APPENDIX A

Matrix Mathematics

A.1 DEFINITIONS

The mathematical description of many physical problems is often simplified by the use of rectangular arrays of scalar quantities of the form

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{A.1}$$

Such an array is known as a *matrix,* and the scalar values that compose the array are the *elements* of the matrix. The position of each element a_{ij} is identified by the *row* subscript i and the *column* subscript j.

The number of rows and columns determine the *order* of a matrix. A matrix having m rows and n columns is said to be of order "m by n" (usually denoted as $m \times n$). If the number of rows and columns in a matrix are the same, the matrix is a *square matrix* and said to be of order n. A matrix having only one row is called a *row matrix* or *row vector*. Similarly, a matrix with a single column is a *column matrix* or *column vector*.

If the rows and columns of a matrix $[A]$ are interchanged, the resulting matrix is known as the *transpose* of $[A]$, denoted by $[A]^T$. For the matrix defined in Equation A.1, the transpose is

$$[A]^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \tag{A.2}$$

and we observe that, if $[A]$ is of order m by n, then $[A]^T$ is of order n by m. For example, if $[A]$ is given by

$$[A] = \begin{bmatrix} 2 & -1 & 3 \\ 4 & 0 & 2 \end{bmatrix}$$

the transpose of $[A]$ is

$$[A]^T = \begin{bmatrix} 2 & 4 \\ -1 & 0 \\ 3 & 2 \end{bmatrix}$$

Several important special types of matrices are defined next. A *diagonal* matrix is a square matrix composed of elements such that $a_{ij} = 0$ and $i \neq j$. Therefore, the only nonzero terms are those on the main diagonal (upper left to lower right). For example,

$$[A] = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

is a diagonal matrix.

An *identity* matrix (denoted $[I]$) is a diagonal matrix in which the value of the nonzero terms is unity. Hence,

$$[A] = [I] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is an identity matrix.

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

A *symmetric* matrix is a square matrix composed of elements such that the nondiagonal values are symmetric about the main diagonal. Mathematically, symmetry is expressed as $a_{ij} = a_{ji}$ and $i \neq j$. For example, the matrix

$$[A] = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 4 & -3 \\ 0 & -3 & 1 \end{bmatrix}$$

is a symmetric matrix. Note that the transpose of a symmetric matrix is the same as the original matrix.

A *skew symmetric* matrix is a square matrix in which the diagonal terms a_{ii} have a value of 0 and the off-diagonal terms have values such that $a_{ij} = -a_{ji}$. An example of a skew symmetric matrix is

$$[A] = \begin{bmatrix} 0 & -2 & 0 \\ 2 & 0 & 3 \\ 0 & -3 & 0 \end{bmatrix}$$

For a skew symmetric matrix, we observe that the transpose is obtained by changing the algebraic sign of each element of the matrix.

A.2 ALGEBRAIC OPERATIONS

Addition and *subtraction* of matrices can be defined only for matrices of the same order. If $[A]$ and $[B]$ are both $m \times n$ matrices, the two are said to be *conformable* for addition or subtraction. The sum of two $m \times n$ matrices is another $m \times n$ matrix having elements obtained by summing the corresponding elements of the original matrices. Symbolically, matrix addition is expressed as

$$[C] = [A] + [B] \tag{A.3}$$

where

$$c_{ij} = a_{ij} + b_{ij} \qquad i = 1, m \qquad j = 1, n \tag{A.4}$$

The operation of matrix subtraction is similarly defined. Matrix addition and subtraction are *commutative* and *associative;* that is,

$$[A] + [B] = [B] + [A] \tag{A.5}$$

$$[A] + ([B] + [C]) = ([A] + [B]) + [C] \tag{A.6}$$

The product of a scalar and a matrix is a matrix in which every element of the original matrix is multiplied by the scalar. If a scalar u multiplies matrix $[A]$, then

