Ivan Markov

INTRODUCTION TO MATRIX CALCULUS

Sofia
2011

1. Introduction to the Linear Algebra

The objective of this chapter is to present the fundamentals of matrices, with emphasis on those aspects that are important in finite element analysis. This is just a review of matrix algebra and should be known to the student from a course in linear algebra.

Of course, only a rather limited discussion of matrices and tensors is given here, but we hope that the focused practical treatment will provide a strong basis for understanding the finite element formulations given later.

1.1 Matrices. Basic Concepts

Further in our exposition with $\mathbb{R}$ will be denoted the set of real numbers, and $\mathbb{R}^n$ is the real n-dimensional vector space. The concept vector space is a fundamental in the theory of matrices but now we use it intuitively, and later we will give its definition. Matrices defined on the field of real numbers $\mathbb{R}$ are mainly used in this book.

From a simplistic point of view, matrices can simply be taken as ordered arrays (tables) of numbers that are subjected to specific rules of addition, multiplication, and so on. It is of course important to be thoroughly familiar with these rules, and we review them in this chapter.

However, by far more interesting aspects of matrices and matrix algebra are recognized when we study how the elements of matrices are derived in the analysis of a physical problem and why the rules of matrix algebra are actually applicable. In this context, the use of tensors and their matrix representations are important and provide a most interesting subject of study.

Of course, only a rather limited discussion of matrices and tensors is given here, but we hope that the focused practical treatment will provide a strong basis for understanding the finite element formulations given later.

Definition: A matrix is an array (table) of ordered numbers. A general matrix consists of *mn* numbers arranged in *m* rows and *n* columns, giving the following array

$$\mathbf{A} \equiv [A] \equiv \mathbf{A}_{m\times n} \equiv [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{1.1}$$

We say that this matrix has order $m \times n$ (*m* by *n*). If $m = n$, the matrix is square with order *n*. With a_{ij} we denote an arbitrary element of the matrix, which belongs to the *i*th row and to the *j*th column. The matrix elements a_{ij} with $i = j$, i.e. a_{ii} are called a ***principal diagonal*** of the matrix **A**. When we have only one row (*m* = 1) or one column (*n* = 1), we also call **A** a vector. Therefore the following one dimensional array

$$\mathbf{x} \equiv \{x\} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \tag{1.2}$$

is a vector. We use the notation $\mathbf{x} \in \mathbb{R}^n$. Further in this book we consider all vectors as columns. The numbers $x_1, x_2, \cdots, x_n$ are components of the vector.

According to the definition we consider that

$$\mathbf{x} \geq 0, \quad \text{if} \;\; x_i \geq 0 \;\; \text{for} \;\; \forall i = 1,2,\cdots n, \tag{1.3}$$

$$\mathbf{x} \leq 0, \quad \text{if} \;\; x_i \leq 0 \;\; \text{for} \;\; \forall i = 1,2,\cdots n. \tag{1.4}$$

The effectiveness of using matrices in practical calculations is readily realized by considering the of a set of linear simultaneous equations such as

$$\begin{aligned} 10x_1 - 4x_2 + 2x_3 \quad &= \;\; 18 \\ -4x_1 + 12x_2 + x_3 \quad &= -14 \\ 2x_1 + x_2 + 15x_3 - x_4 &= \;\; 30 \\ -x_3 + 20x_4 &= \;\; 18 \end{aligned} \tag{1.5}$$

This set of linear algebraic equations can be written as follows

$$\underbrace{\begin{bmatrix} 10 & -4 & 2 & 0 \\ -4 & 12 & 1 & 0 \\ 2 & 1 & 15 & -1 \\ 0 & 0 & -1 & 20 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{Bmatrix}}_{\mathbf{x}} = \underbrace{\begin{Bmatrix} 18 \\ -14 \\ 30 \\ 18 \end{Bmatrix}}_{\mathbf{b}}. \tag{1.6}$$

where the unknowns are x_1, x_2, x_3 and x_4. Using matrix notation, this set of equations is written more compactly as

$$\mathbf{Ax} = \mathbf{b}, \tag{1.7}$$

where **A** is the matrix of the coefficients in the set of linear equations, **x** is the matrix of unknowns, and **b** is the matrix of known quantities.
A comma between subscripts will be used when there is any risk of confusion, e.g., $a_{i+2,j+1}$.

Definition: The ***transpose*** of the $m \times n$ matrix **A**, written as $\mathbf{A}^T$, is obtained by interchanging the rows and columns in **A** If $\mathbf{A}^T = \mathbf{A}$, it follows that the number of rows and columns in **A** are equal and that $a_{ij} = a_{ji}$. Because, $m = n$ we say that **A** is a square matrix of order n, and because $a_{ij} = a_{ji}$, we say that **A** is a symmetric matrix. Note that symmetry implies that **A** is square, but not vice versa; i.e., a square matrix need not be symmetric. Obviously $\left(\mathbf{A}^T\right)^T = \mathbf{A}$.

The process of transposing is clear from the following example:

$$\mathbf{A} = \begin{bmatrix} 4 & -1 & 3 \\ 2 & 15 & -6 \\ 16 & -1 & 9 \\ 3 & 2 & 1 \end{bmatrix}_{4\times 3}, \qquad \mathbf{A}^T = \begin{bmatrix} 4 & 2 & 16 & 3 \\ -1 & 15 & -1 & 2 \\ 3 & -6 & 9 & 1 \end{bmatrix}_{3\times 4}. \tag{1.8}$$

Definition: The matrices **A** and **B** are equal if and only if they have the same number or rows and columns and all corresponding elements are equal, i.e., $a_{ij} = b_{ij}$ for all *i* and *j*.

We have defined matrices to be ordered arrays of numbers and identified them by single symbols. In order to be able to deal with them as we deal with ordinary numbers, it is necessary to define rules corresponding to those which govern equality, addition, subtraction, multiplication, and division of ordinary numbers. We shall simply state the matrix rules and not provide motivation for them.

