Chapter 2

Fundamentals of Linear Algebra

This chapter presents fundamentals of linear algebra that will be necessary in subsequent chapters. Also, the symbols and terminologies that will be used throughout this book are defined here. Since the materials presented here are well established facts or their easy derivatives, theorems and propositions are listed without proofs; readers should refer to standard textbooks on mathematics for the details.

2.1 Vector and Matrix Calculus

2.1.1 Vectors and matrices

Throughout this book, geometric quantities such as vectors and tensors are described with respect to a *Cartesian coordinate system*, the coordinate axes being mutually orthogonal and having the same unit of length[1]. We also assume that the coordinate system is *right-handed*[2].

By a *vector*, we mean a column of real numbers[3]. Vectors are denoted by lowercase boldface letters such as $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{u}$, and $\boldsymbol{v}$; their components are written in the corresponding lowercase italic letters. A vector whose components are a_1, a_2, ..., a_n is also denoted by (a_i), $i = 1, ..., n$; the number n of the components is called the *dimension* of this vector. If the dimension is understood, notations such as (a_i) are used. In the following, an n-dimensional vector is referred to as an n-*vector*. The vector whose components are all 0 is called the *zero vector* and denoted by $\mathbf{0}$ (the dimension is usually implied by the context).

A *matrix* is an array of real numbers. Matrices are denoted by uppercase boldface letters such as $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{S}$, and $\boldsymbol{T}$; their elements are written in the corresponding uppercase italic letters. A matrix is also defined by its elements as (A_{ij}), $i = 1, ..., m$, $j = 1, ..., n$; such a matrix is said to be of *type mn*. In the following, a matrix of type mn is referred to as an *mn-matrix*; if $m = n$, it is also called a *square matrix* or simply *n-dimensional matrix*. If the type is

[1] This is only an intuitive definition, since "orthogonality" and "length" are later defined in terms of coordinates. To be strict, we need to start with axioms of one kind or another (we do not go into the details).

[2] In three dimensions, a Cartesian coordinate system is *right-handed* if the x-, y-, and z-axes have the same orientations as the thumb, the forefinger, and the middle finger, respectively, of a right hand. Otherwise, the coordinate system is *left-handed*. In other dimensions, the handedness, or the *parity*, can be defined arbitrarily: if a coordinate system is right-handed, its mirror image is left-handed (we do not go into the details).

[3] We do not deal with complex numbers in this book.

understood, notations such as (A_{ij}) are used. The matrix whose elements are all 0 is called the *zero matrix* and denoted by $\boldsymbol{O}$ (the type is usually implied by the context). If not explicitly stated, the type is understood to be nn in *this* chapter but 33 *in the rest of this book*.

The *unit matrix* is denoted by $\boldsymbol{I}$; its elements are written as δ_{ij} (not I_{ij}); the dimension is usually implied by the context. The symbol δ_{ij}, which takes value 1 for $i = j$ and 0 otherwise, is called the *Kronecker delta*. Addition and subtraction of matrices and multiplication of a matrix by a scalar, vector, or matrix are defined in the standard way.

The *trace* of nn-matrix $\boldsymbol{A} = (A_{ij})$ is the sum $\sum_{i=1}^{n} A_{ii}$ of its diagonal elements and is denoted by $\mathrm{tr}\boldsymbol{A}$. Evidently, $\mathrm{tr}\boldsymbol{I} = n$. The transpose of a vector or matrix is denoted by superscript $\top$. A matrix $\boldsymbol{A}$ is *symmetric* if (nn) $\boldsymbol{A} = \boldsymbol{A}^\top$. We say that a matrix is of *type* (nn) or an (nn)*-matrix* if it is an n-dimensional symmetric matrix. A matrix $\boldsymbol{A}$ is *antisymmetric* (or *skew-* [nn] *symmetric*) if $\boldsymbol{A} = -\boldsymbol{A}^\top$. We say that a matrix is of *type* $[nn]$ or $[nn]$*-matrix* if it is an n-dimensional antisymmetric matrix. Note the following expression, which is sometimes called the *outer product* of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$:

$$\boldsymbol{a}\boldsymbol{b}^\top = (a_i b_j) = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_n \\ \vdots & \vdots & \cdots & \vdots \\ a_n b_1 & a_n b_2 & \cdots & a_n b_n \end{pmatrix}. \tag{2.1}$$

