B

Linear Algebra: Matrices

TABLE OF CONTENTS

§B.1. Matrices

§B.1.1. Concept

Let us now introduce the concept of a *matrix*. Consider a set of scalar quantities arranged in a rectangular array containing m rows and n columns:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}. \tag{B.1}$$

This array will be called a *rectangular matrix* of order m by n, or, briefly, an $m \times n$ matrix. Not every rectangular array is a matrix; to qualify as such it must obey the operational rules discussed below.

The quantities a_{ij} are called the *entries* or *components* of the matrix. Preference will be given to the latter unless one is talking about the computer implementation. As in the case of vectors, the term "matrix element" will be avoided to lessen the chance of confusion with finite elements. The two subscripts identify the row and column, respectively.

Matrices are conventionally identified by **bold uppercase** letters such as $\mathbf{A}$, $\mathbf{B}$, etc. The entries of matrix $\mathbf{A}$ may be denoted as A_{ij} or a_{ij}, according to the intended use. Occassionally we shall use the short-hand component notation

$$\mathbf{A} = [a_{ij}]. \tag{B.2}$$

Example B.1. The following is a 2×3 numerical matrix:

$$\mathbf{B} = \begin{bmatrix} 2 & 6 & 3 \\ 4 & 9 & 1 \end{bmatrix} \tag{B.3}$$

This matrix has 2 rows and 3 columns. The first row is $(2, 6, 3)$, the second row is $(4, 9, 1)$, the first column is $(2, 4)$, and so on.

In some contexts it is convenient or useful to display the number of rows and columns. If this is so we will write them underneath the matrix symbol.[1] For the example matrix (B.3) we would show

$$\underset{2\times 3}{\mathbf{B}} \tag{B.4}$$

Remark B.1. Matrices should not be confused with determinants.[2] A determinant is a number associated with square matrices ($m = n$), defined according to the rules stated in Appendix C.

[1] A convention introduced in Berkeley courses by Ray Clough. It is particularly useful in blackboard expositions.

[2] This confusion is apparent in the literature of the period 1860–1920.

§B.1.2. Real and Complex Matrices

As in the case of vectors, the components of a matrix may be real or complex. If they are real numbers, the matrix is called *real*, and *complex* otherwise. For the present exposition all matrices will be real.

§B.1.3. Square Matrices

The case $m = n$ is important in practical applications. Such matrices are called *square matrices* of order n. Matrices for which $m \neq n$ are called non-square (the term "rectangular" is also used in this context, but this is fuzzy because squares are special cases of rectangles).

Square matrices enjoy certain properties not shared by non-square matrices, such as the symmetry and antisymmetry conditions defined below. Furthermore many operations, such as taking determinants and computing eigenvalues, are only defined for square matrices.

Example B.2.

$$\mathbf{C} = \begin{bmatrix} 12 & 6 & 3 \\ 8 & 24 & 7 \\ 2 & 5 & 11 \end{bmatrix} \tag{B.5}$$

is a square matrix of order 3.

Consider a square matrix $\mathbf{A} = [a_{ij}]$ of order $n \times n$. Its n components a_{ii} form the *main diagonal*, which runs from top left to bottom right. The *cross diagonal* runs from the bottom left to upper right. The main diagonal of the example matrix (B.5) is {12, 24, 11} and the cross diagonal is {2, 24, 3}.

Entries that run parallel to and above (below) the main diagonal form superdiagonals (subdiagonals). For example, {6, 7} is the first superdiagonal of the example matrix (B.5).

§B.1.4. Symmetry and Antisymmetry

Square matrices for which $a_{ij} = a_{ji}$ are called *symmetric about the main diagonal* or simply *symmetric*.

Square matrices for which $a_{ij} = -a_{ji}$ are called *antisymmetric* or *skew-symmetric*. The diagonal entries of an antisymmetric matrix must be zero.

Example B.3. The following is a symmetric matrix of order 3:

$$\mathbf{S} = \begin{bmatrix} 11 & 6 & 1 \\ 6 & 3 & -1 \\ 1 & -1 & -6 \end{bmatrix}. \tag{B.6}$$

The following is an antisymmetric matrix of order 4:

$$\mathbf{W} = \begin{bmatrix} 0 & 3 & -1 & -5 \\ -3 & 0 & 7 & -2 \\ 1 & -7 & 0 & 0 \\ 5 & 2 & 0 & 0 \end{bmatrix}. \tag{B.7}$$

§B.1.5. Are Vectors a Special Case of Matrices?

Consider the 3-vector $\mathbf{x}$ and a 3×1 matrix $\mathbf{X}$ with the same components:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix}. \tag{B.8}$$

in which $x_1 = x_{11}$, $x_2 = x_{22}$ and $x_3 = x_{33}$. Are $\mathbf{x}$ and $\mathbf{X}$ the same thing? If so we could treat column vectors as one-column matrices and dispense with the distinction.

