Web Appendix R - Vectors and Matrices

R.1 Definitions and Operations

A very common situation in linear system analysis is finding the solution to multiple simultaneous equations which describe complicated systems. Suppose we have N linear equations in the N unknowns $q_1, q_2, \ldots q_N$ of the form

$$\begin{matrix} a_{11}q_1 & a_{12}q_{12} & \cdots & a_{1N}q_N & = \mathrm{x}_1 \\ a_{21}q_1 & a_{21}q_2 & \cdots & a_{2N}q_2 & = \mathrm{x}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N1}q_1 & a_{N2}q_2 & \cdots & a_{NN}q_N & = \mathrm{x}_N \end{matrix} . \tag{R.1}$$

When N becomes large this formulation of the equations becomes clumsy and finding a solution becomes tedious. The use of vectors and matrices is a way of compacting the notation of system of equations and leads to very good techniques of systematically solving them. Formulating systems of equations in matrix form allows us to appreciate the "forest", the overall system of equations, without being distracted by all the detailed relationships between the "trees", the individual equations.

A vector is a one-dimensional ordered array of numbers or variables or functions, given a single name and manipulated as a group. For example, the variables $q_1, q_2, \ldots q_N$ in (R.1) can be written as a single vector variable as

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix} \quad \text{or} \quad \mathbf{q} = \begin{bmatrix} q_1 & q_2 & \cdots & q_N \end{bmatrix}$$

The first form is called a *column vector* and the second form is called a *row vector*. The variables (or numbers or functions) $q_1, q_2, \ldots q_N$ are called the *elements* of the vector. Boldface type **q** distinguishes a vector (and later a matrix) from a scalar variable written in plain face q. A vector can be conceived as a location in a *space*. The number of dimensions of the space is the same as the number of elements in the vector. For example, the vector

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix} \tag{R.2}$$

identifies a position in a two-dimensional space and can be illustrated graphically as in Figure R-1.

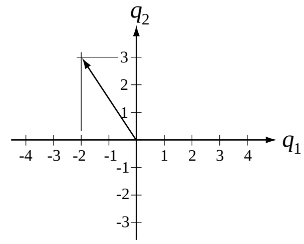


Figure R-1 A two-dimensional vector

A three-dimensional vector can also be represented graphically but higher-dimensional vectors have no convenient graphical representation, even though they are just as real and useful as two- and three-dimensional vectors. Therefore an N-dimensional vector is often spoken of as a location in an *N-space* (short for an N-dimensional space).

A *matrix* is a two-dimensional array of numbers or variables or functions, given a single name and manipulated as a group. For example, the array of coefficients in (R.1) forms the matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}.$$

This is called an $N \times N$ matrix because it has N rows and N columns. By that same terminology, the vector

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix}$$

is an $N \times 1$ matrix and the vector $\mathbf{q} = \begin{bmatrix} q_1 & q_2 & \cdots & q_N \end{bmatrix}$ is a $1 \times N$ matrix. In any specification of the size of a matrix, the notation $m \times n$ means a matrix with m rows and n columns. The notation $m \times n$ is conventionally spoken as "m by n". Sometimes an alternate notation $\left(a_{rc}\right)$ is used to indicate a matrix like **A**. In any reference to a single element of a matrix in the form a_{rc} the first subscript r indicates the row and the second subscript c indicates the column of that element.

If a matrix has the same number of rows and columns it is called a *square* matrix and the number of rows (or columns) is called its *order*. A vector or matrix with only one element is a scalar. The diagonal of a square matrix **A** of order N containing the elements

$a_{11}, a_{22}, \cdots a_{NN}$ is called the *principal diagonal.* If all the elements of the matrix which are not on the principal diagonal are zero the matrix is called a *diagonal matrix*, for example,

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{NN} \end{bmatrix}.$$

If, in addition, the elements on the principal diagonal are all the number 1 the matrix is called a *unit* matrix or an *identity* matrix and is usually indicated by the symbol **I**

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

A matrix whose elements are all zero is called a *zero* matrix or a *null* matrix.

Two matrices are equal if, and only if every element of one is equal to the element in the same row and column in the other. Let two matrices, **A** and **B**, be defined by

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}.$$

Then if $\mathbf{A} = \mathbf{B}$ that implies that $a_{rc} = b_{rc}$ for any *r* and *c*. Obviously the two matrices must have the same number of rows and columns to be equal. The sum of these two matrices is defined by

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix} = \left(a_{rc} + b_{rc} \right)$$

In words, each element in the sum of two matrices is the sum of the corresponding elements in the two matrices being added. Subtraction follows directly from addition,

$$\mathbf{A}-\mathbf{B}=\begin{bmatrix} a_{11}-b_{11} & a_{12}-b_{12} & \cdots & a_{1n}-b_{1n} \\ a_{21}-b_{21} & a_{22}-b_{22} & \cdots & a_{2n}-b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}-b_{m1} & a_{m2}-b_{m2} & \cdots & a_{mn}-b_{mn} \end{bmatrix}=\left(a_{rc}-b_{rc}\right).$$

A matrix can be multiplied by a scalar. Let

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

represent any $m \times n$ matrix and let c represent any scalar. Then

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

In the special case of $c=-1$, $c\mathbf{A}=\mathbf{A}c=\left(-1\right)\mathbf{A}=-\mathbf{A}$ and $\mathbf{B}+\left(-\mathbf{A}\right)=\mathbf{B}-\mathbf{A}$. Using the laws of addition and subtraction and multiplication by a scalar it is easy to show that $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$, $\mathbf{A}+\left(\mathbf{B}+\mathbf{C}\right)=\left(\mathbf{A}+\mathbf{B}\right)+\mathbf{C}$, $c\left(\mathbf{A}+\mathbf{B}\right)=c\mathbf{A}+c\mathbf{B}$, $\left(a+b\right)\mathbf{A}=a\mathbf{A}+b\mathbf{A}$ and $a\left(b\mathbf{A}\right)=\left(ab\right)\mathbf{A}$.

An operation that comes up often in matrix manipulation is the *transpose*. The transpose of a matrix **A** is indicated by the notation $\mathbf{A}^T$. If

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

then

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

In words, the transpose is the matrix formed by making rows into columns and columns into rows. If **A** is an $m \times n$ matrix then $\mathbf{A}^T$ is an $n \times m$ matrix.

