Probability, Random Processes, and Statistical Analysis

Supplementary Materials

(Revision: April 5, 2012)

HISASHI KOBAYASHI, Princeton University
BRIAN L. MARK, George Mason University
WILLIAM TURIN, AT&T Labs Research

CAMBRIDGE
UNIVERSITY PRESS

Contents

1 Selected topics of set theory

To be done.

2 Matrices

In this supplement to our book [5], we provide a review of some topics of matrix algebra that are relevant to the book. This supplement is not a textbook, but rather a collection of mathematical methods illustrated by examples that help engineers and graduate students to understand the material of the book. Throughout this supplement we give references to textbooks that contain in-depth coverage and necessary proofs of the corresponding statements and theorems.

2.1 Matrix Operations

A size $m \times n$ matrix is a table with m rows and n columns:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [a_{ij}]_{m\times n}\,. \tag{2.1}$$

We consider only the matrices whose elements a_{ij} are real or complex numbers. If the elements of the matrix are real, we write $\boldsymbol{A} \in \mathbb{R}^{m\times n}$; if they are complex, we write $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. If the number of rows is equal to the number of columns ($m = n$), $\boldsymbol{A}$ is called a *square matrix*.

A matrix with one row ($m = 1$) $\boldsymbol{a} = [a_1, a_2, \dots, a_n]$ is called a row-vector, and a matrix with one column ($n = 1$)

$$\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix} \tag{2.2}$$

is called a column-vector. Sometimes it is convenient to denote the ij-th element of the matrix $\boldsymbol{A}$ as $(\boldsymbol{A})_{ij} = a_{ij}$. An $\boldsymbol{A}$

A *transpose* $\boldsymbol{A}^\top$ of the matrix $\boldsymbol{A}$ is obtained by replacing its rows with the corresponding columns: $(\boldsymbol{A}^\top)_{ij} \triangleq (\boldsymbol{A})_{ji}$. For example, if

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^\top = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

To save paper, it is convenient to represent a column vector as a transpose of a row vector. For example a column vector of equation (2.2) we can present as $\boldsymbol{b} = [b_1, b_2, \ldots, b_m]^\top$. If $\boldsymbol{A} = \boldsymbol{A}^\top$, the matrix is called *symmetrical*.

For matrices with complex elements a *Hermitian transpose* or *conjugate transpose* matrix is defined as a complex conjugate of its transpose: $\boldsymbol{A}^H = (\bar{\boldsymbol{A}}^\top)$. For example,

$$\begin{bmatrix} 1+i & 2 \\ 3-i & 4+2i \\ 5 & 6 \end{bmatrix}^H = \begin{bmatrix} 1-i & 3+i & 5 \\ 2 & 4-2i & 6 \end{bmatrix}.$$

A matrix is called *Hermitian* if it is equal to its Hermitian transpose: $\boldsymbol{A} = \boldsymbol{A}^H$. Obviously, a real Hermitian matrix is a symmetrical matrix. The *product of a scalar and the matrix* is obtained by multiplying each element of the matrix by this scalar: $(\lambda \boldsymbol{A})_{ij} \triangleq \lambda(\boldsymbol{A})_{ij}$. For example

$$3\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 15 \\ 6 & 12 & 18 \end{bmatrix}$$

The *sum of the matrices* $\boldsymbol{A} + \boldsymbol{B}$ of the same size is the matrix whose elements are the sums of the corresponding elements $(\boldsymbol{A} + \boldsymbol{B})_{ij} \triangleq (\boldsymbol{A})_{ij} + (\boldsymbol{B})_{ij}$. For example

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} + \begin{bmatrix} -11 & 2 & 3 \\ 7 & 3 & -1 \end{bmatrix} = \begin{bmatrix} -10 & 5 & 8 \\ 9 & 7 & 5 \end{bmatrix}$$

If $\boldsymbol{A} = [a_{ij}]_{m\times n}$ and $\boldsymbol{B} = [b_{ij}]_{n\times p}$, then the *product of the matrices* $\boldsymbol{C} = \boldsymbol{AB}$ of the matrices is defined as the matrix $\boldsymbol{C}$ whose element

$$\boxed{c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.} \tag{2.3}$$

For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} -11 & 2 & 3 \\ 7 & 3 & -1 \end{bmatrix} = \begin{bmatrix} -11+14 & 2+6 & 3-2 \\ -33+28 & 6+12 & 9-4 \\ -55+42 & 10+18 & 15-6 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 1 \\ -5 & 18 & 5 \\ -13 & 28 & 9 \end{bmatrix}$$

Note that the product of the matrices is generally non-commutative: $\boldsymbol{AB} \neq \boldsymbol{BA}$. For example, if we exchange the order of multiplication of the matrices of the previous example, we obtain

$$\begin{bmatrix} -11 & 2 & 3 \\ 7 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 10 & 4 \\ 18 & 20 \end{bmatrix}$$

A *lower triangular matrix* is a square matrix whose elements above its main diagonal $a_{ij} = 0$ for $i < j$. An *upper triangular matrix* is defined similarly: $a_{ij} = 0$ for $i > j$. A *diagonal matrix* is a matrix whose off-diagonal elements $a_{ij} = 0$ for $i \neq j$. A diagonal $m \times n$ we denote as

$$\boldsymbol{A} = \text{diag}\{a_1, a_2, \ldots, a_p\} \tag{2.4}$$

where $p = \min(m, n)$. For example, the following matrices are diagonal matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

A square diagonal matrix whose diagonal elements are all ones is called an *identity matrix*:

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} = [\delta_{ij}]_{n\times n} = \text{diag}[1, 1, \dots, 1], \tag{2.5}$$

where δ_{ik} is the Kronecker symbol: $\delta_{ii} = 1$ and $\delta_{ik} = 0$ for $i \neq k$. It is easy to see that $\boldsymbol{AI} = \boldsymbol{A}$ and $\boldsymbol{IA} = \boldsymbol{A}$ for any matrix[1] $\boldsymbol{A}$.

A column matrix $\mathbf{1} = [1 \quad 1 \quad \dots \quad 1]^\top$ is often used in our book [5]. The product $\boldsymbol{A}\mathbf{1}$ is equal to the column matrix whose elements are the sums of rows of the matrix $\boldsymbol{A}$. For example

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 11 \end{bmatrix}$$

Thus defined matrix operations satisfy the following easily verified identities:

1. $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$.
2. $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$.
3. $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}) = \boldsymbol{ABC}$.
4. $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$.
5. $(\boldsymbol{AB})^H = \boldsymbol{B}^H\boldsymbol{A}^H$.
6. $(\boldsymbol{AB})^\top = \boldsymbol{B}^\top\boldsymbol{A}^\top$.
7. A product or sum of lower (upper) triangular matrices is also a lower (upper) triangular matrix.

2.1.0.1 Trace of a matrix

Trace of square matrix $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ is the sum of the elements on its main diagonal:

$$\text{Trace}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}. \tag{2.6}$$

The matrix trace possesses the following properties

1. $\text{Trace}(\boldsymbol{A} + \boldsymbol{B}) = \text{Trace}(\boldsymbol{A}) + \text{Trace}(\boldsymbol{B})$.
2. $\text{Trace}(\lambda\boldsymbol{A}) = \lambda\text{Trace}(\boldsymbol{A})$.
3. If $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{C}^{n\times m}$, then $\text{Trace}(\boldsymbol{AB}) = \text{Trace}(\boldsymbol{BA})$. Indeed,

$$\text{Trace}(\boldsymbol{AB}) = \sum_{i=1}^{m}(\boldsymbol{AB})_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji}.$$

[1] If the matrix $\boldsymbol{A}$ has m rows and n columns, then we should write $\boldsymbol{A}\boldsymbol{I}_{n\times n} = \boldsymbol{A}$ and $\boldsymbol{I}_{m\times m}\boldsymbol{A} = \boldsymbol{A}$, but since the sizes of the identity matrices are obvious, we write in both cases $\boldsymbol{I}$ to simplify notation

Similarly,

$$\mathrm{Trace}(\boldsymbol{BA}) = \sum_{j=1}^{n} (\boldsymbol{BA})_{jj} = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij}.$$

Thus, $\mathrm{Trace}(\boldsymbol{AB}) = \mathrm{Trace}(\boldsymbol{BA})$.

4. $\mathrm{Trace}(\boldsymbol{A}) = \overline{\mathrm{Trace}(\boldsymbol{A}^H)}$ because the Hermite transposition just changes the diagonal elements by their conjugate.

2.1.1 Determinant

Determinant of a square matrix can be defined recursively. The determinant of a scalar (size 1×1 matrix)is equal to the scalar: $\det(a) = a$. The second-order determinant

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{2.7}$$

can be viewed as sum of the products of the elements of a row by the corresponding *adjuncts* or *cofactors* :

$$\det \boldsymbol{A} = a_{11}A_{11} + a_{12}A_{12} \tag{2.8}$$

where $A_{11} = a_{22}$ and $A_{22} = -a_{21}$ are called *adjuncts* or *cofactors* of the elements a_{11} and a_{12}, respectively.

For the third-order matrix, we define

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$$

where the adjuncts are defined as

$$A_{11} = \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \; A_{12} = -\det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \; A_{13} = \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

For the square matrix of size $n \times n$, we define the adjunct as

$$A_{ij} = (-1)^{i+j} M_{ij} \tag{2.9}$$

where M_{ij}, called the *minor* of a_{ij}, is a determinant of the matrix obtained by deleting of i-th row and j-th column from $\boldsymbol{A}$.

For the square matrix of size $n \times n$, we define the determinant recursively as

$$\boxed{\det \boldsymbol{A} = |\boldsymbol{A}| \triangleq \sum_{j=1}^{n} a_{ij}A_{ij}} \tag{2.10}$$

This equation is order-recursive: it defines[2] the determinant of n-th order using the determinants of $(n-1)$-th order which can further be decomposed into determinants of $(n-2)$-th order, and so on. For example,

$$\begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 2 & 1 \\ 2 & 1 \end{vmatrix} - 2 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 3 & 2 \\ 1 & 2 \end{vmatrix} = 1 \cdot 0 - 2 \cdot 2 + 3 \cdot 4 = 8. \tag{2.11}$$

The matrix $\boldsymbol{A}$ is called *singular* if $det\boldsymbol{A} = 0$, otherwise it is called *non-singular*. It is easy to verify the following properties of the determinant (assuming that all the operations are defined).

1. If any raw of a square matrix consists entirely of zeros, the matrix is singular. Indeed, if i-th row consists of zeroes, then equation (2.10) gives $\det \boldsymbol{A} = 0$.
2. Interchanging two rows changing the sign of the determinant. Indeed, this is obvious for the second-order determinants, according to equation (2.7). Then, by induction, all the A_{ij} in equation (2.10) change sign which means that $\det \boldsymbol{A}$ changes sign.
3. If a matrix has two equal rows, it is singular. Indeed, if we exchange the rows, its determinant must change sign. But, since the rows are equal, we have the same matrix. Thus, in this case, $\det \boldsymbol{A} = -\det \boldsymbol{A}$, which is possible only if $\det \boldsymbol{A} = 0$.
4. A sum of products of elements of i-th row of a matrix by the cofactors of another row of the matrix is zero:

$$\sum_{j=1}^{n} a_{ij} A_{kj} = 0 \quad \text{if } i \neq k \tag{2.12}$$

   This property follows from the previous, because the sum is equal to the determinant of the matrix that has two equal rows. This matrix is obtained from the matrix $\boldsymbol{A}$ by replacing its k-th with its i-th row.
5. $\det \boldsymbol{A}^{\top} = \det \boldsymbol{A}$. It follows from this property, that every result that is proven to the columns of the determinant applies to its rows.
6. Combining equations (2.10) and (2.12) we have

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \delta_{ik} \det \boldsymbol{A} \tag{2.13}$$

7. If a a row of a matrix is multiplied by a number, then its determinant is also multiplied by the same number. This property is obtained by multiplying both sides of (2.10) by a number.
8. If two rows of a matrix are proportional, the matrix is singular. Indeed, by factoring the proportionality coefficient we obtain the matrix with two equal rows which is singular.
9. If two matrices differ only by a row, then the determinant of the sum of the matrices is equal to the sum of their determinants. This property follows from equation (2.10) in which $a_{ij} = b_{ij} + c_{ij}$.
10. Addition of a row times any number to the other row of the matrix does not change its determinant. This property is a corollary of the two previous properties since the determinant of the modified matrix can be expressed as a sum of the determinants of the original matrix and the matrix with the proportional rows.
11. For the size $n \times n$ matrix $\boldsymbol{A}$, $\det \lambda \boldsymbol{A} = \lambda^n \det \boldsymbol{A}$. This property follows from property (7): if we multiply $\boldsymbol{A}$ by λ, then each its row is multiplied by λ.

[2] Strictly speaking, we can use this definition using one particular row and then prove (by induction) that the result does not depend on the row selection. We leave the proof to the reader as an exercise.

12. Determinant of a triangular matrix is equal to the product of its diagonal elements. This property follows from equation (2.10) according to which $\det \boldsymbol{A} = a_{11}A_{11}$ where A_{11} is also a triangular matrix. Thus,

$$\det \boldsymbol{A} = a_{11}a_{22}\cdots a_{nn} \tag{2.14}$$

13. Determinant of the the product of the matrices $\det \boldsymbol{AB} = \det \boldsymbol{A} \det \boldsymbol{B}$.
14. Determinant of the Hermite transpose $\det(\boldsymbol{A}^H) = \overline{\det \boldsymbol{A}}$.
15. Determinant of a Hermitian matrix is real: $\det(\boldsymbol{A}^H) = \det \boldsymbol{A} = \overline{\det \boldsymbol{A}}$.

2.1.2 Rank of the Matrix

An order-k minor of the matrix $\boldsymbol{A}$ is a determinant of a $k \times k$ matrix

$$\begin{bmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_k} \\ \dots & \dots & \dots & \dots \\ a_{i_k j_1} & a_{i_k j_2} & \dots & a_{i_k j_k} \end{bmatrix}. \tag{2.15}$$

where $i_1, i_2, ..., i_k$ and $j_1, j_2, ..., j_k$ represent the selected rows and columns of $\boldsymbol{A}$. The *rank* of the matrix is the highest order of all the matrix minors that are different from zero. In other words, there is an order-k minor that is not equal to zero while all the minors of higher orders are zeroes. If the rank of an $m \times n$ matrix is equal to $\min(m, n)$, we say that the matrix has *full rank*.

The rank of the matrix can be found by the following recursive procedure: we try to find consequently the non-zero minors of the order $1, 2, \dots,$ adding at each step just one row and one column to the non-zero minor found in the previous step. The process stops if we found a minor of the order k that is not equal to 0, but all minors of the order $k + 1$, obtained in this process are equal to 0, then the rank of the matrix is k because in this case the minors whose order is larger than $k + 1$ can be expressed using equation (2.10) as a linear combination of the minors of order $k + 1$ which all are zeroes.