Indeed in many contexts a column vector of order n may be treated as a matrix with a single column, i.e., as a matrix of order $n \times 1$. Similarly, a row vector of order m may be treated as a matrix with a single row, i.e., as a matrix of order $1 \times m$.

There are some operations, however, for which the analogy does not carry over, and one has to consider vectors as different from matrices. The dichotomy is reflected in the notational conventions of lower versus upper case. Another important distinction from a practical standpoint is discussed next.

§B.1.6. Where Do Matrices Come From?

Although we speak of "matrix algebra" as embodying vectors as special cases of matrices, in practice the quantities of primary interest to the structural engineer are vectors rather than matrices. For example, an engineer may be interested in displacement vectors, force vectors, vibration eigenvectors, buckling eigenvectors. In finite element analysis even stresses and strains are often arranged as vectors although they are really tensors.

On the other hand, matrices are rarely the quantities of primary interest: they work silently in the background where they are normally engaged in operating on vectors.

§B.1.7. Special Matrices

The *null* matrix, written $\mathbf{0}$, is the matrix all of whose components are zero.

Example B.4. The null matrix of order 2×3 is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{B.9}$$

The *identity matrix*, written $\mathbf{I}$, is a square matrix all of which entries are zero except those on the main diagonal, which are ones.

Example B.5. The identity matrix of order 4 is

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{B.10}$$

A *diagonal matrix* is a square matrix all of which entries are zero except for those on the main diagonal, which may be arbitrary.

Example B.6. The following matrix of order 4 is diagonal:

$$\mathbf{D} = \begin{bmatrix} 14 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}. \tag{B.11}$$

A short hand notation which lists only the diagonal entries is sometimes used for diagonal matrices to save writing space. This notation is illustrated for the above matrix:

$$\mathbf{D} = \mathbf{diag}\;[\;14 \quad -6 \quad 0 \quad 3\;]. \tag{B.12}$$

An *upper triangular* matrix is a square matrix in which all elements underneath the main diagonal vanish. A *lower triangular* matrix is a square matrix in which all entries above the main diagonal vanish.

Example B.7. Here are examples of each kind:

$$\mathbf{U} = \begin{bmatrix} 6 & 4 & 2 & 1 \\ 0 & 6 & 4 & 2 \\ 0 & 0 & 6 & 4 \\ 0 & 0 & 0 & 6 \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 10 & 4 & 0 & 0 \\ -3 & 21 & 6 & 0 \\ -15 & -2 & 18 & 7 \end{bmatrix}. \tag{B.13}$$

§B.2. Elementary Matrix Operations

§B.2.1. Equality

Two matrices $\mathbf{A}$ and $\mathbf{B}$ of same order $m \times n$ are said to be *equal* if and only if all of their components are equal: $a_{ij} = b_{ij}$, for all $i = 1, \ldots m,\ j = 1, \ldots n$. We then write $\mathbf{A} = \mathbf{B}$. If the inequality test fails the matrices are said to be *unequal* and we write $\mathbf{A} \neq \mathbf{B}$.

Two matrices of different order cannot be compared for equality or inequality.

There is no simple test for greater-than or less-than.

§B.2.2. Transposition

The *transpose* of a matrix $\mathbf{A}$ is another matrix denoted by $\mathbf{A}^T$ that has n rows and m columns

$$\mathbf{A}^T = [a_{ji}]. \tag{B.14}$$

The rows of $\mathbf{A}^T$ are the columns of $\mathbf{A}$, and the rows of $\mathbf{A}$ are the columns of $\mathbf{A}^T$.

Obviously the transpose of $\mathbf{A}^T$ is again $\mathbf{A}$, that is, $(\mathbf{A}^T)^T = \mathbf{A}$.

Example B.8.

$$\mathbf{A} = \begin{bmatrix} 5 & 7 & 0 \\ 1 & 0 & 4 \end{bmatrix}, \qquad \mathbf{A}^T = \begin{bmatrix} 5 & 1 \\ 7 & 0 \\ 0 & 4 \end{bmatrix}. \tag{B.15}$$

The transpose of a square matrix is also a square matrix. The transpose of a symmetric matrix $\mathbf{A}$ is equal to the original matrix, *i.e.*, $\mathbf{A} = \mathbf{A}^T$. The negated transpose of an antisymmetric matrix matrix $\mathbf{A}$ is equal to the original matrix, *i.e.* $\mathbf{A} = -\mathbf{A}^T$.

