9 Matrix Algebra and Determinants Fall 2003

This section summarizes several useful definitions and relationships for matrix algebra, and discusses the implementation of matrix operations with Maple.

In general, derivations are not included with this summary. If you need to review the basics of matrix algebra, we recommend Edwards and Penney *Differential Equations and Boundary Value Problems*, 2nd ed., Section 5.1, pp. 284-290.

Review of Matrix Operations

A matrix is a rectangular array of numbers or expressions, which are referred to as the *elements* of the matrix. Most of the applications of matrix algebra in physics use either matrices that are *square*, i.e., that have the same number of rows as columns, or matrices that have only a single row or a single column. We'll confine our attention to these kinds of matrices. The number of rows or columns of a square matrix is called the *order* of the matrix. In these notes, a square matrix will usually be denoted by a boldface Roman capital letter, and a matrix with one column by a boldface Roman lowercase letter.

When the elements of a matrix $\mathbf{A}$ are written as A_{ij}, the first index (i) is *always* the row number and the second (j) is *always* the column number. For a matrix of order n, the indices i and j are integers that each range from 1 to n. A matrix $\mathbf{A}$ with two rows and two columns might be written as $\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$.

Addition and Scalar Multiplication

The sum of two matrices $\mathbf{A}$ and $\mathbf{B}$ is denoted by $\mathbf{A + B}$. Two matrices can be added only when they have the same number of rows and the same number of columns. When they do, the matrix $\mathbf{C = A + B}$ is obtained by adding corresponding elements of $\mathbf{A}$ and $\mathbf{B}$. That is, $C_{ij} = A_{ij} + B_{ij}$. Addition of matrices obeys the commutative and associative rules: For any matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$,

$$\mathbf{A + B = B + A} \qquad \text{and} \qquad \mathbf{(A + B) + C = A + (B + C)} \tag{1}$$

A matrix $\mathbf{A}$ can be multiplied by a scalar c. The product $c\mathbf{A}$ is obtained by multiplying every element of $\mathbf{A}$ by c. That is, $(c\mathbf{A})_{ij} = cA_{ij}$. Multiplication of a matrix by a scalar obeys the commutative and distributive rules: For any matrices $\mathbf{A}$ and $\mathbf{B}$ and any scalars c and d,

$$c\mathbf{A} = \mathbf{A}c, \qquad c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}, \qquad \text{and} \qquad (c + d)\mathbf{A} = c\mathbf{A} + d\mathbf{A}. \tag{2}$$

The vector difference $\mathbf{A - B}$ is defined as

$$\mathbf{A - B} = \mathbf{A} + (-\mathbf{B}) = \mathrm{A} + (-1)\mathbf{A}. \tag{3}$$

Multiplication of Two Matrices

The product of two matrices **A** and **B** is another matrix **C.** The operation is denoted as

$$\mathbf{C} = \mathbf{AB}. \tag{4}$$

The product is defined only when the number of columns of **A** is equal to the number of rows of **B**. In that case, the elements of **C** are defined as

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} . \tag{5}$$

That is, the (i, j) element of C is obtained by multiplying each element in the i'th row of **A** by the corresponding element in the j'th column of **B,** and adding. If **A** and **B** are square 2×2 matrices, then

$$\mathbf{C} = \mathbf{AB}, \qquad \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}. \tag{6}$$

We invite you to check this with the general definition given by Eq. (5).

The product of a square matrix with a single-column matrix is a single-column matrix.

Matrix multiplication obeys the associative and distributive rules; for any matrices **A**, **B**, and **C**,

$$\mathbf{A(BC) = (AB)C} \qquad \text{and} \qquad \mathbf{(B + C) = AB + AC} \tag{7}$$

But in general matrix multiplication *does not* obey the commutative rule. That is, in general $\mathbf{AB} \neq \mathbf{BA}$. (In Eq. (6), if the A's and B's are interchanged, the result is quite different from the original.)

Exception: The product of two matrices whose only nonzero elements are on the main diagonal (called "*diagonal matrices*") is commutative.

The *identity matrix* (also called *unit matrix*), denoted by **I**, is a square matrix with 1's on the main diagonal and zeroes everywhere else. For 2×2 matrices, the identity matrix is

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{8}$$

This matrix has the property that for any matrix **A**, **IA = AI = A.**

The elements of the identity matrix are the Kroneker delta: $I_{ij} = \delta_{ij}$, where

$$\delta_{ij} = 1 \quad \text{when } i = j, \qquad\qquad \delta_{ij} = 0 \quad \text{when } i \neq j\,. \tag{9}$$

The *inverse* of a matrix **A**, denoted by $\mathbf{A}^{-1}$, is a matrix such that

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

Not every matrix has an inverse. A matrix having no inverse (i.e., a matrix **A** for which $\mathbf{A}^{-1}$ does not exist) is called a *singular* matrix.