2.1.3 Inverse Matrix

The matrix $\boldsymbol{A}^{-1}$ is called the inverse of $\boldsymbol{A}$ if

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I} \tag{2.16}$$

The inverse matrix exists if and only if $\boldsymbol{A}$ is non-singular. It follows from this equation that if the inverse matrix exists, then $\boldsymbol{A}$ must be non-singular. Since $\det \boldsymbol{A}^{-1}\boldsymbol{A} = \det \boldsymbol{A}^{-1} \det \boldsymbol{A} = \det \boldsymbol{I} = 1$, we have

$$\det \boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \tag{2.17}$$

To obtain a formula for the inverse matrix, we rewrite equation (2.13) as

$$\frac{1}{\det \boldsymbol{A}} \sum_{j=1}^{n} a_{ij} A_{kj} = \delta_{ik} \tag{2.18}$$

As we can see, the left hand side of this equation represents an element of the product of the matrix $\boldsymbol{A}$ and the transposed matrix $\mathrm{Adj}(\boldsymbol{A})$ of the cofactors (called the *adjugate matrix of A*) multiplied by the scalar $1/\det \boldsymbol{A}$. The right hand side is an element of the identity matrix (2.5). Thus, we can rewrite equation (2.18) in the following

matrix form

$$\boldsymbol{A}\left(\frac{1}{\det \boldsymbol{A}}\,\mathrm{Adj}(\boldsymbol{A})\right)=\boldsymbol{I} \tag{2.19}$$

Comparing this equation with (2.16) we conclude that for any non-singular matrix $\boldsymbol{A}$ the inverse matrix exists and can be written as

$$\boxed{\boldsymbol{A}^{-1}=\frac{1}{\det \boldsymbol{A}}\,\mathrm{Adj}(\boldsymbol{A})} \tag{2.20}$$

As an example of using this equation, let us find the inverse matrix for the second-order matrix

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix} \tag{2.21}$$

We have

$$\det \boldsymbol{A}=a_{11}a_{22}-a_{12}a_{21} \tag{2.22}$$

The matrix of cofactors is

$$\begin{bmatrix} a_{22} & -a_{21}\\ -a_{12} & a_{11}\end{bmatrix} \tag{2.23}$$

The adjugate matrix is the transpose of this matrix. Thus, the inverse matrix

$$\boldsymbol{A}^{-1}=\frac{1}{a_{11}a_{22}-a_{12}a_{21}}\begin{bmatrix} a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{bmatrix} \tag{2.24}$$

To verify the correctness of this equation, we multiply

$$\boldsymbol{A}\boldsymbol{A}^{-1}=\frac{1}{a_{11}a_{22}-a_{12}a_{21}}\begin{bmatrix} a_{11}a_{22}-a_{12}a_{21} & -a_{11}a_{12}+a_{12}a_{11}\\ a_{21}a_{22}-a_{22}a_{21} & -a_{21}a_{12}+a_{22}a_{11}\end{bmatrix}=\begin{bmatrix} 1 & 0\\ 0 & 1\end{bmatrix} \tag{2.25}$$

Caution: Equation (2.20) is used mostly in theoretical derivations. As we will see in the sequel, there are many practical algorithms that are both less computationally expensive and more accurate than (2.20).

2.1.3.1 Unitary and orthogonal matrices

A matrix $\boldsymbol{U}$ is called *unitary* if $\boldsymbol{U}^H=\boldsymbol{U}^{-1}$ i.e.

$$\boldsymbol{U}^H\boldsymbol{U}=\boldsymbol{U}\boldsymbol{U}^H=\boldsymbol{I}. \tag{2.26}$$

A real unitary matrix is called *orthogonal*:

$$\boldsymbol{U}^\top\boldsymbol{U}=\boldsymbol{U}\boldsymbol{U}^\top=\boldsymbol{I}. \tag{2.27}$$

It follows from the properties of the determinants of the Hermite transpose of a matrix that

$$\det \boldsymbol{U}^H \det \boldsymbol{U}=(\det \boldsymbol{U})^2=\det \boldsymbol{I}=1. \tag{2.28}$$

Thus,

$$|\det \boldsymbol{U}|=1 \tag{2.29}$$

and the determinant of a unitary matrix equals either -1 or +1:

$$\det \boldsymbol{U} = \pm 1. \tag{2.30}$$

2.2 Systems of Linear Equations

A system of linear equations has the form

$$\begin{array}{cccc} a_{11}x_1+ & a_{12}x_2+ & \cdots+ & a_{1n}x_n = b_1 \\ a_{21}x_1+ & a_{22}x_2+ & \cdots+ & a_{2n}x_n = b_2 \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}x_1+ & a_{m2}x_2+ & \cdots+ & a_{mn}x_n = b_m \end{array} \tag{2.31}$$

Using matrix notations this system can be written as

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{2.32}$$

where $\boldsymbol{A} = [a_{ij}]_{m\times n}$ is called the *matrix of the system*, vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^\top$ is called the *unknown*, and $\boldsymbol{b} = [b_1, b_2, \ldots, b_n]^\top$ is a known vector.

A *solution* of the system is every vector $\boldsymbol{x}$ that satisfies the system, that is turns (2.32) into an identity. If such vector does not exist, the system does not have a solution. Two systems are called *equivalent* if they have the same set of solutions. Multiplication of both sides of equation (2.32) by a non-singular matrix $\boldsymbol{F}$ leads to an equivalent system $\boldsymbol{F}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{F}\boldsymbol{b}$. Indeed, as we can see, for any matrix $\boldsymbol{F}$, every solution of the original system is a solution of the modified system. If $\boldsymbol{F}$ is non-singular, multiplying the modified system by the inverse matrix $\boldsymbol{F}^{-1}$ leads to the original system (2.32) so that every solution of the modified system is also a solution of the original system.

Consider a system of n equations with n unknowns. In this case, $\boldsymbol{A}$ is a square matrix. If the matrix of the system in non-singular, it has an inverse matrix $\boldsymbol{A}^{-1}$. Multiplying both sides of equation (2.32) by the inverse matrix, we obtain $\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ which gives us the unique solution

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} \tag{2.33}$$

If we use equation (2.20) for the inverse matrix and multiply the adjugate matrix $\operatorname{Adj}(\boldsymbol{A})$ by $\boldsymbol{b}$, we obtain the well known Cramer's rule:

$$x_i = \frac{\det \boldsymbol{A}_i}{\det \boldsymbol{A}} \tag{2.34}$$

where $\boldsymbol{A}_i$ is the matrix formed by replacing the i-th column of $\boldsymbol{A}$ by the column vector $\boldsymbol{b}$. These formulae are rarely used in practice, because the computation of the inverse matrix is more complex then solving the system by Gaussian elimination. The efficient methods for solving linear systems are discussed in the next section.

2.2.1 Row Echelon Form

The process of solving the system of linear equation is based on transforming it into a simpler equivalent system. The following *elementary transformations of the system* are used to solve the system

1. interchanging the equations,
2. multiplying an equation by a non-zero number,
3. adding some equation times a number to the other equation.

It is convenient to perform the elementary transformations of the system using the system's *augmented* matrix

$$\boldsymbol{A}_a = [\boldsymbol{A}\ \boldsymbol{b}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix} \tag{2.35}$$

. Elementary transformations of the system are equivalent to the corresponding transformations of the rows of the augmented matrix: interchanging equations is equivalent to interchanging the corresponding rows, multiplication of an equation by a non-zero number is equivalent to multiplying the corresponding row of the augmented matrix by this number, and adding some equation times a number to the other equation is equivalent to adding the corresponding row of the augmented matrix times the number to the other row. It is easy to see that these operations can be performed by multiplying from the left the augmented matrix of the system by the *elementary matrices* $\boldsymbol{E}$ that are obtained by the same elementary transformations of the identity matrix. Since the elementary matrices are non-singular, therefore, multiplications convert the augmented matrix into an augmented matrix of the equivalent system. In other words, the *elementary transformations of a linear system convert it into an equivalent system*.

The first column of $\boldsymbol{A}_a$ always has a non-zero element (otherwise there would be no variable x_1 in the system). We can always assume that $a_{11} \neq 0$ (this can be achieved by exchanging the rows of the augmented matrix). We call thus obtained coefficient $a_{11} \neq 0$ a *pivot*.

By multiplying the first row by $-a_{i1}/a_{11}$ and adding it to the i-th row we obtain 0 in the place of a_{i1}. Thus, we can create zeroes below a_{11} which means that we eliminate the variable x_1 from the rest $m-1$ equations:

$$\boldsymbol{A}_a^{(1)} = [\boldsymbol{A}^{(1)}, \boldsymbol{b}^{(1)}] = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} & b_2^{(1)} \\ \ldots & \ldots & \ldots & \ldots & \\ 0 & a_{m2}^{(1)} & \ldots & a_{mn}^{(1)} & b_m^{(1)} \end{bmatrix}, \tag{2.36}$$

where

$$\boxed{a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}} \tag{2.37}$$

This process is called the *Gaussian elimination* of the variable x_1. If the rest of the equations contain x_2, i.e. $a_{i2}^{(1)} \neq 0$, we can assume that $a_{22}^{(1)} \neq 0$ (we can always achieve it by interchanging the equations) and use it as a pivot to eliminate x_2 from the rest of $m-2$ equations. Otherwise, it can happen that after the elimination of x_1 from the rest of the equations we also eliminated $x_2, x_3, \ldots, x_{k_2-1}$, but there is an equation containing x_{k_2}. Without loss of generality we can assume that $a_{2k_2} \neq 0$ (this can be achieved by swapping the rows of the matrix) and use it as a pivot to eliminate x_{k_2}, and so on. We can continue the process until we can find pivots in the

following rows. Finally, we obtain the so-called *row echelon form* (REF) for the augmented matrix:

$$\boldsymbol{A}_a^{(r)} = [\boldsymbol{A}_a^{(r)}, \boldsymbol{b}^{(r)}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k_2} & \dots & a_{1k_r} & \dots & a_{1n} & b_1 \\ 0 & 0 & \dots & a^{(1)}_{2k_2} & \dots & a^{(1)}_{2k_r} & \dots & a^{(1)}_{2n} & b^{(1)}_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & a^{(r-1)}_{rk_r} & \dots & a^{(r-1)}_{rn} & b^{(r-1)}_r \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & b^{(r)}_{r+1} \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \end{bmatrix}. \tag{2.38}$$

The system that corresponds to the matrix (2.38) has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1k_2}x_{k_2} + \dots + a_{1k_r}x_{k_r} + \quad \dots + a_{1n}x_n &= b_1 \\ a^{(1)}_{2k_2}x_{k_2} + \dots + a^{(1)}_{2k_r}x_{k_r} + \quad \dots + a^{(1)}_{2n}x_n &= b^{(1)}_2 \\ \dots \qquad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ a^{(r-1)}_{rk_r}x_{k_r} + \dots + a^{(r-1)}_{rn}x_n &= b^{(r-1)}_r \\ 0 &= b^{(r)}_{r+1} \end{aligned} \tag{2.39}$$

The REF system is easy to solve and analyse.

First of all we note that if $b^{(r)}_{r+1} \neq 0$, the last equation $0 = b^{(r)}_{r+1}$ is impossible to satisfy. Thus, the system does not have a solution in this case.

If $b^{(r)}_{r+1} = 0$,we can prove that the system has solutions which can be found by the *back substitution*. Since $a^{(r-1)}_{rk_r} \neq 0$, we can find x_{k_r} from the r-th equation of the system (2.39) by expressing it through $x_{k_r+1}, x_{k_r+2}, \dots, x_n$. Substituting it into the previous equations we find $x_{k_{r-1}}$ and so on. Finally, will express all the unknowns $x_1, x_{k_2}, ..., x_{k_r}$ corresponding to the pivots through the rest of the unknowns

$$\begin{aligned} x_1 &= \alpha_1 + \textstyle\sum_\nu \beta_{1\nu}x_\nu \\ x_{k_2} &= \alpha_2 + \textstyle\sum_\nu \beta_{2\nu}x_\nu \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ x_{k_r} &= \alpha_r + \textstyle\sum_\nu \beta_{r\nu}x_\nu \end{aligned} \tag{2.40}$$

where the index ν in the RHS of this equation can take all the values $\nu \leq n$ except for $1, k_2, ..., k_r$. The unknowns in the RHS of the previous equation are called the *free unknowns*, because we can assign to them any values while the unknowns in the LHS of this equation are the *main* or *basic* unknowns. Their values are uniquely defined from (2.40) by the free unknowns. Any set of values of $n - r$ free unknowns and the corresponding r of the main unknowns define a solutions of the linear system (2.31).

Note that the back substitution can be incorporated into the forward elimination if we eliminate not only the elements below a pivot, but also the elements above the pivot.

Thus, the system has solutions if and only if $b^{(r)}_{r+1} = 0$. In this case, if $r = n$, there is no free variables and the system has the unique solution. If $r < n$, the system has an infinite number of solutions.

It is easy to verify using the properties of determinants that the elementary transformations of the matrix do not change its rank. Therefore, the rank of the matrix is equal to the rank of its REF which is equal to the number of its non-zero rows. Note that $b^{(r)}_{r+1} = 0$ if and only if $\mathrm{Rank}(\boldsymbol{A}) = \mathrm{Rank}(\boldsymbol{A}_a)$. These results are summarized in the following theorem.

Theorem 2.1 (Kronecker-Capelli). *A system of linear equations has a solution if and only if the rank of its coefficient matrix is equal to the rank of its augmented matrix:* $\mathrm{Rank}(\boldsymbol{A}) = \mathrm{Rank}(\boldsymbol{A}_a)$.

A system has the unique solution if and only if $\mathrm{Rank}(\boldsymbol{A}) = \mathrm{Rank}(\boldsymbol{A}_a) = n$.

A system has an infinite number of solutions solution if an only if $\mathrm{Rank}(\boldsymbol{A}) = \mathrm{Rank}(\boldsymbol{A}_a) < n$

Example 2.1: Let us solve the system of linear equations

$$\begin{aligned} 2x\ +y\ \ -z &= \ \ 8 \\ -3x\ -y\ +2z &= -11 \\ -2x\ +y\ +2z &= \ -3 \end{aligned}$$

1. We have the following transformations of the system augmented matrix

$$\begin{bmatrix} 2 & 1 & -1 & 8 \\ -3 & -1 & 2 & -11 \\ -2 & 1 & 2 & -3 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 & -1 & 8 \\ 0 & 0.5 & 0.5 & 1 \\ 0 & 2 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 & -1 & 8 \\ 0 & 0.5 & 0.5 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \tag{2.41}$$

To obtain the second matrix, we multiplied the first row by 3/2 and added it to the second row, then we added first row to the last row. The last matrix (the REF) was obtained by multiplying the second row of the second matrix by -4 and adding it to its third row.

The REF of the system, according to equation (2.41), is

$$\begin{aligned} 2x\ \ +y\ \ \ \ -z &= 8 \\ 0.5y\ +0.5z &= 1\,. \\ -z &= 1 \end{aligned}$$

To find the solution, we perform the back substitution. From the last equation we have $z = -1$. Substituting it into the previous equation, we obtain $y = 3$, and, substituting $y = 3$ and $z = -1$ into the first equation we obtain $x = 2$.

2. Note that the back substitution can be also performed using elementary operations starting from the last equation.

$$\begin{bmatrix} 2 & 1 & -1 & 8 \\ 0 & 0.5 & 0.5 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 & 0 & 7 \\ 0 & 0.5 & 0 & 1.5 \\ 0 & 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 0 & 4 \\ 0 & 0.5 & 0 & 1.5 \\ 0 & 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

The last matrix was obtained by dividing each row of the previous matrix by the pivot (so that the new pivots are equal to 1). The system that corresponds to the last matrix has the form

$$\begin{aligned} x &= 2 \\ y &= 3 \\ z &= -1 \end{aligned}$$

Thus, the last column (2,3,-1) of the previous matrix gives us the solution. The corresponding matrix is called the *reduced row echelon form* (RREF)[3] of the system augmented matrix.

[3] Matlab and the majority of other software packages have functions `rref()` that perform the transformations.

3. The back substitution can be incorporated into the forward part of the Gaussian elimination if in its each step we not only eliminate the matrix elements below its pivots but also above them. Referring to equation (2.41), in addition to the previously performed eliminations, we could have multiplied the second row by -2 and added it to the first row and repeated similar eliminations in the third column as illustrated above. Thus, we could have

$$\begin{bmatrix} 2 & 1 & -1 & 8 \\ -3 & -1 & 2 & -11 \\ -2 & 1 & 2 & -3 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 & -1 & 8 \\ 0 & 0.5 & 0.5 & 1 \\ 0 & 2 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & -2 & 6 \\ 0 & 0.5 & 0.5 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 0 & 4 \\ 0 & 0.5 & 0 & 1.5 \\ 0 & 0 & -1 & 1 \end{bmatrix} \tag{2.42}$$

By dividing each row by its pivot, we obtain the RREF

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

and the solution $(2, 3, -1)$.

