Matrices and Linear Algebra

A matrix is a collection of number, organized in a two-dimensional array with rows and columns. As a concept it is useful in its own right, but the analysis of matrices are particularly prominent in linear algebra, i.e., when dealing with systems of equations. As described in the notation document a matrix can be written in matrix notation or index notation:

$$\mathbf{A} \equiv A_{ij} \tag{1}$$

A special matrix is the identity matrix, which has ones on the diagonal and zeros elsewhere:

$$\mathbf{I} \equiv \delta_{ij} \tag{2}$$

Addition and subtraction of matrices is trivially simple, but requires that the matrices have the same size. Two matrices are added or subtracted by adding or subtracting one element at a time. The same simplicity applies to multiplication of a matrix with a scalar; each element of the matrix is multiplied by the scalar.

Division of matrices is a concept that does not exist. The concept of matrix inverse comes close and is described shortly. Multiplication of matrices is possible and common, but care must be exercised. As a generic example, consider the matrix product

$$\mathbf{A} = \mathbf{BCD} \tag{3}$$

First, the order of multiplication is important. The product in Eq. (3) is not the same as the product **CDB** or other variations of the order. Second, the dimension of the matrices must match in the following sense: The number of columns in a preceding matrix must match the number of rows in the next matrix. This is revealed by the indices in index notation of the same product:

$$A_{il} = B_{ij} C_{jk} D_{kl} \tag{4}$$

where i and l are free indices while j and k are dummy indices that are subjected to the Einstein summation convention. Some additional information is provided in the document on notation.

The transpose of a matrix is possible for square matrices and is written $\mathbf{A}^{\mathrm{T}}$. It simply mirrors the elements of the matrix around the diagonal. Computing the inverse of a matrix, which again is possible only for square matrices, is more costly. For a two-by-two matrix the following simple expression is available

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \cdot \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix} \tag{5}$$

where the determinant, which clearly cannot be zero, is

$$\det(\mathbf{A}) = A_{11}A_{22} - A_{12}A_{21} \tag{6}$$

For higher-dimension matrices the determinant, which is often written |**A**|, is obtained by adding and subtracting products of diagonals, as shown in Figure 1. For higher-dimension matrices the topic of matrix inversion is better dealt with in the context of the problem of solving systems of equations.

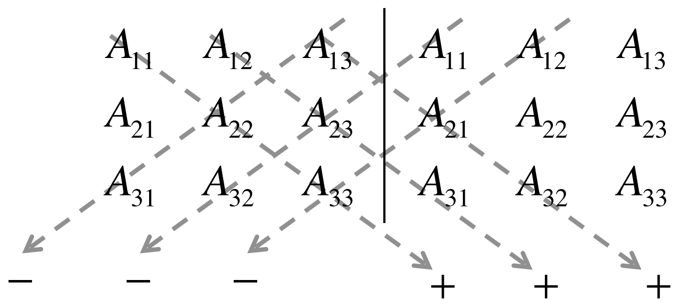

Figure 1: Calculation of matrix determinant.

Systems of Equations

Consider a system of linear equations of the form

$$\mathbf{Ax} = \mathbf{b} \tag{7}$$

where **A** is the coefficient matrix of known constants, **x** is the vector of sought unknowns, and **b** is a vector of known constants. Symbolically, the solution to the system of equations is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \tag{8}$$

In words, the inverse of the coefficient matrix feature prominently, at least as a concept, in the solution of a linear system of equations. As suggested in the discussion of Eq. (6) the inverse of a matrix only exists when the determinant of the coefficient matrix is zero. In fact, the system of linear equations in Eq. (7) has a unique non-trivial solution only if det(**A**)≠0. Conversely, if this determinant is zero then the system has no solution or infinitely many solutions.

Homogeneous Systems

A special version of Eq. (7) is when **b**=**0**. Then the system of equations is said to be homogeneous. A homogeneous system always has the trivial solution **x**=**0**. In fact, if det(**A**)≠0 then the system only has the trivial solution. Conversely, if det(**A**)=0 non-trivial solutions exist.

Eigenvalue Problems

A special version of Eq. (7) of homogeneous systems is

$$\mathbf{Ax} = \lambda\mathbf{x} \quad \Leftrightarrow \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \tag{9}$$

which is called an eigenvalue problem that is written more generally as

$$(\mathbf{A} - \lambda\mathbf{B})\mathbf{x} = \mathbf{0} \tag{10}$$

Like other homogeneous systems, non-trivial solutions exist only when the determinant of the coefficient matrix is zero, i.e., when

$$\det(\mathbf{A} - \lambda\mathbf{B}) = 0 \tag{11}$$

The roots λ_i of Eq. (11) are called eigenvalues. Each eigenvalue has one eigenvector, or eigen-mode, associated with it. The eigenvectors are not unique; they are determined by setting one of the elements of **x** in Eq. (10) equal to unity and solving for the others. This is repeated for each eigenvalue to obtain all eigenvectors. Any scaled version of an eigenvector is also an eigenvector because the choice of unity of one element is arbitrary.

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X} \tag{12}$$

where **X** is a matrix with the eigenvectors of **A** as diagonals.

Diagonalization

Upon computing the eigenvalues of a square matrix **A** it is possible to diagonalize it into a matrix **D** that has the eigenvalues on the diagonal:

Cholesky Decomposition

Within the numerical methods for linear algebra there are several methods for solving linear systems of equations. One approach is LU-factorization, in which the coefficient matrix is decomposed into a lower and an upper triangular matrix:

$$\mathbf{A} = \mathbf{L}\mathbf{U} \tag{13}$$

Substitution into Eq. (7) yields

$$\mathbf{L}\underbrace{\mathbf{U}\mathbf{x}}_{\mathbf{y}} = \mathbf{b} \tag{14}$$

To solve the problem, **y** is first determined and x is next determined from **Ux**=**y**.

Cholesky decomposition is one of the methods to decompose the coefficient matrix. It is based on selecting **U**=**L**$^{\mathrm{T}}$ so that

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T \tag{15}$$

The Cholesky decomposition has broader use than solving systems of equations. It appears in reliability analysis when correlated second-moment random variables are transformed. From Erwin Kreyszig's textbook, the algorithm to determine **L** is

$$\begin{aligned}
L_{11} &= \sqrt{A_{11}} && \\
L_{jj} &= \sqrt{A_{jj} - \sum_{s=1}^{j-1} L_{js}^2} && j = 2,\cdots,n \\
L_{j1} &= \frac{A_{j1}}{L_{11}} && j = 2,\cdots,n \\
L_{jk} &= \frac{1}{L_{kk}}\left(A_{jk} - \sum_{s=1}^{k-1} L_{js}L_{ks}\right) && j = k+1,\cdots,n \qquad k \ge 2
\end{aligned} \tag{16}$$