The real power of matrix methods in linear system analysis comes in the use of the product of two matrices. The product of two matrices is defined in a way that allows us to represent a system of multiple equations as one matrix equation. If a matrix **A** is $m \times n$ and another matrix **B** is $n \times p$, then the matrix product $\mathbf{C} = \mathbf{AB}$ is an $m \times p$ matrix whose elements are given by

$$c_{rc} = a_{r1}b_{1c} + a_{r2}b_{2c} + \cdots + a_{rn}b_{nc} = \sum_{k=1}^{n} a_{rk}b_{kc} .$$

The product **AB** can be described as **A** *postmultiplied* by **B** or as **B** *premultiplied* by **A**. To be able to multiply two matrices the number of columns in the premultiplying matrix (**A** in **AB**) must equal the number of rows in the postmultiplying matrix (**B** in **AB**). The process of computing the *rc* element of **C** can be conceived geometrically as the sum of the products of the elements in the *r*th row of **A** with the corresponding elements in the *c*th column of **B**. Consider the simple case of the premultiplication of a column vector by a row vector. Let

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \cdots & a_N \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} .$$

Then

$$\mathbf{AB} = \begin{bmatrix} a_1 & a_2 & \cdots & a_N \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_N b_N .$$

A is $1 \times N$ and **B** is $N \times 1$ so the product is 1×1 a scalar. This special case in which a row vector premultiplies a column vector of the same length is called a *scalar product* because the product is a scalar. For contrast consider the product **BA**

$$\mathbf{BA} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \cdots & a_N \end{bmatrix} = \begin{bmatrix} b_1 a_1 & b_1 a_2 & \cdots & b_1 a_N \\ b_2 a_1 & b_2 a_2 & \cdots & b_2 a_N \\ \vdots & \vdots & \ddots & \vdots \\ b_N a_1 & b_N a_2 & \cdots & b_N a_N \end{bmatrix} .$$

This result is very different, an $N \times N$ matrix. Obviously matrix multiplication is not (generally) commutative. If **A** and **B** are more general than simply vectors, the process of matrix multiplication can be broken down into multiple scalar products. In the matrix product $\mathbf{C} = \mathbf{AB}$ the *rc* element of **C** is simply the scalar product of the *r*th row of **A** with the *c*th column of **B**. Even though matrix multiplication is not commutative, it is associative and distributive $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$, $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ and $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$. The product of any matrix **A** and the identity matrix **I** (in either order) is the matrix **A**, $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$. One other useful multiplication rule is that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$. In words, the transpose of a product of two matrices in a given order is the product of the transposes of those matrices in the reverse of that order.

Using the rules of matrix multiplication, a system of *N* linear equations in *N* unknowns can be written in matrix form as the single matrix equation $\mathbf{Aq} = \mathbf{x}$ where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}, \ \mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix} \text{ and } \mathbf{x} = \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \\ \vdots \\ \mathrm{x}_N \end{bmatrix}.$$

MATLAB has a rich complement of commands and functions to handle almost any vector or matrix manipulation. The standard arithmetic operators `+`, `-`, `*` and `^`, are all defined for both scalar and vector operands. That is, the multiplication operator `*` performs a true matrix multiplication and the power operator `^` performs a true matrix power operation. These can be modified to perform an *array* operation instead of a matrix operation by the addition of a dot "`.`" before the operator. for example "`.*`" and "`.^`". In these forms the operators simply operate on corresponding pairs of elements in the two matrix operands. (See the MATLAB Tutorial, Appendix B, for more detail). In addition, the operator "`'`" transposes a matrix. (If the matrix elements are complex this operator also conjugates the elements as it transposes the matrix.) The operator "`.'`" transposes without conjugating. Some other useful commands and functions are

```
zeros       - Zeros array.
ones        - Ones array.
rand        - Uniformly distributed random numbers.
randn       - Normally distributed random numbers.
linspace    - Linearly spaced vector.
logspace    - Logarithmically spaced vector.
size        - Size of matrix.
length      - Length of vector.
find        - Find indices of nonzero elements.
```

For example, let two matrices `A` and `B` be defined by

```
»A = round(4*randn(3,3))

A =

    -2     1     5
    -7    -5     0
     1     5     1

»B = round(4*randn(3,3))

B =

     1    -2     0
    -1     9     4
     3    -1     0
```

Then

```
»A+B

ans =

    -1    -1     5
    -8     4     4
     4     4     1

»A-B

ans =

    -3     3     5
    -6   -14    -4
    -2     6     1

»A*B

ans =

    12     8     4
    -2   -31   -20
    -1    42    20

»A^2

ans =

     2    18    -5
    49    18   -35
   -36   -19     6

»A.*B
```

```
ans =

    -2    -2     0
     7   -45     0
     3    -5     0

»A.^2

ans =

     4     1    25
    49    25     0
     1    25     1

»A'

ans =

    -2    -7     1
     1    -5     5
     5     0     1
```

R.2 Orthogonality and Projections

Two vectors are said to be *orthogonal* if their scalar product is zero. In two or three dimensions, orthogonality has the same meaning as geometrically perpendicular. For example, the vectors

$$\mathbf{a} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

are orthogonal and when plotted in (x_1, x_2) space they are obviously perpendicular.

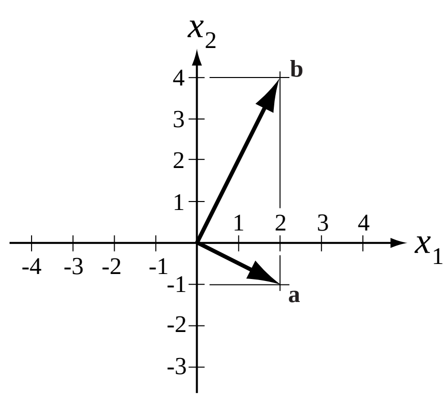


Figure R-2 Two orthogonal vectors in two dimensions

A *unit* vector is a vector of length one. A unit vector in the direction of any vector **a** is $\mathbf{a}/|\mathbf{a}|$, the vector divided by its magnitude. The *projection* **p** of any vector in the direction of a unit vector is their scalar product multiplied by the unit vector . For example, in (x_1, x_2) space a unit vector in the x_1 direction is $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the projection **p** of the vector $\mathbf{a} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ in the direction of that unit vector is

$$\mathbf{p} = \left(\mathbf{a}^T \mathbf{u}_1\right)\mathbf{u}_1 = \left\{\begin{bmatrix} 2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

the x_1 component of **a**. A projection **p** of any vector **a** in the direction of any other vector **b** would be

$$\mathbf{p} = \left(\mathbf{a}^T \mathbf{u}_b\right)\mathbf{u}_b = \left[\mathbf{a}^T\left(\frac{\mathbf{b}}{|\mathbf{b}|}\right)\right]\frac{\mathbf{b}}{|\mathbf{b}|} = \left(\frac{\mathbf{a}^T\mathbf{b}}{|\mathbf{b}|^2}\right)\mathbf{b} \ .$$

Using the fact that $|\mathbf{b}|^2 = \mathbf{b}^T\mathbf{b}$, we can then say that the vector projection **p** of **a** in the direction of **b** is

$$\mathbf{p} = \left(\mathbf{a}^T \mathbf{u}_b\right)\mathbf{u}_b = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{b}^T\mathbf{b}}\mathbf{b} \ .$$

For example, let $\mathbf{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and let $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$. Then the projection of **a** in the **b** direction is

$$\mathbf{p} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{b}^T\mathbf{b}}\mathbf{b} = \frac{\begin{bmatrix} 1 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}}{\begin{bmatrix} 4 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}}\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \frac{7}{17}\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 28/17 \\ 7/17 \end{bmatrix} = \begin{bmatrix} 1.647 \\ 0.412 \end{bmatrix} \quad \text{(Figure R-3)} \ .$$

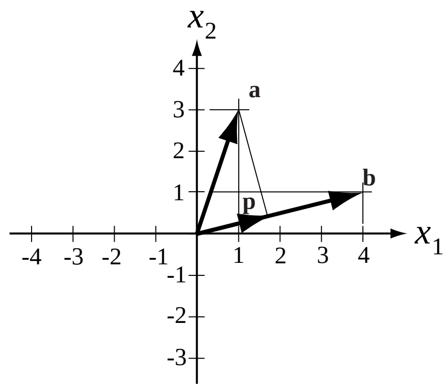


Figure R-3 Two vectors and the projection of one on the other

R.3 Determinants, Cramer's Rule and the Matrix Inverse

Matrices and related concepts can be used to systematically solve systems of linear equations. Consider first a system of two equations and two unknowns

$$\begin{aligned} a_{11}q_1 + a_{12}q_2 &= \mathrm{x}_1 \\ a_{21}q_1 + a_{22}q_2 &= \mathrm{x}_2 \end{aligned} \quad \text{or} \quad \mathbf{Aq} = \mathbf{x}\,.$$