The following identities are very familiar:

$$(\boldsymbol{A}^\top)^\top = \boldsymbol{A}, \qquad (\boldsymbol{A}\boldsymbol{B})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top,$$
$$\mathrm{tr}(\boldsymbol{A}^\top) = \mathrm{tr}\boldsymbol{A}, \qquad \mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A}). \tag{2.2}$$

The *inner product* of vectors $\boldsymbol{a} = (a_i)$ and $\boldsymbol{b} = (b_i)$ is defined by

$$(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a}^\top \boldsymbol{b} = \sum_{i=1}^{n} a_i b_i. \tag{2.3}$$

Evidently, $(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a})$. Vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are said to be *orthogonal* if $(\boldsymbol{a}, \boldsymbol{b}) = 0$. The following identities are easily confirmed:

$$(\boldsymbol{a}, \boldsymbol{T}\boldsymbol{b}) = (\boldsymbol{T}^\top \boldsymbol{a}, \boldsymbol{b}), \qquad \mathrm{tr}(\boldsymbol{a}\boldsymbol{b}^\top) = (\boldsymbol{a}, \boldsymbol{b}). \tag{2.4}$$

The matrix consisting of vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_n$ as its columns in that order is denoted by $(\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n)$. If

$$\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n), \qquad \boldsymbol{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n), \tag{2.5}$$

the following identities hold:

$$\boldsymbol{A}\boldsymbol{B}^\top = \sum_{i=1}^{n} \boldsymbol{a}_i \boldsymbol{b}_i^\top,$$

$$
A^\top B = \begin{pmatrix} (a_1, b_1) & (a_1, b_2) & \cdots & (a_1, b_n) \\ (a_2, b_1) & (a_2, b_2) & \cdots & (a_2, b_n) \\ \vdots & \vdots & \cdots & \vdots \\ (a_n, b_1) & (a_n, b_2) & \cdots & (a_n, b_n) \end{pmatrix}. \tag{2.6}
$$

The *norm*[4] and the *normalization operator* $N[\,\cdot\,]$ are defined as follows:

$$
\|a\| = \sqrt{(a, a)} = \sqrt{\sum_{i=1}^{n} a_i{}^2}, \qquad N[a] = \frac{a}{\|a\|}. \tag{2.7}
$$

A *unit vector* is a vector of unit norm. A set of vectors $\{u_1, ..., u_r\}$ is said to be *orthonormal* if its members are all unit vectors and orthogonal to each other: $(u_i, u_j) = \delta_{ij}$.

The following *Schwarz inequality* holds:

$$
-\|a\| \cdot \|b\| \leq (a, b) \leq \|a\| \cdot \|b\|. \tag{2.8}
$$

Equality holds if vectors a and b are *parallel*, meaning that there exists a real number t such that $a = tb$ or $b = 0$. The Schwarz inequality implies the following *triangle inequality* with the same equality condition:

$$
\|a + b\| \leq \|a\| + \|b\|. \tag{2.9}
$$

2.1.2 Determinant and inverse

The *determinant* of a square matrix $A = (A_{ij})$, denoted by $\det A$ or $|A|$, is defined by

$$
\boxed{\det A = \sum_{i_1, \ldots, i_n = 1}^{n} \epsilon_{i_1 \cdots i_n} A_{1i_1} \cdots A_{ni_n},} \tag{2.10}
$$