Example 2.2: In this example, we consider a system whose matrix has rank that is less than the number of unknowns.

$$\begin{aligned} 2x_1 + 4x_2 - 3x_3 + 5x_4 + x_5 &= 9 \\ 3x_1 + x_2 + x_4 - 3x_5 &= 2 \\ x_1 + 7x_2 - 6x_3 + 9x_4 + 5x_5 &= 16 \end{aligned}$$

Applying the Gaussian elimination to the augmented matrix of the system, we obtain the following RREF

$$\begin{bmatrix} 2 & 4 & -3 & 5 & 1 & 9 \\ 3 & 1 & 0 & 1 & -3 & 2 \\ 1 & 7 & -6 & 9 & 5 & 16 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0.3 & -0.1 & -1.3 & -0.1 \\ 0 & 1 & -0.9 & 1.3 & 0.9 & 2.3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

As we can see, $\mathrm{Rank}(\boldsymbol{A}) = \mathrm{Rank}(\boldsymbol{A}_a) = 2$ which is less than the number of unknowns $n = 5$. Hence, according to the Kronecker-Capelli theorem, the system has an infinite number of solutions that can be expressed using equation (2.40) which in our case takes the form:

$$\begin{aligned} x_1 &= -0.1 - 0.3x_3 + 0.1x_4 + 1.3x_5 \\ x_2 &= 2.3 + 0.9x_3 - 1.3x_4 - 0.9x_5 \end{aligned}$$

where x_3, x_4, x_5 are the free variables that can take any values. For example, if $x_3 = 1, x_4 = 0, x_5 = 0$, then the solution of the system is $\boldsymbol{x} = (-0.4, 3.2, 1, 0, 0)$.

Example 2.3: In this example, we change the last element in the RHS of the system (e.g. write 7 instead of 16), we obtain after the Gaussian elimination

$$\begin{bmatrix} 2 & 4 & -3 & 5 & 1 & 9 \\ 3 & 1 & 0 & 1 & -3 & 2 \\ 1 & 7 & -6 & 9 & 5 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0.3 & -0.1 & -1.3 & -0.1 \\ 0 & 1 & -0.9 & 1.3 & 0.9 & 2.3 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus, $\mathrm{Rank}(\boldsymbol{A}) = 2 < \mathrm{Rank}(\boldsymbol{A}_a) = 3$ and the system does not have solutions.

2.2.2 Block Matrices

It is often convenient to represent a matrix in the block form

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \dots & \boldsymbol{A}_{1N} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \dots & \boldsymbol{A}_{2N} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{A}_{M1} & \boldsymbol{A}_{M2} & \dots & \boldsymbol{A}_{MN} \end{bmatrix} = [\boldsymbol{A}_{ij}]_{M\times N}\,. \tag{2.43}$$

where $\boldsymbol{A}_{ij}$ are sub-matrices of the matrix $\boldsymbol{A}$. The operations between the matrices can be written in the block form. For example, the matrix product $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ of the matrices in the block form is given by

$$\boldsymbol{C}_{ij} = \boldsymbol{A}_{i1}\boldsymbol{B}_{1j} + \boldsymbol{A}_{i2}\boldsymbol{A}_{2j} + \cdots + \boldsymbol{A}_{iN}\boldsymbol{B}_{Nj} = \sum_{k=1}^{N} \boldsymbol{A}_{ik}\boldsymbol{B}_{kj}. \tag{2.44}$$

provided that all the matrix products in this equation are defined.

A block-diagonal matrix is defined as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{22} & \dots & \boldsymbol{0} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{A}_{NN} \end{bmatrix} = \text{block diag}\,[\boldsymbol{A}_{ii}]_{N\times N}\,. \tag{2.45}$$

where all the diagonal blocks are square matrices. It is easy to prove that

$$\det \boldsymbol{A} = \det \boldsymbol{A}_{11} \det \boldsymbol{A}_{22} \cdots \det \boldsymbol{A}_{NN} \tag{2.46}$$

2.3 Linear Vector Space

Definition 2.1 (Linear vector space)**.** *A linear vector space is a set $\mathcal{V}$ whose elements are called vectors (or points) if*

1. *there is a rule according to which for any vectors $\boldsymbol{x} \in \mathcal{V}$ and $\boldsymbol{y} \in \mathcal{V}$ there is a unique vector $\boldsymbol{z} \in \mathcal{V}$ which is called the sum of the vectors and denoted as $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$;*
2. *there is a rule according to which for any vector $\boldsymbol{x} \in \mathcal{V}$ and a number λ from some field of numbers there is a unique vector $\boldsymbol{z} \in \mathcal{V}$ which is called the product of the number and vector and denoted as $\boldsymbol{z} = \lambda\boldsymbol{x}$;*
3. *the defined operations of addition of the vectors and multiplication of a vector by a number satisfy the following axioms (called the axioms of a linear space)*
   - **a)** $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$;
   - **b)** $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$ *for any* $\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{z} \in \mathcal{V}$;
   - **c)** *there is a null-vector* $\boldsymbol{0} \in \mathcal{V}$ *such that* $\boldsymbol{x} + \boldsymbol{0} = \boldsymbol{x}$ *for any* $\boldsymbol{x} \in \mathcal{V}$;
   - **d)** *for any* $\boldsymbol{x} \in \mathcal{V}$ *there is* $\boldsymbol{y} \in \mathcal{V}$ *such that* $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{0}$. *This vector is called the opposite to* $\boldsymbol{x}$ *and denoted as* $-\boldsymbol{x}$;
   - **e)** $1 \cdot \boldsymbol{x} = \boldsymbol{x}$ *for any* $\boldsymbol{x} \in \mathcal{V}$;

f) $\alpha(\beta\boldsymbol{x}) = (\alpha\beta)\boldsymbol{x}$ *for any* $\boldsymbol{x} \in \mathcal{V}$ *and any two numbers* α *and* β*;*
g $(\alpha + \beta)\boldsymbol{x} = (\alpha\boldsymbol{x} + \beta\boldsymbol{x}$ *for any* $\boldsymbol{x} \in \mathcal{V}$ *and any two numbers* α *and* β*;*
h) $\alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha\boldsymbol{x} + \alpha\boldsymbol{y}$ *for any* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$ *and any number* α*.*

This axiomatic definition allows us to apply the theory of linear vector spaces not only to to geometrical vectors but also to various objects (such as polynomials, solutions of linear systems, solutions of the systems of differential equations, random variables, and many other objects that satisfy the axioms of the linear vector space).

We call the vector space real and denote $\mathbb{R}$ if the numbers are from the field of real numbers. We denote the complex vector space as $\mathbb{C}$ if the numbers are from the field of complex numbers.

2.3.1 Linearly Independent Vectors

Definition 2.2 (Linearly dependent and independent vectors)**.** *Let* $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ *be vectors in* $\mathbb{C}$*. We say that these vectors are linearly dependent if there are numbers* $x_1, x_2, \ldots, x_k$ *not all equal to zero, such that*

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k = 0 \tag{2.47}$$

Otherwise (that is if this equation is satisfied only if $x_1 = x_2 = ... = x_k = 0$*), the vectors are called linearly independent.*

If vectors are linearly dependent, then at least one of the coefficients is different from zero. Dividing equation (2.47) by it we can express the corresponding vector as a linear combination of the rest of them.

Definition 2.3 (n-dimensional space)**.** *If in the linear space there are* n *linearly independent vectors and every* $n + 1$ *vectors of the space are linearly dependent, then we call this space* n*-dimensional*

In other words, the maximal number of the linearly independent vectors of the space is called its dimension. If we use real numbers in the definition of the linear space, we denote the n-dimensional space as $\mathbb{R}^n$; the complex n-dimensional space is denoted as $\mathbb{C}^n$.

Definition 2.4 (Basis of n-dimensional space)**.** *An ordered system of* n *linearly independent vectors* $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ *of the* n*-dimensional space is called its basis*

Theorem 2.2. *Any vector* $\boldsymbol{x}$ *of the* n*-dimensional space can be uniquely represented as a linear combination of its basis vectors*

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n. \tag{2.48}$$

Proof. Since any $n + 1$ vectors are linearly dependent

$$\alpha_0\boldsymbol{x} + \alpha_1\boldsymbol{e}_1 + \alpha_2\boldsymbol{e}_2 + \cdots + \alpha_n\boldsymbol{e}_n = 0$$

where at least one of the coefficients is not zero. We claim that $\alpha_0 \neq 0$ because otherwise it would mean that the basis vectors are linearly dependent. Dividing this equation by $\alpha_0 \neq 0$ we obtain (2.48). Uniqueness can be proven by contradiction: if we assume that the representation is not unique, then subtracting an alternative representation we will see that the basis vectors are linearly dependent. □ □

The coefficients $x_1, x_2, \ldots, x_n$ are called *the coordinates* of the vector $\boldsymbol{a}$ in the basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ and we will write $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

It is easy to see that if vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ are presented by their coordinates in the same basis, then $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$ and $\lambda\boldsymbol{x} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$.

2.3.2 Conditions of Linear Dependence

Consider now the conditions of the linear dependence of the vectors in the coordinate form. Let $\boldsymbol{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, m$ be m vectors. They are linearly dependent if equation (2.47) has a non-zero solution. This equation can be written in the following coordinate form

$$\begin{array}{cccc} a_{11}x_1 + & a_{12}x_2 + & \cdots + & a_{1n}x_n = 0 \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + & a_{2n}x_n = 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + & a_{mn}x_n = 0 \end{array} \tag{2.49}$$

which can be written in the following form

$$\boldsymbol{A}\boldsymbol{x} = 0 \tag{2.50}$$

where the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ are columns of the matrix $\boldsymbol{A}$ while $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$. This system is called a *homogeneous system*. It always has a solution $\boldsymbol{x} = 0$ which is called a *trivial solution*. Thus, we can say that the vectors are linearly dependent if and only if the system has a non-trivial solution. As we know, this is possible if and only if $\mathrm{Rank}(\boldsymbol{A}) < m$ (the rank of the system matrix is less than the number of unknowns). Thus, the rank of the matrix is equal to the maximal number of its linearly independent columns (or rows).

For example, if

$$\boldsymbol{a}_1 = (1, 0, 1), \quad \boldsymbol{a}_1 = (0, 1, 0), \quad \boldsymbol{a}_1 = (1, 1, 1),$$

then

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, the $\mathrm{Rank}(\boldsymbol{A}) = 2 < 3$ and the vectors are linearly dependent.

2.3.3 Inner Product Space

Definition 2.5 (Inner product and the norm). *An **inner product** on a complex vector space $\mathbb{C}$ is a function that assigns to each ordered pairs of vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}$ a scalar $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ such that*

(a) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$*;*
(b) $\langle c\boldsymbol{x}, \boldsymbol{y} \rangle = c\langle \boldsymbol{x}, \boldsymbol{y} \rangle$*;*
(c) $\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}$*, the bar denoting complex conjugation;*
(d) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$*, if* $\boldsymbol{x} \neq 0$*.*

The positive square root of $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ denoted by

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

*is called the **norm** of $\boldsymbol{x}$ with respect to the inner product.* □

The axioms (a), (b) and (c) imply

(e) $\langle \boldsymbol{x}, \boldsymbol{y}+\boldsymbol{z}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{x}, \boldsymbol{z}\rangle$.

With (c), we have, for instance,

$$\langle i\boldsymbol{x}, i\boldsymbol{x}\rangle = -i\langle i\boldsymbol{x}, \boldsymbol{x}\rangle = -i^2\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle = \|\boldsymbol{x}\|^2.$$

A complex linear space with the inner product is called **the inner product space** (or a **unitary space**) a real linear space with the inner product is called the **Euclidean space**. Thus, for the Euclidean space, the axiom (c) is replaced by:

(c') $\langle \boldsymbol{y}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$;

Since we can define the distance or **metric** between any elements $\boldsymbol{x}$ and $\boldsymbol{y}$ by $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\|$, the inner product space is a **metric space**.

Definition 2.6 (Orthogonality). *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be vectors in an inner product space $\boldsymbol{V}$. We say that $\boldsymbol{x}$ and $\boldsymbol{y}$ are **orthogonal** if $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0$.* □

As we showed in Sec. 10.1 of our book [5] , the inner product satisfies the Cauchy-Schwarz inequality

$$\frac{|\langle \boldsymbol{x}, \boldsymbol{y}\rangle|}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} \leq 1. \tag{2.51}$$

Thus, we can define the cosine of the angle between the vectors of the Euclidean space as[4]

$$\cos\varphi = \frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}. \tag{2.52}$$

so that we can present the inner product as

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\varphi. \tag{2.53}$$

Consider a right triangle in the unitary space with the orthogonal legs $\boldsymbol{x}$, $\boldsymbol{y}$ ($\boldsymbol{x}^H\boldsymbol{y}$, and hypotenuse $\boldsymbol{z} = \boldsymbol{y}-\boldsymbol{x}$. Then we have

$$\|\boldsymbol{z}\|^2 = \|\boldsymbol{y}-\boldsymbol{x}\|^2 = \langle \boldsymbol{y}-\boldsymbol{x}, \boldsymbol{y}-\boldsymbol{x}\rangle = \|\boldsymbol{x}\|^2 - \langle \boldsymbol{x}, \boldsymbol{y}\rangle - \langle \boldsymbol{y}, \boldsymbol{x}\rangle + \|\boldsymbol{y}\|^2. \tag{2.54}$$

Thus,

$$\|\boldsymbol{z}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 \tag{2.55}$$

which is the Pythagoras theorem for the inner product space. Similarly, we can extend many theorems of the elementary geometry to the Euclidean space geometry. However, we must be careful using these theorems in the general unitary space. For example, we know from the elementary geometry that diagonals of a rhombus (a parallelogram whose sides all have the same length) are always orthogonal. However, this is not necessarily true for a rhombus of a unitary space.

[4] In the unitary space, it is defined as $|\cos\varphi| = \frac{|\langle \boldsymbol{x}, \boldsymbol{y}\rangle|}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}$.

Suppose that vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ are presented by their coordinates in the same basis, then

$$\begin{aligned}\langle \boldsymbol{x}, \boldsymbol{y} \rangle &= \langle x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n, y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + \cdots + y_n\boldsymbol{e}_n \rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle \end{aligned} \tag{2.56}$$

This equation can be written in the following matrix form

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{y}^H \boldsymbol{M} \boldsymbol{x} \tag{2.57}$$

where $\boldsymbol{y}^H$ denotes the conjugate transpose of the column vector $\boldsymbol{y}$ and $\boldsymbol{M}$ is the matrix whose elements are the inner products $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$. Since $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \overline{\langle \boldsymbol{e}_j, \boldsymbol{e}_i \rangle}$, the matrix $\boldsymbol{M}$ is Hermitian: $\boldsymbol{M}^H = \boldsymbol{M}$. (In the real space these matrices are symmetrical: $\boldsymbol{M}^\top = \boldsymbol{M}$.) In addition to being Hermitian, the matrix must be *positive definite* that is

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \boldsymbol{x}^H \boldsymbol{M} \boldsymbol{x} > 0, \quad \forall \boldsymbol{x} \neq \boldsymbol{0} \tag{2.58}$$

according to axiom (d).

We know from geometry, that it is convenient to use the orthonormal basis $\mathbf{i}, \boldsymbol{j}, \boldsymbol{k}$ in $\mathbb{R}^3$. This is also true for other spaces. For that we need the following theorem.