We provide some elementary matrix operations.

a. Two matrices **A** and **B** can be added if and only if they have the same number of rows and columns. ***The addition*** of the matrices is performed by adding all corresponding elements; i.e., if a_{ij} and b_{ij} denote general elements of **A** and **B**, respectively, then $c_{ij} = a_{ij} + b_{ij}$ denotes a general element of **C**, where $\mathbf{C} = \mathbf{A} + \mathbf{B}$. It follows that **C** has the same number of rows and columns as **A** and **B**. If the number of rows and columns of the matrices **A** and **B** is not equal the addition is not defined.

b. A matrix $\mathbf{A}_{m\times n}$ is ***multiplied by the scalar*** $\alpha \in \mathbb{R}$ by multiplying each matrix element by the scalar; i.e., $\mathbf{C} = \alpha\mathbf{A}$ means that $c_{ij} = \alpha a_{ij}$.

c. Two matrices **A** and **B** can be multiplied to obtain the ***matrix product*** $\mathbf{C} = \mathbf{AB}$ if and only if the number of columns in **A** is equal to the number of rows in **B**. Assume that **A** is of order $m \times n$ and **B** is of order $n \times p$. Then for each element in **C** we have

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \quad \Leftrightarrow \quad \mathbf{C} = \mathbf{AB}. \tag{1.9}$$

From (1.9) it follows that to calculate the element c_{ij}, we multiply the elements in the *ith* row of **A** by the elements in the *j*th column of **B** and add all individual products. By taking the product of each row in **A** and each column in **B**, it follows that **C** must be of order $m \times p$.

Example 1.1. Calculate the matrix product $\mathbf{C} = \mathbf{AB}$, where

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & -2 & 3 \\ 1 & 5 & 6 \end{bmatrix}_{3\times 3}, \qquad \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ -5 & 1 \end{bmatrix}_{3\times 2}.$$

We obtain in detail the elements of the first row of the matrix **C** according to (1.9)

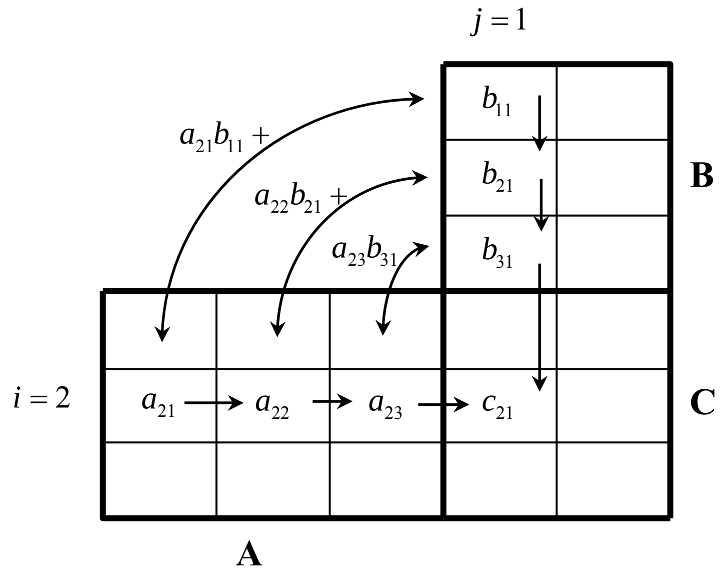
$$c_{11} = \sum_{k=1}^{3} a_{1k} b_{k1} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} = 6,$$

$$c_{12} = \sum_{k=1}^{3} a_{1k} b_{k2} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} = -1.$$

When we calculate all elements of **C**, obtain

$$\mathbf{C} = \begin{bmatrix} 6 & -1 \\ -19 & -3 \\ -19 & 20 \end{bmatrix}_{3\times 2}.$$

The algorithm of multiplication of two matrices $\mathbf{A}_{3\times 3}$ and $\mathbf{B}_{3\times 2}$ is shown in Fig. 1.1.



$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}$$

Fig. 1.1. Matrix multiplication

The procedure for matrix multiplication, shown in Fig. 1.1 is only convenient for “in hand” calculations and also for easily writing of algebraic and differential equations using matrix notations.

The multiplication of a matrix $\mathbf{A}_{m\times n}$ by $\mathbf{x}\in\mathbb{R}^n$ is identical to the multiplication of the matrices $\mathbf{A}_{m\times n}$ и $\mathbf{x}_{n\times 1}$. The given definitions allow to verify the correctness of the transformation of equations (1.5) into the form (1.6).

As is well known, the multiplication of ordinary numbers is ***commutative***; i.e., *ab* = *ba.* We need to investigate if the same holds for matrix multiplication. Let us consider the matrices

$$\mathbf{A}=\begin{Bmatrix}2\\5\end{Bmatrix} \text{ и } \mathbf{B}=\begin{bmatrix}1 & 2\end{bmatrix}.$$

We calculate the matrix products

$$\mathbf{AB}=\begin{bmatrix}2 & 4\\5 & 10\end{bmatrix}, \qquad \mathbf{BA}=\begin{bmatrix}12\end{bmatrix}\equiv 12\,.$$

Therefore, the products **AB** and **BA** are not the same, and it follows that ***matrix multiplication is not commutative***. Indeed, depending on the orders of **A** and **B**, the orders of the two product matrices **AB** and **BA** can be different, and the product **AB** may be defined, whereas the product **BA**

+

may not be calculable (this can be easily inspected for the matrices **A** and **B** in example 1.1.

If the matrices **A** and **B** satisfy $\mathbf{AB}=\mathbf{BA}$, we say that **A** и **B** commute.

The ***distributive law*** is valid and it states that

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}. \tag{1.10}$$

This law holds if the operations in (1.10) are defined.

The ***associative law*** states that

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}, \tag{1.11}$$

in other words, that the order of multiplication is insignificant.

In addition to being noncommutative, matrix algebra differs from scalar algebra in other ways. The cancellation of matrices in matrix equations also cannot be performed, in general, as the cancellation of ordinary numbers. In particular, the equality $\mathbf{AB} = \mathbf{AC}$ does not necessarily imply $\mathbf{B} = \mathbf{C}$, since algebraic summing is involved in forming the matrix products. As another example, if the product of two matrices is a null matrix, that is, $\mathbf{AB} = \mathbf{0}$, the result does not necessarily imply that either **A** or **B** is a null matrix. However, it must be noted that $\mathbf{A} = \mathbf{C}$ if the equation $\mathbf{AB} = \mathbf{CB}$ holds for all possible **B**. Namely, in that case, we simply select **B** to be the identity matrix **I**, and hence $\mathbf{A} = \mathbf{C}$.

The proof of above laws is carried out by using the definition of matrix multiplication in (1.9).

Definition: A ***null matrix*** (also known as a ***zero matrix***) is a matrix of any order in which the value of all elements is 0.