(like a tensor)

where $\epsilon_{i_1 \cdots i_n}$ is the *signature symbol* defined by

$$
\epsilon_{i_1 i_2 \cdots i_n} = \begin{cases} 1 & \text{if } (i_1 i_2 \cdots i_n) \text{ is an even permutation of } (12 \cdots n), \\ -1 & \text{if } (i_1 i_2 \cdots i_n) \text{ is an odd permutation of } (12 \cdots n), \\ 0 & \text{otherwise.} \end{cases} \tag{2.11}
$$

Evidently, $\det I = 1$. The following identity holds:

$$
\det(AB) = \det A \det B. \tag{2.12}
$$

[4] This norm is called the *Euclidean norm* (or the *2-norm*). In general, the norm $\|\mathbf{a}\|$ can be defined arbitrarily as long as (i) $\|\mathbf{a}\| \geq 0$, equality holding if and only if $\mathbf{a} = \mathbf{0}$, (ii) $\|c\mathbf{a}\| = |c| \cdot \|\mathbf{a}\|$ for any scalar c, and (iii) the triangle inequality (2.9) holds. There exist other definitions that satisfy these—the *1-norm* $\|\mathbf{a}\|_1 = \Sigma_{i=1}^n |a_i|$ and the ∞-*norm* $\|\mathbf{a}\|_\infty = \max_i |a_i|$, for instance. They can be generalized into the *Minkowski norm* (or the *p-norm*) $\|\mathbf{a}\|_p = \sqrt[p]{\Sigma_{i=1}^n |a_i|^p}$ for $1 \leq p \leq \infty$; the 1-norm, the 2-norm, and the ∞-norm are special cases of the Minkowski norm for $p = 1, 2, \infty$, respectively.

Replacing A_{ij} by $\delta_{ij} + \varepsilon A_{ij}$ in eq. (2.10) and expanding it in ε, we obtain

$$\det(\boldsymbol{I} + \varepsilon \boldsymbol{A}) = 1 + \varepsilon \mathrm{tr}\boldsymbol{A} + O(\varepsilon^2), \tag{2.13}$$

where the *order symbol* $O(\cdots)$ denotes terms having order the same as or higher than $\cdots$.

Let $\boldsymbol{A}^{(ij)}$ be the matrix obtained from a square matrix $\boldsymbol{A} = (A_{ij})$ by removing the ith row and the jth column. The determinant $\det \boldsymbol{A}$ is expanded in the form

$$\det \boldsymbol{A} = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det \boldsymbol{A}^{(ij)} = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det \boldsymbol{A}^{(ij)}. \tag{2.14}$$

This is called the *cofactor expansion formula*. The *cofactor* (or *adjugate*) *matrix* $\boldsymbol{A}^\dagger = (A^\dagger_{ij})$ of $\boldsymbol{A}$ is defined by

$$A^\dagger_{ij} = (-1)^{i+j} \det \boldsymbol{A}^{(ji)}. \tag{2.15}$$

Eq. (2.14) can be rewritten as

$$\boldsymbol{A}\boldsymbol{A}^\dagger = \boldsymbol{A}^\dagger \boldsymbol{A} = (\det \boldsymbol{A})\boldsymbol{I}. \tag{2.16}$$

The following identity holds:

$$\det(\boldsymbol{A} + \varepsilon \boldsymbol{B}) = \det \boldsymbol{A} + \varepsilon \mathrm{tr}(\boldsymbol{A}^\dagger \boldsymbol{B}) + O(\varepsilon^2). \tag{2.17}$$

The elements of the cofactor matrix $\boldsymbol{A}^\dagger$ of nn-matrix $\boldsymbol{A}$ are all polynomials of degree $n-1$ in the elements of $\boldsymbol{A}$. In three dimensions, the cofactor matrix of $\boldsymbol{A} = (A_{ij})$ has the following form:

$$\boldsymbol{A}^\dagger = \begin{pmatrix} A_{22}A_{33} - A_{32}A_{23} & A_{32}A_{13} - A_{12}A_{33} & A_{12}A_{23} - A_{22}A_{13} \\ A_{23}A_{31} - A_{33}A_{21} & A_{33}A_{11} - A_{13}A_{31} & A_{13}A_{21} - A_{23}A_{11} \\ A_{21}A_{32} - A_{31}A_{22} & A_{31}A_{12} - A_{11}A_{32} & A_{11}A_{22} - A_{21}A_{12} \end{pmatrix}. \tag{2.18}$$

