

An Introduction to Matrix Algebra

E.1 DEFINITIONS

In many situations, we must deal with rectangular arrays of numbers or functions. The rectangular array of numbers (or functions)

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (\text{E.1})$$

is known as a **matrix**. The numbers a_{ij} are called **elements** of the matrix, with the subscript i denoting the row and the subscript j denoting the column.

A matrix with m rows and n columns is said to be a matrix of **order** (m, n) or alternatively called an $m \times n$ (m -by- n) matrix. When the number of the columns equals the number of rows ($m = n$), the matrix is called a **square matrix** of order n . It is common to use boldfaced capital letters to denote an $m \times n$ matrix.

A matrix comprising only one column, that is, an $m \times 1$ matrix, is known as a column matrix or, more commonly, a **column vector**. We will represent a column vector with boldfaced lowercase letters as

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \quad (\text{E.2})$$

Analogously, a **row vector** is an ordered collection of numbers written in a row—that is, a $1 \times n$ matrix. We will use boldfaced lowercase letters to represent vectors. Therefore a row vector will be written as

$$\mathbf{z} = [z_1 \quad z_2 \quad \cdots \quad z_n], \quad (\text{E.3})$$

with n elements.

A few matrices with distinctive characteristics are given special names. A square matrix in which all the elements are zero except those on the principal diagonal, a_{11} , a_{22} , \dots , a_{nn} , is called a **diagonal matrix**. Then, for example, a 3×3 diagonal matrix would be

$$\mathbf{B} = \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}. \quad (\text{E.4})$$

If all the elements of a diagonal matrix have the value 1, then the matrix is known as the **identity matrix** \mathbf{I} , which is written as

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (\text{E.5})$$

When all the elements of a matrix are equal to zero, the matrix is called the **zero**, or **null matrix**. When the elements of a matrix have a special relationship so that $a_{ij} = a_{ji}$, it is called a **symmetrical** matrix. Thus, for example, the matrix

$$\mathbf{H} = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 6 & 4 \\ 1 & 4 & 8 \end{bmatrix} \quad (\text{E.6})$$

is a symmetrical matrix of order (3, 3).

E.2 ADDITION AND SUBTRACTION OF MATRICES

The addition of two matrices is possible only for matrices of the same order. The sum of two matrices is obtained by adding the corresponding elements. Thus if the elements of \mathbf{A} are a_{ij} and the elements of \mathbf{B} are b_{ij} , and if

$$\mathbf{C} = \mathbf{A} + \mathbf{B}, \quad (\text{E.7})$$

then the elements of \mathbf{C} that are c_{ij} are obtained as

$$c_{ij} = a_{ij} + b_{ij}. \quad (\text{E.8})$$

For example, the matrix addition for two 3×3 matrices is as follows:

$$\mathbf{C} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 3 \\ 0 & 6 & 2 \end{bmatrix} + \begin{bmatrix} 8 & 2 & 1 \\ 1 & 3 & 0 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 3 & 1 \\ 2 & 2 & 3 \\ 4 & 8 & 3 \end{bmatrix}. \quad (\text{E.9})$$

From the operation used for performing the operation of addition, we note that the process is **commutative**; that is,

$$\boxed{\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.} \quad (\text{E.10})$$

Also we note that the addition operation is **associative**, so that

$$\boxed{(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}).} \quad (\text{E.11})$$

To perform the operation of subtraction, we note that if a matrix \mathbf{A} is multiplied by a constant α , then every element of the matrix is multiplied by this constant. Therefore we can write

$$\alpha \mathbf{A} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{12} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}. \quad (\text{E.12})$$

Then to carry out a subtraction operation, we use $\alpha = -1$, and $-\mathbf{A}$ is obtained by multiplying each element of \mathbf{A} by -1 . For example,

$$\mathbf{C} = \mathbf{B} - \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} - \begin{bmatrix} 6 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 1 & 1 \end{bmatrix}. \quad (\text{E.13})$$

E.3 MULTIPLICATION OF MATRICES

The multiplication of two matrices \mathbf{AB} requires that the number of columns of \mathbf{A} be equal to the number of rows of \mathbf{B} . Thus if \mathbf{A} is of order $m \times n$ and \mathbf{B} is of order $n \times q$, then the product is of order $m \times q$. The elements of a product

$$\mathbf{C} = \mathbf{AB} \quad (\text{E.14})$$

are found by multiplying the i th row of \mathbf{A} and the j th column of \mathbf{B} and summing these products to give the element c_{ij} . That is,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{iq}b_{qj} = \sum_{k=1}^q a_{ik}b_{kj}. \quad (\text{E.15})$$