Using non-matrix methods we can solve for the unknowns by any convenient method and the answers are found to be

$$q_1 = \frac{a_{22}\,\mathrm{x}_1 - a_{12}\,\mathrm{x}_2}{a_{11}a_{22} - a_{21}a_{12}} \quad \text{and} \quad q_2 = \frac{a_{11}\,\mathrm{x}_2 - a_{21}\,\mathrm{x}_1}{a_{11}a_{22} - a_{21}a_{12}}\,. \tag{R.3}$$

Notice that the two denominators in (R.3) are the same. This denominator which is common to both solutions is called the *determinant* of this system of equations and is conventionally indicated by the notation

$$\Delta_{\mathbf{A}} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \left|\mathbf{A}\right|. \tag{R.4}$$

The notations $\Delta_{\mathbf{A}}$ and $\left|\mathbf{A}\right|$ mean the same thing and can be used interchangeably. The determinant of a 2×2 system of equations written in the standard form of (R.4) is a scalar found by forming the product of the 11 and 22 elements in the matrix of coefficients **A** and subtracting from that the product of the 21 and 12 elements. The numerators in (R.3) can be interpreted as determinants also,

$$\Delta_1 = \mathrm{x}_1\,a_{22} - \mathrm{x}_2\,a_{12} = \begin{vmatrix} \mathrm{x}_1 & a_{12} \\ \mathrm{x}_2 & a_{22} \end{vmatrix} \quad \text{and} \quad \Delta_2 = a_{11}\,\mathrm{x}_2 - a_{21}\,\mathrm{x}_1 = \begin{vmatrix} a_{11} & \mathrm{x}_1 \\ a_{21} & \mathrm{x}_2 \end{vmatrix}.$$

Then, using this notation, the solutions of the two linear equations can be written in a very compact form

$$q_1 = \frac{\Delta_1}{\Delta_{\mathbf{A}}} \text{ and } q_2 = \frac{\Delta_2}{\Delta_{\mathbf{A}}} \ , \ \Delta_{\mathbf{A}} \neq 0 \ . \quad \text{(R.6)}$$

As indicated in (R.6), these solutions only exist if the determinant is not zero. If the determinant is zero, that is an indication that the equations are not independent.

In preparation for extending this technique to larger systems of equations we will formalize and generalize the process of finding a determinant by defining the terms *minor* and *cofactor*. In any square matrix, the minor of any element of the matrix is defined as the determinant of the matrix found by eliminating all the elements in the same row and the same column as the element in question. For example, the minor of element a_{11} in **A** is the determinant of the single-element matrix a_{22} a scalar. The determinant of a scalar is just the scalar itself. The cofactor of any particular element a_{rc} of a matrix is the product of the minor of that element and the factor $(-1)^{r+c}$. So the cofactor of the element a_{12} is $-a_{21}$. The determinant Δ can be found by choosing any row or column of the matrix and, for each element in that row or column, forming the product of the element and the determinant of its cofactor and then adding all such products for that row or column. For example, expanding along the second column of **A** we would calculate the determinant to be

$$\Delta_{\mathbf{A}} = a_{12}(-a_{21}) + a_{22}a_{11}$$

which is the same determinant we got before. Expanding along the bottom row,

$$\Delta_{\mathbf{A}} = a_{21}(-a_{12}) + a_{22}a_{11}.$$

This general procedure can be extended to larger systems of equations. Applying this technique to a 3×3 system,

$$\begin{aligned} a_{11}q_1 + a_{12}q_2 + a_{13}q_3 &= \mathrm{x}_1 \\ a_{21}q_1 + a_{22}q_2 + a_{23}q_3 &= \mathrm{x}_2 \\ a_{31}q_1 + a_{32}q_2 + a_{33}q_3 &= \mathrm{x}_3 \end{aligned} \qquad \text{or} \qquad \mathbf{Aq} = \mathbf{x}$$

$$q_1 = \frac{\Delta_1}{\Delta_{\mathbf{A}}} \ , \ q_2 = \frac{\Delta_2}{\Delta_{\mathbf{A}}} \ , q_3 = \frac{\Delta_3}{\Delta_{\mathbf{A}}}$$

where one way (among many equivalent ways) of expressing the determinant is

$$\Delta_{\mathbf{A}} = \left|\mathbf{A}\right| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

and

$$\Delta_1 = \begin{vmatrix} \mathrm{x}_1 & a_{12} & a_{13} \\ \mathrm{x}_2 & a_{22} & a_{23} \\ \mathrm{x}_3 & a_{32} & a_{33} \end{vmatrix} \;,\; \Delta_2 = \begin{vmatrix} a_{11} & \mathrm{x}_1 & a_{13} \\ a_{21} & \mathrm{x}_2 & a_{23} \\ a_{31} & \mathrm{x}_3 & a_{33} \end{vmatrix} \;,\; \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & \mathrm{x}_1 \\ a_{21} & a_{22} & \mathrm{x}_2 \\ a_{31} & a_{32} & \mathrm{x}_3 \end{vmatrix}.$$

This method of finding solutions to systems of linear equations is called *Cramer's Rule*. It is very handy, especially in symbolic solutions of systems of equations. In the actual numerical solution of systems of equations on a computer, Cramer's rule is less efficient than other techniques, like Gaussian elimination for example.

Here are some other properties of determinants that are sometimes useful in vector and matrix analysis of signals and systems:

1. If any two rows or columns of a matrix are exchanged the determinant changes sign (but not magnitude).

$$\begin{vmatrix} \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{q1} & a_{q2} & a_{q3} & \cdots & a_{qn} \\ \vdots & \vdots & \vdots & \cdots & \vdots \end{vmatrix} = - \begin{vmatrix} \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{q1} & a_{q1} & a_{q3} & \cdots & a_{qn} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots & \cdots & \vdots \end{vmatrix}$$

2. The determinant of the identity matrix is 1.

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1$$

3. If any two rows or two columns of a matrix are equal the determinant is zero.

$$\begin{vmatrix} \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots & \cdots & \vdots \end{vmatrix} = 0$$

4. A matrix with a row or column of zeros has a determinant of zero.

$$\begin{vmatrix} \vdots & 0 & \cdots & \vdots \\ \vdots & 0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & \vdots & \vdots \end{vmatrix} = 0$$

5. The determinant of the product of two matrices is the product of the determinants.

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$$

6. Transposing a matrix does not change its determinant.

$$|\mathbf{A}| = |\mathbf{A}^T|$$

One other important operation is the inverse of a matrix. The inverse of a matrix $\mathbf{A}$ is defined as the matrix $\mathbf{A}^{-1}$ which when premuliplied or postmultiplied by $\mathbf{A}$ yields the identity matrix

