2.14 Analysis and Design of Feedback Control Systems

Introduction to Matrices

Derek Rowell
October 2002

Modern system dynamics is based upon a matrix representation of the dynamic equations governing the system behavior. A basic understanding of elementary matrix algebra is essential for the analysis of state-space formulated systems. A full discussion of linear algebra is beyond the scope of this note and only a brief summary is presented here. The reader is referred to a text on linear algebra, such as Strang (1980), for a detailed explanation and expansion of the material summarized here.

1 Definition

A matrix is a two dimensional array of numbers or expressions arranged in a set of rows and columns. An $m \times n$ matrix $\mathbf{A}$ has m rows and n columns and is written

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

where the element a_{ij}, located in the ith row and the jth column, is a *scalar* quantity; a numerical constant, or a single valued expression. If $m = n$, that is there are the same number of rows as columns, the matrix is *square*, otherwise it is a *rectangular* matrix.

A matrix having either a single row ($m = 1$) or a single column ($n = 1$) is defined to be a *vector* because it is often used to define the coordinates of a point in a multi-dimensional space. (In this note the convention has been adopted of representing a vector by a lower case "bold-face" letter such as $\mathbf{x}$, and a general matrix by a "bold-face" upper case letter such as $\mathbf{A}$.) A vector having a single row, for example

$$\mathbf{x} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \end{bmatrix} \tag{2}$$

is defined to be a *row vector*, while a vector having a single column is defined to be a column vector

$$\mathbf{y} = \begin{bmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{m1} \end{bmatrix}. \tag{3}$$

Two special matrices are the square *identity* matrix, **I**, which is defined to have all of its elements equal to zero except those on the *main diagonal* (where $i = j$) which have a value of one:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \tag{4}$$

and the *null* matrix **0**, which has the value of zero for all of its elements:

$$\mathbf{0} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{5}$$

2 Elementary Matrix Arithmetic

2.0.1 Matrix Addition

The operation of addition of two matrices is only defined when both matrices have the same dimensions. If **A** and **B** are both $(m \times n)$, then the sum

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \tag{6}$$

is also $(m \times n)$ and is defined to have each element the sum of the corresponding elements of **A** and **B**, thus

$$c_{ij} = a_{ij} + b_{ij}. \tag{7}$$

Matrix addition is both associative, that is

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}, \tag{8}$$

and commutative

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}. \tag{9}$$

The subtraction of two matrices is similarly defined; if **A** and **B** have the same dimensions, then the difference

$$\mathbf{C} = \mathbf{A} - \mathbf{B} \tag{10}$$

implies that the elements of **C** are

$$c_{ij} = a_{ij} - b_{ij}. \tag{11}$$

2.0.2 Multiplication of a Matrix by a Scalar Quantity

If **A** is a matrix and k is a scalar quantity, the product $\mathbf{B} = \mathrm{k}\mathbf{A}$ is defined to be the matrix of the same dimensions as **A** whose elements are simply all scaled by the constant k,

$$b_{ij} = k \times a_{ij}. \tag{12}$$

2.0.3 Matrix Multiplication

Two matrices may be multiplied together only if they meet conditions on their dimensions that allow them to *conform*. Let $\mathbf{A}$ have dimensions $m \times n$, and $\mathbf{B}$ be $n \times p$, that is $\mathbf{A}$ has the same number as columns as the number of rows in $\mathbf{B}$, then the product

$$\mathbf{C} = \mathbf{AB} \tag{13}$$

is defined to be an $m \times p$ matrix with elements

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{14}$$

The element in position ij is the sum of the products of elements in the ith row of the first matrix ($\mathbf{A}$) and the corresponding elements in the jth column of the second matrix ($\mathbf{B}$). Notice that the product $\mathbf{AB}$ is not defined unless the above condition is satisfied, that is the number of columns of the first matrix must equal the number of rows in the second.