Thus we obtain c_{11} , the first element of \mathbf{C} , by multiplying the first row of \mathbf{A} by the first column of \mathbf{B} and summing the products of the elements. We should note that, in general, matrix multiplication is not commutative; that is

$$\boxed{\mathbf{AB} \neq \mathbf{BA}.} \quad (\text{E.16})$$

Also we note that the multiplication of a matrix of $m \times n$ by a column vector (order $n \times 1$) results in a column vector of order $m \times 1$.

A specific example of multiplication of a column vector by a matrix is

$$\mathbf{x} = \mathbf{Ay} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} (a_{11}y_1 + a_{12}y_2 + a_{13}y_3) \\ (a_{21}y_1 + a_{22}y_2 + a_{23}y_3) \end{bmatrix}. \quad (\text{E.17})$$

Note that \mathbf{A} is of order 2×3 , and \mathbf{y} is of order 3×1 . Therefore the resulting matrix \mathbf{x} is of order 2×1 , which is a column vector with two rows. There are two elements of \mathbf{x} , and

$$x_1 = (a_{11}y_1 + a_{12}y_2 + a_{13}y_3) \quad (\text{E.18})$$

is the first element obtained by multiplying the first row of \mathbf{A} by the first (and only) column of \mathbf{y} .

Another example, which the reader should verify, is

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 7 & 6 \\ -5 & -6 \end{bmatrix}. \quad (\text{E.19})$$

For example, the element c_{22} is obtained as $c_{22} = -1(2) + 2(-2) = -6$.

Now we are able to use this definition of multiplication in representing a set of simultaneous linear algebraic equations by a matrix equation. Consider the following set of algebraic equations:

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= u_1, \\ 2x_1 + x_2 + 6x_3 &= u_2, \\ 4x_1 - x_2 + 2x_3 &= u_3. \end{aligned} \quad (\text{E.20})$$

We can identify two column vectors as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{and} \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}. \quad (\text{E.21})$$

Then we can write the matrix equation

$$\mathbf{Ax} = \mathbf{u}, \quad (\text{E.22})$$

where

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 6 \\ 4 & -1 & 2 \end{bmatrix}.$$

We immediately note the utility of the matrix equation as a compact form of a set of simultaneous equations.

The multiplication of a row vector and a column vector can be written as

$$\mathbf{xy} = [x_1 \quad x_2 \quad \cdots \quad x_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n. \quad (\text{E.23})$$

Thus we note that the multiplication of a row vector and a column vector results in a number that is a sum of a product of specific elements of each vector.

As a final item in this section, we note that the multiplication of any matrix by the identity matrix results in the original matrix, that is, $\mathbf{AI} = \mathbf{A}$.

E.4 OTHER USEFUL MATRIX OPERATIONS AND DEFINITIONS

The **transpose** of a matrix \mathbf{A} is denoted in this text as \mathbf{A}^T . One will often find the notation \mathbf{A}' for \mathbf{A}^T in the literature. The transpose of a matrix \mathbf{A} is obtained by interchanging the rows and columns of \mathbf{A} . For example, if

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

then

$$\mathbf{A}^T = \begin{bmatrix} 6 & 1 & -2 \\ 0 & 4 & 3 \\ 2 & 1 & -1 \end{bmatrix}. \quad (\text{E.24})$$

Therefore we are able to denote a row vector as the transpose of a column vector and write

$$\mathbf{x}^T = [x_1 \quad x_2 \quad \cdots \quad x_n]. \quad (\text{E.25})$$

Because \mathbf{x}^T is a row vector, we obtain a matrix multiplication of \mathbf{x}^T by \mathbf{x} as follows:

$$\mathbf{x}^T \mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1^2 + x_2^2 + \cdots + x_n^2. \quad (\text{E.26})$$

Thus the multiplication $\mathbf{x}^T \mathbf{x}$ results in the sum of the squares of each element of \mathbf{x} .