Example B.9.

$$\mathbf{A} = \begin{bmatrix} 4 & 7 & 0 \\ 7 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix} = \mathbf{A}^T, \qquad \mathbf{W} = \begin{bmatrix} 0 & 7 & 0 \\ -7 & 0 & -2 \\ 0 & 2 & 0 \end{bmatrix} = -\mathbf{W}^T \tag{B.16}$$

§B.2.3. Addition and Subtraction

The simplest operation acting on two matrices is *addition*. The sum of two matrices of the same order, $\mathbf{A}$ and $\mathbf{B}$, is written $\mathbf{A}+\mathbf{B}$ and defined to be the matrix

$$\mathbf{A}+\mathbf{B} \stackrel{\text{def}}{=} [a_{ij} + b_{ij}]. \tag{B.17}$$

Like vector addition, matrix addition is commutative: $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$, and associative: $\mathbf{A}+(\mathbf{B}+\mathbf{C}) = (\mathbf{A}+\mathbf{B})+\mathbf{C}$. For $n=1$ or $m=1$ the operation reduces to the addition of two column or row vectors, respectively.

For matrix subtraction, replace $+$ by $-$ in the definition (**?**).

Example B.10. The sum of

$$\mathbf{A} = \begin{bmatrix} 1 & -3 & 0 \\ 4 & 2 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 6 & 3 & -3 \\ 7 & -2 & 5 \end{bmatrix} \quad \text{is} \quad \mathbf{A}+\mathbf{B} = \begin{bmatrix} 7 & 0 & -3 \\ 11 & 0 & 4 \end{bmatrix}. \tag{B.18}$$

§B.2.4. Scalar Multiplication

Multiplication of a matrix $\mathbf{A}$ by a scalar c is defined by means of the relation

$$c\,\mathbf{A} \stackrel{\text{def}}{=} [ca_{ij}] \tag{B.19}$$

That is, each entry of the matrix is multiplied by c. This operation is often called *scaling* of a matrix. If $c=0$, the result is the null matrix. Division of a matrix by a nonzero scalar c is equivalent to multiplication by $(1/c)$.

Example B.11.

$$\text{If} \quad \mathbf{A} = \begin{bmatrix} 1 & -3 & 0 \\ 4 & 2 & -1 \end{bmatrix}, \qquad 3\,\mathbf{A} = \begin{bmatrix} 3 & -9 & 0 \\ 12 & 6 & -3 \end{bmatrix}. \tag{B.20}$$

§B.3. Matrix Products

§B.3.1. Matrix by Vector Product

Before describing the general matrix product of two matrices, let us treat the particular case in which the second matrix is a column vector. This so-called *matrix-vector product* merits special attention because it occurs very frequently in the applications. Let $\mathbf{A} = [a_{ij}]$ be an $m \times n$ matrix, $\mathbf{x} = \{x_j\}$ a column vector of order n, and $\mathbf{y} = \{y_i\}$ a column vector of order m. The matrix-vector product is symbolically written

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{B.21}$$

to mean the linear transformation

$$y_i \stackrel{\text{def}}{=} \sum_{j=1}^{n} a_{ij} x_j \stackrel{\text{sc}}{=} a_{ij} x_j, \qquad i = 1, \ldots, m. \tag{B.22}$$

Example B.12. The product of a 2×3 matrix and a vector of order 3 is a vector of order 2:

$$\begin{bmatrix} 1 & -3 & 0 \\ 4 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -5 \\ 5 \end{bmatrix} \tag{B.23}$$

This product definition is not arbitrary but emanates from the analytical and geometric properties of entities represented by matrices and vectors.

For the product definition to make sense, the column dimension of the matrix $\mathbf{A}$ (called the pre-multiplicand) must equal the dimension of the vector $\mathbf{x}$ (called the post-multiplicand). For example, the reverse product $\mathbf{x}\mathbf{A}$ does not make sense unless $m = n = 1$.

If the row dimension m of $\mathbf{A}$ is one, the matrix formally reduces to a row vector, and the matrix-vector product reduces to the inner product defined by equation (A.15) of Appendix A. The result of this operation is a one-dimensional vector or scalar. We thus see that the present definition properly embodies previous cases.

The associative and commutative properties of the matrix-vector product fall under the rules of the more general matrix-matrix product discussed next.

§B.3.2. Matrix by Matrix Product

We now pass to the most general matrix-by-matrix product, and consider the operations involved in computing the product $\mathbf{C}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$:

$$\mathbf{C} = \mathbf{A}\,\mathbf{B}. \tag{B.24}$$

Here $\mathbf{A} = [a_{ij}]$ is a matrix of order $m \times n$, $\mathbf{B} = [b_{jk}]$ is a matrix of order $n \times p$, and $\mathbf{C} = [c_{ik}]$ is a matrix of order $m \times p$. The entries of the result matrix $\mathbf{C}$ are defined by the formula

$$c_{ik} \stackrel{\text{def}}{=} \sum_{j=1}^{n} a_{ij} b_{jk} \stackrel{\text{sc}}{=} a_{ij} b_{jk}, \quad i = 1, \ldots, m, \quad k = 1, \ldots, p. \tag{B.25}$$

