Chapter 1

Vectors and Matrices

1.1 Introduction

Vectors and matrices are used extensively throughout this text. Both are essential as one cannot derive and analyze laws of physics and physical measurements without vectors and one cannot process these measurements in a digital computer without matrices. Further, each has powerful features that when combined can result in considerable derivation simplifications. We shall highlight the main similarities and the subtle differences between them.

In this context we will be concerned with Euclidean vector spaces for which the inner product is defined. In n-dimensional Euclidean vector spaces it is possible to construct a set of n orthogonal unit vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$. As such, an arbitrary vector $\mathbf{v}$ in this space is represented by

$$\mathbf{v} = v_1\mathbf{r}_1 + v_2\mathbf{r}_2 + \cdots + v_n\mathbf{r}_n \tag{1.1}$$

where $v_1, v_2, \ldots, v_n$ are the scalar coordinates of $\mathbf{v}$.

In the special case of three dimensional vector space, the unit vectors $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ can be graphically represented by a set of three orthogonal axes. Hence, v_1, v_2, v_3 (the coordinates of the 3-dimensional vector $\mathbf{v}$) are the projections of $\mathbf{v}$ along $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$.

A matrix on the other hand, is an array of n rows and m columns, of $n \times m$ numbers [1,2]. For example, a 3×3 matrix $\mathbf{A}$ is represented by

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The transpose of a $i \times j$ matrix **B** is a $j \times i$ matrix, denoted by **B′**, in which the rows and columns of **B** trade places. For example, the transpose of the above matrix **A** is given by

$$\mathbf{A}' = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

In particular, a $n \times 1$ array is a column matrix and a $1 \times n$ array is a row matrix. Transposing a row matrix makes it a column matrix and vice versa.

To adapt vectors to matrix manipulations it is common to use row or column matrices whose components are the vector coordinates. When there is no confusion, column matrices will be used to denote vectors. So vector **v**, in Eq. (1.1), can be represented in matrix form by

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{1.2}$$

Implicit in the above equation that the components $v_1, v_2, \ldots, v_n$ are the coordinates of **v** along the vectors $\mathbf{r}_1$, $\mathbf{r}_2, \ldots, \mathbf{r}_n$.

1.2 Vector Inner Product

The inner product of the two real vectors **u** and **v** is defined by [3,4]

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \tag{1.3}$$

where $\{u_1, u_2, \ldots, u_n\}$ and $\{v_1, v_2, \ldots, v_n\}$ are the coordinates of **u** and **v**. The length of vector **u**, also called its norm, is defined by the inner product

$$\text{norm}(\mathbf{u}) = |\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} \tag{1.4}$$

We call **u** and **v** orthogonal if their inner product equals zero; if, in addition, each is of unit length then they are orthonormal. In particular the set of vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$, introduced above, are orthonormal because

$$\mathbf{r}_i \cdot \mathbf{r}_i = 1 \quad \text{and} \quad \mathbf{r}_i \cdot \mathbf{r}_j = 0, \quad i, j = 1, 2, \ldots, n; \; i \neq j \tag{1.5}$$

Thus it can be seen from Eqs. (1.1) and (1.5) that the coordinates of vector **v** are given by

$$v_i = \mathbf{v} \cdot \mathbf{r}_i, \quad i = 1, 2, \ldots, n \tag{1.6}$$

In a three dimensional space the inner product has a physical significance. The cosine of the angle between vectors **u** and **v**, denoted by cos(**u, v**) is determined by the cosine law (see Appendix A)

$$\cos(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \tag{1.7}$$

If **v** is a unit vector we get from Eqs. (1.6) and (1.7)

$$v_{\mathrm{i}} = \mathbf{v} \cdot \mathbf{r}_i = \cos(\mathbf{v}, \mathbf{r}_i), \quad i = 1, 2, \ldots, n \tag{1.8}$$