Definition: An ***identity*** (or ***unit) matrix*** **I**, which is a square matrix of order *n* with only zero elements except for its diagonal entries, which are unity. For example, the identity matrix of order 3 is

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.12}$$

It can be easily seen that arbitrary matrices **A** and **B** satisfy the following equations

$$\mathbf{A0} = \mathbf{0A} = \mathbf{0}, \qquad \mathbf{AI} = \mathbf{A}, \qquad \mathbf{IB} = \mathbf{B}, \tag{1.13}$$

if the corresponding matrix operations are defined.

It can be easily shown that the transpose of the product of two matrices **A** and **B** is equal to the product of the transposed matrices in reverse order; i.e.,

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T. \tag{1.14}$$

The proof that (1.14) does hold is obtained using the definition for the evaluation of a matrix product given in (1.9).

Definition: A ***symmetric*** matrix is a square matrix $\mathbf{A} = \left[a_{ij}\right]$ composed of elements such that the nondiagonal values are symmetric about the principal diagonal. Mathematically, symmetry is expressed as $a_{ij} = a_{ji,}$ for $\forall i, j$. If the matrix **A** is symmetric then $\mathbf{A}^T = \mathbf{A}$.

For an arbitrary matrix **A** the products $\mathbf{A}^T\mathbf{A}$ и $\mathbf{A}\mathbf{A}^T$ are symmetric matrices. Това се вижда от равенството

$$\left(\mathbf{A}\mathbf{A}^T\right)^T = \left(\mathbf{A}^T\right)^T \mathbf{A}^T = \mathbf{A}\mathbf{A}^T . \tag{1.15}$$

For example, if

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \qquad \text{then} \quad \mathbf{A}\mathbf{A}^T = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix} \qquad \text{and} \quad \mathbf{A}^T\mathbf{A} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix}.$$

It should be noted that although **A** and **B** may be symmetric, **AB** is, in general, not symmetric. The following example demonstrates this observation

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}, \qquad \mathbf{AB} = \begin{bmatrix} 5 & -1 \\ 0 & 2 \end{bmatrix}.$$

If **A** is symmetric, the matrix $\mathbf{B}^T\mathbf{D}\mathbf{B}$ is always symmetric. The proof follows using (1.14):

$$\left(\mathbf{B}^T\mathbf{D}\mathbf{B}\right)^T = \left(\mathbf{D}\mathbf{B}\right)\left(\mathbf{B}^T\right)^T = \mathbf{B}^T\mathbf{D}^T\mathbf{B},$$

But, because $\mathbf{D}^T = \mathbf{D}$, we have

$$\left(\mathbf{B}^T\mathbf{D}\mathbf{B}\right)^T = \mathbf{B}^T\mathbf{D}\mathbf{B}. \tag{1.16}$$

and hence $\mathbf{B}^T\mathbf{D}\mathbf{B}$ is symmetric. In particular, we can obtain (1.15) from (1.16) in case of $\mathbf{D} = \mathbf{I}$.

If the elements of the matrix $\mathbf{A}_{m\times n}$ are differentiable functions of one or several variables the differentiation of the matrix is performed according to the rule

$$\frac{\partial \mathbf{A}}{\partial x} = \left[\frac{\partial a_{ij}}{\partial x}\right], \tag{1.17}$$

and the integration is defined as

$$\int_a^b \mathbf{A}dx = \left[\int_a^b a_{ij}dx\right]. \tag{1.18}$$

It is clear from (1.17) and (1.18) that the matrix differentiation and the integration are operations applied to each matrix element.

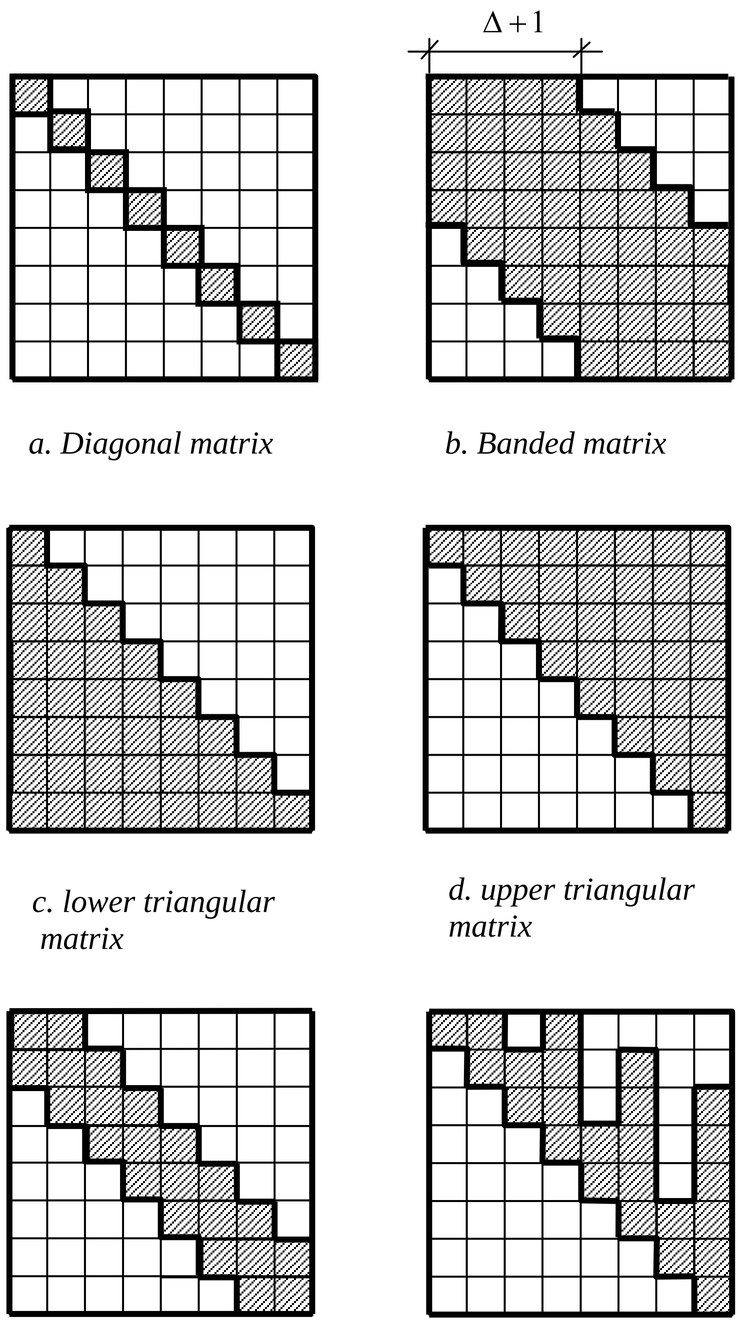
1.2 Special types of matrices

Whenever the elements of a matrix obey a certain law, we can consider the matrix to be of ***special form***. A ***real matrix*** is a matrix whose elements are all real. A ***complex matrix*** has elements that may be complex. We shall deal only with real matrices.