We see that the $(i,k)^{th}$ entry of $\mathbf{C}$ is computed by taking the *inner product* of the i^{th} row of $\mathbf{A}$ with the k^{th} column of $\mathbf{B}$. For this definition to work and the product be possible, *the column dimension of* $\mathbf{A}$ *must be the same as the row dimension of* $\mathbf{B}$. Matrices that satisfy this rule are said to be *product-conforming*, or *conforming* for short. If the two matrices do not conform, their product is undefined. The following mnemonic notation often helps in remembering this rule:

$$\underset{m\times p}{\mathbf{C}} = \underset{m\times n}{\mathbf{A}}\ \underset{n\times p}{\mathbf{B}} \tag{B.26}$$

For the matrix-by-vector case treated in the preceding subsection, $p = 1$.

Matrix $\mathbf{A}$ is called the pre-multiplicand and is said to *premultiply* $\mathbf{B}$. Matrix $\mathbf{B}$ is called the post-multiplicand and is said to *postmultiply* $\mathbf{A}$. This careful distinction on which matrix comes first is a consequence of the absence of commutativity: even if $\mathbf{BA}$ exists (it only does if $m = n$), it is not generally the same as $\mathbf{AB}$.

For *hand* computations, the matrix product is most conveniently organized by the so-called Falk's scheme:

$$\begin{array}{cc} & \begin{bmatrix} b_{11} & \cdots & b_{ik} & \cdots & b_{1p} \\ \vdots & \ddots & \downarrow & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{np} \end{bmatrix} \\ \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \rightarrow & a_{in} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} & \begin{bmatrix} & & & & \\ & & \vdots & & \\ & \cdots & c_{ik} & & \\ & & & & \\ & & & & \end{bmatrix} \end{array}. \tag{B.27}$$

Each entry in row i of $\mathbf{A}$ is multiplied by the corresponding entry in column k of $\mathbf{B}$ (note the arrows), and the products are summed and stored in the $(i,k)^{th}$ entry of $\mathbf{C}$.

Example B.13. To illustrate Falk's scheme, let us form the product $\mathbf{C} = \mathbf{AB}$ of the following matrices

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 2 \\ 4 & -1 & 5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 2 & 1 & 0 & -5 \\ 4 & 3 & -1 & 0 \\ 0 & 1 & -7 & 4 \end{bmatrix} \tag{B.28}$$

The matrices are conforming because the column dimension of $\mathbf{A}$ and the row dimension of $\mathbf{B}$ are the same (3). We arrange the computations as shown below:

$$\begin{array}{cccc} & & \begin{bmatrix} 2 & 1 & 0 & -5 \\ 4 & 3 & -1 & 0 \\ 0 & 1 & -7 & 4 \end{bmatrix} & = \mathbf{B} \\ \mathbf{A} = & \begin{bmatrix} 3 & 0 & 2 \\ 4 & -1 & 5 \end{bmatrix} & \begin{bmatrix} 6 & 5 & -14 & -7 \\ 4 & 6 & -34 & 0 \end{bmatrix} & = \mathbf{C} = \mathbf{AB} \end{array}. \tag{B.29}$$

Here $3 \times 2 + 0 \times 4 + 2 \times 0 = 6$ and so on.

§B.3.3. Matrix Powers

If $\mathbf{A} = \mathbf{B}$, the product $\mathbf{AA}$ is called the *square* of $\mathbf{A}$ and is denoted by $\mathbf{A}^2$. Note that for this definition to make sense, $\mathbf{A}$ must be a square matrix; else the factors would not be conforming.

Similarly, $\mathbf{A}^3 = \mathbf{AAA} = \mathbf{A}^2\mathbf{A} = \mathbf{AA}^2$. Other positive-integer powers can be defined in an analogous manner.

This definition does not encompass negative powers. For example, $\mathbf{A}^{-1}$ denotes the *inverse* of matrix $\mathbf{A}$, which is studied in Appendix C. The general power $\mathbf{A}^m$, where m can be a real or complex scalar, can be defined with the help of the matrix spectral form and requires the notion of eigensystem covered in Appendix D.

A square matrix $\mathbf{A}$ that satisfies $\mathbf{A} = \mathbf{A}^2$ is called *idempotent*. We shall see later that this condition characterizes the so-called projector matrices.

A square matrix $\mathbf{A}$ whose p^{th} power is the null matrix is called *p-nilpotent*.

§B.3.4. Matrix Product Properties

Associativity. The associative law is verified:

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}. \tag{B.30}$$

Hence we may delete the parentheses and simply write $\mathbf{ABC}$.

Distributivity. The distributive law also holds: If $\mathbf{B}$ and $\mathbf{C}$ are matrices of the same order, then

$$\mathbf{A}\,(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}, \qquad \text{and} \qquad (\mathbf{B}+\mathbf{C})\,\mathbf{A} = \mathbf{BA}+\mathbf{CA}. \tag{B.31}$$

Commutativity. The commutativity law of scalar multiplication does not generally hold. If $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same order, then the products $\mathbf{AB}$ and $\mathbf{BA}$ are both possible but in general $\mathbf{AB} \neq \mathbf{BA}$.