Meaning of Matrix Product

Here is an example, using 2×2 matrices, of the motivation for defining the product of two matrices as we have done in Eqs. (5) and (6),. (The discussion can be generalized easily to square and single-column matrices of any order.) Suppose two (or more) variables y_1 and y_2 are related to two (or more) other variables x_1 and x_2 by a linear transformation in the form

$$\begin{aligned} y_1 &= A_{11}x_1 + A_{12}x_2, \\ y_2 &= A_{21}x_1 + A_{22}x_{2,} \end{aligned} \tag{10}$$

where the A's are constants that describe the transformation. We define the matrices

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \qquad \text{Then } \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{11}$$

The transformation can then be written compactly in matrix language as $\mathbf{y = Ax}$.

A second linear transformation relates the y's to a third set of variables, the z's:

$$\begin{aligned} z_1 &= B_{11}y_1 + B_{12}y_2, \\ z_2 &= B_{21}y_1 + B_{22}y_2, \end{aligned} \tag{12}$$

We define

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}. \quad \text{Then } \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \text{ or } \mathbf{z} = \mathbf{By}. \tag{13}$$

These two transformations can be combined into a single transformation from **x** to **z:**

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \text{or} \qquad \mathbf{z} = \mathbf{Cx}.. \tag{14}$$

How are the elements of **C** related to the elements of **A** and **B**? The answer is that **C** is simply the matrix product of **B** and **A,** in that order:

$$\begin{pmatrix} C_{11} & C_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad \text{or simply } \mathbf{C = BA.} \tag{15}$$

The essential point is that the matrix product **BA** is *defined* so that the matrices representing linear transformations combine in this way. We invite you to verify that the two successive transformations described above are produced by the matrix

$$\mathbf{C} = \mathbf{BA}, \quad \text{or} \quad \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} B_{11}A_{11} + B_{12}A_{21} & B_{11}A_{12} + B_{12}A_{22} \\ B_{21}A_{11} + B_{22}A_{21} & B_{21}A_{12} + B_{22}A_{22} \end{pmatrix}. \tag{16}$$

and that this is the matrix product **BA**, as given by Eqs. (5) and (6) (with the A's and B's interchanged).

Note that multiplication of matrices is in general *not* commutative; if the two linear transformations are carried out in the reverse order, the results are different.

Determinants and Linear Equations

The set of equations

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 &= b_1 \\ a_{21}y_1 + a_{22}y_2 &= b_2 \end{aligned} \tag{17}$$

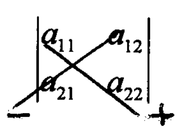
are referred to as a set of *simultaneous linear equations* for the two unknowns y_1 and y_2. It is assumed that all the a's and b's are known. (In this discussion we will assume that the number of equations is the same as the number of unknowns, in this case, two.) The solution of these equations, that is, the values of y_1 and y_2 that satisfy these equations, is given (except for some special cases to be discussed later) by

$$y_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad y_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{18}$$

The determinant $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ is defined as

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \tag{19}$$

and the other determinants are defined similarly.



The diagram shows a scheme for evaluating 2×2 determinants: Multiply the elements along each diagonal and add the results, with the indicated signs.

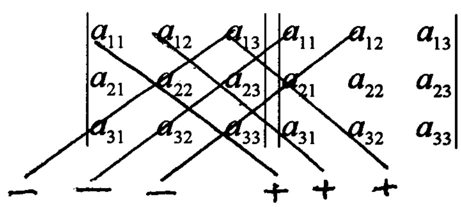
If there are three simultaneous equations for three unknowns, the corresponding expressions are:

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + a_{13}y_3 &= b_1, \\ a_{21}y_1 + a_{22}y_2 + a_{23}y_3 &= b_2, \\ a_{31}y_1 + a_{32}y_2 + a_{33}y_3 &= b_3, \end{aligned} \tag{20}$$

$$y_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad y_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad y_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}. \qquad (21)$$

The 3×3 determinant $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ is defined as

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \qquad (22)$$



The diagram shows a scheme for evaluating a 3×3 determinant: Multiply the elements along each diagonal line and add the results, with the signs as indicated. You should check to verify that this process leads to the sum of six products of a's given above.

This formulation can be extended to any number n of simultaneous equations with an equal number of unknowns, and a general $n \times n$ determinant can be defined. General methods for evaluating a determinant of any order are given in the section of Edwards and Penney cited above. We won't discuss this generalization further in these notes. But note that a determinant is always a single number, not an array of numbers, even though it is computed from an array of numbers.