Matrix multiplication is associative, that is

$$\mathbf{A}\left(\mathbf{BC}\right) = \left(\mathbf{AB}\right)\mathbf{C}, \tag{15}$$

but is not commutative in general

$$\mathbf{AB} \neq \mathbf{BA}, \tag{16}$$

in fact unless the two matrices are square, reversing the order in the product will cause the matrices to be nonconformal. The order of the terms in the product is therefore very important. In the product $\mathbf{C} = \mathbf{AB}$, $\mathbf{A}$ is said to *pre-multiply* $\mathbf{B}$, while $\mathbf{B}$ is said to *post-multiply* $\mathbf{A}$.

It is interesting to note in passing that unlike the scalar case, the fact that $\mathbf{AB} = \mathbf{0}$ does not imply that either $\mathbf{A} = \mathbf{0}$ or that $\mathbf{B} = \mathbf{0}$.

3 Representing Systems of Equations in Matrix Form

3.0.4 Linear Algebraic Equations

The rules given above for matrix arithmetic allow a set of linear algebraic equations to be written compactly in matrix form. Consider a set of n independent linear equations in the variables x_i for $i = 1, \ldots, n$

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & \vdots & & \ddots & & \vdots & & \vdots \\
a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n
\end{array} \tag{17}$$

and write the coefficients a_{ij} in a square matrix $\mathbf{A}$ of dimension n

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

the unknowns x_{ij} in a column vector $\mathbf{x}$ of length N

$$\mathbf{x} = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix},$$

and the constants on the right-hand side in a column vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

then equations may be written as the product

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{18}$$

which may be written compactly as

$$\mathbf{Ax} = \mathbf{b}. \tag{19}$$

3.0.5 State Equations

The modeling procedures described in this note generate a set of first-order linear differential equations

$$\begin{array}{ccccccccccccccc} \frac{d}{dt}x_1 & = & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & + & b_{11}u_1 & + & \cdots & b_{1m}u_m \\ \frac{d}{dt}x_2 & = & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & + & b_{21}u_1 & + & \cdots & b_{2m}u_m \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots & & \ddots & \vdots \\ \frac{d}{dt}x_n & = & a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & + & b_{n1}u_1 & + & \cdots & b_{nm}u_m \end{array} \tag{20}$$

If the derivative of a matrix is defined to be a matrix consisting of the derivatives of the elements of the original matrix, the above equations may be written

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ b_{21} & \cdots & b_{2m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \tag{21}$$

or in the standard matrix form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{22}$$

where

$$\dot{\mathbf{x}} = \frac{d}{dt}\mathbf{x} \tag{23}$$

4 Functions of a Matrix

4.1 The Transpose of a Matrix

The transpose of an $m \times n$ matrix $\mathbf{A}$, written $\mathbf{A}^T$, is the $n \times m$ matrix formed by interchanging the rows and columns of A. For example, if

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \tag{24}$$

then in terms of the elements of the above matrix, the transpose is

$$\mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}. \tag{25}$$

Notice that the transpose of a row vector produces a column vector, and similarly the transpose of a column vector produces a row vector. The transpose of the *product* of two matrices is the *reversed* product of the transpose of the two individual matrices,

$$(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T\mathbf{A}^T. \tag{26}$$

The rules of matrix multiplication show that the product of a vector and its transpose is the sum of the squares of all of the elements

$$\mathbf{x}(\mathbf{x}^T) = \sum_{i=1}^{n}(x_i)^2. \tag{27}$$

4.2 The Determinant

The determinant of an $n \times n$ square matrix, written $\det \mathbf{A}$ is an important scalar quantity that is a function of the elements of the matrix. When written using the elements of the matrix, the determinant is enclosed between vertical bars, for example

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \tag{28}$$

The determinant of a matrix of size $n \times n$ is defined recursively in terms of lower order determinants ($(n-1) \times (n-1)$) as follows. The *minor* of an element a_{ij}, written M_{ij}, is another determinant of order $(n-1)$ that is formed by deleting the ith row and the jth column from the original determinant. In the above example,

$$M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}. \tag{29}$$

The *cofactor* α_{ij} of an element a_{ij} in a determinant is simply its minor M_{ij} multiplied by either $+1$ or -1, depending upon its position, that is

$$\begin{aligned} \text{cof } a_{ij} &= \alpha_{ij} \\ &= (-1)^{i+j} M_{ij}. \end{aligned} \tag{30}$$