If $\mathbf{AB} = \mathbf{BA}$, the matrices $\mathbf{A}$ and $\mathbf{B}$ are said to *commute*. One important case is when $\mathbf{A}$ and $\mathbf{B}$ are diagonal. In general $\mathbf{A}$ and $\mathbf{B}$ commute if they share the same eigensystem.

Example B.14. Matrices

$$\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} a-\beta & b \\ b & c-\beta \end{bmatrix}, \tag{B.32}$$

commute for any a, b, c, β. More generally, $\mathbf{A}$ and $\mathbf{B} = \mathbf{A} - \beta\mathbf{I}$ commute for any square matrix $\mathbf{A}$.

Transpose of a Product. The transpose of a matrix product is equal to the product of the transposes of the operands taken in reverse order:

$$(\mathbf{A}\,\mathbf{B})^T = \mathbf{B}^T\,\mathbf{A}^T. \tag{B.33}$$

The general transposition formula for an arbitrary product sequence is

$$(\mathbf{A}\,\mathbf{B}\,\mathbf{C}\ldots\mathbf{M}\,\mathbf{N})^T = \mathbf{N}^T\,\mathbf{M}^T\ldots\mathbf{C}^T\,\mathbf{B}^T\,\mathbf{A}^T. \tag{B.34}$$

Congruential Transformation. If $\mathbf{B}$ is a *symmetric* matrix of order m and $\mathbf{A}$ is an arbitrary $m \times n$ matrix, then

$$\mathbf{S} = \mathbf{A}^T \, \mathbf{B} \, \mathbf{A}. \tag{B.35}$$

is a symmetric matrix of order n. Such an operation is called a congruential transformation. It occurs very frequently in finite element analysis when changing coordinate bases because such a transformation preserves energy.

Loss of Symmetry. The product of two symmetric matrices is not generally symmetric.

Null Matrices may have Non-null Divisors. The matrix product $\mathbf{AB}$ can be zero although $\mathbf{A} \neq \mathbf{0}$ and $\mathbf{B} \neq \mathbf{0}$. Likewise it is possible that $\mathbf{A} \neq \mathbf{0}$, $\mathbf{A}^2 \neq \mathbf{0}$, $\ldots$, but $\mathbf{A}^p = \mathbf{0}$.

§B.4. Bilinear and Quadratic Forms

Let $\mathbf{x}$ and $\mathbf{y}$ be two column vectors of order n, and $\mathbf{A}$ a real square $n \times n$ matrix. Then the following triple product produces a *scalar* result:

$$s = \underset{1\times n}{\mathbf{y}^T} \; \underset{n\times n}{\mathbf{A}} \; \underset{n\times 1}{\mathbf{x}} \tag{B.36}$$

This is called a *bilinear form*. Matrix $\mathbf{A}$ is called the *kernel* of the form.

Transposing both sides of (B.36) and noting that the transpose of a scalar does not change, we obtain the result

$$s = \mathbf{x}^T \mathbf{A}^T \mathbf{y} = \mathbf{y}^T \mathbf{A} \mathbf{x}. \tag{B.37}$$

If $\mathbf{A}$ is symmetric and vectors $\mathbf{x}$ and $\mathbf{y}$ coalesce, *i.e.*

$$\mathbf{A}^T = \mathbf{A}, \qquad \mathbf{x} = \mathbf{y}, \tag{B.38}$$

the bilinear form becomes a *quadratic form*

$$s = \mathbf{x}^T \mathbf{A} \mathbf{x}. \tag{B.39}$$

Transposing both sides of a quadratic form reproduces the same equation.

Example B.15. The kinetic energy of a dynamic system consisting of three point masses m_1, m_2, m_3 moving in one dimension with velocities v_1, v_2 and v_3, respectively, is

$$T = \tfrac{1}{2}(m_1 v_1^2 + m_2 v_2^2 + m_3 v_3^2). \tag{B.40}$$

This can be expressed as the quadratic form

$$T = \tfrac{1}{2} \mathbf{v}^T \, \mathbf{M} \, \mathbf{v}, \tag{B.41}$$

in which

$$\mathbf{M} = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}. \tag{B.42}$$

Here $\mathbf{M}$ denotes the system mass matrix whereas $\mathbf{v}$ is the system velocity vector.

§B.5. Matrix Orthogonality

Let $\mathbf{A}$ and $\mathbf{B}$ be two product-conforming real matrices. For example, $\mathbf{A}$ is $k \times m$ whereas $\mathbf{B}$ is $m \times n$. If their product is the null matrix

$$\mathbf{C} = \mathbf{A}\,\mathbf{B} = \mathbf{0}, \tag{B.43}$$

the matrices are said to be *orthogonal*. This is the generalization of the notions of vector orthogonality discussed in the previous Appendix.