A ***diagonal*** matrix is a square matrix composed of elements such that $a_{ij} = 0$ if $i \neq j$. Therefore, the only nonzero terms are those on the main diagonal (upper left to lower right). Sometimes the diagonal matrix is written as

$$\mathbf{D} = diag\left(d_{ii}\right).$$

Schematically the diagonal matrix is shown in Fig. 1.2.a. All matrix elements which are not hatched are zeroes.



a. Diagonal matrix

b. Banded matrix

c. lower triangular matrix

d. upper triangular matrix

e. three diagonal matrix

f. "sky line" matrix

Fig. 1.2 Special types of matrices

A square matrix **A** is ***banded*** (Fig. 1.2.b) if it is satisfy the following conditions

$$a_{ij} = 0 \quad \text{за } \forall j > i + \Delta \quad \text{и за } \forall i > j + \Delta , \tag{1.19}$$

where $2\Delta + 1$ is the bandwidth of **A**. As an example, the following matrix is a symmetric banded matrix of order 5. The half-bandwidth Δ is 2:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & -2 & 0 & 0 \\ 1 & 5 & -1 & 1 & 0 \\ -2 & -1 & 12 & -1 & 1 \\ 0 & 1 & -1 & 9 & 2 \\ 0 & 0 & 1 & 2 & 10 \end{bmatrix}.$$

If the half-bandwidth of a matrix is zero, we have nonzero elements only on the diagonal of the matrix and denote it as a ***diagonal matrix***. For example, the identity matrix is a diagonal matrix.

The matrices shown in Fig. 1.2.c and in Fig. 1.2.d are called a lower ***triangular*** matrix and a upper ***triangular*** matrix.

A ***symmetric*** matrix in Fig. 1.2.f is "***skyline***" matrix. All elements which are outside the skyline are zeroes. The finite element method problems very often use large scale skyline matrices.

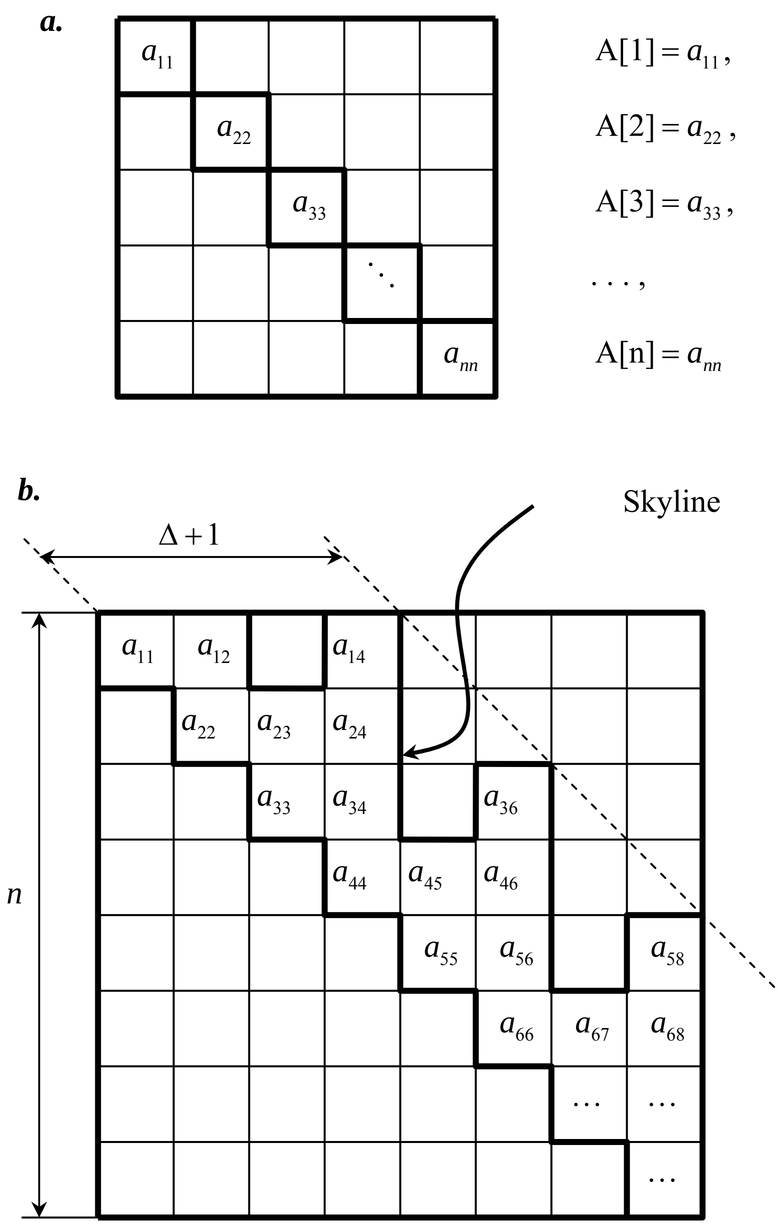
In computer calculations with matrices, we need to use a scheme of storing the elements of the matrices in computer memory. An obvious way of storing the elements of a matrix **A** of order $m \times n$ is simply to dimension in the program an array $\mathrm{A}[\mathrm{m}][\mathrm{n}]$, and store each matrix element a_{ij} in the storage location $\mathrm{A}[\mathrm{i}][\mathrm{j}]$. However, in many calculations we store in this way unnecessarily many zero elements of **A**, which are never needed in the calculations. Also, if **A** is symmetric, we should probably take advantage of it and store only the upper half of the matrix, including the diagonal elements. In general, only a restricted number of high-speed storage locations are available, and it is necessary to use an effective storage scheme in order to be able to take into high-speed core the maximum matrix size possible. If the matrix is too large to be contained in high-speed storage, the solution process will involve reading and writing on secondary storage, which can add significantly to the solution cost. Fortunately, in finite element analysis, the system matrices are symmetric and banded. Therefore, with an effective storage scheme, rather large-order matrices can be kept in high-speed core.