Particular Cases

1) If the determinant in the denominator of the expressions for the y's is *not* zero, and if at least one of the b's is non-zero, then the above equations have exactly one solution, that is, one set of y's that satisfy all the equations. This solution is said to be *unique*.

2) If *all* the b's are zero, the equations are said to be *homogeneou*s because each term in every equation contains exactly one of the y's. In this case, the equations *always* have the solution $y_1 = y_2 = y_3 = \cdots = 0$. This is called a *trivial* solution. If the denominator determinant is not zero, this is the *only* solution of the set of equations.

3) If all the b's are zero and the denominator determinant *is* zero, the set of equations has infinitely many solutions. That is, *a set of simultaneous homogeneous linear equations has non-trivial solutions if, and only if, the denominator determinant in the above expressions is zero.*

Summary of Matrix Definitions

For any square matrix **A:**

Symbol	Term	Definition
$\mathbf{A}^{-1}$	inverse	$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.
$\mathbf{A}^{T}$ or $\tilde{\mathbf{A}}$	transpose	Interchange rows and columns. $(\mathbf{A}^{T})_{ij} = A_{ji}$.
$\mathbf{A}^{+}$ or $\mathbf{A}^{*}$	adjoint	Interchange rows and columns and take complex conjugate of each element (also called Hermitean conjugate or Hermitean transpose). $(\mathbf{A}^{+})_{ij} = A_{ji}^{*}$.

A *symmetric* matrix is one for which $\mathbf{A} = \mathbf{A}^{T}$.

A *skew symmetric* (or *antisymmetric*) matrix is one for which $\mathbf{A} = -\mathbf{A}^{T}$.

An *orthogonal* matrix is one for which $\mathbf{A}^{T}\mathbf{A} = \mathbf{AA}^{T} = \mathbf{I}$, or $\mathbf{A}^{T} = \mathbf{A}^{-1}$.

A *Hermitean* matrix is one for which $\mathbf{A} = \mathbf{A}^{+}$. (I.e., it is *self-adjoint*.)

A *unitary* matrix is one for which $\mathbf{A}^{+}\mathbf{A} = \mathbf{AA}^{+} = \mathbf{I}$, or $\mathbf{A}^{+} = \mathbf{A}^{-1}$.

A *unimodular* matrix is one for which $|\mathbf{A}| = 1$.

A *singular* matrix is one with determinant zero: $|\mathbf{A}| = 0$. Such a matrix has no inverse.

Summary of Matrix Properties

Addition of matrices obeys the associative and commutative rules.

Multiplication of a matrix by a scalar obeys the distributive and commutative rules.

Multiplication of matrices obeys the associative and distributive laws, but in general it *does not* obey the commutative law. Exception: The product of two matrices that have only diagonal elements (called "*diagonal* matrices") is commutative.

The identity matrix (or unit matrix), denoted by **I,** is a square matrix with ones on the diagonal and zeros everywhere else. The elements of the identity matrix are the Kroneker delta: $I_{ij} = \delta_{ij}$. For any matrix **A,** $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$.

The inverse of a product of any number of matrices equals the product of the inverses in the reverse order: $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

The transpose of a product of any number of matrices equals the product of the transposes in the reverse order: $(\mathbf{ABC})^{T} = \mathbf{C}^{T}\mathbf{B}^{T}\mathbf{A}^{T}$. Also, $(\mathbf{ABC})^{+} = \mathbf{C}^{+}\mathbf{B}^{+}\mathbf{A}^{+}$.

The determinant of a product of matrices equals the product of the determinants of the matrices: $|\mathbf{ABC}| = |\mathbf{A}| \cdot |\mathbf{B}| \cdot |\mathbf{C}|$.

Matrix Algebra with Maple

Maple has many useful capabilities for matrix algebra. Most of them are included in a package called "Linear Algebra." As with other packages you have used (such as plots and DEtools), this has to be loaded explicitly before you can do anything else. To do this, enter the command with(linalg); This package contains an enormous repertoire of over 100 functions, only a few of which we need to use in this course. If you end with(linalg) with a *semicolon,* Maple lists them all on the screen. You'll soon tire of this; when you do, end the command instead with a *colon* instead of a semicolon; this loads all the same functions but suppresses the list.

Entering a Matrix

There are several ways to enter a matrix into a Maple worksheet. They all start with the word matrix. One way is to specify the number of *rows* and number of *columns* (always in that order), followed by a *list* of elements of the matrix, enclosed in square brackets. For example, the command matrix(3,3, [1, 2, 3, 3, 4, 5, 5, 6, 9]); produces a matrix with three rows and three columns, with the elements filled in by rows from left to right and then from top to bottom: $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 5 & 6 & 9 \end{pmatrix}$. If you want to call this matrix A, then use

```
A := matrix(3,3, [1, 2, 3, 3, 4, 5, 5, 6, 9]);
```

An alternative is to designate *each row* of the matrix as a list, and the entire matrix as a *list of lists* (i.e., a list of the rows). In this case the numbers of rows and columns are determined by the lists and don't have to be entered explicitly. In the above example, we could use A := matrix([[1, 2, 3], [3, 4, 5], [5, 6, 9]]); Note the commas and the nested square brackets.