The transpose of the product of two matrices is the product in reverse order of their transposes, so that

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T. \quad (\text{E.27})$$

The sum of the main diagonal elements of a square matrix \mathbf{A} is called the **trace** of \mathbf{A} , written as

$$\text{tr } \mathbf{A} = a_{11} + a_{22} + \dots + a_{nn}. \quad (\text{E.28})$$

The **determinant** of a square matrix is obtained by enclosing the elements of the matrix \mathbf{A} within vertical bars; for example,

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (\text{E.29})$$

If the determinant of \mathbf{A} is equal to zero, then the determinant is said to be **singular**. The value of a determinant is determined by obtaining the minors and cofactors of the determinants. The **minor** of an element a_{ij} of a determinant of order n is a determinant of order $(n - 1)$ obtained by removing the row i and the column j of the original determinant. The cofactor of a given element of a determinant is the minor of the element with either a plus or minus sign attached; hence

$$\text{cofactor of } a_{ij} = \alpha_{ij} = (-1)^{i+j} M_{ij},$$

where M_{ij} is the minor of a_{ij} . For example, the cofactor of the element a_{23} of

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad (\text{E.30})$$

is

$$\alpha_{23} = (-1)^5 M_{23} = - \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}. \quad (\text{E.31})$$

The value of a determinant of second order (2×2) is

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}a_{22} - a_{21}a_{12}). \quad (\text{E.32})$$

The general n th-order determinant has a value given by

$$\boxed{\det \mathbf{A} = \sum_{j=1}^n a_{ij}\alpha_{ij}} \quad \text{with } i \text{ chosen for one row,} \quad (\text{E.33})$$

or

$$\boxed{\det \mathbf{A} = \sum_{i=1}^n a_{ij}\alpha_{ij}} \quad \text{with } j \text{ chosen for one column.} \quad (\text{E.33})$$

That is, the elements a_{ij} are chosen for a specific row (or column), and that entire row (or column) is expanded according to Eq. (E.33). For example, the value of a specific 3×3 determinant is

$$\begin{aligned} \det \mathbf{A} &= \det \begin{bmatrix} 2 & 3 & 5 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \\ &= 2 \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} - 1 \begin{vmatrix} 3 & 5 \\ 1 & 0 \end{vmatrix} + 2 \begin{vmatrix} 3 & 5 \\ 0 & 1 \end{vmatrix} \\ &= 2(-1) - (-5) + 2(3) = 9, \end{aligned} \quad (\text{E.34})$$

where we have expanded in the first column.

The **adjoint matrix** of a square matrix \mathbf{A} is formed by replacing each element a_{ij} by the cofactor α_{ij} and transposing. Therefore

$$\text{adjoint } \mathbf{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix}^T = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} \end{bmatrix}. \quad (\text{E.35})$$

E.5 MATRIX INVERSION

The inverse of a square matrix \mathbf{A} is written as \mathbf{A}^{-1} and is defined as satisfying the relationship

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}. \quad (\text{E.36})$$

The **inverse** of a matrix \mathbf{A} is

$$\boxed{\mathbf{A}^{-1} = \frac{\text{adjoint of } \mathbf{A}}{\det \mathbf{A}}} \quad (\text{E.37})$$

when the $\det \mathbf{A}$ is not equal to zero. For a 2×2 matrix we have the adjoint matrix

$$\text{adjoint } \mathbf{A} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}, \quad (\text{E.38})$$

and the $\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21}$. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 0 & -1 & 1 \end{bmatrix}. \quad (\text{E.39})$$

The determinant has a value $\det \mathbf{A} = -7$. The cofactor α_{11} is

$$\alpha_{11} = (-1)^2 \begin{vmatrix} -1 & 4 \\ -1 & 1 \end{vmatrix} = 3. \quad (\text{E.40})$$

In a similar manner we obtain

$$\mathbf{A}^{-1} = \frac{\text{adjoint } \mathbf{A}}{\det \mathbf{A}} = \left(-\frac{1}{7}\right) \begin{bmatrix} 3 & -5 & 11 \\ -2 & 1 & 2 \\ -2 & 1 & -5 \end{bmatrix}. \quad (\text{E.41})$$

E.6 MATRICES AND CHARACTERISTIC ROOTS

A set of simultaneous linear algebraic equations can be represented by the matrix equation