Theorem 2.3. *Mutually orthogonal non-zero vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ are linearly independent.*

Proof. Indeed, if we assume that these vectors are linearly dependent, then equation (2.47) has a non-zero solution $x_1, x_2, \ldots, x_k$. If we assume that $x_i \neq 0$, then by multiplying both sides of equation (2.47) by $\boldsymbol{a}_i$ and using the orthogonality $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = 0$ for $i \neq j$, we obtain the contradiction: $x_i\langle \boldsymbol{a}_i, \boldsymbol{a}_i \rangle = 0$.□ □

Thus, any n non-zero orthogonal vectors of the n-dimensional space represent a basis of the space. If we divide each vector by its length, we obtain the *orthonormal basis*. In the orthonormal basis, the matrix $\boldsymbol{M} = \boldsymbol{I}$ is the identity matrix and the inner product takes especially simple form

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{y}^H \boldsymbol{x} = \sum_{i=1}^{n} x_i\bar{y}_i \tag{2.59}$$

Example 2.4: Consider a set of all matrices $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ with the previously defined matrix addition an multiplication by a complex number. Obviously, this set represents a linear space. For any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of this space define their inner product as

$$\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \mathrm{Trace}(\boldsymbol{B}^H \boldsymbol{A}) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\bar{b}_{ij}. \tag{2.60}$$

As we can see, this definition is equivalent to the definition of the inner product of the vectors in $\mathbb{C}^{m\times n}$ that are constructed by concatenating columns of the matrices. Therefore, the inner product of the matrices satisfies all the axioms of the unitary space.

The norm of $\boldsymbol{A}$ induced by this definition

$$\|\boldsymbol{A}\| = \sqrt{\langle \boldsymbol{A}, \boldsymbol{B} \rangle} = \sqrt{\mathrm{Trace}(\boldsymbol{A}^H \boldsymbol{A})} = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2. \tag{2.61}$$

is called the *Frobenius norm.*

2.3.3.1 Gram-Schmidt orthogonalization

Each inner product space has an orthonormal basis. It can be constructed by the so-called Gram-Schmidt orthogonalization of any basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$. The first vector of the new basis we define as $\boldsymbol{g}_1 = \boldsymbol{e}_1/\|\boldsymbol{e}_1\|$, next we construct a vector as $\boldsymbol{h}_2 = \boldsymbol{e}_2 + \alpha_{11}\boldsymbol{g}_1$ and find α_{11} such that $\boldsymbol{g}_1$ and $\boldsymbol{h}_2$ were orthogonal:

$$\langle \boldsymbol{g}_1, \boldsymbol{h}_2 \rangle = \langle \boldsymbol{g}_1, \boldsymbol{e}_2 \rangle + \alpha_{11}\langle \boldsymbol{g}_1, \boldsymbol{g}_1 \rangle = 0$$

Thus,

$$\alpha_{11} = -\langle \boldsymbol{g}_1, \boldsymbol{e}_2 \rangle.$$

since $\langle \boldsymbol{g}_1, \boldsymbol{g}_1 \rangle = 1$. Therefore, we have

$$\boldsymbol{h}_2 = \boldsymbol{e}_2 - \langle \boldsymbol{e}_2, \boldsymbol{g}_1 \rangle \boldsymbol{g}_1.$$

Dividing $\boldsymbol{h}_2$ by its length we obtain the second vector $\boldsymbol{g}_2 = \boldsymbol{h}_2/\|\boldsymbol{h}_2\|$ of the orthonormal basis. If $n = 2$, we have the orthonormal basis $\{\boldsymbol{g}_1, \boldsymbol{g}_2\}$, otherwise the process continues. Assuming that $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_{n-1}$ have been constructed, we find that

$$\boldsymbol{h}_n = \boldsymbol{e}_n - \langle \boldsymbol{e}_n, \boldsymbol{g}_{n-1} \rangle \boldsymbol{g}_{n-1} - \cdots - \langle \boldsymbol{e}_n, \boldsymbol{g}_1 \rangle \boldsymbol{g}_1 \tag{2.62}$$

is orthogonal to $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_{n-1}$. By normalizing it we obtain $\boldsymbol{g}_n = \boldsymbol{h}_n/\|\boldsymbol{h}_n\|$[5] the orthonormal basis $\boldsymbol{g}_1, \boldsymbol{g}_2, ..., \boldsymbol{g}_n$. Since we have constructed an orthonormal basis, the algorithm is often called the *Gram-Schmidt orthonormalization.*

2.3.3.2 Matrix form of the Gram-Schmidt orthogonalization

Equation (2.62) can be written as

$$\boldsymbol{e}_n = \|\boldsymbol{h}_n\|\boldsymbol{g}_n + \langle \boldsymbol{e}_n, \boldsymbol{g}_{n-1} \rangle \boldsymbol{g}_{n-1} + \cdots + \langle \boldsymbol{e}_n, \boldsymbol{g}_1 \rangle \boldsymbol{g}_1 = \sum_{i=1}^{n} r_{in}\boldsymbol{g}_i \tag{2.63}$$

where we substituted $\boldsymbol{h}_n = \|\boldsymbol{h}_n\|\boldsymbol{g}_n$ and denoted as r_{in} the coefficient of $\boldsymbol{g}_i$.

Starting from $n = 1$, we can rewrite this equation as

$$\begin{aligned}
\boldsymbol{e}_1 &= r_{11}\boldsymbol{g}_1 \\
\boldsymbol{e}_2 &= r_{12}\boldsymbol{g}_1 + r_{22}\boldsymbol{g}_2 && (2.64)\\
\boldsymbol{e}_3 &= r_{13}\boldsymbol{g}_1 + r_{23}\boldsymbol{g}_2 + r_{33}\boldsymbol{g}_3 && (2.65)\\
&\ldots\ldots\ldots\ldots\ldots\ldots && (2.66)\\
\boldsymbol{e}_n &= r_{1n}\boldsymbol{g}_1 + r_{2n}\boldsymbol{g}_2 + \cdots + r_{nn}\boldsymbol{g}_n.
\end{aligned}$$

[5] We would like to point out that in the general unitary space the normalization can be achieved in the infinite number of ways:$\boldsymbol{g}_n = \exp(i\phi)\boldsymbol{h}_n/\|\boldsymbol{h}_n\|$.

If we denote as $\boldsymbol{A} = [\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n]$ the matrix whose columns are the vectors of the original basis, then we can present the previous equation in the following matrix form

$$\boldsymbol{A} = [\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = \boldsymbol{Q}_n \boldsymbol{R}_n \tag{2.67}$$

where we denoted as $\boldsymbol{Q}_n = [\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n]$ and

$$\boldsymbol{R}_n = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} \tag{2.68}$$

is an upper triangular matrix. Since, by construction, the vectors $\boldsymbol{g}_i$ represent an orthonormal basis

$$\boldsymbol{g}_i^H \boldsymbol{g}_j = \delta_{ij}. \tag{2.69}$$

In other words, the Hermite transpose of $\boldsymbol{Q}_n$ is the inverse matrix: $\boldsymbol{Q}_n^H = \boldsymbol{Q}_n^{-1}$. The matrices that possess this property are called the *unitary* matrices. The real unitary matrices are called the *orthogonal* matrices, they are defined by $\boldsymbol{Q}_n^\top = \boldsymbol{Q}_n^{-1}$.

Thus, in the matrix form, the Gram-Schmidt orthogonalization is represented by the so-called the *QR-decomposition*

$$\boldsymbol{A} = \boldsymbol{Q}_n \boldsymbol{R}_n \tag{2.70}$$

where $\boldsymbol{Q}_n$ is a unitary matrix and $\boldsymbol{R}_n$ is the upper triangular matrix.

The direct Gram-Schmidt orthogonalization is in general an unstable process which is prone to propagation of errors, because in each step we divide by the norm of a vector which can be small. On the other hand, if we can find a stable algorithm for the QR-decomposition, then the columns of $\boldsymbol{Q}_n$ will give us the desired orthonormal basis. These algorithms are considered in [3].

Note that, in our development, the diagonal elements of $\boldsymbol{R}_n$ are all positive. However, in general, the QR-decomposition by other algorithms might not satisfy this condition. Geometrically, this corresponds to selecting the vector $-\boldsymbol{g}_i$ instead of $\boldsymbol{g}_i$. By changing the direction of the corresponding vectors we obtain the QR-decomposition with with positive diagonal elements r_{ii}. Thus, if we obtained the QR-decomposition using a general algorithm, we must write

$$\|\boldsymbol{e}_i\| = |r_{ii}|. \tag{2.71}$$

The Matlab `qr()` function performs the QR decomposition.

Example 2.5: Let $\boldsymbol{a} = (1, 2, 0)$, $\boldsymbol{b} = (2, -1, 3)$, $\boldsymbol{c} = (-1, 1, -1)$.

The Matlab function `qr(A)`, where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -1 & 1 \\ 0 & 3 & -1 \end{bmatrix}$$

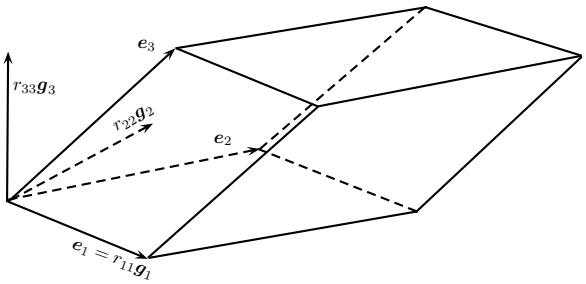

Figure 2.1 Gram-Schmidt orthonormalization.

gives us

$$\boldsymbol{Q} = \begin{bmatrix} -0.4472 & 0.5345 & -0.7171 \\ -0.8944 & -0.2673 & 0.3586 \\ 0 & 0.8018 & 0.5976 \end{bmatrix}, \quad \boldsymbol{R} = \begin{bmatrix} -2.2361 & 0 & -0.4472 \\ 0 & 3.7417 & -1.6036 \\ 0 & 0 & 0.4781 \end{bmatrix}.$$

As we can see $r_{11} = -2.2361$ is negative. Multiplying first row of $\boldsymbol{Q}$ by (-1), we obtain the QR-decomposition

$$\boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{R}_1 = \begin{bmatrix} 0.4472 & -0.5345 & 0.7171 \\ -0.8944 & -0.2673 & 0.3586 \\ 0 & 0.8018 & 0.5976 \end{bmatrix} \begin{bmatrix} 2.2361 & 0 & -0.4472 \\ 0 & 3.7417 & -1.6036 \\ 0 & 0 & 0.4781 \end{bmatrix}.$$

in which all the diagonal elements of $\boldsymbol{R}_1$ are positive.

Note that if the original vectors are linearly dependent, the QR-decomposition is still possible (with some $r_{ii} = 0$). In this case, the QR-decomposition allows us to obtain the basis of the column space of the matrix $\boldsymbol{A}$.

2.3.3.3 Geometrical interpretation of the Gram-Schmidt orthogonalization

Suppose that vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ from $\mathbb{R}^3$ are linearly independent. Then, since $\boldsymbol{g}_1$ is a unit vector, it follows from (2.64) that $\|\boldsymbol{e}_1\| = |r_{11}|$. Equation (2.65) decomposes $\boldsymbol{e}_2$ into two orthogonal components. Since $\boldsymbol{g}_2$ is a unit vector, $|r_{22}|$ is the altitude of the parallelogram whose sides are the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. Therefore, the area of this parallelogram $V_2 = |r_{11}r_{22}|$. Equation (2.66) tells us that r_{33} is the altitude of the parallelepiped with the sides $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$. Thus, its volume is $V_3 = |r_{11}r_{22}r_{33}|$ (see Fig. 2.3.3.3).

Based on these results, it is natural to define a volume of the parallelepiped whose sides represent linearly independent vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ as

$$\boxed{V_n = |r_{11}r_{22}\cdots r_{nn}|.} \tag{2.72}$$

where r_{ii} are the coefficients of $\boldsymbol{g}_i$ in the Gram-Schmidt decomposition of $\boldsymbol{e}_i$. This definition gives us the length of the vector in $\mathbb{R}^1$, the area of the parallelogram in $\mathbb{R}^2$, and the volume of the parallelepiped in $\mathbb{R}^3$. I also agrees with the general definition of measure in the Euclidean space.

To prove that this definition does not depend on the order of selecting vectors for the orthonormalization, we use equation (2.70) according to which $\det \boldsymbol{A}_n = \det \boldsymbol{U}_n \det \boldsymbol{R}_n$. Since $|\det \boldsymbol{U}_n| = 1$ and $\det \boldsymbol{R}_n = r_{11}r_{22}\cdots r_{nn}$,

we conclude that

$$\boxed{V_n = |\det \boldsymbol{A}_n|} \tag{2.73}$$

i.e. the volume of the parallelepiped whose sides are the columns of the matrix $\boldsymbol{A}_n$ is the absolute value of the determinant of the matrix.

2.3.4 Linear Operator

Let $\mathcal{X}$ and $\mathcal{Y}$ be two linear spaces.

Definition 2.7. *A linear operator $\mathcal{Y} = \tilde{\boldsymbol{A}}\mathcal{X}$ is a rule that maps every vector $\boldsymbol{x} \in \mathcal{X}$ to a vector $\boldsymbol{y} \in \mathcal{Y}$ satisfying the following conditions*

1. $\tilde{\boldsymbol{A}}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \tilde{\boldsymbol{A}}\boldsymbol{x}_1 + \tilde{\boldsymbol{A}}\boldsymbol{x}_2$ *for any* $\boldsymbol{x}_1 \in \mathcal{X}$ *and* $\boldsymbol{x}_2 \in \mathcal{X}$*;*
2. $\tilde{\boldsymbol{A}}(\lambda\boldsymbol{x}) = \lambda\tilde{\boldsymbol{A}}\boldsymbol{x}$ *for any* $\boldsymbol{x} \in \mathcal{X}$ *and number* λ*.*

The space of all $\boldsymbol{y} = \tilde{\boldsymbol{A}}\boldsymbol{x}$, $\boldsymbol{x} \in \mathcal{X}$ is called the $operator\ range$ or the *image* of $\mathcal{X}$ and is denoted as

$$\mathcal{R}(\tilde{\boldsymbol{A}}) = \left\{ \tilde{\boldsymbol{A}}\boldsymbol{x} : \boldsymbol{x} \in \mathcal{X} \right\}. \tag{2.74}$$

The dimension the operator range is called the *operator rank*

$$\mathrm{Rank}(\tilde{\boldsymbol{A}}) \triangleq \dim \mathcal{R}(\tilde{\boldsymbol{A}}) \tag{2.75}$$

The space of all $\boldsymbol{x} \in \mathcal{X}$ that the operator $\tilde{\boldsymbol{A}}$ maps onto the null-space is called the $operator\ kernel$ and is denoted as

$$\mathcal{N}(\tilde{\boldsymbol{A}}) = \left\{ \boldsymbol{x} \in \mathcal{X} : \tilde{\boldsymbol{A}}(\boldsymbol{x}) = 0 \right\}. \tag{2.76}$$

The dimension the operator kernel is called the *operator nullity*

$$\mathrm{Null}(\tilde{\boldsymbol{A}}) \triangleq \dim \mathcal{N}(\tilde{\boldsymbol{A}}) \tag{2.77}$$

The rank and nullity of an operator in $\mathbb{C}^n$ are related by [3, p.50]

$$\mathrm{Rank}(\tilde{\boldsymbol{A}}) + \mathrm{Null}(\tilde{\boldsymbol{A}}) = \dim \mathcal{X} \tag{2.78}$$

The $\tilde{\boldsymbol{A}}$ is invertible if there exists an operator $\tilde{\boldsymbol{A}}^{-1}$ with domain $\mathcal{Y}$ and range $\mathcal{X}$ such that

$$\tilde{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{y} \text{ if and only if } \tilde{\boldsymbol{A}}^{-1}\boldsymbol{y} = \boldsymbol{x} \text{ for any } \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in \mathcal{Y} \tag{2.79}$$

The operator $\tilde{\boldsymbol{A}}^{-1}$ is called the inverse to $\tilde{\boldsymbol{A}}$.