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

The inverse of a matrix can be found by multiple methods. One formula for the inverse of a matrix is

$$\mathbf{A}^{-1} = \frac{\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{n2} & \cdots & A_{NN} \end{bmatrix}^T}{|\mathbf{A}|} .$$

where A_{rc} is the cofactor of the element a_{rc} in the matrix **A**. In words, the inverse of a matrix is the transpose of the matrix of cofactors of the elements of **A**, divided by the determinant of **A**. A term that is useful here is the *adjoint* of a matrix. The adjoint of **A** adj **A** is transpose of the matrix of cofactors. Therefore, the inverse of a matrix is the adjoint of the matrix, divided by the determinant of the matrix

$$\mathbf{A}^{-1} = \frac{\text{adj}\,\mathbf{A}}{|\mathbf{A}|} \quad . \qquad \text{(R.7)}$$

Of course, if the determinant is zero, the inverse of **A** is undefined. In that case it does not have an inverse. One use of the inverse of a matrix can be seen in the solution of a system of equations written in matrix form as $\mathbf{Aq} = \mathbf{x}$. If we premultiply both sides of this matrix equation by $\mathbf{A}^{-1}$ we get $\mathbf{A}^{-1}(\mathbf{Aq}) = (\mathbf{A}^{-1}\mathbf{A})\mathbf{q} = \mathbf{Iq} = \mathbf{q} = \mathbf{A}^{-1}\mathbf{x}$. So a straightforward way of solving a matrix equation of the form, $\mathbf{Aq} = \mathbf{x}$, is to premultiply both sides by the inverse of **A**. This directly yields the solution, $\mathbf{q} = \mathbf{A}^{-1}\mathbf{x}$, if $|\mathbf{A}| \neq 0$. The determinant of the inverse of a matrix is the reciprocal of the determinant of the matrix

$$|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|} .$$

If we have a scalar equation $aq = 0$ we know that either a or q or both must be zero to satisfy the equation. If a is a constant and q is a variable and we want to find a non-zero value of q that satisfies the equation, a must be zero. A very common situation in matrix analysis of systems of differential or difference equations is a matrix equation of the form $\mathbf{Aq} = 0$ where $\mathbf{A}$ is the $N \times N$ matrix of coefficients of N independent differential or difference equations and $\mathbf{q}$ is an $N \times 1$ vector of variables. This equation has non-zero solutions for $\mathbf{q}$ only if the determinant of $\mathbf{A}$ is zero. If $\mathbf{A}$ is the zero matrix, its determinant is zero and any vector, $\mathbf{q}$, will satisfy the equation. But $\mathbf{A}$ need not be a zero matrix, it must only have a determinant of zero. If $\mathbf{A}$ is not the zero matrix and its determinant is zero, then there are only certain particular vectors $\mathbf{q}$ that can solve the equation. If $\mathbf{q}$ is a solution to $\mathbf{Aq} = 0$ then for any scalar c $c\mathbf{q}$ is also a solution. If $\mathbf{q}_1$ and $\mathbf{q}_2$ are both solutions to $\mathbf{Aq} = 0$, then any linear combination of $\mathbf{q}_1$ and $\mathbf{q}_2$ is also a solution. These properties will be significant when we come to eigenvalues and eigenvectors.

MATLAB can find the determinant of a matrix and the inverse of a matrix, using the `det` and `inv` functions. For example,

```
»A = round(4*randn(3,3))

A =

     0    -5    -3
    -3     3     3
     1     6     5

»det(A)

ans =

   -27

»inv(A)

ans =

    0.1111   -0.2593    0.2222
   -0.6667   -0.1111   -0.3333
    0.7778    0.1852    0.5556

»inv(A)*A

ans =

    1.0000   -0.0000   -0.0000
         0    1.0000         0
         0   -0.0000    1.0000
```

R.4 Derivatives and Differences

The derivative of a matrix is simply the matrix of the derivatives of the corresponding elements of the matrix. For example, if

$$\mathbf{A} = \begin{bmatrix} t^2 & \text{y}t \\ \sin(t) & e^{5t} \end{bmatrix}$$

then

$$\frac{d}{dt}(\mathbf{A}) = \begin{bmatrix} 2t & \text{y} \\ \cos(t) & 5e^{5t} \end{bmatrix}.$$

Some common differentiation rules which follow from this definition are

$$\frac{d}{dt}(\mathbf{A} + \mathbf{B}) = \frac{d}{dt}(\mathbf{A}) + \frac{d}{dt}(\mathbf{B})$$

$$\frac{d}{dt}(c\mathbf{A}) = c\frac{d}{dt}(\mathbf{A})$$

$$\frac{d}{dt}(\mathbf{AB}) = \mathbf{A}\frac{d}{dt}(\mathbf{B}) + \mathbf{B}\frac{d}{dt}(\mathbf{A})$$

which are formally just like their scalar counterparts.

In an analogous manner the first forward difference of a matrix is the matrix of first forward differences of the corresponding elements . If

$$\mathbf{A} = \begin{bmatrix} \alpha^n & e^{-\frac{n}{2}} \\ n^2 + 3n & \cos\left(\frac{2\pi n}{N_0}\right) \end{bmatrix}$$

then

$$\Delta(\mathbf{A}) = \begin{bmatrix} \alpha^{n+1} - \alpha^n & e^{-\frac{n+1}{2}} - e^{-\frac{n}{2}} \\ (n+1)^2 + 3(n+1) - (n^2 + 3n) & \cos\left(\frac{2\pi(n+1)}{N_0}\right) - \cos\left(\frac{2\pi n}{N_0}\right) \end{bmatrix}$$

or

$$\Delta(\mathbf{A}) = \begin{bmatrix} \alpha^n(\alpha - 1) & e^{-\frac{n}{2}}\left(e^{-\frac{1}{2}} - 1\right) \\ 2n + 4 & \cos\left(\frac{2\pi(n+1)}{N_0}\right) - \cos\left(\frac{2\pi n}{N_0}\right) \end{bmatrix}.$$

Don't confuse $\Delta_{\mathbf{A}}$, the determinant of the matrix **A** with $\Delta(\mathbf{A})$, the first forward difference of the matrix **A**. The difference is usually clear in context.

R.5 Eigenvalues and Eigenvectors

It is always possible to write a system of linear, independent, constant-coefficient, ordinary differential equations as a single matrix equation of the form

$$\mathbf{q}' = \mathbf{A}\mathbf{q} + \mathbf{B}\mathbf{x} \ . \tag{R.9}$$

where $\mathbf{q}$ is the vector of solution functions and $\mathbf{x}$ is the vector of forcing functions and **A** and **B** are coefficient matrices. (This is the form that is used in state variable analysis.) Thus (R.9) is a linear, constant-coefficient, ordinary, first-order matrix, differential equation. Using an example from Appendix I, the system of differential equations

$$\begin{aligned} y_1' + 5y_1 + 2y_2 &= 10 \\ y_2' + 3y_2 + y_1 &= 0 \end{aligned}$$

with initial conditions, $y_1(0) = 1$ and $y_2(0) = 0$, can be written as $\mathbf{q}' = \mathbf{A}\mathbf{q} + \mathbf{B}\mathbf{x}$ where ,

$$\mathbf{A} = \begin{bmatrix} -5 & -2 \\ -1 & -3 \end{bmatrix} \ , \ \mathbf{q} = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \ , \ \mathbf{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \ , \ \mathbf{x} = 10$$

and

$$\mathbf{q}' = \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \frac{d}{dt}\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \frac{d}{dt}(\mathbf{q})$$

with an initial-condition vector, $\mathbf{q}_0 = \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The corresponding homogeneous equation is