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad (\text{E.42})$$

where the \mathbf{y} vector can be considered as a transformation of the vector \mathbf{x} . We might ask whether it may happen that a vector \mathbf{y} may be a scalar multiple of \mathbf{x} . Trying $\mathbf{y} = \lambda\mathbf{x}$, where λ is a scalar, we have

$$\lambda\mathbf{x} = \mathbf{A}\mathbf{x}. \quad (\text{E.43})$$

Alternatively Eq. (E.43) can be written as

$$\lambda\mathbf{x} - \mathbf{A}\mathbf{x} = (\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}, \quad (\text{E.44})$$

where \mathbf{I} = identity matrix. Thus the solution for \mathbf{x} exists if and only if

$\det (\lambda\mathbf{I} - \mathbf{A}) = 0.$

(E.45)

This determinant is called the characteristic determinant of \mathbf{A} . Expansion of the determinant of Eq. (E.45) results in the **characteristic equation**. The characteristic equation is an n th-order polynomial in λ . The n roots of this characteristic equation are called the **characteristic roots**. For every possible value λ_i ($i = 1, 2, \dots, n$) of the n th-order characteristic equation, we can write

$$(\lambda_i\mathbf{I} - \mathbf{A})\mathbf{x}_i = \mathbf{0}. \quad (\text{E.46})$$

The vector \mathbf{x}_i is the **characteristic vector** for the i th root. Let us consider the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}. \quad (\text{E.47})$$

The characteristic equation is found as follows:

$$\det \begin{bmatrix} (\lambda - 2) & -1 & -1 \\ -2 & (\lambda - 3) & -4 \\ 1 & 1 & (\lambda + 2) \end{bmatrix} = (-\lambda^3 + 3\lambda^2 + \lambda - 3) = 0. \quad (\text{E.48})$$

The roots of the characteristic equation are $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 3$. When $\lambda = \lambda_1 = 1$, we find the first characteristic vector from the equation

$$\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \quad (\text{E.49})$$

and we have $\mathbf{x}_1^T = k [1 \ -1 \ 0]$, where k is an arbitrary constant usually chosen equal to 1. Similarly, we find

$$\mathbf{x}_2^T [0 \ 1 \ -1],$$

and

$$\mathbf{x}_3^T = [2 \ 3 \ -1]. \quad (\text{E.50})$$

E.7 THE CALCULUS OF MATRICES

The derivative of a matrix $\mathbf{A} = \mathbf{A}(t)$ is defined as

$$\frac{d}{dt} [\mathbf{A}(t)] = \begin{bmatrix} da_{11}(t)/dt & da_{12}(t)/dt & \cdots & da_{1n}(t)/dt \\ \vdots & \vdots & \ddots & \vdots \\ da_{n1}(t)/dt & da_{n2}(t)/dt & \cdots & da_{nn}(t)/dt \end{bmatrix}. \quad (\text{E.51})$$

That is, the derivative of a matrix is simply the derivative of each element $a_{ij}(t)$ of the matrix.

The **matrix exponential function** is defined as the power series

$$\exp [\mathbf{A}] = e^{\mathbf{A}} = \mathbf{I} + \frac{\mathbf{A}}{1!} + \frac{\mathbf{A}^2}{2!} + \cdots + \frac{\mathbf{A}^k}{k!} + \cdots = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}, \quad (\text{E.52})$$

where $\mathbf{A}^2 = \mathbf{A}\mathbf{A}$, and, similarly, \mathbf{A}^k implies \mathbf{A} multiplied k times. This series can be shown to be convergent for all square matrices. Also a matrix exponential that is a function of time is defined as

$$e^{\mathbf{A}t} = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!}. \quad (\text{E.53})$$

If we differentiate with respect to time, then we have

$$\frac{d}{dt}(e^{\mathbf{A}t}) = \mathbf{A}e^{\mathbf{A}t}. \quad (\text{E.54})$$

Therefore for a differential equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}, \quad (\text{E.55})$$

we might postulate a solution $\mathbf{x} = e^{\mathbf{A}t}\mathbf{c} = \phi\mathbf{c}$, where the matrix ϕ is $\phi = e^{\mathbf{A}t}$, and \mathbf{c} is an unknown column vector. Then we have

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}, \quad (\text{E.56})$$

or

$$\mathbf{A}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t}, \quad (\text{E.57})$$

and we have in fact satisfied the relationship, Eq. (E.55). Then the value of \mathbf{c} is simply $\mathbf{x}(0)$, the initial value of \mathbf{x} , because when $t = 0$, we have $\mathbf{x}(0) = \mathbf{c}$. Therefore the solution to Eq. (E.55) is

$$\boxed{\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0).} \quad (\text{E.58})$$