2.3.4.1 Matrix of a Linear Operator

Let $\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n$ and $\boldsymbol{g}_1, \boldsymbol{e}_2, ..., \boldsymbol{g}_m$ be bases of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then any vector $\boldsymbol{x} \in \mathcal{V}$ can be expressed as

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n \tag{2.80}$$

and

$$\boldsymbol{y} = \tilde{\boldsymbol{A}}\boldsymbol{x} = x_1\tilde{\boldsymbol{A}}\boldsymbol{e}_1 + x_2\tilde{\boldsymbol{A}}\boldsymbol{e}_2 + \cdots + x_n\tilde{\boldsymbol{A}}\boldsymbol{e}_n \tag{2.81}$$

Thus, it is sufficient to define images of basis vectors to define a linear operator. Since $\tilde{\boldsymbol{A}}\boldsymbol{e}_i \in \mathcal{V}_1$, we can write

$$\begin{aligned}
\tilde{\boldsymbol{A}}\boldsymbol{e}_1 &= a_{11}\boldsymbol{g}_1 + a_{21}\boldsymbol{g}_2 + \cdots + a_{m1}\boldsymbol{g}_m \\
\tilde{\boldsymbol{A}}\boldsymbol{e}_2 &= a_{12}\boldsymbol{g}_1 + a_{22}\boldsymbol{g}_2 + \cdots + a_{m2}\boldsymbol{g}_m \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\tilde{\boldsymbol{A}}\boldsymbol{e}_n &= a_{1n}\boldsymbol{g}_1 + a_{2n}\boldsymbol{g}_2 + \cdots + a_{mn}\boldsymbol{g}_m
\end{aligned} \tag{2.82}$$

The matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2.83}$$

whose columns, according to equation (2.82), are coordinates of the images of the basis vectors is called the *matrix of linear operator*.

Substituting (2.82) into (2.81) we obtain

$$\boldsymbol{y} = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\boldsymbol{g}_i = \sum_{i=1}^{m} \boldsymbol{g}_i \sum_{j=1}^{n} a_{ij}x_j \tag{2.84}$$

Thus, vector $\boldsymbol{y}$ has the following coordinates

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, ..., m \tag{2.85}$$

This equation represents the relationship between the coordinates of the image $\boldsymbol{y}$ through the coordinates of *its pre-image* $\boldsymbol{x}$ and can be expressed in the following matrix form

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{2.86}$$

Thus, if the bases of both spaces are fixed, every linear operator has a unique matrix and every matrix defines a unique linear operator. This relationship is similar to the relationship between a vector and its coordinates.

In a special case, when an operator maps $\mathcal{X}$ onto itself, both bases are the same (i.e. $\boldsymbol{g}_i = \boldsymbol{e}_i$), the matrix of an operator is a *square matrix*.

Example 2.6: The projection operator Let us find a matrix of the operator that projects every vector of $\mathbb{R}^2$ onto some line in the orthonormal basis $\boldsymbol{e}_1 = \mathbf{i}$, $\boldsymbol{e}_2 = \boldsymbol{j}$. Suppose that the line has angle α with the vector $\boldsymbol{e}_1$.

We need to find the coordinates of the images of the basis vectors. $\tilde{\boldsymbol{A}}\boldsymbol{e}_1$ is the projection of $\boldsymbol{e}_1$ onto the line that is the vector along the line whose length is $\cos\alpha$. Its coordinates in the selected basis are its projections onto the basis vectors:

$$\cos\alpha(\cos\alpha, \sin\alpha) = (\cos^2\alpha, \cos\alpha\sin\alpha)$$

Thus,

$$\tilde{\boldsymbol{A}}\boldsymbol{e}_1 = \cos^2\alpha\ \boldsymbol{e}_1 + \cos\alpha\sin\alpha\ \boldsymbol{e}_2$$

. We find similarly

$$\tilde{\boldsymbol{A}}\boldsymbol{e}_2 = \cos\alpha \sin\alpha\, \boldsymbol{e}_1 + \sin^2\alpha\, \boldsymbol{e}_2$$

and, therefore, the matrix of the operator in the selected basis is

$$\boldsymbol{A} = \begin{bmatrix} \cos^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & \sin^2\alpha \end{bmatrix}$$

Note that if we select a basis vector $\boldsymbol{e}_1^*$ as a unit vector along the projection line and $\boldsymbol{e}_2^*$ as a unit vector perpendicular to it, then we would have

$$\tilde{\boldsymbol{A}}\boldsymbol{e}_1^* = \boldsymbol{e}_1^* \quad \text{and} \quad \tilde{\boldsymbol{A}}\boldsymbol{e}_2^* = \boldsymbol{0}$$

Thus, in this basis, the matrix of the projection operator is

$$\boldsymbol{A}^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

2.3.5 Transformation of Coordinates

Let us find the relationship between the coordinates of a vector $\boldsymbol{x}$ in different bases. Suppose that along with the "old" basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ of $\mathbb{C}^n$ we select a "new" basis $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n$. Then we can write

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n$$

$$\boldsymbol{x} = x_1^*\boldsymbol{g}_1 + x_2^*\boldsymbol{g}_2 + \cdots + x_n^*\boldsymbol{g}_n$$

Thus, we have

$$\sum_{i=1}^{n} x_i\boldsymbol{e}_i = \sum_{j=1}^{n} x_j^*\boldsymbol{g}_j \tag{2.87}$$

The vectors of the new basis can be expresses as linear combinations of the vectors of the old basis:

$$\boldsymbol{g}_j = \sum_{i=1}^{n} t_{ij}\boldsymbol{e}_i \;\; i = 1, 2, \ldots, n. \tag{2.88}$$

Substituting these equations into (2.87) and grouping coefficients of the same basis vectors, we obtain

$$\sum_{i=1}^{n} x_i\boldsymbol{e}_i = \sum_{i=1}^{n} \boldsymbol{e}_i \sum_{j=1}^{n} t_{ij}x_j^* \tag{2.89}$$

Since basis vectors are linearly independent, the coefficients of the same basis vectors on both sides of this equation must be equal:

$$x_i = \sum_{j=1}^{n} t_{ij}x_j^*. \tag{2.90}$$

This equation expresses the coordinates of the vector in the old basis through the coordinates of the vector in the new basis. It can be written in the following matrix form

$$\boldsymbol{x} = \boldsymbol{T}\boldsymbol{x}^* \tag{2.91}$$

where, according to equation (2.88), $\boldsymbol{T} = [t_{ij}]_{n\times n}$ is the matrix whose columns are the coordinates of the new basis vectors in the old basis. Since the matrix $\boldsymbol{T}$ is non-singular, we can express the new coordinates through the old ones as

$$\boldsymbol{x}^* = \boldsymbol{T}^{-1}\boldsymbol{x}. \tag{2.92}$$

Example 2.7: Let $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ be two orthogonal unit vectors in the plane ($\mathbb{R}^2$) and the new basis $\boldsymbol{g}_1$ and $\boldsymbol{g}_1$ is obtained by rotating the old basis counter clockwise by angle α. Then it is easy to see that

$$\boldsymbol{g}_1 = \quad \cos\alpha\, \boldsymbol{e}_1 + \sin\alpha\, \boldsymbol{e}_2, \tag{2.93}$$

$$\boldsymbol{g}_2 = -\sin\alpha\, \boldsymbol{e}_1 + \cos\alpha\, \boldsymbol{e}_2. \tag{2.94}$$

Therefore,

$$\boldsymbol{T} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

and we have the following relations between the old an new coordinates

$$x_1 = x_1^* \cos\alpha - x_2^* \sin\alpha, \tag{2.95}$$

$$x_2 = x_1^* \sin\alpha + x_2^* \cos\alpha. \tag{2.96}$$

2.3.6 Similarity Transformation

Suppose that we have a linear operator in the matrix form $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the column matrices of coordinates of the corresponding vectors in some basis. We would like to find the relationship between the coordinates of these vectors in some other basis.

Using equation (2.90), we can write

$$\boldsymbol{T}\boldsymbol{y}^* = \boldsymbol{A}\boldsymbol{T}\boldsymbol{x}^* \tag{2.97}$$

or

$$\boldsymbol{y}^* = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}\boldsymbol{x}^*. \tag{2.98}$$

It follows from this equation that the matrix of the operator in the new basis has the form

$$\boldsymbol{A}^* = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}. \tag{2.99}$$

The formula in the RHS of this equation is called the *similarity transformation* of the matrix $\boldsymbol{A}$. Thus, we can say that matrices of the same operator in different bases are similar to each other.

2.3.6.1 Geometrical meaning of an operator determinant

Since

$$\det \boldsymbol{A}^* = \det \boldsymbol{T}^{-1} \det \boldsymbol{A} \det \boldsymbol{T} = \frac{1}{\det \boldsymbol{T}} \det \boldsymbol{A} \det \boldsymbol{T} = \det \boldsymbol{A}$$

we conclude that the determinant of the linear operator matrix does not depend on the selected basis and, therefore, can be called the *operator determinant*.

Suppose that vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are mapped onto vectors $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$ by the linear operator with the matrix $\boldsymbol{A}$ in some basis. We can express this as

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}, \text{ where } \boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n], \boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n]. \tag{2.100}$$

Since the determinant of the matrix product equals to the product of the determinants, we can write

$$\det \boldsymbol{Y} = \det \boldsymbol{A} \det \boldsymbol{X} \tag{2.101}$$

As we know, the absolute value of the determinant of a matrix equals to the volume of the parallelepiped whose sides are the columns of the matrix. Thus, it follows from the previous equation that

$$V_y = |\det \boldsymbol{A}| V_x \tag{2.102}$$

Thus, we can say that the absolute value of the determinant of an operator $\tilde{\boldsymbol{A}}$ is the coefficient

$$|\det \boldsymbol{A}| = \frac{V_y}{V_x} \tag{2.103}$$

of contraction (expansion) of the volume of the parallelepiped after mapping by the operator. We use this result in developing the variable substitution in multidimensional integrals.

In the two-dimensional space the volumes are replaced by the areas and the previous formula takes the form

$$|\det \boldsymbol{A}| = \frac{S_y}{S_x} \tag{2.104}$$

We can now ask a question: In which basis does the matrix of a linear operator has the simplest form? The answer to this question will be given in the next section.

2.3.7 Eigenvalues and Eigenvectors

Definition 2.8. *A vector $\boldsymbol{x} \neq 0$ is called the eigenvector (or characteristic vector) of a linear operator $\tilde{\boldsymbol{A}}$ if there is a number λ called eigenvalue (or characteristic number) the such that*

$$\tilde{\boldsymbol{A}}\boldsymbol{x} = \lambda \boldsymbol{x}. \tag{2.105}$$

In other words, an operator transforms its eigenvector $\boldsymbol{x}$ into its collinear vector $\lambda \boldsymbol{x}$.

To find the eigenvectors, we rewrite equation (2.105) in a coordinate form in some basis

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = 0. \tag{2.106}$$

As we know, this homogeneous equation has a non-trivial solution if and only if its determinant is equal to zero:

$$\Delta(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{2.107}$$

This equation is called the *characteristic equation* for the matrix $\boldsymbol{A}$. As we can see, it is order-n algebraic equation.

Definition 2.9. *The multiplicity k_i of the root λ_i of the characteristic polynomial $\Delta(\lambda)$ is called the algebraic multiplicity of the eigenvalue.*

This means that the characteristic polynomial can be factored as

$$\Delta(\lambda) = (\lambda - \lambda_i)^{k_i} q(\lambda),\ q(\lambda_i) \neq 0. \tag{2.108}$$

Definition 2.10. *The maximal number of linearly independent eigenvectors corresponding to the eigenvalue λ_i is called its geometric multiplicity.*

According to equation (2.106), the geometric multiplicity of λ_i is equal to the dimension of the null space: $\dim \mathcal{N}(\boldsymbol{A} - \lambda_i \boldsymbol{I}) = \mathrm{Null}(\boldsymbol{A} - \lambda_i \boldsymbol{I}) = n - \mathrm{Rank}(\boldsymbol{A} - \lambda_i \boldsymbol{I})$.

An operator is said to have a *simple structure* [2] in $\mathbb{C}^n$ if it has n linearly independent eigenvectors. Let an operator opA have a simple structure with linearly independent eigenvectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ corresponding to eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. We can use these vectors as basis vectors of $\mathbb{C}^n$. In this basis, which is called an *eigenbasis*, because of the relationship $\tilde{\boldsymbol{A}}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$, the operator has a diagonal matrix

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \tag{2.109}$$

Since matrices of the same operator are similar, we can write

$$\boldsymbol{\Lambda} = \boldsymbol{T}^{-1} \boldsymbol{A} \boldsymbol{T}. \tag{2.110}$$

A matrix of a simple structure operator is called a *simple structure matrix*. Thus, we can say that a simple structure matrix is similar to a diagonal matrix.

The converse statement is also true. Indeed, equation (2.110) can be written as

$$\boldsymbol{T}\boldsymbol{\Lambda} = \boldsymbol{A}\boldsymbol{T}. \tag{2.111}$$

which shows that the columns of the matrix $\boldsymbol{T}$ are the eigenvectors of $\boldsymbol{A}$. Since $\boldsymbol{T}^{-1}$ exists, $\det \boldsymbol{T} \neq 0$ so that the eigenvectors are linearly independent. Hence, a matrix has a simple structure if and only if it is *diagonalizable* (i.e. is similar to a diagonal matrix).

Equation (2.110) can be rewritten as

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1} = \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i \boldsymbol{y}_i^{\top} \tag{2.112}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ denote columns of $\boldsymbol{T}$ and $\boldsymbol{T}^{-1}$, respectively. Since $\boldsymbol{T} \cdot \boldsymbol{T}^{-1} = \boldsymbol{I}$, we have

$$\boldsymbol{x}_i \boldsymbol{y}_j^{\top} = \delta_{ij} \tag{2.113}$$

which means that the vectors are orthogonal if they are real. However, they are not orthogonal in the general case since the inner product of the complex vectors is $\boldsymbol{x} \cdot \boldsymbol{y}^H$. Equation (2.112) is called the *spectral decomposition* or *eigendecomposition* of the matrix $\boldsymbol{A}$ that has a simple structure.

Note, that if a simple structure matrix $\boldsymbol{A}$ is a real, the spectral decomposition in the real space $\mathbb{R}^n$ is possible only if all the eigenvalues are real.

Consider now some conditions for the matrices to have a simple structure [1, p.98].

Theorem 2.4. *A matrix has a simple structure if and only if for each eigenvalue λ_i its algebraic multiplicity k_i equals its geometric multiplicity* $\mathrm{Null}(\boldsymbol{A} - \lambda_i \boldsymbol{I}) = n - \mathrm{Rank}(\boldsymbol{A} - \lambda_i \boldsymbol{I})$.

Corollary 2.1. *If all the eigenvalues of a matrix are different, the matrix has a simple structure.*

Proof. In this case $k_i = 1$, $i = 1, 2, ..., n$, and, therefore, $\mathrm{Null}(\boldsymbol{A} - \lambda_i \boldsymbol{I}) = k_i = 1$ and we have n linearly independent eigenvectors.□ □

Note, that this corollary gives only a *sufficient* condition for the matrix to have a simple structure. The converse is not true as we saw in the previous theorem. If a matrix has a simple structure, some eigenvalues can be the same as long as the conditions of theorem 2.4 are satisfied.

Theorem 2.5. *A matrix has a simple structure if and only if, for each eigenvalue λ_i of algebraic multiplicity k_i, the* $\mathrm{Rank}\,(\boldsymbol{A} - \lambda_i \boldsymbol{I}) = n - k_i$.