$$\mathbf{q}' = \mathbf{A}\mathbf{q} \tag{R.10}$$

and we know that the solution of the homogeneous equation is a linear combination of solutions of the form $\mathbf{q}_h = \mathbf{K}_h e^{\lambda t}$, where $\mathbf{K}_h$ is a 2×1 vector of arbitrary constants instead of the single arbitrary constant we would have if we were solving a first-order scalar differential equation. Therefore we know that

$$\mathbf{q}_h' = \mathbf{K}_h \lambda e^{\lambda t} = \lambda \mathbf{q} \tag{R.11}$$

and, equating (R.11) and (R.10), that the solution of the homogeneous system of equations is the solution of the matrix equation $\mathbf{Aq} = \lambda\mathbf{q}$. This can be rearranged into $\mathbf{Aq} - \lambda\mathbf{q} = \left(\mathbf{A} - \lambda\mathbf{I}\right)\mathbf{q} = 0$. For a non-trivial solution ($\mathbf{q} \neq 0$) $\left|\mathbf{A} - \lambda\mathbf{I}\right| = 0$ or

$$\left|\begin{bmatrix} -5 & -2 \\ -1 & -3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right| = \begin{vmatrix} -5-\lambda & -2 \\ -1 & -3-\lambda \end{vmatrix} = 0$$

or $\left(-5-\lambda\right)\left(-3-\lambda\right) - 2 = 0$ which is equivalent to the system of characteristic equations

$$\begin{aligned} \left(\lambda+5\right)K_{h1} + 2K_{h2} &= 0 \\ K_{h1} + \left(\lambda+3\right)K_{h2} &= 0 \end{aligned} \quad . \tag{R.12}$$

The formulation, $\left|\mathbf{A} - \lambda\mathbf{I}\right| = 0$, is the *matrix characteristic equation* for a matrix differential or difference equation. From (R.12) we get $\lambda^2 + 8\lambda + 13 = 0$. The eigenvalues are completely determined by the coefficient matrix $\mathbf{A}$. For each eigenvalue there is a corresponding *eigenvector* $\mathbf{q}$ which, together with the eigenvalue, solves $\left(\mathbf{A} - \lambda\mathbf{I}\right)\mathbf{q} = 0$. For λ_1 we have the equality

$$\left(\mathbf{A} - \lambda_1\mathbf{I}\right)\mathbf{q} = 0 .$$

Any row of $\mathbf{A} - \lambda_1\mathbf{I}$ can be use to determine the direction of an eigenvector. For example, using the first row,

$$\left(-5-\lambda_1\right)q_1 - 2q_2 = 0$$

or

$$\frac{q_1}{q_2} = \frac{y_1\left(t\right)}{y_2\left(t\right)} = \frac{K_{h1}e^{\lambda_1 t}}{K_{h2}e^{\lambda_1 t}} = \frac{K_{h1}}{K_{h2}} = -\frac{2}{5+\lambda_1}$$

This sets the ratio of the components of the eigenvector and therefore its direction (but not any unique magnitude, yet) in the *N*-space, where, in this case, $N = 2$. Therefore this eigenvector would be

$$\mathbf{q}_1 = K_1 \begin{bmatrix} -\dfrac{2}{5+\lambda_1} \\ 1 \end{bmatrix} e^{\lambda_1 t} = K_1 \begin{bmatrix} -0.732 \\ 1 \end{bmatrix} e^{-2.268t}$$

where K_1 is an arbitrary constant. If we had used the second row we would have gotten exactly the same vector direction. Using λ_2 we would get

$$\mathbf{q}_2 = K_2 \begin{bmatrix} -\dfrac{2}{5+\lambda_2} \\ 1 \end{bmatrix} e^{\lambda_2 t} = K_2 \begin{bmatrix} 2.732 \\ 1 \end{bmatrix} e^{-5.732t} .$$

Notice that we have arbitrarily set the q_2 component of **q** to 1 in both cases. This is just a convenience. We can do it because we only know the ratio of the two components of **q**, not their exact values (yet). Since the exact values of the two vector components are not known, only their ratio, they are often chosen so as to make the length of **q** one, making **q** a unit vector. Writing the vectors as unit vectors we get

$$\mathbf{q}_1 = \begin{bmatrix} -0.5907 \\ 0.807 \end{bmatrix} e^{-2.268t} \text{ and } \qquad \mathbf{q}_2 = \begin{bmatrix} 0.9391 \\ 0.3437 \end{bmatrix} e^{-5.732t} .$$

The most general homogeneous solution is a linear combination of the eigenvectors of the form,

$$\mathbf{q} = K_{h1}\mathbf{q}_1 + K_{h2}\mathbf{q}_2 = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} \begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix}$$

where the two arbitrary constants, K_{h1} and K_{h2}, must be chosen to satisfy initial conditions.

The next solution step is to find the particular solution

$$\mathbf{q}_p = \begin{bmatrix} y_{1p}(t) \\ y_{2p}(t) \end{bmatrix} .$$

Since the forcing function is a constant $\mathbf{x} = 10$ the particular solution is a vector of constants

$$\mathbf{q}_p = \begin{bmatrix} K_{p1} \\ K_{p2} \end{bmatrix} .$$

Substituting into (R.9) $\mathbf{q}_p' = \mathbf{A}\mathbf{q}_p + \mathbf{B}\mathbf{x}$. Since $\mathbf{q}_p$ is a vector of constants, $\mathbf{q}_p' = 0$ and $\mathbf{A}\mathbf{q}_p = -\mathbf{B}\mathbf{x}$. Solving, $\mathbf{q}_p = -\mathbf{A}^{-1}\mathbf{B}\mathbf{x}$. The inverse of **A** is

$$\mathbf{A}^{-1} = \frac{1}{13}\begin{bmatrix} -3 & 2 \\ 1 & -5 \end{bmatrix}.$$

Therefore

$$\mathbf{q}_p = -\frac{1}{13}\begin{bmatrix} -3 & 2 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}10 = -\frac{1}{13}\begin{bmatrix} -3 \\ 1 \end{bmatrix}10 = \frac{1}{13}\begin{bmatrix} 30 \\ -10 \end{bmatrix}.$$

Now we know that the total solution is

$$\mathbf{q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix}\begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} + \mathbf{q}_p$$

or

$$\mathbf{q} = \begin{bmatrix} -0.5907e^{-2.268t} & 0.9391e^{-5.732t} \\ 0.807e^{-2.268t} & 0.3437e^{-5.732t} \end{bmatrix}\begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} + \begin{bmatrix} 2.308 \\ -0.769 \end{bmatrix}.$$

The only task left is to solve for the arbitrary constants K_{h1} and K_{h2}. The vector of initial conditions (at time $t = 0$) is

$$\mathbf{q}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then

$$\begin{bmatrix} -0.5907 & 0.9391 \\ 0.807 & 0.3437 \end{bmatrix}\begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} + \begin{bmatrix} 2.308 \\ -0.769 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Solving,

$$\begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} = \begin{bmatrix} -0.5907 & 0.9391 \\ 0.807 & 0.3437 \end{bmatrix}^{-1}\begin{bmatrix} -1.308 \\ 0.769 \end{bmatrix} = \begin{bmatrix} 1.2194 \\ -0.6258 \end{bmatrix}$$

and, finally,

$$\mathbf{q} = \begin{bmatrix} \mathrm{y}_1(t) \\ \mathrm{y}_2(t) \end{bmatrix} = \begin{bmatrix} -0.5907e^{-2.268t} & 0.9391e^{-5.732t} \\ 0.807e^{-2.268t} & 0.3437e^{-5.732t} \end{bmatrix}\begin{bmatrix} 1.2194 \\ -0.6258 \end{bmatrix} + \begin{bmatrix} 2.308 \\ -0.769 \end{bmatrix}$$

or

$$\mathbf{q} = \begin{bmatrix} \mathrm{y}_1(t) \\ \mathrm{y}_2(t) \end{bmatrix} = \begin{bmatrix} -0.7203e^{-2.268t} - 0.5877e^{-5.732t} \\ 0.9841e^{-2.268t} - 0.2151e^{-5.732t} \end{bmatrix} + \begin{bmatrix} 2.308 \\ -0.769 \end{bmatrix}.$$