$$[B] = u[A] \tag{A.7}$$

where the elements of $[B]$ are given by

$$b_{ij} = ua_{ij} \qquad i = 1, m \qquad j = 1, n \tag{A.8}$$

Matrix multiplication is defined in such a way as to facilitate the solution of simultaneous linear equations. The *product* of two matrices $[A]$ and $[B]$ denoted

$$[C] = [A][B] \tag{A.9}$$

exists only if the number of columns in $[A]$ is the equal to the number of rows in $[B]$. If this condition is satisfied, the matrices are said to be *conformable for multiplication*. If $[A]$ is of order $m \times p$ and $[B]$ is of order $p \times n$, the matrix product $[C] = [A][B]$ is an $m \times n$ matrix having elements defined by

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} \tag{A.10}$$

Thus, each element c_{ij} is the sum of products of the elements in the ith row of $[A]$ and the corresponding elements in the jth column of $[B]$. When referring to the matrix product $[A][B]$, matrix $[A]$ is called the *premultiplier* and matrix $[B]$ is the *postmultiplier*.

In general, matrix multiplication is *not* commutative; that is,

$$[A][B] \neq [B][A] \tag{A.11}$$

Matrix multiplication does satisfy the *associative* and *distributive* laws, and we can therefore write

$$\begin{aligned}([A][B])[C] &= [A]([B][C])\\ [A]([B]+[C]) &= [A][B]+[A][C]\\ ([A]+[B])[C] &= [A][C]+[B][C]\end{aligned} \tag{A.12}$$

In addition to being noncommutative, matrix algebra differs from scalar algebra in other ways. For example, the equality $[A][B] = [A][C]$ does not necessarily imply $[B] = [C]$, 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, $[A][B] = [0]$, the result does not necessarily imply that either $[A]$ or $[B]$ is a null matrix.

A.3 DETERMINANTS

The *determinant* of a square matrix is a *scalar* value that is unique for a given matrix. The determinant of an $n \times n$ matrix is represented symbolically as

$$\det[A] = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{A.13}$$

and is evaluated according to a very specific procedure. First, consider the 2×2 matrix

$$[A] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{A.14}$$

for which the determinant is defined as

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \equiv a_{11}a_{22} - a_{12}a_{21} \tag{A.15}$$

Given the definition of Equation A.15, the determinant of a square matrix of any order can be determined.

Next, consider the determinant of a 3×3 matrix

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{A.16}$$

defined as

$$|A| = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \tag{A.17}$$

Note that the expressions in parentheses are the determinants of the second-order matrices obtained by striking out the first row and the first, second, and third columns, respectively. 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. The minor obtained by removing row i and column j is denoted $|M_{ij}|$. Using this notation, Equation A.17 becomes

$$|A| = a_{11}|M_{11}| - a_{12}|M_{12}| + a_{13}|M_{13}| \tag{A.18}$$

and the determinant is said to be expanded in terms of the *cofactors* of the first row. The cofactors of an element a_{ij} are obtained by applying the appropriate algebraic sign to the minor $|M_{ij}|$ as follows. If the sum of row number i and column number j is even, the sign of the cofactor is positive; if $i + j$ is odd, the sign of the cofactor is negative. Denoting the cofactor as C_{ij} we can write

$$C_{ij} = (-1)^{i+j}|M_{ij}| \tag{A.19}$$

The determinant given in Equation A.18 can then be expressed in terms of cofactors as

$$|A| = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \tag{A.20}$$

The determinant of a square matrix of *any* order can be obtained by expanding the determinant in terms of the cofactors of any row i as

$$|A| = \sum_{j=1}^{n} a_{ij}C_{ij} \tag{A.21}$$

or any column j as

$$|A| = \sum_{i=1}^{n} a_{ij}C_{ij} \tag{A.22}$$

Application of Equation A.21 or A.22 requires that the cofactors C_{ij} be further expanded to the point that all minors are of order 2 and can be evaluated by Equation A.15.