§B.5.1. Matrix Orthogonalization Via Projectors

The *matrix orthogonalization* problem can be stated as follows. Product conforming matrices $\mathbf{A}$ and $\mathbf{B}$ are given but their product is not zero. How can $\mathbf{A}$ be orthogonalized with respect to $\mathbf{B}$ so that (B.43) is verified? Suppose that $\mathbf{B}$ is $m \times n$ with $m \ge n$ and that $\mathbf{B}^T\mathbf{B}$ is nonsingular (equivalently, $\mathbf{B}$ has full rank).[3] Then form the $m \times m$ *orthogonal projector matrix*, or simply *projector*

$$\boxed{\mathbf{P}_B = \mathbf{I} - \mathbf{B}\,(\mathbf{B}^T\,\mathbf{B})^{-1}\,\mathbf{B}^T.} \tag{B.44}$$

in which $\mathbf{I}$ is the $m \times m$ identity matrix. Since $\mathbf{P}_B = \mathbf{P}_B^T$, the projector is square symmetric.

Note that

$$\mathbf{P}_B\,\mathbf{B} = \mathbf{B} - \mathbf{B}\,(\mathbf{B}^T\,\mathbf{B})^{-1}\,(\mathbf{B}^T\mathbf{B}) = \mathbf{B} - \mathbf{B} = \mathbf{0}. \tag{B.45}$$

It follows that $\mathbf{P}_B$ projects $\mathbf{B}$ onto its null space. Likewise $\mathbf{B}^T\,\mathbf{P}_B = \mathbf{0}$. Postmultiplying $\mathbf{A}$ by $\mathbf{P}_B$ yields

$$\boxed{\tilde{\mathbf{A}} = \mathbf{A}\,\mathbf{P}_B = \mathbf{A} - \mathbf{A}\,\mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\,\mathbf{B}^T.} \tag{B.46}$$

Matrix $\tilde{\mathbf{A}}$ is called the *projection* of $\mathbf{A}$ onto the null space of $\mathbf{B}$.[4] It is easily verified that $\tilde{\mathbf{A}}$ and $\mathbf{B}$ are orthogonal:

$$\tilde{\mathbf{A}}\,\mathbf{B} = \mathbf{A}\,\mathbf{B} - \mathbf{A}\,\mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\,(\mathbf{B}^T\mathbf{B}) = \mathbf{A}\,\mathbf{B} - \mathbf{A}\,\mathbf{B} = \mathbf{0}. \tag{B.47}$$

Consequently, forming $\tilde{\mathbf{A}}$ via (B.44) and (B.46) solves the orthogonalization problem.

If $\mathbf{B}$ is square and nonsingular, $\tilde{\mathbf{A}} = \mathbf{0}$, as may be expected. If $\mathbf{B}$ has more columns than rows, that is $m < n$, the projector (B.44) cannot be constructed since $\mathbf{B}\,\mathbf{B}^T$ is necessarily singular. A similar difficulty arises if $m \ge n$ but $\mathbf{B}^T\mathbf{B}$ is singular. Such cases require treatment using generalized inverses, which is a topic beyond the scope of this Appendix.[5]

In some applications, notably FEM, matrix $\mathbf{A}$ is square symmetric and it is desirable to preserve symmetry in $\tilde{\mathbf{A}}$. That can be done by pre-and postmultiplying by the projector:

$$\boxed{\tilde{\mathbf{A}} = \mathbf{P}_B\,\mathbf{A}\,\mathbf{P}_B.} \tag{B.48}$$

Since $\mathbf{P}_B^T = \mathbf{P}_B$, the operation (B.48) is a congruential transformation, which preserves symmetry.

---

[3] If you are not sure what “singular”, “nonsingular” and “rank” mean or what $(.)^{-1}$ stands for, please read §D.4.

[4] In contexts such as control and signal processing, $\mathbf{P}_B$ is called a *filter* and the operation (B.46) is called *filtering*.

[5] See e.g., the textbooks [77,596].

§B.5.2. Orthogonal Projector Properties

The following properties of the projector (B.44) are useful when checking out computations. Forming its square as

$$
\begin{aligned}
\mathbf{P}_B^2 &= \mathbf{P}_B\,\mathbf{P}_B = \mathbf{I} - 2\mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T + \mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\,\mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T \\
&= \mathbf{I} - 2\mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T + \mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \mathbf{I} - \mathbf{B}\,(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \mathbf{P}_B,
\end{aligned}
\tag{B.49}
$$

shows that the projector matrix is idempotent. Repeating the process one sees that $\mathbf{P}_B^n = \mathbf{P}_B$, in which n is an arbitrary nonnegative integer.