Just as was true with differential equations, it is always possible to write a system of linear, independent, constant-coefficient, ordinary difference equations as a single matrix equation of the form,

$$\mathbf{q}[n+1] = \mathbf{A}\mathbf{q}[n] + \mathbf{B}\mathbf{x}[n] \quad . \tag{R.13}$$

where $\mathbf{q}$ is the vector of solution functions, $\mathbf{x}$ is the vector of forcing functions and $\mathbf{A}$ and $\mathbf{B}$ are coefficient matrices. Thus (R.13) is a linear, constant-coefficient, ordinary, first-order, matrix, difference equation. Consider the example from Appendix I, the system of difference equations,

$$3y_1[n] + 2y_1[n-1] + y_2[n] = 0$$
$$4y_2[n] + 2y_2[n-1] + y_1[n] = 5$$

with initial conditions, $y_1[0] = 0$ and $y_2[0] = 2$. It can be rearranged into the form,

$$3y_1[n] + y_2[n] = -2y_1[n-1]$$
$$4y_2[n] + y_1[n] = -2y_2[n-1] + 5$$

or the equivalent form,

$$\begin{aligned} 3y_1[n+1] + y_2[n+1] &= -2y_1[n] \\ y_1[n+1] + 4y_2[n+1] &= -2y_2[n] + 5 \end{aligned} \quad . \tag{R.14}$$

There seems to be a problem here. How do we arrange this equation into the form, $\mathbf{q}[n+1] = \mathbf{Aq}[n] + \mathbf{Bx}[n]$? The answer lies in redefining the functions we are solving for. Let $q_1[n] = 3y_1[n] + y_2[n]$ and $q_2[n] = y_1[n] + 4y_2[n]$ implying that

$$y_1[n] = \frac{4q_1[n] - q_2[n]}{11} \quad \text{and} \quad y_2[n] = -\frac{q_1[n] - 3q_2[n]}{11} . \tag{R.15}$$

Then (R.14) can then be written as

$$\begin{aligned} q_1[n+1] &= \frac{-8q_1[n] + 2q_2[n]}{11} \\ q_2[n+1] &= \frac{2q_1[n] - 6q_2[n]}{11} + 5 \end{aligned} \quad . \tag{R.16}$$

We can express (R.16) in the standard matrix form $\mathbf{q}[n+1] = \mathbf{Aq}[n] + \mathbf{Bx}[n]$ where

$$\mathbf{A} = \frac{1}{11}\begin{bmatrix} -8 & 2 \\ 2 & -6 \end{bmatrix} , \ \mathbf{q} = \begin{bmatrix} q_1[n] \\ q_2[n] \end{bmatrix} , \ \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} , \ \mathbf{x} = 5$$

with an initial-condition vector,

$$\mathbf{q}_0 = \begin{bmatrix} 3y_1[0] + y_2[0] \\ y_1[0] + 4y_2[0] \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} .$$

The corresponding homogeneous equation is

$$\mathbf{q}[n+1] = \mathbf{A}\mathbf{q}[n] \tag{R.17}$$

and we know that the solution of the homogeneous equation is a linear combination of solutions of the form $\mathbf{q}_h = \mathbf{K}_h \alpha^n$ where $\mathbf{K}_h$ is a 2×1 vector of arbitrary constants. Therefore we know that

$$\mathbf{q}_h[n+1] = \mathbf{K}_h \alpha^{n+1} = \alpha \mathbf{q}_h \tag{R.18}$$

and, substituting (R.18) into (R.17), that the solution of the homogeneous system of equations is the solution of the matrix equation,

$$\mathbf{A}\mathbf{q} = \alpha \mathbf{q} \tag{R.19}$$

which can be rearranged into $\mathbf{A}\mathbf{q} - \alpha\mathbf{q} = (\mathbf{A} - \alpha\mathbf{I})\mathbf{q} = 0$. For a non-trivial solution, $\mathbf{q} \neq 0$ $|\mathbf{A} - \alpha\mathbf{I}| = 0$ or

$$\left| \frac{1}{11}\begin{bmatrix} -8 & 2 \\ 2 & -6 \end{bmatrix} - \alpha \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \frac{1}{11}\begin{vmatrix} -8-11\alpha & 2 \\ 2 & -6-11\alpha \end{vmatrix} = 0$$

or

$$(8+11\alpha)(6+11\alpha) - 4 = 0 . \tag{R.20}$$

The formulation $|\mathbf{A} - \alpha\mathbf{I}| = 0$ is the matrix characteristic equation for a matrix difference equation. From (R.20), we get $121\alpha^2 + 154\alpha + 44 = 0$ and the eigenvalues are $\alpha_1 = -0.4331$ and $\alpha_2 = -0.8396$ as we found in the previous solution of this system of difference equations. Notice that the redefinition of the functions we are solving for did not change the eigenvalues. The eigenvalues are completely determined by the coefficient matrix $\mathbf{A}$. For each eigenvalue there is a corresponding eigenvector $\mathbf{q}$ which, together with the eigenvalue, solves (R.19). For α_1 we have the equality $[\mathbf{A} - \alpha_1\mathbf{I}]\mathbf{q} = 0$. Any row of $\mathbf{A} - \alpha_1\mathbf{I}$ can be use to determine the direction of an eigenvector. For example, using the first row, $(-8 - 11\alpha_1)q_1 + 2q_2 = 0$ or

$$\frac{q_1[n]}{q_2[n]} = \frac{K_{h1}\alpha_1^n}{K_{h2}\alpha_1^n} = \frac{K_{h1}}{K_{h2}} = \frac{2}{8+11\alpha_1} .$$

Therefore a unit eigenvector would be

$$\mathbf{q}_1 = K_1 \begin{bmatrix} \dfrac{2}{8+11\alpha_1} \\ 1 \end{bmatrix} \alpha_1^n = \begin{bmatrix} 0.5257 \\ 0.8507 \end{bmatrix} (-0.4331)^n .$$

Using α_2 we would get the other unit eigenvector

$$\mathbf{q}_2 = \begin{bmatrix} \dfrac{2}{8+11\alpha_2} \\ 1 \end{bmatrix} \alpha_2^n = \begin{bmatrix} -0.8507 \\ 0.5257 \end{bmatrix} \left(-0.8396\right)^n .$$

Again the most general homogeneous solution is a linear combination of the eigenvectors of the form

$$\mathbf{q} = K_{h1}\mathbf{q}_1 + K_{h2}\mathbf{q}_2 = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} \begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix}$$

where the two arbitrary constants K_{h1} and K_{h2} must be chosen to satisfy initial conditions.