In the case of the (3×3) example above

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

At the limit, the determinant of a 1×1 matrix (a scalar) is the value of the scalar itself. The determinant of a high order matrix is found by expanding in terms of the elements of any selected row or column and the cofactors of those elements. If the expansion is done by selecting the *ith* row, the determinant is defined as a sum of order $(n-1)$ determinants as

$$\det \mathbf{A} = \sum_{j=1}^{n} a_{ij}\alpha_{ij}, \tag{32}$$

that is the sum of products of the elements of the row and their cofactors. Similarly, the expansion in terms of the elements of the *jth* column of the determinant is

$$\det \mathbf{A} = \sum_{i=1}^{n} a_{ij}\alpha_{ij}. \tag{33}$$

If a (2×2) determinant is expanded by the top row, the result is

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\alpha_{11} + a_{12}\alpha_{12} = a_{11}a_{22} - a_{12}a_{21} \tag{34}$$

If $\mathbf{A}$ is a (3×3) matrix, and the expansion is done by the first column, then

$$\begin{aligned} \det \mathbf{A} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \end{aligned} \tag{35}$$

4.3 The Matrix Inverse

Matrix division is not defined. For a square matrix, multiplication by an inverse may be thought of as an analogous operation to the process of multiplication of a scalar quantity by its reciprocal. For a square $n \times n$ matrix $\mathbf{A}$, define its inverse, written $\mathbf{A}^{-1}$, as that matrix (if it exists) that pre- or post-multiplies $\mathbf{A}$ to give the identity matrix $\mathbf{I}$,

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{A} &= \mathbf{I} \\ \mathbf{A}\mathbf{A}^{-1} &= \mathbf{I}. \end{aligned} \tag{36}$$

The importance of the inverse matrix can be seen from the solution of a set of algebraic linear equations such as

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{37}$$

If the inverse $\mathbf{A}^{-1}$ exists then pre-multiplying both sides gives

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \tag{38}$$

$$\mathbf{I}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \tag{39}$$

and since pre-multiplying a column vector of length n by the nth order identity matrix $\mathbf{I}$ does not affect its value, this process gives an explicit solution to the set of equations

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \tag{40}$$

The inverse of a square matrix does not always exist. If the inverse exists, the matrix is defined to be *non-singular*, if it does not exist the matrix is *singular*.

The *adjoint* matrix, adj $\mathbf{A}$, of a square matrix $\mathbf{A}$ is defined as the transpose of the matrix of cofactors of the elements of $\mathbf{A}$, that is

$$\text{adj}\,\mathbf{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix}^T = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} \end{bmatrix}. \tag{41}$$

The inverse of $\mathbf{A}$ is found from the determinant and the adjoint of $\mathbf{A}$,

$$\mathbf{A}^{-1} = \frac{\text{adj}\,\mathbf{A}}{\det \mathbf{A}}. \tag{42}$$

Notice that the condition for the inverse to exist, that is for $\mathbf{A}$ to be non-singular, is that $\det \mathbf{A} \neq 0$.

For a (2×2) matrix

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \tag{43}$$

For higher order matrices the elements in the adjoint matrix must be found by expanding out the cofactors of each element. For numerical matrices of order four and greater it is usually expedient to use one of the many computer matrix manipulation packages to compute the inverse.

5 Eigenvalues and Eigenvectors

In the study of dynamic systems we are frequently confronted with a set of linear algebraic equations of the form

$$\lambda \mathbf{x} = \mathbf{A}\mathbf{x} \tag{44}$$

where $\mathbf{x}$ is a column vector of length n, $\mathbf{A}$ is an $n \times n$ square matrix and λ is a scalar quantity. The problem to be solved is to find the values of λ satisfying this equation, and the corresponding vectors $\mathbf{x}$. The values of λ are known as the *eigenvalues* of the matrix $\mathbf{A}$, and the corresponding vectors are the *eigenvectors*.