A.4 MATRIX INVERSION

The *inverse* of a square matrix $[A]$ is a square matrix denoted by $[A]^{-1}$ and satisfies

$$[A]^{-1}[A] = [A][A]^{-1} = [I] \tag{A.23}$$

that is, the product of a square matrix and its inverse is the identity matrix of order n. The concept of the inverse of a matrix is of prime importance in solving simultaneous linear equations by matrix methods. Consider the algebraic system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= y_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= y_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= y_3 \end{aligned} \tag{A.24}$$

which can be written in matrix form as

$$[A]\{x\} = \{y\} \tag{A.25}$$

where

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{A.26}$$

is the 3 × 3 *coefficient matrix,*

$$\{x\} = \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} \tag{A.27}$$

is the 3 × 1 column matrix (vector) of unknowns, and

$$\{y\} = \begin{Bmatrix} y_1 \\ y_2 \\ y_3 \end{Bmatrix} \tag{A.28}$$

is the 3 × 1 column matrix (vector) representing the right-hand sides of the equations (the "forcing functions").

If the inverse of matrix $[A]$ can be determined, we can multiply both sides of Equation A.25 by the inverse to obtain

$$[A]^{-1}[A]\{x\} = [A]^{-1}\{y\} \tag{A.29}$$

Noting that

$$[A]^{-1}[A]\{x\} = ([A]^{-1}[A])\{x\} = [I]\{x\} = \{x\} \tag{A.30}$$

the solution for the simultaneous equations is given by Equation A.29 directly as

$$\{x\} = [A]^{-1}\{y\} \tag{A.31}$$

While presented in the context of a system of three equations, the result represented by Equation A.31 is applicable to any number of simultaneous algebraic equations and gives the unique solution for the system of equations.

The inverse of matrix $[A]$ can be determined in terms of its cofactors and determinant as follows. Let the *cofactor matrix* $[C]$ be the square matrix having as elements the cofactors defined in Equation A.19. The *adjoint* of $[A]$ is defined as

$$\operatorname{adj}[A] = [C]^T \tag{A.32}$$

The inverse of $[A]$ is then formally given by

$$[A]^{-1} = \frac{\operatorname{adj}[A]}{|A|} \tag{A.33}$$

If the determinant of $[A]$ is 0, Equation A.33 shows that the inverse does not exist. In this case, the matrix is said to be *singular* and Equation A.31 provides no solution for the system of equations. Singularity of the coefficient matrix indicates one of two possibilities: (1) no solution exists or (2) multiple (nonunique) solutions exist. In the latter case, the algebraic equations are not linearly independent.

Calculation of the inverse of a matrix per Equation A.33 is cumbersome and not very practical. Fortunately, many more efficient techniques exist. One such technique is the *Gauss-Jordan reduction* method, which is illustrated using a 2×2 matrix:

$$[A] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{A.34}$$

The gist of the Gauss-Jordan method is to perform simple row and column operations such that the matrix is reduced to an identity matrix. The sequence of operations required to accomplish this reduction produces the inverse. If we divide the first row by a_{11}, the operation is the same as the multiplication

$$[B_1][A] = \begin{bmatrix} \dfrac{1}{a_{11}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & \dfrac{a_{12}}{a_{11}} \\ a_{21} & a_{22} \end{bmatrix} \tag{A.35}$$

Next, multiply the first row by a_{21} and subtract from the second row, which is equivalent to the matrix multiplication

$$[B_2][B_1][A] = \begin{bmatrix} 1 & 0 \\ -a_{21} & 1 \end{bmatrix} \begin{bmatrix} 1 & \dfrac{a_{12}}{a_{11}} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & \dfrac{a_{12}}{a_{11}} \\ 0 & a_{22} - \dfrac{a_{12}}{a_{11}} a_{21} \end{bmatrix} = \begin{bmatrix} 1 & \dfrac{a_{12}}{a_{11}} \\ 0 & \dfrac{|A|}{a_{11}} \end{bmatrix} \tag{A.36}$$