The *inverse* $\boldsymbol{A}^{-1}$ of a square matrix $\boldsymbol{A}$ is defined by

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}, \tag{2.19}$$

if such an $\boldsymbol{A}^{-1}$ exists. A square matrix is *singular* if its inverse does not exist, and *nonsingular* (or *of full rank*) otherwise. Eq. (2.16) implies that if $\boldsymbol{A}$ is nonsingular, its inverse $\boldsymbol{A}^{-1}$ is given by

$$\boxed{\boldsymbol{A}^{-1} = \frac{\boldsymbol{A}^\dagger}{\det \boldsymbol{A}}.}$$

This is the reason for doing cofactors.

If we define $\boldsymbol{A}^0 = \boldsymbol{I}$, the following identities hold for nonsingular matrices (k is a nonnegative integer):

$$(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}, \quad (\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, \quad (\boldsymbol{A}^{-1})^k = (\boldsymbol{A}^k)^{-1},$$

$$(\boldsymbol{A}^\top)^{-1} = (\boldsymbol{A}^{-1})^\top, \qquad \det \boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}}. \tag{2.21}$$

The third identity implies that matrix $(\boldsymbol{A}^{-1})^k$ can be unambiguously denoted by $\boldsymbol{A}^{-k}$. Note that the determinant and the inverse are defined only for square matrices.

Let $\boldsymbol{A}$ be a nonsingular nn-matrix, and $\boldsymbol{B}$ a nonsingular mm-matrix. Let $\boldsymbol{S}$ and $\boldsymbol{T}$ be nm-matrices. The following *matrix inversion formula* holds, provided that the inverses involved all exist:

$$\boxed{(\boldsymbol{A} + \boldsymbol{S}\boldsymbol{B}\boldsymbol{T}^\top)^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{S}(\boldsymbol{B}^{-1} + \boldsymbol{T}^\top\boldsymbol{A}^{-1}\boldsymbol{S})^{-1}\boldsymbol{T}^\top\boldsymbol{A}^{-1}.} \tag{2.22}$$

If $m = 1$, the nm-matrices $\boldsymbol{S}$ and $\boldsymbol{T}$ are n-vectors, and the mm-matrix $\boldsymbol{B}$ is a scalar. If we let $\boldsymbol{B} = 1$ and write $\boldsymbol{S}$ and $\boldsymbol{T}$ as $\boldsymbol{s}$ and $\boldsymbol{t}$, respectively, the above formula reduces to

$$(\boldsymbol{A} + \boldsymbol{s}\boldsymbol{t}^\top)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{s}\boldsymbol{t}^\top\boldsymbol{A}^{-1}}{1 + (\boldsymbol{t}, \boldsymbol{A}^{-1}\boldsymbol{s})}. \tag{2.23}$$

useful special cases

For $\boldsymbol{A} = \boldsymbol{I}$, we obtain

$$(\boldsymbol{I} + \boldsymbol{s}\boldsymbol{t}^\top)^{-1} = \boldsymbol{I} - \frac{\boldsymbol{s}\boldsymbol{t}^\top}{1 + (\boldsymbol{s}, \boldsymbol{t})}. \tag{2.24}$$

2.1.3 Vector product in three dimensions

In three dimensions, the signature symbol defined by eq. (2.11) is often referred to as the *Eddington epsilon*[5]. It satisfies the following identity:

$$\sum_{m=1}^{3} \epsilon_{ijm}\epsilon_{klm} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}. \tag{2.25}$$