Proof. All eigenvectors corresponding to λ_i represent the null space $N(\boldsymbol{A} - \lambda_i \boldsymbol{I})$ whose dimension is $n - r_i$ where $r_i = \mathrm{Rank}(\boldsymbol{A} - \lambda_i \boldsymbol{I})$. According to the previous theorem, $\boldsymbol{A}$ has a simple structure if and only if this dimension is k_i. Thus, $n - r_i = k_i$.□ □

Example 2.8: Let

$$\boldsymbol{A} = \begin{bmatrix} 0.5 & 0.5 \\ 0.25 & 0.75 \end{bmatrix} \tag{2.114}$$

Its characteristic equation is

$$\det[\boldsymbol{A} - \lambda \boldsymbol{I}] = \det \begin{bmatrix} 0.25 - \lambda & 0.75 \\ 0.25 & 0.75 - \lambda \end{bmatrix} = \lambda^2 - 0.75\lambda - 0.25 = 0 \tag{2.115}$$

It has two solutions (eigenvalues) $\lambda_1 = 0.25$ and $\lambda_2 = 1$. Since the eigenvalues are different, the matrix has a simple structure. The corresponding eigenvectors are found from (2.106). For λ_1 this system takes the form

$$\begin{bmatrix} 0.5 - \lambda_1 & 0.5 \\ 0.25 & 0.75 - \lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.25 & 0.5 \\ 0.25 & 0.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \tag{2.116}$$

This system is equivalent to one equation

$$0.25x_1 + 0.5x_2 = 0 \tag{2.117}$$

This equation has an infinite number of solutions. One of them is $\boldsymbol{x}_1 = (2, -1)$. The eigenvector corresponding to λ_2 is found similarly: $\boldsymbol{x}_2 = (1, 1)$. Thus,

$$\boldsymbol{T} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}, \boldsymbol{T}^{-1} = \frac{1}{3}\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}. \tag{2.118}$$

The spectral decomposition has the following form

$$\boldsymbol{A} = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 0.25 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 2 \\ -1 \end{bmatrix}[1\ -1] + \frac{1}{3}\begin{bmatrix} 1 \\ 1 \end{bmatrix}[1\ 2] \tag{2.119}$$

2.3.8 Jordan Canonical Form

If a matrix $\boldsymbol{A}$ does not have a simple structure, there is not enough linearly independent eigenvectors to construct an eigenbasis of $\mathbb{C}^n$. Let λ_i, $i = 1, 2, ..., m$ be all different eigenvalues of $\boldsymbol{A}$ and we need to use some other vectors to construct a basis of $\mathbb{C}^n$.

Definition 2.11. ***(Generalized eigenvector)*** *Vector $\boldsymbol{x}$ is called the generalized eigenvectors of k-th order of an operator $\tilde{\boldsymbol{A}}$ if there is a number λ such that*

$$(\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})^{k-1}\boldsymbol{x} \neq 0 \quad (\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})^{k}\boldsymbol{x} = 0 \tag{2.120}$$

It is clear that previously defined eigenvectors are also generalized eigenvectors of the first order:

$$(\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})^{0}\boldsymbol{x} = \boldsymbol{x} \neq 0 \quad (\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})\boldsymbol{x} = 0. \tag{2.121}$$

If an operator $\tilde{\boldsymbol{A}}$ does not have a simple structure, its matrix is not diagonalizable. In this case, the simplest matrix that the operator can have is the so-called the *Jordan canonical form*. This matrix is obtained if we select the generalized eigenvectors of $\tilde{\boldsymbol{A}}$ to construct a basis of $\mathbb{C}^n$. Denote the pre-image of $\boldsymbol{x}$ in equation (2.120) as $(\boldsymbol{x}^{(1)} : \boldsymbol{x} = (\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})\boldsymbol{x}^{(1)}$, the pre-image of $\boldsymbol{x}^{(1)}$ as $\boldsymbol{x}^{(2)} : \boldsymbol{x}^{(1)} = (\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})\boldsymbol{x}^{(2)}$, and so on. Thus we have the sequence $\boldsymbol{x}, \boldsymbol{x}^{(1)}, ..., \boldsymbol{x}^{(k-1)}$ of vectors constituting the *Jordan chain of vectors* satisfying

$$(\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})\boldsymbol{x} = 0,\ (\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})\boldsymbol{x}^{(1)} = \boldsymbol{x}, \ldots, (\tilde{\boldsymbol{A}} - \lambda\boldsymbol{I})\boldsymbol{x}^{(k-1)} = \boldsymbol{x}^{(k-2)} \tag{2.122}$$

After opening parentheses in these equations we obtain

$$\tilde{\boldsymbol{A}}\boldsymbol{x} = \lambda\boldsymbol{x},\ \tilde{\boldsymbol{A}}\boldsymbol{x}^{(1)} = \lambda\boldsymbol{x}^{(1)} + \boldsymbol{x}, \ldots, \tilde{\boldsymbol{A}}\boldsymbol{x}^{(k-1)} = \lambda\boldsymbol{x}^{(k-1)} + \boldsymbol{x}^{(k-2)} \tag{2.123}$$

If we include these vectors into a basis of $\mathbb{C}^n$, then, according to definition of the matrix of a linear operator, it will have the following form

$$\begin{bmatrix} \boldsymbol{J} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{D} \end{bmatrix} \tag{2.124}$$

where

$$\boldsymbol{J} = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix} \tag{2.125}$$

is called a *Jordan block* of the matrix.

If we can find enough Jordan chains such that their vectors can be selected as a basis of the $\mathbb{C}^n$, then, in this basis, the matrix of the operator (according to equation (2.124)) will be a block-diagonal matrix, called the *Jordan canonical form*, whose diagonal blocks are the Jordan blocks (2.125).

In other words, a basis of $\mathbb{C}^n$ can be constructed using Jordan chains of an operator $\tilde{\boldsymbol{A}}$. The structure of this basis (and, therefore, the corresponding Jordan canonical form) can be described as follows For every eigenvalue λ_i of the algebraic multiplicity k_i,

- there are
$$h_1 = n - 2\,\mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I}) + \mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^2 \quad \text{blocks of size } 1, \tag{2.126}$$
- there are
$$h_2 = \mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I}) - 2\,\mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^2 + \mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^3 \quad \text{second} - \text{order blocks} \tag{2.127}$$
and so on,
- there are
$$h_k = \mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^{k-1} - 2\,\mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^k + \mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^{k+1} \quad \text{blocks of size } k \tag{2.128}$$
- there are
$$h_{m_i+1} = \mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^{m_i-1} - \mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^{m_i} \text{ blocks of the largest size } m_i + 1 \tag{2.129}$$
where m_i is the smallest number such that $\mathrm{Rank}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})^{m_i+1} = \mathrm{Rank}(\tilde{\boldsymbol{A}} = \lambda_i \boldsymbol{I})^{m_i+2}$.

For each eigenvalue λ_i, the total number of the Jordan chains equals to the geometric multiplicity of the eigenvalue:

$$g_i = \sum_{j=1}^{m_i+1} h_j = n - r_i, \tag{2.130}$$

and the total number of vectors in these chains equals to the algebraic multiplicity of λ_i:

$$k_i = \sum_{j=1}^{m_i+1} (m_i + 1)\, h_j. \tag{2.131}$$

Thus, we proved that the number of Jordan chains corresponding to the eigenvalue λ is equal to its geometric multiplicity g and the total number of vectors in these chains equals to the algebraic multiplicity k of λ. As we have seen before, there is a Jordan matrix block corresponding to each Jordan chain of vectors. We construct Jordan chains for all different eigenvalues. It is not difficult to prove that these vectors are linearly independent. Since the sum of all the algebraic multiplicities of the eigenvalues equals to the degree n of the characteristic

polynomial (which is also the dimension of $\mathbb{C}^n$) we can always construct the *Jordan basis of the* $\mathbb{C}^n$ using the Jordan chains of vectors. The matrix of $\tilde{\boldsymbol{A}}$ in this basis is the Jordan canonical form which is a block-diagonal matrix whose diagonal blocks have the form of equation (2.125):

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_p \end{bmatrix}, \text{ where } \boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix} \tag{2.132}$$

where p is the number of Jordan chains which is equal to the maximal number of the linearly independent eigenvectors of $\tilde{\boldsymbol{A}}$.

Thus, every matrix is similar to a Jordan canonical form:

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1} \text{ or } \boldsymbol{J} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} \tag{2.133}$$

The columns of the transformation matrix $\boldsymbol{T}$, according to equation (2.91), are the coordinates of the vectors of the Jordan basis in the "old" basis in which $\boldsymbol{A}$ is the matrix of the operator $\tilde{\boldsymbol{A}}$. Thus, to obtain the matrix, we need to find all the generalized eigenvectors of the matrix. Alternatively, if we know the structure of the Jordan canonical form (e.g. by using equations (2.126-2.129)), we can obtain the matrix by solving for $\boldsymbol{T}$ equation (2.133) which is a non-zero solution of

$$\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}\boldsymbol{J}. \tag{2.134}$$

Example 2.9: Let

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{bmatrix} \tag{2.135}$$

Its characteristic equation

$$\Delta(\lambda) = \det(\boldsymbol{A} - \lambda\boldsymbol{I}) = \det \begin{bmatrix} 3-\lambda & 2 & -3 \\ 4 & 10-\lambda & -12 \\ 3 & 6 & -7-\lambda \end{bmatrix} = 0 \tag{2.136}$$

after simplifications becomes $(2-\lambda)^3 = 0$. Its root $\lambda = 2$ has the algebraic multiplicity $k = 3$. We have

$$r = \mathrm{Rank}(\boldsymbol{A} - 2\boldsymbol{I}) = \mathrm{Rank} \begin{bmatrix} 1 & 2 & -3 \\ 4 & 8 & -12 \\ 3 & 6 & -9 \end{bmatrix} = 1 \tag{2.137}$$

because all the columns of this matrix are proportional to the first column. Since the geometric multiplicity of λ is $g = \dim \mathcal{N}(\boldsymbol{A} - 2\boldsymbol{I}) = n - r = 2 < k = 3$, this matrix has only two linearly independent eigenvectors and, therefore, does not have a simple structure. Because $(\boldsymbol{A} - 2\boldsymbol{I})^2 = 0$, $\mathrm{Rank}(\boldsymbol{A} - 2\boldsymbol{I})^2 = 0$. Then, according to (2.126) and (2.127),

$$h_1 = 3 - 2 + 0 = 1$$

and

$$h_2 = 1 - 2 \cdot 0 + 0 = 1$$

. The corresponding Jordan canonical form contains one size-one block and one size-two block:

$$\boldsymbol{J} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

To find the transformation matrix $\boldsymbol{T}$, we need to find two Jordan chains of sizes two and one. The length-two chain vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_1^{(1)}$, representing two first columns of $\boldsymbol{T}$, satisfy equation (2.122) which in our case takes the form

$$(\tilde{\boldsymbol{A}} - 2\boldsymbol{I})\boldsymbol{t}_1 = 0, \ (\tilde{\boldsymbol{A}} - 2\boldsymbol{I})\boldsymbol{t}_1^{(1)} = \boldsymbol{t}_1 \tag{2.138}$$

These equations can be written in the following matrix form

$$\begin{bmatrix} \boldsymbol{A} - 2\boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{I} & \boldsymbol{A} - 2\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_1^{(1)} \end{bmatrix} = 0 \ \text{ or } \ \begin{bmatrix} 1 & 2 & -3 & 0 & 0 & 0 \\ 4 & 8 & -12 & 0 & 0 & 0 \\ 3 & 6 & -9 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 2 & -3 \\ 0 & -1 & 0 & 4 & 8 & -12 \\ 0 & 0 & -1 & 3 & 6 & -9 \end{bmatrix} \begin{bmatrix} t_{11} \\ t_{21} \\ t_{31} \\ t_{12} \\ t_{22} \\ t_{32} \end{bmatrix} = 0 \tag{2.139}$$

The REF of this system (which can be obtained using the Matlab function `rref()`) is

$$\begin{aligned} t_{11} - t_{12} - 2t_{22} + 3t_{32} &= 0 \\ t_{21} - 4t_{12} - 8t_{22} + 12t_{32} &= 0 \\ t_{31} - 3t_{12} - 6t_{22} + 9t_{32} &= 0 \end{aligned}$$

Here t_{12}, t_{22}, and t_{32} are free variables. To find a non-zero solution, we let $t_{12} = 1$, $t_{22} = 0$, and $t_{32} = 0$ to obtain

$$\boldsymbol{t}_1 = (1, 4, 3), \quad \boldsymbol{t}_1^{(1)} = (1, 0, 0).$$

In order to complete the Jordan basis we need to find an eigenvector $\boldsymbol{t}_2$ which is linearly independent of $\boldsymbol{t}_1$. All the eigenvectors are found as solutions of $(\boldsymbol{A} - 2\boldsymbol{I})\boldsymbol{x} = 0$ which is equivalent to

$$x_1 + 2x_2 - 3x_3 = 0.$$

By setting $x_3 = 0$ we find $\boldsymbol{t}_2 = (-2, 1, 0)$. Thus, we have a Jordan basis $\boldsymbol{e}_1 = \boldsymbol{t}_1$, $\boldsymbol{e}_2 = \boldsymbol{t}_1^{(1)}$, $\boldsymbol{e}_3 = \boldsymbol{t}_2$ and the transformation matrix takes the following form

$$\boldsymbol{T} = \begin{bmatrix} 1 & 1 & -2 \\ 4 & 0 & 1 \\ 3 & 0 & 0 \end{bmatrix} \tag{2.140}$$

The inverse matrix

$$\boldsymbol{T}^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{3} \\ 1 & 2 & -3 \\ 0 & 1 & -\frac{4}{3} \end{bmatrix} \tag{2.141}$$

Equation (2.133) takes the form

$$\boldsymbol{TJT}^{-1} = \begin{bmatrix} 1 & 1 & -2 \\ 4 & 0 & 1 \\ 3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{3} \\ 1 & 2 & -3 \\ 0 & 1 & -\frac{4}{3} \end{bmatrix} = \begin{bmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{bmatrix} \tag{2.142}$$

or

$$\boldsymbol{T}^{-1}\boldsymbol{AT} = \begin{bmatrix} 0 & 0 & \frac{1}{3} \\ 1 & 2 & -3 \\ 0 & 1 & -\frac{4}{3} \end{bmatrix} \begin{bmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 \\ 4 & 0 & 1 \\ 3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{2.143}$$

In conclusion, we would like to point out that there are several other methods for constructing the Jordan canonical form. One of them uses the theory of the elementary polynomials and elementary quotients of polynomial matrices [2], the other method is based on the minimal polynomial of the matrix [1, pp. 151-152].

2.4 Adjoint Operators

Hermitian matrices play an important role in applications (see, for example, Chapter 13 of our book [5]). In this section, we study important properties of Hermitian matrices. These matrices are closely related to the self-ajoint operators which we will study first.

Definition 2.12. *Let $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{B}}$ be linear operators in an unitary space $\mathcal{U}$. If for any two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{U}$*

$$\langle \tilde{\boldsymbol{A}}\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \tilde{\boldsymbol{B}}\boldsymbol{y}\rangle, \tag{2.144}$$

*the operators are called the **adjoint** operators.*

*If $\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{B}}$ the operator is called **self-adjoint**. In this case*

$$\langle \tilde{\boldsymbol{A}}\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \tilde{\boldsymbol{A}}\boldsymbol{y}\rangle, \tag{2.145}$$

The relationship between the self-adjoint operators and Hermitian matrices is established by the following theorem

Theorem 2.6. *An operator $\tilde{\boldsymbol{A}}$ is self-adjoint if and only if its matrix in an orthonormal basis is Hermitian.*

Proof. Let $\boldsymbol{A}$ be the matrix of the operator $\tilde{\boldsymbol{A}}$ in some orthonormal basis. To find the matrix of the adjoint operator, we express the inner product in its coordinate form as

$$\langle \tilde{\boldsymbol{A}}\boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{y}^H\boldsymbol{Ax} = (\boldsymbol{A}^H\boldsymbol{y})^H\boldsymbol{x} = \langle \boldsymbol{x}, \tilde{\boldsymbol{A}}^H\boldsymbol{y}\rangle. \tag{2.146}$$

where we denoted as $\tilde{\boldsymbol{A}}^H$ the adjoint operator because, as we can see, the matrix of the adjoint operator in the orthonormal basis is the Hermite transpose of the matrix $\boldsymbol{A}$. Thus, the operator is self-adjoint if and only if $\boldsymbol{A} = \boldsymbol{A}^H$ i.e. the matrix is Hermitian. □ □

For operators in $\mathbb{R}^n$ the matrix of the adjoint operator is the transpose matrix $\boldsymbol{A}^\top$ ($\boldsymbol{A}^H = \boldsymbol{A}^\top$for real matrices) while the matrix of the self-adjoint operator is symmetric ($\boldsymbol{A} = \boldsymbol{A}^\top$).