The above equation may be written as a set of homogeneous equations

$$\lambda \mathbf{x} - \mathbf{A}\mathbf{x} = (\lambda \mathbf{I} - \mathbf{A})\,\mathbf{x} = \mathbf{0}, \tag{45}$$

where $\mathbf{0}$ is a null column vector, and

$$(\lambda \mathbf{I} - \mathbf{A}) = \begin{bmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{m1} & -a_{m2} & \cdots & \lambda - a_{mn} \end{bmatrix}. \tag{46}$$

A theorem of linear algebra states that for a set of homogeneous linear equations a nontrivial set of solutions exists only if the coefficient matrix is singular, that is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0. \tag{47}$$

The determinant, $\det(\lambda \mathbf{I} - \mathbf{A})$, is known as the *characteristic determinant* of the matrix $\mathbf{A}$. Expansion of the determinant results in an nth order polynomial in λ, known as the *characteristic polynomial* of $\mathbf{A}$. The n roots of the *characteristic equation*, formed by equating the characteristic polynomial to zero, will define those values of λ that make the matrix $(\lambda \mathbf{I} - \mathbf{A})$ singular. These values are known as the *eigenvalues* of the matrix $\mathbf{A}$.

For example, to find the eigenvalues of the matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix},$$

the characteristic equation is given by

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \begin{vmatrix} \lambda - 5 & -4 \\ -1 & \lambda - 2 \end{vmatrix} = \lambda^2 - 7\lambda + 6 = 0$$

and has roots (and hence eigenvalues) $\lambda_1 = 6$ and $\lambda_2 = 1$. The eigenvectors corresponding to each of these values may be found by substituting back into the original equations, for example if $\lambda = 6$ the equations become

$$\begin{aligned} -x_1 + 4x_2 &= 0 \\ x_1 - 4x_2 &= 0 \end{aligned}$$

which has the non-unique solution $x_1 = 4x_2$. Any vector maintaining this ratio between its two elements is therefore an eigenvector corresponding to the eigenvalue $\lambda = 6$. In general, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ then so is $k\mathbf{x}$ for any scalar value k.

6 Cramer's Method

Cramer's method is a convenient method for manually solving low-order non-homogeneous sets of linear equations. If the equations are written in matrix form

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{48}$$

then the ith element of the vector $\mathbf{x}$ may be found directly from a ratio of determinants

$$x_i = \frac{\det \mathbf{A}_{(i)}}{\det \mathbf{A}} \tag{49}$$

where $\mathbf{A}_{(i)}$ is the matrix formed by replacing the ith column of $\mathbf{A}$ with the column vector $\mathbf{b}$. For example, solve

$$\begin{array}{rcrcrcr} 2x_1 & - & x_2 & + & 2x_3 & = & 2 \\ x_1 & + & 10x_2 & - & 3x_3 & = & 5 \\ -x_1 & + & x_2 & + & x_3 & = & -3 \end{array} .$$

Then

$$\det \mathbf{A} = \begin{vmatrix} 2 & -1 & 2 \\ 1 & 10 & -3 \\ -1 & 1 & 1 \end{vmatrix} = 46$$

and

$$\begin{aligned} x_1 &= \frac{1}{46} \begin{vmatrix} 2 & -1 & 2 \\ 5 & 10 & -3 \\ -3 & 1 & 1 \end{vmatrix} \\ &= 2 \end{aligned}$$

$$\begin{aligned} x_2 &= \frac{1}{46} \begin{vmatrix} 2 & 2 & 2 \\ 1 & 5 & -3 \\ -1 & -3 & 1 \end{vmatrix} \\ &= 0 \end{aligned}$$

$$\begin{aligned} x_3 &= \frac{1}{46} \begin{vmatrix} 2 & -1 & 2 \\ 1 & 10 & 5 \\ -1 & 1 & -3 \end{vmatrix} \\ &= -1 \end{aligned}$$