When a vector set $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ is represented in column matrices, then they are orthonormal if

$$\mathbf{r}_i' \mathbf{r}_i = 1 \quad \text{and} \quad \mathbf{r}_i' \mathbf{r}_j = 0, \quad i, j = 1, 2, \ldots, n; \ i \neq j \tag{1.9}$$

The Hermitian inner product of the complex vectors **u** and **v** is defined by

$$\mathbf{u} \cdot \mathbf{v}^* = \mathbf{u}' \mathbf{v}^* = u_1 v_1^* + u_2 v_2^* + \ldots + u_n v_n^* \tag{1.10}$$

where the prime denotes the matrix transpose and the * denotes the complex conjugate. A matrix is called square if its number of rows is the same as its number of columns. The identity matrix is a real square whose diagonal components are all ones and the rest are all zeros. For example the 3×3 identity matrix is

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

1.3 Vector Cross Products and Skew Symmetric Matrix Algebra

Suppose that $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ is an orthonormal vector set for a 3-dimensional vector space in which the two vectors **u** and **v** are given by

$$\mathbf{u} = u_1\mathbf{r}_1 + u_2\mathbf{r}_2 + u_3\mathbf{r}_3$$
$$\mathbf{v} = v_1\mathbf{r}_1 + v_2\mathbf{r}_2 + v_3\mathbf{r}_3$$

then the vector cross product of **u** and **v**, denoted by $\mathbf{w} = \mathbf{u} \times \mathbf{v}$, is defined by [5,6]

$$\mathbf{w} = (u_2v_3 - u_3v_2)\mathbf{r}_1 + (u_3v_1 - u_1v_3)\mathbf{r}_2 + (u_1v_2 - u_2v_1)\mathbf{r}_3 \quad (1.11)$$

Notice that the cross product of the two vectors **u** and **v** is another vector that is orthogonal to both **u** and **v**. In matrix notation **w** is given by

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$$

It is shown in Appendix A that if **u** and **v** are unit vectors then the magnitude of their vector cross product, **w,** is the sine of the angle between them. Therefore if **k** is the unit vector along **w** then

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \sin(\mathbf{u}, \mathbf{v})\mathbf{k} \quad (1.12)$$

A skew symmetric matrix **B** is a matrix with the property of **B′** = -**B**. A 3-dimensional skew symmetric matrix emulates the cross product operation and enables it to be expressed in matrix notation. The vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

corresponds to the skew symmetric matrix defined by

$$\mathrm{S}(\mathbf{b}) = \tilde{\mathbf{b}} = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix} \tag{1.13}$$

The operator S and the tilde notation are identical and will be used interchangeably to denote skew symmetric matrices, even though the latter will be used whenever possible.

Properties of the Skew Symmetric Matrix

In the following it is assumed that **a**, **b** and **w** are three dimensional arbitrary vectors (column matrix) and α and β are arbitrary scalars.

1. Correspondence to vector cross product:

$$\tilde{\mathbf{b}}\mathbf{w} = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} b_2 w_3 - b_3 w_2 \\ b_3 w_1 - b_1 w_3 \\ b_1 w_2 - b_2 w_1 \end{bmatrix} = \mathbf{b} \times \mathbf{w} \tag{1.14}$$

Thus operating the matrix product of $\tilde{\mathbf{b}}$ and **w** corresponds to the vector cross product of the **b** and **w**.

2. Linearity:

$$\alpha\,\mathrm{S}(\mathbf{a}) + \beta\,\mathrm{S}(\mathbf{b}) = \alpha \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} + \beta \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix} \Rightarrow$$

$$\alpha\,\mathrm{S}(\mathbf{a}) + \beta\,\mathrm{S}(\mathbf{b}) = \begin{bmatrix} 0 & -\alpha a_3 - \beta b_3 & \alpha a_2 + \beta b_2 \\ \alpha a_3 + \beta b_3 & 0 & -\alpha a_1 - \beta b_1 \\ -\alpha a_2 - \beta b_2 & \alpha a_1 + \beta b_1 & 0 \end{bmatrix} \Rightarrow$$

$$\alpha \mathrm{S}(\mathbf{a}) + \beta \mathrm{S}(\mathbf{b}) = \mathrm{S}(\alpha \mathbf{a} + \beta \mathbf{b}) \tag{1.15}$$