If $\mathbf{B}$ is $m \times n$ with $m \geq n$ and full rank n, $\mathbf{P}_B$ has $m - n$ unit eigenvalues and n zero eigenvalues. This is shown in the paper [235], in which various applications of orthogonal projectors and orthogonalization to multilevel FEM computations are covered in detail.

Homework Exercises for Appendix B: Matrices

EXERCISE B.1 Given the three matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 1 & 0 \\ -1 & 2 & 3 & 1 \\ 2 & 5 & -1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & -2 \\ 1 & 0 \\ 4 & 1 \\ -3 & 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 0 & 2 \end{bmatrix} \tag{EB.1}$$

compute the product $\mathbf{D} = \mathbf{ABC}$ by hand using Falk's scheme. *Hint:* do **BC** first, then premultiply that by **A**.

EXERCISE B.2 Given the square matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ -4 & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & 0 \\ 1 & -2 \end{bmatrix} \tag{EB.2}$$

verify by direct computation that $\mathbf{AB} \neq \mathbf{BA}$.

EXERCISE B.3 Given the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 2 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & -1 & 4 \\ -1 & 2 & 0 \\ 4 & 0 & 0 \end{bmatrix} \tag{EB.3}$$

(note that **B** is symmetric) compute $\mathbf{S} = \mathbf{A}^T\mathbf{BA}$, and verify that **S** is symmetric.

EXERCISE B.4 Given the square matrices

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 0 & 3 \\ 3 & -2 & -5 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & -6 & -3 \\ 7 & -14 & -7 \\ -1 & 2 & 1 \end{bmatrix} \tag{EB.4}$$

verify that $\mathbf{AB} = \mathbf{0}$ although $\mathbf{A} \neq \mathbf{0}$ and $\mathbf{B} \neq \mathbf{0}$. Is **BA** also null?

EXERCISE B.5 Given the square matrix

$$\mathbf{A} = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \tag{EB.5}$$

show by direct computation that $\mathbf{A}^2 \neq \mathbf{0}$ but $\mathbf{A}^3 = \mathbf{0}$.

EXERCISE B.6 Can a diagonal matrix be antisymmetric?

EXERCISE B.7 (Tougher) Prove the matrix product transposition rule (B.33). *Hint:* call $\mathbf{C} = (\mathbf{AB})^T$, $\mathbf{D} = \mathbf{B}^T\mathbf{A}^T$, and use the matrix product definition (B.25) to show that the generic entries of **C** and **D** agree.

EXERCISE B.8 If **A** is an arbitrary $m \times n$ matrix, show: (a) both products $\mathbf{A}^T\mathbf{A}$ and $\mathbf{AA}^T$ are possible, and (b) both products are square and symmetric. *Hint*: for (b) use the symmetry condition $\mathbf{S} = \mathbf{S}^T$ and (B.31).

EXERCISE B.9 Show that $\mathbf{A}^2$ only exists if and only if **A** is square.

EXERCISE B.10 If **A** is square and antisymmetric, show that $\mathbf{A}^2$ is symmetric. *Hint*: start from $\mathbf{A} = -\mathbf{A}^T$ and apply the results of Exercise B.8.

Homework Exercises for Appendix B - Solutions

EXERCISE B.1

$$\begin{array}{rr} & \begin{bmatrix} 1 & -3 & 2 \\ 2 & 0 & 2 \end{bmatrix} = \mathbf{C} \\ \mathbf{B} = \begin{bmatrix} 2 & -2 \\ 1 & 0 \\ 4 & 1 \\ -3 & 2 \end{bmatrix} & \begin{bmatrix} -2 & -6 & 0 \\ 1 & -3 & 2 \\ 6 & -12 & 10 \\ 1 & 9 & -2 \end{bmatrix} = \mathbf{BC} \\ \mathbf{A} = \begin{bmatrix} 2 & 4 & 1 & 0 \\ -1 & 2 & 3 & 1 \\ 2 & 5 & -1 & 2 \end{bmatrix} & \begin{bmatrix} 6 & -36 & 18 \\ 23 & -27 & 32 \\ -3 & 3 & -4 \end{bmatrix} = \mathbf{ABC} = \mathbf{D} \end{array} . \tag{EB.6}$$

EXERCISE B.2

$$\mathbf{AB} = \begin{bmatrix} 6 & -6 \\ -10 & -4 \end{bmatrix} \neq \mathbf{BA} = \begin{bmatrix} 3 & 9 \\ 9 & -1 \end{bmatrix} \tag{EB.7}$$

EXERCISE B.3

$$\mathbf{S} = \mathbf{A}^T\mathbf{BA} = \begin{bmatrix} 23 & -6 \\ -6 & 8 \end{bmatrix}, \tag{EB.8}$$

which is symmetric, like **B**.