After a matrix has been entered, individual elements can be changed. For example, to change the element in the third row and second column of **A** from 6 to 8, enter A[3,2] := 8; Then to redisplay the matrix, enter evalm(A);

Note: You'll be using several Maple commands for which the output is a matrix but for which Maple doesn't display the matrix explicitly. To display it, enter the command evalm(…), with the appropriate expression inside the parentheses. "evalm" is of course an abbreviation for "evaluate matrix."

We often think of a matrix with only one row or one column as a vector. However, Maple treats vectors and matrices somewhat differently. For our purposes, it is usually best to think of vectors as one-column or one-row matrices (with two indices).

Matrix Addition

Two matrices can be added only when they have the same number of rows and the same number of columns. When they do, we just add corresponding elements. That is, if **A + B = C**, then $A_{ij} + B_{ij} = C_{ij}$. The matrix difference **A − B** is defined as **A − B = A + (−1)B**. The Maple commands are C := A + B; for the sum; C := A – B; for the difference.

Multiplication by a Scalar

When a matrix is multiplied by a scalar, each element of the matrix is multiplied by the scalar. Maple uses * for this operation. For the matrix A above, we could write B := 2*A; When we do this, Maple just repeats the command; it doesn't show us the result. To display the resulting matrix, we use the command evalm. Thus evalm(B); or evalm(2*A); displays the actual matrix result. The scalar multiplier can be a number or any scalar function or algebraic expression.

Multiplication of Two Matrices

The product of two matrices is denoted in Maple by the special symbol &*, with no space between the & and the *. (Note that &* is the shift of two adjacent keys on the top row of the keyboard.) If we have defined matrices A and B as described above, the product is denoted as C := A &* B; Maple does not display the matrix C but simply repeats the operation, as it did with multiplication by a scalar. To display C, use evalm(C); or evalm(A &* B); Remember that matrix multiplication is in general *not* commutative; A &* B; is different from B &* A; Thus you shouldn't expect Maple to understand an expression such as A * B (where A and B are matrices), and it doesn't. An exception is the case of multiplication of a matrix such as A by itself. In that case Maple *does* accept the expression A * A.

Inverse of a Matrix

The inverse of a matrix **A** is usually denoted by $\mathbf{A}^{-1}$ The Maple command is inverse(A); Or we can give it a name: B := inverse(A); Then we can check whether **B** really is the inverse of **A** by computing B &* A and A &* B and verifying that the result in each case is the identity matrix (or unit matrix). Once again, you need to use the command evalm(B &* A); to display the result. If a matrix has zero determinant, it is said to be *singular*; such a matrix has no inverse. In that case, when you try to find the inverse, Maple gives you an error message.

If you need to enter the identity matrix as part of an expression, the Maple notation is &*() This seems complicated, but note that it's just the shift of four successive keys in the top row, so it's just like playing a scale on the piano. Note that you can't use I for the identity matrix because Maple already uses I as the imaginary unit, $I = \sqrt{-1}$. If you want to use a more intuitive notation for the identity matrix, such as ID, then include the command alias(ID = &*()); at the beginning of your worksheet. For more information on aliases, read the help file ?alias.

Transpose of a Matrix

The transpose of a matrix $\mathbf{A}$ is a matrix with the rows and columns of the original matrix interchanged. It is denoted in various ways; two common ways are $\mathbf{A}^{T}$ and $\tilde{\mathbf{A}}$. The Maple command for the transpose of $\mathbf{A}$ is transpose(A);

When the elements of a matrix are complex, it is sometimes useful to consider a matrix with rows and columns interchanged from the original matrix, and with each element replaced by its complex conjugate. This is denoted as $\mathbf{A}^{*}$ or $\mathbf{A}^{+}$. It goes by various names, including the *adjoint* matrix, the *Hermitean conjugate*, and the *Hermitean transpose*. Maple computes this matrix with the command htranspose(A); (Note: Maple uses a different definition for *adjoint*; the definition given here is the one used in most physics literature.)

Determinant of a Matrix

The Maple command for the determinant $|\mathbf{A}|$ of a matrix $\mathbf{A}$, is det(A); Or you can replace A by any matrix expression, such as det(4*A + B &* C); Remember that the determinant of a matrix is a *scalar* quantity, a single number, function, or algebraic expression.