}_{\min}\|\mathbf{y}\|.$$

Since

$$\|\mathbf{x}\| = \|\widehat{\mathsf{V}}\mathbf{y}\| = \|\mathbf{y}\|,$$

we conclude that

$$\overline{\sigma}_{\min}\|\mathbf{x}\| \le \|\mathsf{B}\mathbf{x}\| \ \text{ for any } \ \mathbf{x} \in \mathrm{Im}\mathsf{B}^T, \tag{2.33}$$

or, equivalently,

$$\overline{\sigma}_{\min}\|\mathbf{x}\| \le \|\mathsf{B}^T\mathbf{x}\| \ \text{ for any } \ \mathbf{x} \in \mathrm{Im}\mathsf{B}. \tag{2.34}$$

The SVD (2.27) can be used to introduce the *Moore–Penrose generalized inverse* of an $m \times n$ matrix B by

$$\mathsf{B}^{\dagger} = \mathsf{V}\mathsf{S}^{\dagger}\mathsf{U}^T,$$

where $\mathsf{S}^{\dagger}$ is the diagonal matrix with the entries $[\mathsf{S}^{\dagger}]_{ii} = 0$ if $\sigma_i = 0$ and $[\mathsf{S}^{\dagger}]_{ii} = \sigma_i^{-1}$ otherwise. It is easy to check that

$$\mathsf{B}\mathsf{B}^{\dagger}\mathsf{B} = \mathsf{U}\mathsf{S}\mathsf{V}^T\mathsf{V}\mathsf{S}^{\dagger}\mathsf{U}^T\mathsf{U}\mathsf{S}\mathsf{V}^T = \mathsf{U}\mathsf{S}\mathsf{V}^T = \mathsf{B}, \tag{2.35}$$

so that the Moore–Penrose generalized inverse is a generalized inverse. If B is a full row rank matrix, then it may be checked directly that

$$\mathsf{B}^{\dagger} = \mathsf{B}^T(\mathsf{B}\mathsf{B}^T)^{-1}.$$

If B is a singular matrix and $\mathbf{c} \in \mathrm{Im}\mathsf{B}$, then $\mathbf{x}_{\mathrm{LS}} = \mathsf{B}^{\dagger}\mathbf{c}$ is a solution of the system of linear equations $\mathsf{B}\mathbf{x} = \mathbf{c}$, i.e.,

$$\mathsf{B}\mathbf{x}_{\mathrm{LS}} = \mathbf{c}.$$

Notice that $\mathbf{x}_{\mathrm{LS}} \in \mathrm{Im}\mathsf{B}^T$, so that if $\overline{\mathbf{x}}$ is any other solution, then $\overline{\mathbf{x}} = \mathbf{x}_{\mathrm{LS}} + \mathbf{d}$, where $\mathbf{d} \in \mathrm{Ker}\mathsf{B}$, $\mathbf{x}_{\mathrm{LS}}^T\mathbf{d} = 0$, and

$$\|\mathbf{x}_{\mathrm{LS}}\|^2 \leq \|\overline{\mathbf{x}}_{\mathrm{LS}}\|^2 + \|\mathbf{d}\|^2 = \|\overline{\mathbf{x}}\|^2. \tag{2.36}$$

The vector $\mathbf{x}_{\mathrm{LS}}$ is called the *least square solution* of $\mathsf{B}\mathbf{x} = \mathbf{c}$.

Obviously

$$\|\mathsf{B}^{\dagger}\| = \overline{\sigma}_{\min}^{-1}, \tag{2.37}$$

where $\overline{\sigma}_{\min}$ denotes the least nonzero singular value of B, so that

$$\|\mathbf{x}_{\mathrm{LS}}\| = \|\mathsf{B}^{\dagger}\mathbf{c}\| \leq \overline{\sigma}_{\min}^{-1}\|\mathbf{c}\|. \tag{2.38}$$