It is not difficult to prove [1, p. 116] that

- All eigenvalues of a self-adjoint operator are real.
- Eigenvectors of a self-adjoint operator corresponding to different eigenvalues are orthogonal.
- A self-adjoint operator has a simple structure.

It follows from these properties that any Hermitian matrix is similar to a diagonal matrix (2.109):

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1} \tag{2.147}$$

whose elements are real numbers. Moreover, we can construct an orthonormal eigenbasis of $\mathbb{C}^n$ consisting of eigenvectors of a self-adjoint operator. This can be achieved by constructing orthonormal eigenbases in subspaces $\mathcal{N}(\tilde{\boldsymbol{A}} - \lambda_i \boldsymbol{I})$ of the eigenvectors corresponding to each eigenvalue (for example, using the Gram-Schmidt orthogonalization).

We know that the columns of the transformation matrix $\boldsymbol{T}$ in equation (2.147) are the coordinates of the basis eigenvectors. If the basis is orthonormal, we denote this matrix as $\boldsymbol{U} = [u_{ij}]_{n\times n}$. The orthonormality of the vectors can be expressed as

$$\sum_{k=1}^{n} \bar{u}_{ki} u_{kj} = \boldsymbol{u}_i^H \boldsymbol{v}_j = \delta_{ij} \tag{2.148}$$

where $\boldsymbol{u}_i$ are the columns of the matrix $\boldsymbol{U}$. This equation shows that $\boldsymbol{U}^H$ is the inverse matrix of $\boldsymbol{U}$ (see equation (2.18)). Thus, $\boldsymbol{U}$ is an unitary matrix:

$$\boldsymbol{U}^H = \boldsymbol{U}^{-1} \tag{2.149}$$

Hence, equation (2.147) can be rewritten as

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H \tag{2.150}$$

Matrices satisfying this equation are called *unitary equivalent*. Thus, any Hermitian matrix is unitary equivalent to a diagonal matrix. The spectral decomposition of a Hermitian matrix has the form

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^H \tag{2.151}$$

Real symmetrical matrices represent a special case of the Hermitian matrices. Thus, all the previous results are also valid for the symmetrical matrices in $\mathbb{R}^n$ with the replacement of the Hermite transpose $\boldsymbol{A}^H$ with the real transpose $\boldsymbol{A}^\top$. For the real symmetrical matrices the spectral decomposition (2.151) can be written as

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\top = \sum_{i=1}^{n} \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^\top. \tag{2.152}$$

2.5 Hermitian and Quadratic Forms

Definition 2.13. *A bivariate Hermitian form is a second-order homogeneous polynomial*

$$H(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^H \boldsymbol{A}\boldsymbol{y} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x}_i y_j \tag{2.153}$$

where $\boldsymbol{x} \in \mathbb{C}^n$, $\boldsymbol{y} \in \mathbb{C}^n$, and $\boldsymbol{A}$ is a Hermitian matrix which is called the matrix of the Hermitian form.

According to equation (2.57) the inner product of two vectors is an example of the Hermitian form.

If $\boldsymbol{x} = \boldsymbol{y}$, the Hermitian form

$$H(\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \overline{x}_i x_j \tag{2.154}$$

is called the *univariate Hermitian form* or simply the Hermitian form. This form represents a real function of the complex variables. Indeed, since $H(\boldsymbol{x}, \boldsymbol{x})$ is a number, its transpose is the same number. Using Property 5 (Section 2.1) of the Hermite transpose of the product of matrices, we obtain

$$\overline{H(\boldsymbol{x}, \boldsymbol{x})} = [H(\boldsymbol{x}, \boldsymbol{x})]^H = \boldsymbol{x}^H \boldsymbol{A}^H \boldsymbol{x} = \boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x} = H(\boldsymbol{x}, \boldsymbol{x}). \tag{2.155}$$

Thus, $\overline{H(\boldsymbol{x}, \boldsymbol{x})} = H(\boldsymbol{x}, \boldsymbol{x})$ and, therefore, $H(\boldsymbol{x}, \boldsymbol{x})$ is real.

Definition 2.14. *A bivariate quadratic form is a second-order homogeneous polynomial of real variables*

$$Q(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{y} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j \tag{2.156}$$

where $\boldsymbol{A}$ is a real symmetrical matrix. An univariate quadratic form is defined by

$$Q(\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{2.157}$$

If the matrix is not symmetrical, we can always make it symmetrical by replacing $a_{ij} x_i x_j + a_{ji} x_j x_i$ with $a_{ij}^{(1)} x_i x_j + a_{ij}^{(1)} x_j x_i$, where $a_{ij}^{(1)} = 0.5(a_{ij} + a_{ji})$ without changing the quadratic form.

2.5.0.1 Equivalent Hermitian forms

A linear variable substitution $\boldsymbol{y} = \boldsymbol{T} \boldsymbol{x}$ transforms a Hermitian form as

$$H(\boldsymbol{y}, \boldsymbol{y}) = (\boldsymbol{T} \boldsymbol{x})^H \boldsymbol{A} (\boldsymbol{T} \boldsymbol{x}) = \boldsymbol{x}^H \boldsymbol{T}^H \boldsymbol{A} \boldsymbol{T} \boldsymbol{x} = \boldsymbol{x}^H \boldsymbol{B} \boldsymbol{x}. \tag{2.158}$$

where

$$\boldsymbol{B} = \boldsymbol{T}^H \boldsymbol{A} \boldsymbol{T} \tag{2.159}$$

is the matrix of the Hermitian form in new variables. Matrix $\boldsymbol{B}$ defined by the previous equation is called *congruent* to the matrix $\boldsymbol{A}$.

Two Hermitian forms $H(\boldsymbol{x}, \boldsymbol{x})$ and $H_1(\boldsymbol{x}, \boldsymbol{x})$ are called *equivalent* if there exists a non-singular matrix $\boldsymbol{T}$ such that $H(\boldsymbol{x}, \boldsymbol{x}) = H_1(\boldsymbol{T}\boldsymbol{x}, \boldsymbol{T}\boldsymbol{x})$. In other words, their matrices are related by equation (2.159) where $\det \boldsymbol{T} \neq 0$. In the sequel we will use only non-singular matrices $\boldsymbol{T}$.

It follows from equation (2.159) that

$$\det(\boldsymbol{B} - \lambda \boldsymbol{I}) = \det \boldsymbol{T}^H \det(\boldsymbol{A} - \lambda \boldsymbol{I}) \det \boldsymbol{T} = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) |\det \boldsymbol{T}|^2 \tag{2.160}$$

which means that the eigenvalues of the matrix of the Hermitian form do not change after a linear variable substitution with the non-singular matrix $\boldsymbol{T}$. Therefore, matrices of all equivalent Hermitian forms have the same eigenvalues which are called the *eigenvalues of the Hermitian form*. Since the rank of the matrix is equal to the

number of its non-zero eigenvalues, it is also the same for all equivalent Hermitian forms an it is called the *rank of the Hermitian form.*

If the transformation matrix is a unitary matrix, then the matrices of the equivalent Hermitian forms are *unitary equivalent* and equation (2.159) can be written as

$$\boldsymbol{B} = \boldsymbol{U}^H \boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}. \tag{2.161}$$

Since any Hermitian matrix $\boldsymbol{A}$ is unitary equivalent to a diagonal matrix $\boldsymbol{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$ whose diagonal elements are the eigenvalues of $\boldsymbol{A}$:

$$\boldsymbol{\Lambda} = \boldsymbol{U}^H \boldsymbol{A}\boldsymbol{U}, \tag{2.162}$$

we conclude that any Hermitian form can be transformed to

$$H(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{x}) = \lambda_1|x_1|^2 + \lambda_2|x_2|^2 + \cdots + \lambda_n|x_n|^2. \tag{2.163}$$

The RHS of this equation represents the so-called the *diagonal Hermitian form.*

Note that a unitary transformation of vectors does not change their inner product:

$$(\boldsymbol{U}\boldsymbol{y})^H(\boldsymbol{U}\boldsymbol{x}) = \boldsymbol{y}^H\boldsymbol{U}^H\boldsymbol{U}\boldsymbol{x} = \boldsymbol{y}^H\boldsymbol{I}\boldsymbol{x} = \boldsymbol{y}^H\boldsymbol{x}. \tag{2.164}$$

Therefore, they do not change lengths of the vectors and angles between the vectors and represent linear isometries: rotations, reflections, and their combinations. A quadratic form with the real coefficients is a special case of a Hermitian form, therefore, any quadratic form is orthogonally equivalent to the diagonal quadratic form is

$$Q(\boldsymbol{T}\boldsymbol{x}, \boldsymbol{T}\boldsymbol{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2 \tag{2.165}$$

where $\boldsymbol{T}$ is the orthogonal matrix. Since a transformation with the orthogonal matrix maps an orthogonal basis into an orthogonal basis, this transformation has numerous applications in geometry.

Example 2.10: Consider a second-order curve that is defined by the following equation

$$5x^2 + 4xy + 8y^2 - 32x - 56y + 80 = 0. \tag{2.166}$$

We would like to find the canonical form of the equation of this curve.

The quadratic form in the LHS of equation (2.166) is

$$5x^2 + 4xy + 8y^2 = [x\ y]\begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = [x\ y]\boldsymbol{A}\begin{bmatrix} x \\ y \end{bmatrix}. \tag{2.167}$$

The characteristic polynomial of the matrix of this quadratic form is

$$\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = \begin{bmatrix} 5-\lambda & 2 \\ 2 & 8-\lambda \end{bmatrix} = \lambda^2 - 13\lambda + 36. \tag{2.168}$$

Its roots $\lambda_1 = 4$ and $\lambda_2 = 9$ are the eigenvalues of $\boldsymbol{A}$. Thus, the diagonal form of the quadratic form is $4x_1^2 + 9y_1^2$.

To transform the linear terms of (2.166) to the new coordinates, we need to find the matrix $\boldsymbol{T}$ of the transformation of coordinates. The columns of this matrix are the orthonormal eigenvalues of $\boldsymbol{A}$.

The coordinates of the eigenvectors are found from the equation

$$\begin{aligned} (5-\lambda)v_1 + 2v_2 &= 0 \\ 2v_1 + (8-\lambda)v_2 &= 0 \end{aligned} \tag{2.169}$$

in which λ is one of the eigenvalues.

For $\lambda = 4$ this equation has the form

$$\begin{aligned} v_1 + 2v_2 &= 0 \\ 2v_1 + 4v_2 &= 0 \end{aligned} \tag{2.170}$$

A non-zero solution of this system can be found by letting $v_2 = -1$ and finding $v_1 = 2$. Thus, we have an eigenvector (2,-1). Dividing this vector by its length, we obtain the normalized vector

$$\left(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right).$$

Next, we find similarly the normalized eigenvector corresponding to $\lambda = 9$:

$$\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right).$$

These vectors are orthogonal. Therefore, the coordinate transformation matrix has the form

$$\boldsymbol{T} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}. \tag{2.171}$$

Hence, the old coordinates x, y and the new coordinates x_1, y_1 are related by the following formulae

$$\begin{aligned} x &= \frac{2}{\sqrt{5}}\, x_1 + \frac{1}{\sqrt{5}}\, y_1 \\ y &= -\frac{1}{\sqrt{5}}\, x_1 + \frac{2}{\sqrt{5}}\, y_2. \end{aligned} \tag{2.172}$$

Substituting these equations into (2.166) we obtain after grouping similar terms

$$4\, x_1^2 + 9\, y_1^2 - \frac{8}{\sqrt{5}}\, x_1 - \frac{144}{\sqrt{5}}\, y_1 + 80 = 0. \tag{2.173}$$

After completing the squares we can rewrite this equation as

$$4(x_1 - \frac{1}{\sqrt{5}})^2 + 9(y_1 - \frac{8}{\sqrt{5}})^2 - 36 = 0 \tag{2.174}$$

or

$$\frac{(x_1 - \frac{1}{\sqrt{5}})^2}{9} + \frac{(y_1 - \frac{8}{\sqrt{5}})^2}{4} = 1. \tag{2.175}$$

As we can see, the curve is an ellipse.

2.5.1 Definite Hermitian forms

Definition 2.15. *A Hermitian form $H(\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{x}^H \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{x} \in \mathbb{C}^n$, and its matrix $\boldsymbol{A}$ are called*

Positive definite *if $H(\boldsymbol{x}, \boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq 0$. In this case we write $\boldsymbol{A} > 0$.*

Positive semidefinite or nonnegative definite *if $H(\boldsymbol{x}, \boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}$ and there is $\boldsymbol{x} \neq 0$ such that $H(\boldsymbol{x}, \boldsymbol{x}) = 0$. In this case we write $\boldsymbol{A} \geq 0$.*

Negative definite *if $-H(\boldsymbol{x}, \boldsymbol{x})$ is positive definite. In this case we write $\boldsymbol{A} < 0$.*

Negative semidefinite or nonpositive definite *if $-H(\boldsymbol{x}, \boldsymbol{x})$ is positive semidefinite. In this case we write $\boldsymbol{A} \leq 0$.*

Since the negative (semi)definite forms are defined using the positive (semi)definite forms, it is sufficient to study the latter.

Theorem 2.7. *A Hermitian form is positive definite (semidefinite) if and only if its all eigenvalues are positive (non-negative).*

Proof. A Hermitian form $H(\boldsymbol{x}, \boldsymbol{x})$ is positive definite (semidefinite) if and only if its equivalent diagonal form $H(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{x})$ is positive definite (semidefinite). The latter, according to equation (2.163) is positive definite if and only if its all eigenvalues are positive (non-negative). □ □

As we pointed out previously, finding eigenvalues of a matrix is a difficult problem especially for large matrices. There are simpler methods for deciding whether a general Hermitian form is positive (semi)definite, because a Hermitian form $H(\boldsymbol{x}, \boldsymbol{x})$ is positive definite (semidefinite) if and only if its equivalent Hermitian form $H(\boldsymbol{T}\boldsymbol{x}, \boldsymbol{T}\boldsymbol{x})$ is positive definite (semidefinite) where $\boldsymbol{T}$ is any non-singular matrix (not necessarily a unitary matrix). In other words, it is sufficient to find a diagonal congruent matrix $\boldsymbol{T}^H \boldsymbol{A}\boldsymbol{T}$ (and not necessarily a similar matrix $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$) that lets us to solve the problem.

Any Hermitian form can be presented by linear transformation of its variables (in an infinite number of ways) as an equivalent diagonal Hermitian form

$$H(\boldsymbol{T}\boldsymbol{x}, \boldsymbol{T}\boldsymbol{x}) = a_1|x_1|^2 + a_2|x_2|^2 + \cdots + a_n|x_n|^2. \tag{2.176}$$

It is often convenient to decide if a matrix is positive-definite by using the Cholesky decomposition which we consider in the next section.