3. Skewness

$$\tilde{\mathbf{b}}' = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}' = \begin{bmatrix} 0 & b_3 & -b_2 \\ -b_3 & 0 & b_1 \\ b_2 & -b_1 & 0 \end{bmatrix} = -\tilde{\mathbf{b}}$$

Thus

$$\tilde{\mathbf{b}}' = -\tilde{\mathbf{b}} \tag{1.16}$$

4. Operating on self vector

$$\tilde{\mathbf{b}}\mathbf{b} = \mathbf{b} \times \mathbf{b} = \mathbf{0} \tag{1.17}$$

Lemma 1.1

$$\mathrm{S}(\mathbf{b} \times \mathbf{w}) = \tilde{\mathbf{b}}\tilde{\mathbf{w}} - \tilde{\mathbf{w}}\tilde{\mathbf{b}} \tag{1.18}$$

Proof: Since

$$\begin{aligned} \tilde{\mathbf{b}}\tilde{\mathbf{w}} &= \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -b_2 w_2 - b_3 w_3 & b_2 w_1 & b_3 w_1 \\ b_1 w_2 & -b_3 w_3 - b_1 w_1 & b_3 w_2 \\ b_1 w_3 & b_2 w_3 & -b_1 w_1 - b_2 w_2 \end{bmatrix} \end{aligned} \tag{1.19}$$

The RHS matrix can be arranged in the form of

$$\tilde{\mathbf{b}}\tilde{\mathbf{w}} = -(\mathbf{w}'\mathbf{b})\mathbf{I} + \mathbf{w}\mathbf{b}' \tag{1.20}$$

Likewise

$$\tilde{\mathbf{w}}\tilde{\mathbf{b}} = -(\mathbf{b}'\mathbf{w})\mathbf{I} + \mathbf{b}\mathbf{w}' \tag{1.21}$$

Subtracting Eq. (1.21) from Eq. (1.20) gives

$$\begin{aligned}
\mathbf{w}\mathbf{b}' - \mathbf{b}\mathbf{w}' &= \tilde{\mathbf{b}}\tilde{\mathbf{w}} - \tilde{\mathbf{w}}\tilde{\mathbf{b}} \\
&= \begin{bmatrix} 0 & b_2 w_1 - b_1 w_2 & b_3 w_1 - b_1 w_3 \\ b_1 w_2 - b_2 w_1 & 0 & b_3 w_2 - b_2 w_3 \\ b_1 w_3 - b_3 w_1 & b_2 w_3 - b_3 w_2 & 0 \end{bmatrix} \quad (1.22) \\
&= \mathrm{S}\begin{bmatrix} b_2 w_3 - b_3 w_2 \\ b_3 w_1 - b_1 w_3 \\ b_1 w_2 - b_2 w_1 \end{bmatrix} = \mathrm{S}(\mathbf{b} \times \mathbf{w})
\end{aligned}$$

Before we explore further properties of the skew matrix we now introduce the orthonormal matrix. A matrix $\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \ldots \quad \mathbf{r}_n]$ is called orthonormal if its columns are mutually orthonomal, or equivalently satisfies Eq. (1.9). A three-dimensional orthonormal matrix

$$\mathbf{C} = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{c}_3] \tag{1.23}$$

have these properties

$$\begin{aligned} \mathbf{c}_1 \times \mathbf{c}_2 &= \mathbf{c}_3, \\ \mathbf{c}_2 \times \mathbf{c}_3 &= \mathbf{c}_1, \\ \mathbf{c}_3 \times \mathbf{c}_1 &= \mathbf{c}_2 \end{aligned} \tag{1.24}$$