A ***diagonal matrix*** **A** of order *n* is stored in one-dimensional array

$$\mathrm{A}[\mathrm{i}] = a_{ii}, \quad i = 1, 2, \cdots, n . \tag{1.20}$$

This way of storage is shown in Fig. 1.3.a.

Consider a banded matrix as shown in Fig. 1.3.b. The zero elements within the "skyline" of the matrix may be changed to nonzero elements in the solution process; for example, a_{46} may be a zero element but becomes nonzero during the solution process. The zero elements that are outside the skyline are always zeroes after the transformation has been performed. Therefore we allocate storage locations to zero elements within the skyline but do not need to store zero elements that are outside the skyline. The storage scheme that will be used in the finite element solution process is indicated in Fig. 1.3.b. Of course, there are different schemes for matrix storage in the FEM.



A[1]=a_{11}, A[2]=a_{22}, A[3]=a_{12}, A[4]=a_{33}, A[5]=a_{23},
A[6]=a_{44}, A[7]=a_{34}, A[8]=a_{24}, A[9]=a_{14}, A[10]=a_{55},
A[11]=a_{45},…, A[MaxArr]=a_{nn}

Figure 1.3. Storage of matrix A in a one-dimensional array. a. diagonal matrix; b. symmetric banded matrix; b. symmetric matrix "skyline"

In case of storage shown in Fig. 1.3.b the addresses of the diagonal elements should be stored in one-dimensional array as follows

$$\begin{array}{lll} \text{DiagAddr[1]=1}, & \text{DiagAddr[2]=2}, & \text{DiagAddr[3]=4}, \\ \text{DiagAddr[4]=6}, & \text{DiagAddr[5]=10}, \ldots, & \text{DiagAddr[n]=MaxArr}, \\ \text{DiagAddr[n+1]=MaxArr+1}, & & \end{array} \tag{1.21}$$

where MaxArr is the total number of elements within skyline.

1.3 Determinant and Trace of a Matrix

A mathematical object consisting of many components usually is characterized (evaluated) by means of one number (scalar). Typical examples are the ***trace*** and ***determinant of a matrix***. The trace and determinant of a matrix are defined only if the matrix is square. Both quantities are single numbers, which are evaluated from the elements of the matrix and are therefore functions of the matrix elements.

Mainly two approaches use in the determinant definition. Here we represent the definition according to the determinant of the matrix $\mathbf{A}_{n\times n}$ is calculating by using the determinants of lower order matrices.

Definition:

a. The determinant of a matrix of order 1 (n=1) is simply the element of the matrix; i.e., if $\mathbf{A}=[a_{11}]$, then $\det\mathbf{A}=a_{11}$.

b. If i is an arbitrary row of the matrix $\mathbf{A}_{n\times n}$, then the determinant $\det\mathbf{A}$ is calculated as follows

$$\det\mathbf{A}=\sum_{j=1}^{n}(-1)^{i+j}\,a_{ij}\det\mathbf{A}_{ij}\,, \tag{1.22}$$

where $\mathbf{A}_{ij}$ is the matrix of order $(n-1)\times(n-1)$ resulting from the deletion of the row i and the column j of $\mathbf{A}$. The equation (1.22) is known as ***Laplace decomposition*** along ith row. By analogy if we use Laplace decomposition along jth column we obtain

$$\det\mathbf{A}=\sum_{i=1}^{n}(-1)^{i+j}\,a_{ij}\det\mathbf{A}_{ij}\,. \tag{1.23}$$

Note that $\det\mathbf{A}_{ij}$ are the determinants of the $(n-1)$ order matrices obtained by striking out the ith row and the jth column for given j. These are known as ***minors***. A minor of a determinant is another determinant formed by removing an equal number of rows and columns from the original determinant.

Example: Evaluate the determinant of **A**, where

$$\mathbf{A}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$$

Using the relation (1.22), we obtain

$$\mathbf{A}_{11}=[a_{22}],\qquad \mathbf{A}_{12}=[a_{21}],$$

$$\det\mathbf{A}=(-1)^{1+1}a_{11}\det\mathbf{A}_{11}+(-1)^{1+2}a_{12}\det\mathbf{A}_{12}\,,$$

but

$$\det\mathbf{A}_{11}=a_{22}\,,\qquad \det\mathbf{A}_{12}=a_{21}\,.$$

Hence

$$\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21} . \tag{1.24}$$

This relation is the general formula for the determinant of a 2×2 matrix.

Using the recurrence relation (1.22) along the row 1 (*i*=1) for the matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

we obtain

$$\mathbf{A}_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \qquad \mathbf{A}_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \qquad \mathbf{A}_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

According to formula (1.24)

$$\det \mathbf{A}_{11} = a_{22}a_{33} - a_{23}a_{32} ,$$

$$\det \mathbf{A}_{12} = a_{21}a_{33} - a_{23}a_{31} ,$$

$$\det \mathbf{A}_{13} = a_{21}a_{32} - a_{22}a_{31} .$$

Then from (1.22) we can write

$$\det \mathbf{A} = \left(-1\right)^{1+1} a_{11} \det \mathbf{A}_{11} + \left(-1\right)^{1+2} a_{12} \det \mathbf{A}_{12} + \left(-1\right)^{1+3} a_{13} \det \mathbf{A}_{13} .$$

After substituting $\det \mathbf{A}_{ij}$ into the previous equation we obtain

$$\begin{aligned} \det \mathbf{A} = {} & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned} . \tag{1.25}$$

In such a way we have obtained known formulae (1.24) and (1.25) for the determinant of 3×3 matrix.

We shall remember some fundamental properties of the determinants without proof.

1. $\det \mathbf{A}^T = \det \mathbf{A}$.
2. $\det \mathbf{I} = 1$.
3. The determinant of the matrix product is

$$\det\left(\mathbf{AB}\right) = \det \mathbf{A} \det \mathbf{B} . \tag{1.26}$$

4. The determinant of a ***diagonal matrix*** $\mathbf{D} = diag\left(d_{ii}\right)$ is

$$\det \mathbf{D} = d_{11}d_{22} \ldots d_{nn} .$$

5. The determinant of the lower ***triangular*** matrix or of the upper ***triangular*** matrix is equal to the product of diagonal terms.