The next solution step is to find the particular solution

$$\mathbf{q}_p = \begin{bmatrix} q_{1p}[n] \\ q_{2p}[n] \end{bmatrix}.$$

Since the forcing function is a constant, $\mathbf{x} = 5$, the particular solution is a vector of constants

$$\mathbf{q}_p = \begin{bmatrix} K_{p1} \\ K_{p2} \end{bmatrix}.$$

Substituting into (R.16), $\mathbf{q}_p[n+1] = \mathbf{A}\mathbf{q}_p[n] + \mathbf{B}\mathbf{x}[n]$. Since $\mathbf{q}_p$ is a vector of constants, $\mathbf{q}_p[n+1] = \mathbf{q}_p[n]$ and $\left(\mathbf{I}-\mathbf{A}\right)\mathbf{q}_p[n] = \mathbf{B}\mathbf{x}[n]$. Solving, $\mathbf{q}_p[n] = \left(\mathbf{I}-\mathbf{A}\right)^{-1}\mathbf{B}\mathbf{x}[n]$. The inverse of $\left(\mathbf{I}-\mathbf{A}\right)$ is

$$\left(\mathbf{I}-\mathbf{A}\right)^{-1} = \begin{bmatrix} 0.5862 & 0.0690 \\ 0.0690 & 0.6552 \end{bmatrix}.$$

Therefore

$$\mathbf{q}_p = \begin{bmatrix} 0.5862 & 0.0690 \\ 0.0690 & 0.6552 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} 5 = \begin{bmatrix} 0.345 \\ 3.276 \end{bmatrix}.$$

Now we know that the total solution is

$$\mathbf{q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} \begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} + \mathbf{q}_p$$

or

$$\mathbf{q} = \begin{bmatrix} 0.5257(-0.4331)^n & -0.8507(-0.8396)^n \\ 0.8507(-0.4331)^n & 0.5257(-0.8396)^n \end{bmatrix} \begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} + \begin{bmatrix} 0.345 \\ 3.276 \end{bmatrix} . \quad \text{(R.21)}$$

The only task left is to solve for the arbitrary constants, K_{h1} and K_{h2}. The vector of initial conditions (at time, $n = 0$) is

$$\mathbf{q}_0 = \begin{bmatrix} 2 \\ 8 \end{bmatrix}.$$

Then

$$\begin{bmatrix} 0.5257 & -0.8507 \\ 0.8507 & 0.5257 \end{bmatrix} \begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} + \begin{bmatrix} 0.345 \\ 3.276 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}.$$

Solving,

$$\begin{bmatrix} K_{h1} \\ K_{h2} \end{bmatrix} = \begin{bmatrix} 0.5257 & -0.8507 \\ 0.8507 & 0.5257 \end{bmatrix}^{-1} \begin{bmatrix} 1.655 \\ 4.724 \end{bmatrix} = \begin{bmatrix} 4.8885 \\ 1.0754 \end{bmatrix} \quad \text{(R.22)}$$

and, finally, combining (R.21) and(R.22),

$$\mathbf{q} = \begin{bmatrix} 2.57(-0.4331)^n - 0.915(-0.8396)^n \\ 4.1585(-0.4331)^n + 0.5655(-0.8396)^n \end{bmatrix} + \begin{bmatrix} 0.345 \\ 3.276 \end{bmatrix} . \quad \text{(R.23)}$$

This solution should be equivalent to the previous solution of this system of difference equations in Appendix I

$$y_1[n] = 0.557(-0.4331)^n - 0.3841(-0.8396)^n - 0.1724$$

and

$$y_2[n] = 0.9005(-0.4331)^n + 0.2374(-0.8396)^n + 0.8621.$$

Using (R.15),

$$y_1[n] = \frac{4q_1[n] - q_2[n]}{11} \quad \text{and} \quad y_2[n] = -\frac{q_1[n] - 3q_2[n]}{11}.$$

with (R.23) we get

$$\begin{aligned} y_1[n] &= 0.557(-0.4331)^n - 0.3841(-0.8396)^n - 0.1724 \\ y_2[n] &= 0.9005(-0.4331)^n + 0.2374(-0.8396)^n + 0.8621 \end{aligned} \tag{R.24}$$

confirming that the two solution techniques agree.

Exercises

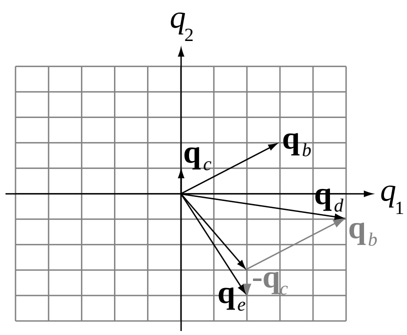
(On each exercise, the answers listed are in random order.)

1. Graph these **q** vectors in the "$q_1 - q_2$" plane.

(a) $\mathbf{q}_a = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$ (b) $\mathbf{q}_b = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ (c) $\mathbf{q}_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(d) $\mathbf{q}_d = \mathbf{q}_a + \mathbf{q}_b$ (e) $\mathbf{q}_e = \mathbf{q}_a - \mathbf{q}_c$

Answers:



2. Let

$$\mathbf{A} = \begin{bmatrix} 2 & -5 & 1 \\ 8 & 7 & -3 \\ -9 & 4 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 3 & 2 & 9 \\ 7 & 6 & 4 \\ -3 & -1 & 8 \end{bmatrix} \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} -4 & 1 & 2 \\ 5 & -2 & 7 \\ -2 & -8 & 3 \end{bmatrix}$$

Find these matrices.

(a) $\mathbf{A} + \mathbf{B}$ (b) $\mathbf{A} + \mathbf{C}$ (c) $\mathbf{B} - \mathbf{C}$

Answers: $\begin{bmatrix} 7 & 1 & 7 \\ 2 & 8 & -3 \\ -1 & 7 & 5 \end{bmatrix}$ $\begin{bmatrix} 5 & -3 & 10 \\ 15 & 13 & 1 \\ -12 & 3 & 10 \end{bmatrix}$ $\begin{bmatrix} -2 & -4 & 3 \\ 13 & 5 & 4 \\ -11 & -4 & 5 \end{bmatrix}$

3. Using the **A**, **B** and **C** matrices of Exercise 2, find these matrices.

(a) $-\mathbf{A}$ (b) $3\mathbf{B} + 2\mathbf{C}$ (c) $\mathbf{C} - 2\mathbf{B}$

Answers: $\begin{bmatrix} 1 & 8 & 31 \\ 31 & 14 & 26 \\ -13 & -19 & 30 \end{bmatrix}$ $\begin{bmatrix} -10 & -3 & -16 \\ -9 & -14 & -1 \\ 4 & -6 & -13 \end{bmatrix}$ $\begin{bmatrix} -2 & 5 & -1 \\ -8 & -7 & 3 \\ 9 & -4 & -2 \end{bmatrix}$

4. Find the transposes of these matrices.

(a) $\mathbf{A} = \begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \end{bmatrix}$ (b) $\mathbf{B} = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$ (c) $\mathbf{C} = \begin{bmatrix} 4 & 9 & -3 & -1 \end{bmatrix}$

(d) $\mathbf{D} = \mathbf{A} + \mathbf{C}^T$ (e) $\mathbf{E} = \mathbf{B} + \mathbf{B}^T$