The *vector* (or *exterior*) *product* of 3-vectors $\boldsymbol{a} = (a_i)$ and $\boldsymbol{b} = (b_i)$ is defined by

cross product

$$\boxed{\boldsymbol{a} \times \boldsymbol{b} = \Big(\sum_{j,k=1}^{3} \epsilon_{ijk} a_j b_k\Big) = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}}. \tag{2.26}$$

Evidently,

$$\begin{gathered} \boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}, \qquad \boldsymbol{a} \times \boldsymbol{a} = \boldsymbol{0}, \\ (\boldsymbol{b}, \boldsymbol{a} \times \boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{a} \times \boldsymbol{b}) = 0. \end{gathered} \tag{2.27}$$

The following identities, known as the *Lagrange formulae*, are direct consequences of eq. (2.25):

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a}, \boldsymbol{b})\boldsymbol{c},$$

[5] Some authors use different terminologies such as the *Levi-Civita symbol*.

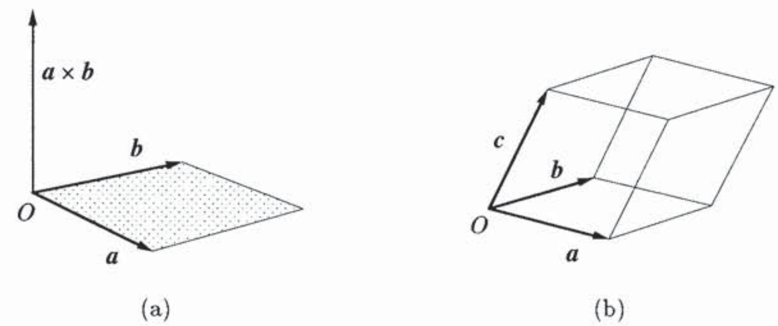


Fig. 2.1. (a) Vector product. (b) Scalar triple product.

$$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = (\boldsymbol{a}, \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{b}, \boldsymbol{c})\boldsymbol{a}. \tag{2.28}$$

The expressions $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ and $(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}$ are called *vector triple products*. The following identities also hold:

$$(\boldsymbol{a} \times \boldsymbol{b}, \boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a}, \boldsymbol{c})(\boldsymbol{b}, \boldsymbol{d}) - (\boldsymbol{a}, \boldsymbol{d})(\boldsymbol{b}, \boldsymbol{c}), \tag{2.29}$$

$$\|\boldsymbol{a} \times \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2 - (\boldsymbol{a}, \boldsymbol{b})^2. \tag{2.30}$$

If 3-vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ make angle θ, we have

$$(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\| \cos\theta, \qquad \|\boldsymbol{a} \times \boldsymbol{b}\| = \|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\| \sin\theta. \tag{2.31}$$

Eq. (2.30) states the well-known trigonometric identity $\cos^2\theta + \sin^2\theta = 1$. From eq. (2.26), the third of eqs. (2.27), and the second of eqs. (2.31), we can visualize $\boldsymbol{a} \times \boldsymbol{b}$ as a vector normal to the plane defined by $\boldsymbol{a}$ and $\boldsymbol{b}$; the length of $\boldsymbol{a} \times \boldsymbol{b}$ equals the area of the parallelogram made by $\boldsymbol{a}$ and $\boldsymbol{b}$ (Fig. 2.1a).

The *scalar triple product* $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}|$ of 3-vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ is the determinant of the matrix $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ having $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ as its columns in that order. We say that three 3-vectors $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ are a *right-handed* system if $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}| > 0$ and a *left-handed* system if $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}| < 0$. The scalar triple product $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}|$ equals the signed volume of the parallelepiped defined by $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ (Fig. 2.1b); the volume is positive if the three vectors are a right-handed system in that order and negative if they are a left-handed system. The equality $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}| = 0$ holds if and only if $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are *coplanar*, i.e., if they all lie on a common plane.