EXERCISE B.4

$$\begin{array}{rr} & \begin{bmatrix} 3 & -6 & -3 \\ 7 & -14 & -7 \\ -1 & 2 & 1 \end{bmatrix} = \mathbf{B} \\ \mathbf{A} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 0 & 3 \\ 3 & -2 & -5 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{AB} = \mathbf{0} \end{array} . \tag{EB.9}$$

However,

$$\mathbf{BA} = \begin{bmatrix} -6 & 3 & 3 \\ -14 & 7 & 7 \\ 2 & -1 & -1 \end{bmatrix} \neq \mathbf{0}. \tag{EB.10}$$

EXERCISE B.5

$$\mathbf{A}^2 = \mathbf{AA} = \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{A}^3 = \mathbf{AAA} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{0} \tag{EB.11}$$

EXERCISE B.6 Only if it is the null matrix.

EXERCISE B.7 To avoid "indexing indigestion" let us carefully specify the dimensions of the given matrices and their transposes:

$$\underset{m\times n}{\mathbf{A}} = [a_{ij}], \qquad \underset{n\times m}{\mathbf{A}^T} = [a_{ji}]. \tag{EB.12}$$

$$\underset{n\times p}{\mathbf{B}} = [b_{jk}], \qquad \underset{p\times n}{\mathbf{B}^T} = [b_{kj}] \tag{EB.13}$$

Indices i, j and k run over $1\ldots m$, $1\ldots n$ and $1\ldots p$, respectively. Now call

$$\underset{p\times m}{\mathbf{C}} = [c_{ki}] = (\mathbf{AB})^T. \tag{EB.14}$$

$$\underset{p\times m}{\mathbf{D}} = [d_{ki}] = \mathbf{B}^T \mathbf{A}^T. \tag{EB.15}$$

From the definition of matrix product,

$$c_{ki} = \sum_{j=1}^{n} a_{ij} b_{jk}. \tag{EB.16}$$

$$d_{ki} = \sum_{j=1}^{n} b_{jk} a_{ij} = \sum_{j=1}^{n} a_{ij} b_{jk} = c_{ki}. \tag{EB.17}$$

Hence $\mathbf{C} = \mathbf{D}$ for any $\mathbf{A}$ and $\mathbf{B}$, and the statement is proved.

EXERCISE B.8

(a) If $\mathbf{A}$ is $m \times n$, $\mathbf{A}^T$ is $n \times m$. Next we write the two products to be investigated:

$$\underset{n\times m}{\mathbf{A}^T}\ \underset{m\times n}{\mathbf{A}}, \qquad \underset{m\times n}{\mathbf{A}}\ \underset{n\times m}{\mathbf{A}^T} \tag{EB.18}$$

In both cases the column dimension of the premultiplicand is equal to the row dimension of the postmultiplicand. Therefore both products are possible.

(b) To verify symmetry we use three results. First, the symmetry test: transpose equals original; second, transposing twice gives back the original; and, finally, the transposed-product formula proved in Exercise B.7.

$$(\mathbf{A}^T \mathbf{A})^T = \mathbf{A}^T (\mathbf{A}^T)^T = \mathbf{A}^T \mathbf{A}. \tag{EB.19}$$

$$(\mathbf{A}\mathbf{A}^T)^T = (\mathbf{A}^T)^T \mathbf{A}^T = \mathbf{A}\mathbf{A}^T. \tag{EB.20}$$

Or, to do it more leisurely, call $\mathbf{B} = \mathbf{A}^T$, $\mathbf{B}^T = \mathbf{A}$, $\mathbf{C} = \mathbf{A}\mathbf{B}$, and let's go over the first one again:

$$\mathbf{C}^T = (\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T = \mathbf{A}\mathbf{A}^T = \mathbf{A}\mathbf{B} = \mathbf{C}. \tag{EB.21}$$

Since $\mathbf{C} = \mathbf{C}^T$, $\mathbf{C} = \mathbf{A}\mathbf{A}^T$ is symmetric. Same mechanics for the second one.

EXERCISE B.9 Let $\mathbf{A}$ be $m \times n$. For $\mathbf{A}^2 = \mathbf{A}\mathbf{A}$ to exist, the column dimension n of the premultiplicand $\mathbf{A}$ must equal the row dimension m of the postmultiplicand $\mathbf{A}$. Hence $m = n$ and $\mathbf{A}$ must be square.

EXERCISE B.10 Premultiply both sides of $\mathbf{A} = -\mathbf{A}^T$ by $\mathbf{A}$ (which is always possible because $\mathbf{A}$ is square):

$$\mathbf{A}^2 = \mathbf{A}\mathbf{A} = -\mathbf{A}\mathbf{A}^T. \tag{EB.22}$$

But from Exercise B.8 we know that $\mathbf{A}\mathbf{A}^T$ is symmetric. Since the negated of a symmetric matrix is symmetric, so is $\mathbf{A}^2$.