The solution of a system of linear algebraic equations uses the specific decomposition $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^T$, where **L** is a lower unit triangular matrix (the diagonal elements of **L** are equal to 1 and $\det \mathbf{L} = 1$) and **D** is a diagonal matrix. In that case,

$$\det \mathbf{A} = \det \mathbf{L} \det \mathbf{D} \det \mathbf{L}^T, \tag{1.27}$$

and because $\det \mathbf{L} = 1$, we have

$$\det \mathbf{A} = \det \mathbf{D} = d_{11} d_{22} \dots d_{nn}. \tag{1.28}$$

6. If α is a scalar, then

$$\det(\alpha \mathbf{A}) = \alpha^n \det \mathbf{A}. \tag{1.29}$$

Definition: The trace of the square matrix **A** of order *n* is denoted as *tr*(**A**) and is equal to

$$tr(\mathbf{A}) = Sp(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}. \tag{1.30}$$

1.4 Matrix Inversion

Definition: The ***inverse*** of a matrix **A** is denoted by $\mathbf{A}^{-1}$. Assume that the inverse exists; then the elements of $\mathbf{A}^{-1}$ are such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}, \tag{1.31}$$

where **I** is the identity matrix. The inverse of a matrix does not need to exist. A trivial example is the null matrix.

Definition: A square matrix **A** is said to be ***nonsingular*** if $\det(\mathbf{A}) \neq 0$. A matrix **A** is a ***singular matrix*** if $\det(\mathbf{A}) = 0$.

Theorem: Each nonsingular matrix possesses an inverse.

We shall describe some ***properties of the inverse:***

1. ***Determinant of the inverse of matrix***. It can be easily obtained by using the definition (1.31) and the relation (1.26)

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad \Rightarrow \quad \det \mathbf{A}^{-1} \det \mathbf{A} = \det \mathbf{I} = 1.$$

Therefore

$$\det \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}}. \tag{1.32}$$

2. ***The inverse of the product of matrices***

$$\left(\mathbf{AB}\right)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{1.33}$$

3. ***The inverse of the transposed matrix***

$$\left(\mathbf{A}^{-1}\right)^{T} = \left(\mathbf{A}^{T}\right)^{-1}. \tag{1.34}$$

4. The inverse of a symmetric matrix is also symmetric.

If the inverse of matrix **A** can be determined, we can multiply both sides of Equation (1.7) by the inverse $\mathbf{A}^{-1}$ to obtain the solution for the simultaneous equations directly as

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \tag{1.35}$$

However, the inversion of **A** is very costly, and it is much more effective to only solve the equations in (1.7) without inverting **A**. *Indeed. although we may write symbolically that* $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$, *to evaluate* **x** *we actually only solve the equations.*

One way of calculating the inverse $\mathbf{A}^{-1}$ of a matrix **A** of order *n* is in terms of its determinant. If β_{ij} are the elements of the inverse, i.e., $\boldsymbol{\beta} \equiv \mathbf{A}^{-1} = \left[\beta_{ij}\right]$, then they are calculated as follows

$$\beta_{ij} = \left(-1\right)^{i+j}\frac{\det \mathbf{A}_{ji}}{\det \mathbf{A}} = \frac{\alpha_{ji}}{\det \mathbf{A}}, \qquad \left(i, j = 1,2,\ldots,n\right), \tag{1.36}$$

where $\mathbf{A}_{ji}$ is the submatrix of order $\left(n-1\right)\times\left(n-1\right)$ resulting from the deletion of the row *j* and the column *i* of **A**. We have to notice that the quantities $\alpha_{ji} = \left(-1\right)^{i+j}\det \mathbf{A}_{ji}$ are known as ***adjoints*** of **A**.

Calculation of the inverse of a matrix per Equation (1.36) is cumbersome and not very practical. Formulae (1.36) are convenient for matrices of low order ($n \le 3$). That is why we represent here a general algorithm for calculating the inverse $\mathbf{A}^{-1}$. Let **A** and $\boldsymbol{\beta}$ are square matrices of order *n* that satisfy the following matrix equation

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{I}, \tag{1.37}$$

where **I** is an identity matrix. Then $\boldsymbol{\beta} = \mathbf{A}^{-1}$. For the solution of each system of equations in (1.37), we can use the well known algorithms for solution of a system of linear algebraic equations. Let $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$ are vector columns of the matrix $\boldsymbol{\beta}$ and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are vector columns of the unite matrix, i.e.,

$$\boldsymbol{\beta}_i = \begin{Bmatrix} \beta_{1i} \\ \beta_{2i} \\ \vdots \\ \beta_{ni} \end{Bmatrix}, \qquad \mathbf{e}_i = \begin{Bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{Bmatrix} \rightarrow \mathit{row}\ i,$$

then each column $\boldsymbol{\beta}_i$ of the inverse $\boldsymbol{\beta}$ is obtained as a result of solution of the system of equations

$$\mathbf{A}\boldsymbol{\beta}_i = \mathbf{e}_i, \quad (i = 1, 2, \ldots, n). \tag{1.38}$$

After the solution the *i*th system of equations (1.38) is performed, the solution vector $\boldsymbol{\beta}_i$ is put in the *i*th column of the inverse $\mathbf{A}^{-1} = \boldsymbol{\beta}$

Example 1.2. Calculate the inverse $\boldsymbol{\beta} = \mathbf{A}^{-1}$ where

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

The identity matrix **I** of the third order and its vector-columns $\mathbf{e}_i$ are respectively

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix}, \qquad \mathbf{e}_1 = \begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix}, \qquad \mathbf{e}_2 = \begin{Bmatrix} 0 \\ 1 \\ 0 \end{Bmatrix}, \qquad \mathbf{e}_3 = \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix}.$$

Using a computer program for each $i = 1, 2, 3$ we solve the systems of equations

$$\mathbf{A}\boldsymbol{\beta}_i = \mathbf{e}_i$$

and obtain

$$\boldsymbol{\beta}_1 = \begin{Bmatrix} 3 \\ 2 \\ 1 \end{Bmatrix}, \qquad \boldsymbol{\beta}_2 = \begin{Bmatrix} 2 \\ 2 \\ 1 \end{Bmatrix}, \qquad \boldsymbol{\beta}_3 = \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix}.$$

After the solution is made we arrange $\boldsymbol{\beta}_i \ (i = 1, 2, 3)$ in the columns of $\boldsymbol{\beta} = \mathbf{A}^{-1}$, and hence

$$\boldsymbol{\beta} = \mathbf{A}^{-1} = \begin{bmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