We can also write

$$|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}| = (\boldsymbol{a} \times \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{b} \times \boldsymbol{c}, \boldsymbol{a}) = (\boldsymbol{c} \times \boldsymbol{a}, \boldsymbol{b}). \tag{2.32}$$

Since $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} \times \boldsymbol{b}| = \|\boldsymbol{a} \times \boldsymbol{b}\|^2$, the vector product $\boldsymbol{a} \times \boldsymbol{b}$ is oriented, if it is not $\boldsymbol{0}$, in such a way that $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} \times \boldsymbol{b}\}$ form a right-handed system (Fig. 2.1a).

The following identity also holds:

$$(\boldsymbol{a} \times \boldsymbol{b}) \times (\boldsymbol{c} \times \boldsymbol{d}) = |\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{d}|\boldsymbol{c} - |\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}|\boldsymbol{d} = |\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{d}|\boldsymbol{b} - |\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}|\boldsymbol{a}. \tag{2.33}$$

Taking the determinant of $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})^\top$ (see eq. (2.12)), we obtain

$$|\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}|^2 = \begin{vmatrix} \|\boldsymbol{a}\|^2 & (\boldsymbol{a},\boldsymbol{b}) & (\boldsymbol{a},\boldsymbol{c}) \\ (\boldsymbol{b},\boldsymbol{a}) & \|\boldsymbol{b}\|^2 & (\boldsymbol{b},\boldsymbol{c}) \\ (\boldsymbol{c},\boldsymbol{a}) & (\boldsymbol{c},\boldsymbol{b}) & \|\boldsymbol{c}\|^2 \end{vmatrix}. \tag{2.34}$$

The *vector* (or *exterior*) *product* of 3-vector $\boldsymbol{a}$ and 33-matrix $\boldsymbol{T} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3)$ is defined by

$$\boxed{\boldsymbol{a} \times \boldsymbol{T} = (\boldsymbol{a} \times \boldsymbol{t}_1, \boldsymbol{a} \times \boldsymbol{t}_2, \boldsymbol{a} \times \boldsymbol{t}_3).}$$

cross product with matrices: handy to have for non-GA people

From this definition, the following identities are obtained:

$$\boldsymbol{a} \times (\boldsymbol{T}\boldsymbol{b}) = (\boldsymbol{a} \times \boldsymbol{T})\boldsymbol{b},$$

a^{x}

$$\boxed{\boldsymbol{a} \times \boldsymbol{I} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}}, \qquad (\boldsymbol{a} \times \boldsymbol{I})^\top = -\boldsymbol{a} \times \boldsymbol{I}. \tag{2.36}$$

The matrix $\boldsymbol{a} \times \boldsymbol{I}$ is called the *antisymmetric matrix associated with* the 3-vector $\boldsymbol{a}$. The following identity is an alternative expression to the Lagrange formulae (2.28):

$$(\boldsymbol{a} \times \boldsymbol{I})(\boldsymbol{b} \times \boldsymbol{I})^\top = (\boldsymbol{a},\boldsymbol{b})\boldsymbol{I} - \boldsymbol{b}\boldsymbol{a}^\top. \tag{2.37}$$

The vector (or exterior) product of 33-matrix $\boldsymbol{T}$ and 3-vector $\boldsymbol{b}$ is defined by

$$\boxed{\boldsymbol{T} \times \boldsymbol{b} = \boldsymbol{T}(\boldsymbol{b} \times \boldsymbol{I})^\top} \tag{2.38}$$

This definition implies the following identities: for future reference...

$$(\boldsymbol{a} \times \boldsymbol{T})^\top = \boldsymbol{T}^\top \times \boldsymbol{a}, \qquad (\boldsymbol{T} \times \boldsymbol{b})^\top = \boldsymbol{b} \times \boldsymbol{T}^\top,$$

$$(\boldsymbol{T} \times \boldsymbol{b})\boldsymbol{c} = \boldsymbol{T}(\boldsymbol{c} \times \boldsymbol{b}). \tag{2.39}$$