Answers: $\begin{bmatrix} 4 \\ 9 \\ -3 \\ -1 \end{bmatrix}$ $\begin{bmatrix} 2 & -1 & 3 & 1 \end{bmatrix}$ $\begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$ $\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$ $\begin{bmatrix} 6 \\ 8 \\ 0 \\ 0 \end{bmatrix}$

5. Find these matrix products.

(a) $\mathbf{A} = \begin{bmatrix} 2 & 3 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ (b) $\mathbf{B} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}\begin{bmatrix} 2 & 3 \end{bmatrix}$

(c) $\mathbf{C} = \begin{bmatrix} 3 & -4 \\ 1 & 9 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ (d) $\mathbf{D} = \mathbf{B}\mathbf{B}^T$

(e) $\mathbf{E} = \mathbf{C}^T\mathbf{C}$ (f) $\mathbf{F} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}^T \begin{bmatrix} 3 & -4 \\ 1 & 9 \end{bmatrix}^T$

(g) $\mathbf{G} = \begin{bmatrix} 2 & -5 & 1 \\ -2 & 20 & 10 \\ 7 & -1 & 1 \end{bmatrix}\begin{bmatrix} -2 & 1 & 5 \\ -7 & -5 & 0 \\ 1 & 5 & 1 \end{bmatrix}$

(h) $\mathbf{H} = \begin{bmatrix} -2 & -7 & 1 \\ 1 & -5 & 5 \\ 5 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & -2 & 7 \\ -5 & 20 & -1 \\ 1 & 10 & 1 \end{bmatrix}$

Answers: $\begin{bmatrix} 32 & 32 & 11 \\ -126 & -52 & 0 \\ -6 & 17 & 36 \end{bmatrix}$ $\begin{bmatrix} 32 & -126 & -6 \\ 32 & -52 & 17 \\ 11 & 0 & 36 \end{bmatrix}$ $\begin{bmatrix} 81 & -252 \\ -252 & 784 \end{bmatrix}$

0 $\begin{bmatrix} -9 \\ 28 \end{bmatrix}$ $\begin{bmatrix} -9 & 28 \end{bmatrix}$ $\begin{bmatrix} 117 & -78 \\ -78 & 52 \end{bmatrix}$ $\begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix}$

6. Solve these systems of equations using Cramer's rule, if it is possible. If it is not possible, state why.

(a)
$$\begin{aligned} 2q_1 + 7q_2 &= -4 \\ -q_1 - 4q_2 &= 9 \end{aligned}$$

(b)
$$\begin{bmatrix} -1 & -12 & 0 \\ -7 & 6 & 8 \\ 3 & 15 & 11 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} -8 \\ -7 \\ 3 \end{bmatrix}$$

(c)
$$\begin{bmatrix} -4 & 12 & 0 & -10 \\ 6 & 6 & -1 & 13 \\ 7 & 11 & -14 & -7 \\ 6 & -11 & 2 & 5 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ -7 \\ 0 \end{bmatrix}$$

Answers:

$$q_1 = -\frac{2544}{16254} \quad q_2 = -\frac{861}{16254} \quad q_3 = \frac{6999}{16254} \quad q_4 = -\frac{1641}{16254}$$

$$q_1 = 47 \quad q_2 = -14$$

$$q_1 = \frac{780}{1158} \quad q_2 = \frac{707}{1158} \quad q_3 = -\frac{861}{1158}$$

7. Invert these matrices, if it is possible. If it is not possible, state why.

(a) $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ -9 & -3 \end{bmatrix}$ (b) $\mathbf{B} = \begin{bmatrix} -1 & -6 \\ 0 & 22 \end{bmatrix}$

(c) $\mathbf{C} = \begin{bmatrix} -9 & 15 & 3 \\ 6 & 5 & -9 \\ 5 & -6 & 0 \end{bmatrix}$ (d) $\mathbf{D} = \begin{bmatrix} -1 & 3 \\ 7 & 2 \\ -8 & -3 \end{bmatrix}$

(e) $\mathbf{E} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (f) $\mathbf{F} = \mathbf{AB}$

(g) $\mathbf{G} = \mathbf{D}^T\mathbf{D}$

Answers: Not invertible Not invertible $\begin{bmatrix} -1 & -\frac{3}{11} \\ 0 & \frac{1}{22} \end{bmatrix}$

$-\frac{1}{372}\begin{bmatrix} -54 & -18 & -150 \\ -45 & -15 & -63 \\ -61 & 21 & -135 \end{bmatrix}$ Not invertible $\frac{1}{1283}\begin{bmatrix} 22 & -35 \\ -35 & 114 \end{bmatrix}$ $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

8. Where possible solve the equations of Exercise 6 using a matrix inverse.

Answers: (See Exercise 6)

9. Solve these systems of differential or difference equations.

(a) $\begin{aligned} y_1'(t) &= -3y_1(t) + 2y_2(t) + 5 \\ y_2'(t) &= 4y_1(t) + y_2(t) - 2 \end{aligned}$, $y_1(0) = 3$, $y_2(0) = -1$

(b) $\begin{aligned} y_1'(t) &= 4y_1(t) + 10y_2'(t) + 5e^{-t} \\ y_2'(t) &= -y_1(t) + 3y_2(t) - 7e^{-3t} \end{aligned}$, $y_1(0) = 0$, $y_2(0) = 9$

(c) $\begin{aligned} y_1[n+1] &= 4y_1[n] - 12y_2[n] + 3 \\ y_2[n+1] &= -y_1[n] + 7y_2[n] - 1 \end{aligned}$, $y_1[0] = 1$, $y_2[0] = 0$

(d) $$\begin{aligned} y_1[n+1] &= 6y_1[n] + 8y_2[n] - \left(\frac{1}{2}\right)^n \\ y_2[n+1] &= 5y_1[n+1] - 3y_2[n] + \left(\frac{3}{4}\right)^n \end{aligned} \quad , \; y_1[0] = -6 \quad , \; y_2[0] = 1$$

Answers:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 1.642e^{-4.4641t} + 0.539e^{2.4641t} \\ -1.202e^{-4.4641t} + 1.475e^{2.4641t} \end{bmatrix} + \frac{1}{11}\begin{bmatrix} 9 \\ -14 \end{bmatrix}$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = e^{-1.5t}\begin{bmatrix} -15.5\cos(3.1225t) + 74.37\sin(3.1225t) \\ 5.4166\cos(3.1225t) + 12.7688\sin(3.1225t) \end{bmatrix} + \begin{bmatrix} -2e^{-t} + 17.5e^{-3t} \\ -0.5e^{-t} + 4.0833e^{-3t} \end{bmatrix}$$

$$\begin{bmatrix} y_1[n] \\ y_2[n] \end{bmatrix} = \begin{bmatrix} -2.651(0.5422)^n - 0.9488(-0.0922)^n \\ -7.746(0.5422)^n + 3.246(-0.0922)^n \end{bmatrix} + \begin{bmatrix} 4.6 \\ 4.5 \end{bmatrix}$$

$$\begin{bmatrix} y_1[n] \\ y_2[n] \end{bmatrix} = \begin{bmatrix} -0.7147(43.4146)^n - 5.214(43.4146)^n \\ -3.3425(43.4146)^n + 4.171(-0.4146)^n \end{bmatrix} + \begin{bmatrix} 0.0892\left(\frac{1}{2}\right)^n - 0.1610\left(\frac{3}{4}\right)^n \\ 0.0638\left(\frac{1}{2}\right)^n + 0.1057\left(\frac{3}{4}\right)^n \end{bmatrix}$$