Multiply the second row by $a_{11}/|A|$:

$$[B_3][B_2][B_1][A] = \begin{bmatrix} 1 & 0 \\ 0 & \dfrac{a_{11}}{|A|} \end{bmatrix} \begin{bmatrix} 1 & \dfrac{a_{12}}{a_{11}} \\ 0 & \dfrac{|A|}{a_{11}} \end{bmatrix} = \begin{bmatrix} 1 & \dfrac{a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} \tag{A.37}$$

Finally, multiply the second row by a_{12}/a_{11} and subtract from the first row:

$$[B_4][B_3][B_2][B_1][A] = \begin{bmatrix} 1 & -\dfrac{a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \dfrac{a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [I] \tag{A.38}$$

Considering Equation A.23, we see that

$$[A]^{-1} = [B_4][B_3][B_2][B_1] \tag{A.39}$$

and carrying out the multiplications in Equation A.39 results in

$$[A]^{-1} = \frac{1}{|A|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{A.40}$$

This application of the Gauss-Jordan procedure may appear cumbersome, but the procedure is quite amenable to computer implementation.

A.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. Consider a system of n linear algebraic equations governing n unknowns x_i expressed in matrix form as

$$[A]\{x\} = \{f\} \tag{A.41}$$

in which we want to eliminate the first p unknowns. The matrix equation can be written in partitioned form as

$$\begin{bmatrix} [A_{11}] & [A_{12}] \\ [A_{21}] & [A_{22}] \end{bmatrix} \begin{Bmatrix} \{X_1\} \\ \{X_2\} \end{Bmatrix} = \begin{Bmatrix} \{F_1\} \\ \{F_2\} \end{Bmatrix} \tag{A.42}$$

where the orders of the submatrices are as follows

$$\begin{aligned} [A_{11}] &\Rightarrow p \times p \\ [A_{12}] &\Rightarrow p \times (n-p) \\ [A_{21}] &\Rightarrow (n-p) \times p \\ [A_{22}] &\Rightarrow (n-p) \times (n-p) \\ \{X_1\}, \{F_1\} &\Rightarrow p \times 1 \\ \{X_2\}, \{F_{12}\} &\Rightarrow (n-p) \times 1 \end{aligned} \tag{A.43}$$

The complete set of equations can now be written in terms of the matrix partitions as

$$\begin{aligned} [A_{11}]\{X_1\} + [A_{12}]\{X_2\} &= \{F_1\} \\ [A_{21}]\{X_1\} + [A_{22}]\{X_2\} &= \{F_2\} \end{aligned} \tag{A.44}$$

The first p equations (the upper partition) are solved as

$$\{X_1\} = [A_{11}]^{-1}(\{F_1\} - [A_{12}]\{X_2\}) \tag{A.45}$$

(implicitly assuming that the inverse of A_{11} exists). Substitution of Equation A.45 into the remaining $n - p$ equations (the lower partition) yields

$$\left([A_{22}] - [A_{21}]\lfloor A_{11}\rfloor^{-1}[A_{12}]\right)\{X_2\} = \{F_2\} - [A_{21}]\lfloor A_{11}\rfloor^{-1}]\{F_1\} \tag{A.46}$$

Equation A.46 is the reduced set of $n - p$ algebraic equations representing the original system and containing all the effects of the first p equations. In the context of finite element analysis, this procedure is referred to as *static condensation*.

As another application (commonly encountered in finite element analysis), we consider the case in which the partitioned values $\{X_1\}$ are known but the corresponding right-hand side partition $\{F_1\}$ is unknown. In this occurrence, the lower partitioned equations are solved directly for $\{X_2]$ to obtain

$$\{X_2\} = [A_{22}]^{-1}(\{F_2\} - [A_{21}]\{X_1\}) \tag{A.47}$$

The unknown values of $\{F_1]$ can then be calculated directly using the equations of the upper partition.