It is easy to confirm that

$$(\boldsymbol{a} \times \boldsymbol{T}) \times \boldsymbol{b} = \boldsymbol{a} \times (\boldsymbol{T} \times \boldsymbol{b}), \tag{2.40}$$

which can be written unambiguously as $\boldsymbol{a} \times \boldsymbol{T} \times \boldsymbol{b}$. We also have

$$(\boldsymbol{a} \times \boldsymbol{T} \times \boldsymbol{b})^\top = \boldsymbol{b} \times \boldsymbol{T}^\top \times \boldsymbol{a}. \tag{2.41}$$

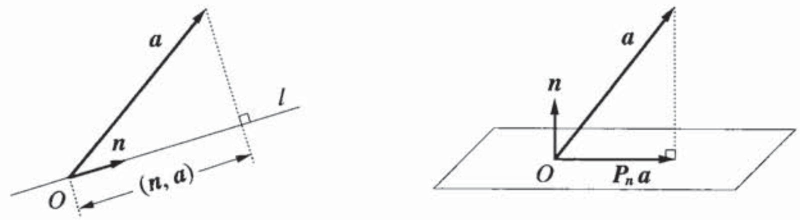
Eq. (2.37) now reads

$$\boldsymbol{a} \times \boldsymbol{I} \times \boldsymbol{b} = (\boldsymbol{a},\boldsymbol{b})\boldsymbol{I} - \boldsymbol{b}\boldsymbol{a}^\top. \tag{2.42}$$

The following identities are also important:

$$(\boldsymbol{a} \times \boldsymbol{b})(\boldsymbol{c} \times \boldsymbol{d})^\top = \boldsymbol{a} \times (\boldsymbol{b}\boldsymbol{d}^\top) \times \boldsymbol{c} = \boldsymbol{b} \times (\boldsymbol{a}\boldsymbol{c}^\top) \times \boldsymbol{d}, \tag{2.43}$$

$$(\boldsymbol{a} \times \boldsymbol{b}, \boldsymbol{T}(\boldsymbol{c} \times \boldsymbol{d})) = (\boldsymbol{a}, (\boldsymbol{b} \times \boldsymbol{T} \times \boldsymbol{d})\boldsymbol{c}) = (\boldsymbol{b}, (\boldsymbol{a} \times \boldsymbol{T} \times \boldsymbol{c})\boldsymbol{d}). \tag{2.44}$$



(a) (b)

Fig. 2.2. (a) Projection onto a line. (b) Projection onto a plane.

The *exterior product* $[\boldsymbol{A} \times \boldsymbol{B}]$ of 33-matrices $\boldsymbol{A} = (A_{ij})$ and $\boldsymbol{B} = (B_{ij})$ is a 33-matrix defined as follows[6]:

$$[\boldsymbol{A} \times \boldsymbol{B}]_{ij} = \sum_{k,l,m,n=1}^{3} \epsilon_{ikl}\epsilon_{jmn}A_{km}B_{ln}. \tag{2.45}$$

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both symmetric, their exterior product $[\boldsymbol{A} \times \boldsymbol{B}]$ is also symmetric.