1.5 Matrix Partitioning

Any matrix can be subdivided or ***partitioned*** into a number of submatrices of lower order. The concept of matrix partitioning is most useful in reducing the size of a system of equations and accounting for specified values of a subset of the dependent variables. A ***submatrix*** is a matrix that is obtained from the original matrix by including only the elements of certain rows and columns. The idea is demonstrated using a specific case in which the dashed lines are the lines of partitioning

$$\mathbf{A}=\left[\begin{array}{c|ccc|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{array}\right], \tag{1.39}$$

It should be noted that each of the partitioning lines must run completely across the original matrix. Using the partitioning, matrix **A** is written as

$$\mathbf{A}=\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}, \tag{1.40}$$

where

$$\mathbf{A}_{11}=\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \quad \mathbf{A}_{12}=\begin{bmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \end{bmatrix}, \qquad \mathbf{A}_{13}=\begin{bmatrix} a_{15} & a_{16} \\ a_{25} & a_{26} \end{bmatrix}, \ldots \tag{1.41}$$

$\mathbf{A}_{ij}$ are called ***submatrices*** of the matrix **A**.

The partitioning of matrices can be of advantage in saving computer storage; namely, if submatrices repeat, it is necessary to store the submatrix only once. The same applies in arithmetic. Using submatrices, we may identify a typical operation that is repeated many times. We then carry out this operation only once and use the result whenever it is needed.

The rules to be used in calculations with partitioned matrices follow from the definition of matrix addition, subtraction, and multiplication. Using partitioned matrices we can add, subtract, or multiply as if the submatrices were ordinary matrix elements, provided the original matrices have been partitioned in such a way that it is permissible to perform the individual submatrix additions, subtractions, or multiplications.

The rules for partitioned matrices are analogous to rules for ordinary matrices. It should be noted that the partitioning of the original matrices is only an approach to facilitate matrix manipulations and does not change any results.

Example 1.3. Calculate the matrix product $\mathbf{C = AB}$ in Example 1.1, by using the following partitioning

$$\mathbf{A}=\left[\begin{array}{cc|c} 4 & 1 & 0 \\ 0 & -2 & 3 \\ \hline 1 & 5 & 6 \end{array}\right]=\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \qquad \mathbf{B}=\left[\begin{array}{cc} 1 & -1 \\ 2 & 3 \\ \hline -5 & 1 \end{array}\right]=\begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix}.$$

By using the definition (1.9), or the multiplication scheme shown in Fig. 1.1, we obtain

$$\mathbf{AB}=\begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21} \\ \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21} \end{bmatrix}. \tag{1.42}$$

Then we calculate the products in (1.42)

$$\mathbf{A}_{11}\mathbf{B}_{11}=\begin{bmatrix} 4 & 1 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}=\begin{bmatrix} 6 & -1 \\ -4 & -6 \end{bmatrix},$$

$$\mathbf{A}_{12}\mathbf{B}_{21} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \begin{bmatrix} -5 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -15 & 3 \end{bmatrix},$$

$$\mathbf{A}_{21}\mathbf{B}_{11} = \begin{bmatrix} 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 14 \end{bmatrix}$$

$$\mathbf{A}_{22}\mathbf{B}_{21} = \begin{bmatrix} 6 \end{bmatrix} \begin{bmatrix} -5 & 1 \end{bmatrix} = \begin{bmatrix} -30 & 6 \end{bmatrix}.$$

After substituting these products into (1.42), it follows

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 6 & -1 \\ -19 & -3 \\ -19 & 20 \end{bmatrix}.$$

It should be emphasized that the partitioning have to be performed in such a way that the corresponding matrix operations are possible.

Transposing of a partitioned matrix is carried out according to the rules for transposing of an ordinary matrix taking into account that the submatrices should be also transposed. For the matrix (1.40) we have

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}, \qquad \mathbf{A}^T = \begin{bmatrix} \mathbf{A}_{11}^T & \mathbf{A}_{21}^T \\ \mathbf{A}_{12}^T & \mathbf{A}_{22}^T \\ \mathbf{A}_{13}^T & \mathbf{A}_{23}^T \end{bmatrix}.$$

Definition: A square matrix $\mathbf{A}$ is an ***orthogonal matrix*** if its inverse $\mathbf{A}^{-1}$ is equal to its transpose $\mathbf{A}^T$, i.e.,

$$\mathbf{A}^T = \mathbf{A}^{-1}, \tag{1.43a}$$

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}. \tag{1.43b}$$

Example 1.4. If $\alpha, \beta \in \mathbb{R}$ и $|\alpha| + |\beta| \neq 0$, then the following matrix

$$\mathbf{R} = \frac{1}{\sqrt{\alpha^2 + \beta^2}} \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \tag{1.44}$$

is orthogonal. It can be easily found out that the matrix $\mathbf{R}$ satisfy relations (1.43b).

Orthogonal matrices have the following ***properties***:

1. Let us write the orthogonal matrix $\mathbf{A}$ in the form $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_n \end{bmatrix}$, where $\mathbf{A}_i \left(i = 1, 2, \ldots n \right)$ are the vector-columns of the matrix $\mathbf{A}$, i.e., $\mathbf{A}_i = \begin{bmatrix} a_{1i} & a_{2i} & \cdots & a_{ni} \end{bmatrix}^T$. They satisfy the following relations

$$\mathbf{A}_i^T\mathbf{A}_j = \mathbf{A}_j^T\mathbf{A}_i = \delta_{ij} = \begin{cases} 1, \text{if } i = j \\ 0, \text{if } i \neq j \end{cases}, \tag{1.45}$$

i.e., the vectors $\mathbf{A}_i$ are ***orthonormal***. It means that the vectors $\mathbf{A}_i$ are orthogonal to each other. δ_{ij} is called the ***Kronecker delta***.

2. ***The determinant*** of the orthogonal matrix **A** is equal to ± 1, i.e., $\det \mathbf{A} = \pm 1$.
3. If **A** is orthogonal then both $\mathbf{A}^T$ and $\mathbf{A}^{-1}$ are orthogonal.
4. If **A** and **B** are orthogonal matrices of order *n* the product **AB** is orthogonal. It is obviously from the following relationships

$$(\mathbf{AB})^T(\mathbf{AB}) = \mathbf{B}^T\mathbf{A}^T\mathbf{AB} = \mathbf{B}^T\mathbf{IB} = \mathbf{B}^T\mathbf{B} = \mathbf{I}. \tag{1.46}$$

5. Each orthogonal matrix of the second order can be written in the form

$$\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\varepsilon\sin\varphi & \varepsilon\cos\varphi \end{bmatrix},$$

where $\varepsilon = \pm 1$ and $\varphi \in [0, 2\pi]$ is an arbitrary angle.