Lemma 1.2 If **C** is 3×3 orthonormal matrix, then

$$\mathbf{C}\,\mathrm{S}(\mathbf{b})\mathbf{C}' = \mathrm{S}(\mathbf{C}\mathbf{b}) \tag{1.25}$$

Proof: Since

$$\mathbf{C}\tilde{\mathbf{b}}\mathbf{C}' = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix} \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix} \mathbf{C}'$$

Multiplying **C** by $\tilde{\mathbf{b}}$ and expanding $\mathbf{C}'$ yields

$$\mathbf{C}\tilde{\mathbf{b}}\mathbf{C}' = \begin{bmatrix} b_3\mathbf{c}_2 - b_2\mathbf{c}_3 & b_1\mathbf{c}_3 - b_3\mathbf{c}_1 & b_2\mathbf{c}_1 - b_1\mathbf{c}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1' \\ \mathbf{c}_2' \\ \mathbf{c}_3' \end{bmatrix}$$

Multiplying the two right hand side matrices and collecting terms gives

$$\mathbf{C}\tilde{\mathbf{b}}\mathbf{C}' = b_1(\mathbf{c}_3\mathbf{c}_2' - \mathbf{c}_2\mathbf{c}_3') + b_2(\mathbf{c}_1\mathbf{c}_3' - \mathbf{c}_3\mathbf{c}_1') + b_3(\mathbf{c}_2\mathbf{c}_1' - \mathbf{c}_1\mathbf{c}_2')$$

Applying Eq. (1.22) to the above equation yields

$$\mathbf{C}\tilde{\mathbf{b}}\mathbf{C}' = b_1\,\mathrm{S}(\mathbf{c}_2 \times \mathbf{c}_3) + b_2\,\mathrm{S}(\mathbf{c}_3 \times \mathbf{c}_1) + b_3\,\mathrm{S}(\mathbf{c}_1 \times \mathbf{c}_2)$$

Substituting from Eq. (1.24) in the above equation gives

$$\mathbf{C}\tilde{\mathbf{b}}\mathbf{C}' = b_1 \mathrm{S}(\mathbf{c}_1) + b_2 \mathrm{S}(\mathbf{c}_2) + b_3 \mathrm{S}(\mathbf{c}_3)$$

By virtue of the linearity of the S operator it follows that

$$\mathbf{C}\tilde{\mathbf{b}}\mathbf{C}' = \mathrm{S}(b_1\mathbf{c}_1 + b_2\mathbf{c}_2 + b_3\mathbf{c}_3) = \mathrm{S}\left(\begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3\end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right) = \mathrm{S}(\mathbf{Cb})$$

In this chapter we explored the little differences between the vector and matrix notations. In general, a vector represents a direction in the three dimensional Euclidean space and a magnitude. A matrix is a set of elements that are arranged in a specific manner. A vector cross product is a special operation pertains to vectors but can be emulated in matrix notations using the skew symmetric matrix. One advantage with matrix products is their associativity, $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ which is not the case for vector cross products, $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} \neq \mathbf{a}\times(\mathbf{b}\times\mathbf{c})$. Awareness of these distinctions will allow us to move from one notation to the other as desired. In the following chapter, the usefulness of these tools will be very vivid as they allow us to describe vector rotations (that are given in vector notations) in terms of transformation matrices.

References

1. D. E. Bourne, P. C. Kendall, Vector Analysis, Allyn and Bacon Inc., Boston, MA, 1967.
2. R. Bellman, Introduction to Matrix Analysis, McGraw Hill, New York, New York, 1970.
3. R. Larson, R. Hostetler, B. Edwards, Calculus with Analytic Geometry, D. C. Heath and Company, Lexington, Ma, 1994.
4. D. Varberg, E. Purcell, Calculus with Analytic Geometry, Prentice Hall, Englewood Cliffs, New Jersey, 1992.