It can be verified directly that

$$\left(\mathsf{B}^{\dagger}\right)^T = \left(\mathsf{B}^T\right)^{\dagger}.$$

2.9 Graphs, Walks, and Adjacency Matrices

We shall need some simple results of graph theory and linear algebra. Let us recall that the *vertices* V_i and the *edges* of the *graph of the mesh of the triangulation* $\mathscr{T} = \{\tau_i\}$ of a polyhedral domain Ω are the nodes of the mesh and their adjacent couples V_iV_j, respectively. Recall that the edges V_i and V_j are *adjacent* if there is an element $\tau_k \in \mathscr{T}$ such that V_iV_j is the edge of τ_k. The graph is fully described by the *adjacency matrix* D with the nonzero entries d_{ij} equal to one if the nodes V_i and V_j are adjacent. Since the graph of the mesh is not oriented and does not contain loops, the adjacency matrix is symmetric and its diagonal entries d_{ii} are equal to zero. Let us also recall that the *walk of length k* in the mesh of $\mathscr{T}$ is a sequence of the distinct nodes $V_{i_1}, \ldots, V_{i_k}$ such that the edges $V_{i_j}V_{i_{j+1}}$, $j = 1, 2, \ldots, k-1$ belong to the graph of the mesh. Thus

$$d_{i_j i_{j+1}} = 1, \quad j = 1, \ldots, k-1.$$

The walk $(V_{i_1}, \ldots, V_{i_k})$ starts at V_{i_1} and ends at V_{i_k}. Moreover, we call a walk between nodes V_i and V_k an (i, k)-*walk*. We use the following well-known observation [5].

Lemma 2.3 *Let* D *be the adjacency matrix of the mesh of* $\mathscr{T}$ *and* $\mathsf{B} = \mathsf{D}^k$. *Then each entry* b_{ij} *of* B *gives the number of distinct* (i, j)*-walks of length* k.

Proof To see why our lemma holds, we use induction on k. For $k = 1$ the claim follows immediately from the definition of D. Suppose that for some $k \geq 1$ the entry b_{ij} in $\mathsf{B} = \mathsf{D}^k$ gives the number of distinct (i, j)-walks of length k. For convenience we denote $\mathsf{C} = \mathsf{D}^{k+1}$, so that $\mathsf{C} = \mathsf{B}\mathsf{D}$. The entries of C are given by

$$c_{ij} = \sum_{\ell=1}^{n} b_{i\ell} d_{\ell j},$$

where the number $b_{i\ell}$ gives the number of distinct (i, ℓ)-walks of length k and $d_{\ell j} = 0$ or $d_{\ell j} = 1$. If a particular edge $V_\ell V_j$ is not in the mesh (graph) then $d_{\ell j} = 0$ and $b_{i\ell} d_{\ell j} = 0$. Thus, there is no (i, j)-walk of length $k + 1$ with the last-but-one node ℓ. On the other hand, if $d_{\ell, j} = 1$ then we can prolong each (i, ℓ)-walk to an (i, j)-walk. Thus c_{ij} gives the number of all distinct (i, j)-walks of length $k + 1$. □

The following corollary follows easily from Lemma 2.3.

Corollary 2.1 *Let* D *denote the adjacency matrix of a given mesh and* $\mathbf{e} = [e_i]$, $e_i = 1$, $i = 1, 2, \ldots, n$. *Then the number* $w(i, k)$ *of distinct walks of length* k *starting at node* i *is given by*

$$w(i, k) = [\mathsf{D}^k \mathbf{e}]_i.$$

If the mesh is approximately regular, we expect that more walks of length k originate from the nodes that are near a center of the mesh than from the vertices that are far from it. It simply follows that the node with the index i which satisfies $w(i, k) \geq w(j, k)$, $j = 1, 2, \ldots, n$, for sufficiently large k is in a sense near to the center of the mesh.

References

1. Demmel, J.W.: Applied Numerical Linear Algebra. SIAM, Philadelphia (1997)
2. Laub, A.J.: Matrix Analysis for Scientists and Engineers. SIAM, Philadelphia (2005)
3. Golub, G.H., Van Loan, C.F.: Matrix Computations, 2nd edn. Johns Hopkins University Press, Baltimore (1989)
4. Dostál, Z., Kozubek, T., Markopoulos, A., Menšík, M.: Cholesky factorization of a positive semidefinite matrix with known kernel. Appl. Math. Comput. **217**, 6067–6077 (2011)
5. Diestel, R.: Graph Theory. Springer, Heidelberg (2005)

http://www.springer.com/978-1-4939-6832-9

Scalable Algorithms for Contact Problems

Dostál, Z.; Kozubek, T.; Sadowska, M.; Vondrák, V.

2016, XIX, 340 p. 80 illus., 23 illus. in color., Hardcover
ISBN: 978-1-4939-6832-9