The proof of these properties can be easily performed by using the definition (1.43).

2. Bilinear and Quadratic Forms

Let $\mathbf{x} \in \mathbb{R}^n$ be a column vector of order *n*, and **A** a real square $n \times n$ matrix. Then the following triple product produces a *scalar* result:

$$F(\mathbf{x}) = \mathbf{x}^T\mathbf{Ax} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \tag{1}$$

This is called a ***quadratic form***. The matrix **A** is known as a matrix of the quadratic form. Usually this matrix is symmetric. Now imagine that we calculate all possible values if *F* as follows. Let coefficients x_i in **x** be allowed to assume independently any and all real values except for all x_i simultaneously zero.

Definition: The quadratic form $F(\mathbf{x})$ is

- positive definite if $F > 0$ for all $\mathbf{x} \in \mathbb{R}^n$
- positive semidefinite if $F \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$
- negative semidefinite if $F \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$
- negative definite if $F < 0$ for all $\mathbf{x} \in \mathbb{R}^n$

and simply ***indefinite*** if *F* can be either positive or negative.
A symmetric matrix **A** is ***positive definite*** if the quadratic form $F(\mathbf{x}) = \mathbf{x}^T\mathbf{Ax}$ is positive definite.
For example, the following two matrices are respectively positive definite and positive semidefinite

$$\begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \qquad \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

If $F(\mathbf{x}) = F(x_1, x_2, ..., x_n)$ is a smooth real-valued function defined on $\mathbb{R}^n$ then the ***gradient*** of $F(\mathbf{x})$ is

$$\nabla F = \frac{\partial F}{\partial \mathbf{x}} = \begin{Bmatrix} \dfrac{\partial F}{\partial x_1} \\ \dfrac{\partial F}{\partial x_2} \\ \vdots \\ \dfrac{\partial F}{\partial x_n} \end{Bmatrix}. \tag{2}$$

Let $\mathbf{x} = [x_1, x_2, ..., x_n]^T$ and let **A** be an arbitrary *n* by *n* square matrix that does not depend on the x_i. We consider the quadratic form

$$\varphi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{A}\mathbf{x}. \tag{3}$$

In detail the quadratic form (3) can be written as follows

$$\begin{aligned} \varphi(\mathbf{x}) = \frac{1}{2}\Big(& a_{11}x_1^2 + a_{12}x_1x_2 + ... + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + ... + a_{2n}x_2x_n \\ & .... \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + ... + a_{nn}x_n^2 \Big). \end{aligned} \tag{3a}$$

By means of differentiation of (3a) with respect to each x_i it can be easily found out that

$$\nabla \varphi = \frac{\partial \varphi}{\partial \mathbf{x}} = \begin{Bmatrix} \dfrac{\partial \varphi}{\partial x_1} \\ \dfrac{\partial \varphi}{\partial x_2} \\ \vdots \\ \dfrac{\partial \varphi}{\partial x_n} \end{Bmatrix} = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^T\right)\mathbf{x}. \tag{4}$$

If **A** is a *symmetric matrix*, i.e., $\mathbf{A}^T = \mathbf{A}$, then the relation (4) is

$$\nabla \varphi = \frac{\partial \varphi}{\partial \mathbf{x}} = \mathbf{A}\mathbf{x} \qquad \text{и} \qquad \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = a_{ij} \text{ за } \forall i, j. \tag{5}$$

As a special case, if $\mathbf{A} \equiv \mathbf{I}$, then $\varphi(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$ and

$$\nabla \varphi = \frac{\partial \varphi}{\partial \mathbf{x}} = \mathbf{x}. \tag{6}$$

Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T$ and **A** be an arbitrary *n* by *n* matrix that does not depend on **x** and **y**. Let us consider the following scalar

$$\psi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y}, \tag{7}$$

The function (7) is called ***bilinear form***. It can be easily proved that

$$\frac{\partial \psi}{\partial \mathbf{x}} = \mathbf{A}\mathbf{y}. \tag{8}$$

Obviously we can write

$$\psi^T = \psi = \mathbf{y}^T \mathbf{A}^T \mathbf{x}.$$

By using (8) we obtain

$$\frac{\partial \psi}{\partial \mathbf{y}} = \mathbf{A}^T \mathbf{x}. \tag{9}$$

Particularly, if **A** is a unit matrix, then $\psi = \mathbf{x}^T \mathbf{y}$ and hence

$$\frac{\partial \psi}{\partial \mathbf{x}} = \mathbf{y} \qquad \text{и} \qquad \frac{\partial \psi}{\partial \mathbf{y}} = \mathbf{x}. \tag{10}$$

Example 1: Calculate the gradient of the function

$$\Pi(\mathbf{Z}) = \frac{1}{2} \mathbf{Z}^T \mathbf{K} \mathbf{Z} - \mathbf{Z}^T \mathbf{F}, \tag{11}$$

where **K** is a symmetric matrix and **F** is a vector that does not depend on **Z**.

The first term of (11) is a quadratic form and the second term is a linear function of **x**. Using (5) and (10) we obtain

$$\frac{\partial \Pi}{\partial \mathbf{Z}} = \mathbf{K}\mathbf{Z} - \mathbf{F}. \tag{12}$$

We have to remind that the necessary condition for extremum of $\Pi(\mathbf{Z})$ is

$$\frac{\partial \Pi}{\partial \mathbf{Z}} = \mathbf{K}\mathbf{Z} - \mathbf{F} = 0 \qquad \text{or} \qquad \mathbf{K}\mathbf{Z} = \mathbf{F}. \tag{13}$$

If the matrix **K** is positive definite then the function $\Pi(\mathbf{Z})$ has a minimum for **Z**, that satisfied (13).

Example 2: Calculate the first derivative of the function

$$u(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon}^T \mathbf{D}\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^T \mathbf{D}\boldsymbol{\varepsilon}_0 + \boldsymbol{\varepsilon}^T \boldsymbol{\sigma}_0$$

with respect of $\boldsymbol{\varepsilon}$.

By using the above rules we obtain

$$\frac{\partial u}{\partial \boldsymbol{\varepsilon}} = \mathbf{D}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_0) + \boldsymbol{\sigma}_0 .$$