2.5.2 Cholesky Decomposition

Theorem 2.8 (The Cholesky decomposition)**.** *A Hermitian matrix $\boldsymbol{A}$ is positive definite if and only if there exists a non-singular lower triangular matrix $\boldsymbol{C}$ with positive real diagonal entries such that*

$$\boldsymbol{A} = \boldsymbol{C}\boldsymbol{C}^H. \tag{2.177}$$

This formula is known as the *Cholesky decomposition* of $\boldsymbol{A}$, and $\boldsymbol{C}$ is called the Cholesky factor or a square root of $\boldsymbol{A}$.

Proof. If Hermitian matrix is positive definite, then all its eigenvalues are positive and we can rewrite equation (2.150) as

$$\boldsymbol{A} = \boldsymbol{C}\boldsymbol{C}^H \tag{2.178}$$

where

$$\boldsymbol{C} = \boldsymbol{U}\ \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}) = \boldsymbol{U} \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{bmatrix}. \tag{2.179}$$

The matrix $\boldsymbol{C}$ is non-singular since

$$\det \boldsymbol{C} = \det \boldsymbol{U} \det \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}) = \pm\sqrt{\lambda_1}\sqrt{\lambda_2}\cdots\sqrt{\lambda_n}) \neq 0$$

Conversely, if $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{C}^H$ where $\boldsymbol{C}$ is non-singular, then for any $\boldsymbol{x} \neq 0$

$$\boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x} = \boldsymbol{x}^H \boldsymbol{C}\boldsymbol{C}^H \boldsymbol{x} = \|\boldsymbol{C}^H \boldsymbol{x}\|^2 > 0.$$

□ □

We obtained equation (2.178) using the matrix spectral decomposition. However, the Cholesky decomposition can be obtained without finding the eigenvalues of $\boldsymbol{A}$. Indeed, we can rewrite equation (2.178) as

$$\begin{bmatrix} a_{11} & \overline{a}_{21} & \cdots & \overline{a}_{n1} \\ a_{21} & a_{22} & \cdots & \overline{a}_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} c_{11} & 0 & \cdots & 0 \\ c_{21} & c_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \begin{bmatrix} c_{11} & \overline{c}_{21} & \cdots & \overline{c}_{n1} \\ 0 & c_{22} & \cdots & \overline{c}_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c_{nn} \end{bmatrix} \tag{2.180}$$

Multiplying matrices in the RHS of this equation and comparing the results with the corresponding elements of $\boldsymbol{A}$ in the LHS we obtain a system of equations from which it is easy to find elements of the Cholesky factor. For the first column of $\boldsymbol{A}$, we have

$$a_{11} = c_{11}^2,\ a_{21} = c_{11}c_{21},\ \ldots, a_{n1} = c_{11}c_{n1}. \tag{2.181}$$

From this system we find

$$c_{11} = \sqrt{a_{11}},\ c_{21} = a_{21}/c_{11},\ \ldots, c_{n1} = a_{n1}/c_{11}. \tag{2.182}$$

Substituting this solution into (2.180), we can obtain similarly the second column of $\boldsymbol{C}$ as

$$c_{22} = \sqrt{a_{22} - |c_{21}|^2},\ c_{32} = (a_{32} - \overline{c}_{21}c_{31})/c_{22},\ \ldots, c_{n2} = (a_{n2} - \overline{c}_{21}c_{n1})/c_{22}. \tag{2.183}$$

Continuing the process, we obtain the following recursive equations

$$\begin{aligned} c_{ii} &= \sqrt{a_{ii} - \textstyle\sum_{k=1}^{i-1} |c_{ik}|^2} \\ c_{ij} &= \tfrac{1}{c_{ii}}(a_{ij} - \textstyle\sum_{k=1}^{j-1} c_{ik}\overline{c}_{jk}), \quad i = 1, 2, \ldots, n. \end{aligned} \tag{2.184}$$

We can now rephrase Theorem 2.8: a Hermitian matrix is positive definite if and only if the recursive equations in (2.184) are satisfied (i.e. all the expressions under the square root must be positive).

The other criterion of positive-definiteness of a Hermitian form (or matrix) is given by the following theorem

Theorem 2.9 (Sylvester's criterion)**.** *[4, p. 404], [1, p. 136] A Hermitian form is positive-definite if and only if its all leading diagonal minors are positive:*

$$D_1 = a_{11} > 0, D_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} > 0, \ldots, D_r = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} > 0. \tag{2.185}$$

Let us consider now some other applications of the Cholesky decomposition.

2.5.2.1 Generation of multidimensional Gaussian variates

Consider an affine transformation

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b} \tag{2.186}$$

of a real random vector (RV) $\boldsymbol{X}$ with the mean $E[\boldsymbol{X}] = \boldsymbol{\mu}_x$ and the variance matrix $\boldsymbol{\Sigma}_x = E[(\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{X} - \boldsymbol{\mu}_x)^\top]$. Applying the expectation to both sides of the previous equation, we obtain

$$\boldsymbol{\mu}_y = E[\boldsymbol{Y}] = E[\boldsymbol{A}\boldsymbol{X}] = \boldsymbol{A}\boldsymbol{\mu}_x + \boldsymbol{b}$$

and

$$\begin{aligned}\boldsymbol{\Sigma}_y &= E[(\boldsymbol{Y} - \boldsymbol{\mu}_y)(\boldsymbol{Y} - \boldsymbol{\mu}_y)^\top] = E[\boldsymbol{A}(\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{A}(\boldsymbol{X} - \boldsymbol{\mu}_x))^\top] \\ &= \boldsymbol{A}E[(\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{X} - \boldsymbol{\mu}_x)^\top]\boldsymbol{A}^\top = \boldsymbol{A}\boldsymbol{\Sigma}_x\boldsymbol{A}^\top.\end{aligned}$$

Thus, the mean and the variance matrix of the $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$ are given by

$$\boxed{\begin{aligned}\boldsymbol{\mu}_y &= \boldsymbol{A}\boldsymbol{\mu}_x + \boldsymbol{b} \\ \boldsymbol{\Sigma}_y &= \boldsymbol{A}\boldsymbol{\Sigma}_x\boldsymbol{A}^\top\end{aligned}} \tag{2.187}$$

As we can see, the variance matrix $\boldsymbol{\Sigma}_y$ is congruent to $\boldsymbol{\Sigma}_x$. In particular, if the components of the RV $\boldsymbol{X}$ are uncorrelated i.e. $\boldsymbol{\Sigma}_x = \boldsymbol{I}$, then $\boldsymbol{\Sigma}_y = \boldsymbol{A}\boldsymbol{\Sigma}_x\boldsymbol{A}^\top$. Therefore, if a RV has a positive definite variance matrix, we can express it as

$$\boldsymbol{Y} = \boldsymbol{C}\boldsymbol{U} + \boldsymbol{\mu}_y \tag{2.188}$$

where $\boldsymbol{U}$ is a random vector whose components are zero-mean and unit variance uncorrelated RVs, $\boldsymbol{C}$ is the Cholesky factor of $\boldsymbol{\Sigma}_y$:

$$\boldsymbol{\Sigma}_y = \boldsymbol{C}\boldsymbol{C}^\top, \tag{2.189}$$

Equation (2.188) in the one-dimensional case represents the normalised RV $U = (Y - a)/\sigma$.

As we know, the uncorrelated Gaussian RVs are independent. Therefore, equation (2.188) can be used to generate multidimensional Gaussian variates with the mean $\boldsymbol{\mu}_y$ and the variance matrix $\boldsymbol{\Sigma}_y$ from the independent Gaussian variables with zero mean and unit variance that was considered In Sec. 5.4.3 of our book [5].

2.5.2.2 Multivariate Gaussian conditional distribution

In our book [5] , we offered a derivation of equations (4.124) and (4.125) as an exercise. Let us derive these equations using the Cholesky decomposition.

Let

$$\boldsymbol{X} = \begin{bmatrix}\boldsymbol{X}_a \\ \boldsymbol{X}_b\end{bmatrix} \tag{2.190}$$

be a Gaussian RV with the mean and the variance matrices

$$\boldsymbol{\mu} = \begin{bmatrix}\boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b\end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix}\boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb}\end{bmatrix}. \tag{2.191}$$

We would like to find the conditional PDF of $\boldsymbol{X}_a$ given $\boldsymbol{X}_b = \boldsymbol{x}_b$.

For this purpose, we use equation (2.188) to present $\boldsymbol{X}$ as a function of independent Gaussian RVs:

$$\boldsymbol{X} = \boldsymbol{C}\boldsymbol{U} + \boldsymbol{\mu} \tag{2.192}$$

where $\boldsymbol{C}$ is the Cholesky factor of $\boldsymbol{\Sigma}$:

$$\begin{bmatrix}\boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb}\end{bmatrix} = \begin{bmatrix}\boldsymbol{C}_{aa} & 0 \\ \boldsymbol{C}_{ba} & \boldsymbol{C}_{bb}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_{aa}^\top & \boldsymbol{C}_{ba}^\top \\ 0 & \boldsymbol{C}_{bb}^\top\end{bmatrix} \tag{2.193}$$

Using these notations, equation can be rewritten in the following block form

$$\boldsymbol{X}_a = \boldsymbol{C}_{aa}\boldsymbol{U}_a + \boldsymbol{\mu}_a \tag{2.194}$$

$$\boldsymbol{X}_b = \boldsymbol{C}_{ba}\boldsymbol{U}_a + \boldsymbol{C}_{bb}\boldsymbol{U}_b + \boldsymbol{\mu}_b. \tag{2.195}$$

If $\boldsymbol{X}_b = \boldsymbol{x}_b$ is fixed, we can solve first of these equations for $\boldsymbol{U}_a$ and substitute it into the second equation to obtain

$$\boldsymbol{X}_b = \boldsymbol{C}_{ba}\boldsymbol{C}_{aa}^{-1}(\boldsymbol{x}_a - \boldsymbol{\mu}_a) + \boldsymbol{C}_{bb}\boldsymbol{U}_b + \boldsymbol{\mu}_b. \tag{2.196}$$

This equation shows that the $\boldsymbol{X}_b$, given $\boldsymbol{X}_b = \boldsymbol{x}_b$, is a Gaussian RV whose mean vector and the variance matrix can be obtained using equation (2.187):

$$\boldsymbol{\mu}_{b|a} = \boldsymbol{\mu}_b + \boldsymbol{C}_{ba}\boldsymbol{C}_{aa}^{-1}(\boldsymbol{x}_a - \boldsymbol{\mu}_a) \tag{2.197}$$

$$\boldsymbol{\Sigma}_{b|a} = \boldsymbol{C}_{bb}\boldsymbol{C}_{bb}^{\top}. \tag{2.198}$$

Comparing matrices on both sides of equation (2.193), we find that

$$\boldsymbol{C}_{ba}\boldsymbol{C}_{aa}^{-1} = \boldsymbol{\Sigma}_{ba}\boldsymbol{C}_{aa}^{-H}\boldsymbol{C}_{aa}^{-1} = \boldsymbol{\Sigma}_{ba}(\boldsymbol{C}_{aa}\boldsymbol{C}_{aa}^{-H})^{-1} = \boldsymbol{\Sigma}_{ba}\boldsymbol{\Sigma}_{aa}^{-1} \tag{2.199}$$

and

$$\boldsymbol{C}_{bb}\boldsymbol{C}_{bb}^{H} = \boldsymbol{\Sigma}_{bb} - \boldsymbol{C}_{ba}\boldsymbol{C}_{ba}^{H} = \boldsymbol{\Sigma}_{bb} - \boldsymbol{\Sigma}_{ba}\boldsymbol{\Sigma}_{aa}^{-1}\boldsymbol{\Sigma}_{ba}^{H}. \tag{2.200}$$

Using these equations, we can rewrite equations (2.197) as

$$\boxed{\begin{aligned}\boldsymbol{\mu}_{b|a} &= \boldsymbol{\mu}_b + \boldsymbol{\Sigma}_{ba}\boldsymbol{\Sigma}_{aa}^{-1}(\boldsymbol{x}_a - \boldsymbol{\mu}_a)\\ \boldsymbol{\Sigma}_{b|a} &= \boldsymbol{\Sigma}_{bb} - \boldsymbol{\Sigma}_{ba}\boldsymbol{\Sigma}_{aa}^{-1}\boldsymbol{\Sigma}_{ab}\end{aligned}} \tag{2.201}$$

References

[1] E. L. Bloch, L. I. Loshinsky, and W. Y. Turin. *Basics of Linear Algebra and Its Applications*. Vishaya Shcola Publishing House, Moscow, USSR, 1971.

[2] F. R. Gantmacher. *The Theory of Matrices*. Chelsea Publishing Co., New York, 1959.

[3] G. H. Golub and C. F. V. Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London, 1996.

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

[5] H. Kobayashi, B. L. Mark, and W. Turin. *Probability Random Processes and Statistical Analysis*. Cambridge University Press, Cambridge UK, 2012.

3 The Dirac delta function

To be done.

4 Stieltjes and Lebesgue integrals, and dF notation

To be done.

5 Selected topics in measure theory

To be done.

6 Interchanging limit and integral

To be done.

7 Differentiating integrals and sums

To be done.

8 Complex analysis: Contour integral and the residue theorem

To be done.

9 Functional transformation and Jacobians

To be done.

10 Stirling's Formula

Two functions $f(x)$ and $g(x)$ are called *asymptotically equivalent* as $x \to x_0$ if

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 1. \tag{10.1}$$

In this case, we write $f(x) \sim g(x)$ and we call $g(x)$ an asymptotic representation of $f(x)$.

The Stirling's formula gives us an asymptotic representation of the factorial. Since $n! = \Gamma(n+1)$, we derive a more general the asymptotic representation for the gamma-function $\Gamma(p+1)$:

$$\Gamma(p+1) \sim p^p \exp(-p)\sqrt{2\pi\, p}. \tag{10.2}$$

which means that

$$\lim_{p\to\infty} \frac{\Gamma(p+1)}{p^p \exp(-p)\sqrt{2\pi\, p}} = 1.$$

To prove this formula, we use the following integral representation of the gamma-function

$$\Gamma(p+1) = \int_0^\infty x^p \exp(-x)\, dx = \int_0^\infty \exp(p\ln x - x)\, dx. \tag{10.3}$$

To find the asymptotic formula, we note that the integrand in the RHS of this equation attains its maximum $p^p \exp(-p)$ at $x = p$. As p grows, this maximal value becomes dominant and we expect the dominant contribution to the integral from the small vicinity of $x = p$. After the substitution $x = t + p$ we obtain

$$\Gamma(p+1) = \int_{-p}^\infty (p+t)^n \exp(-p-t)\, dt = p^p \exp(-p) \int_{-p}^\infty \left(1 + \frac{t}{p}\right)^p \exp(-t)\, dt. \tag{10.4}$$

The logarithm of the last integrand can be written as

$$p\ln(1+\frac{t}{p}) - t = -\frac{t^2}{2p} + \cdots \tag{10.5}$$

where we used the Taylor series for the natural logarithm. Thus, the integrand behaves as $\exp(-\frac{t^2}{2p})$ in the vicinity of the maximum. So we can write

$$\Gamma(p+1) \sim p^p \exp(-p) \int_{-p}^\infty \exp\left(-\frac{t^2}{2p}\right)\, dt \sim p^p \exp(-p) \int_{-\infty}^\infty \exp\left(-\frac{t^2}{2p}\right)\, dt = p^p \exp(-p)\sqrt{2\pi\, p} \tag{10.6}$$

where we used the Gaussian integral formula

$$\int_{-\infty}^\infty \exp\left(-\frac{t^2}{2\sigma^2}\right)\, dt = \sqrt{2\pi}\,\sigma$$

derived in Section 4.2.4 of our book [1].

Thus, the Stirling's formula for the factorial has the form

$$\boxed{n! \sim n^n \exp(-n)\sqrt{2\pi\, n}.} \tag{10.7}$$

References

[1] H. Kobayashi, B. L. Mark, and W. Turin. *Probability Random Processes and Statistical Analysis*. Cambridge University Press, Cambridge UK, 2012.