2.1.4 *Projection matrices*

If a vector $\boldsymbol{a}$ is projected orthogonally onto a line l that extends along a unit vector $\boldsymbol{n}$, it defines on l a segment of signed length $(\boldsymbol{n}, \boldsymbol{a})$ (Fig. 2.2a); it is positive in the direction $\boldsymbol{n}$ and negative in the direction $-\boldsymbol{n}$. The vector $\boldsymbol{a}$ is decomposed into the component $(\boldsymbol{n}, \boldsymbol{a})\boldsymbol{n}$ parallel to l and the component $\boldsymbol{a} - (\boldsymbol{n}, \boldsymbol{a})\boldsymbol{n}$ $(= (\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^\top)\boldsymbol{a})$ orthogonal it. Let $\{\boldsymbol{n}\}_L$ be the one-dimensional subspace defined by unit vector $\boldsymbol{n}$, and $\{\boldsymbol{n}\}_L^\perp$ its *orthogonal complement*—the set of all vectors orthogonal to $\boldsymbol{n}$. The projection of a vector $\boldsymbol{a}$ onto $\{\boldsymbol{n}\}_L^\perp$ is written as $\boldsymbol{P_n a}$ (Fig. 2.2b). The matrix $\boldsymbol{P_n}$ is defined by

projection on (dual) plane ***n***

$$\boxed{\boldsymbol{P_n} = \boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^\top} \tag{2.46}$$

and called the *projection matrix* onto the plane orthogonal to $\boldsymbol{n}$, or the projection matrix *along* $\boldsymbol{n}$. The following identities are easily confirmed:

$$\boldsymbol{P_n} = \boldsymbol{P_n}^\top, \qquad \boldsymbol{P_n}^2 = \boldsymbol{P_n},$$
$$\det \boldsymbol{P_n} = 0, \quad \mathrm{tr}\boldsymbol{P_n} = n - 1, \quad \|\boldsymbol{P_n}\| = \sqrt{n-1}. \tag{2.47}$$

Here, the matrix norm $\|\cdot\|$ is defined by $\|\boldsymbol{A}\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}{}^2}$ for mn-matrix $\boldsymbol{A} = (A_{ij})$. In three dimensions, eq. (2.42) implies the following identity for unit vector $\boldsymbol{n}$:

$$\boldsymbol{n} \times \boldsymbol{I} \times \boldsymbol{n} = (\boldsymbol{n} \times \boldsymbol{I})(\boldsymbol{n} \times \boldsymbol{I})^\top = \boldsymbol{P_n}. \tag{2.48}$$

[6]For example, $[\mathbf{A} \times \mathbf{B}]_{11} = A_{22}B_{33} - A_{32}B_{23} - A_{23}B_{32} + A_{33}B_{22}$.

The projection matrix can be generalized as follows. Let the symbol $\mathcal{R}^n$ denote the n-dimensional space of all n-vectors. Let $\mathcal{S}$ be an m-dimensional subspace of $\mathcal{R}^n$, and $\mathcal{N}$ $(= \mathcal{S}^\perp)$ its orthogonal complement—the set of all vectors that are orthogonal to every vector in $\mathcal{S}$. The *orthogonal projection*[7] $\boldsymbol{P}_\mathcal{N}$ onto $\mathcal{S}$ is a linear mapping such that for an arbitrary vector $\boldsymbol{v} \in \mathcal{R}^n$

$$\boldsymbol{P}_\mathcal{N}\boldsymbol{v} \in \mathcal{S}, \qquad \boldsymbol{v} - \boldsymbol{P}_\mathcal{N}\boldsymbol{v} \in \mathcal{N}. \tag{2.49}$$

In other words, $\boldsymbol{P}_\mathcal{N}$ is the operator that *removes* the component in $\mathcal{N}$. We also use an alternative notation $\boldsymbol{P}^\mathcal{S}$ when we want to indicate the space to be projected explicitly. Let $\{\boldsymbol{n}_1, ..., \boldsymbol{n}_m\}$ be an orthonormal basis of $\mathcal{N}$. The orthogonal projection $\boldsymbol{P}_\mathcal{N}$ has the following matrix expression:

$$\boxed{\boldsymbol{P}_\mathcal{N} = \boldsymbol{I} - \sum_{i=1}^{m} \boldsymbol{n}_i\boldsymbol{n}_i^\top.} \tag{2.50}$$

Eqs. (2.47) can be generalized as follows:

$$\boldsymbol{P}_\mathcal{N} = \boldsymbol{P}_\mathcal{N}^\top, \qquad \boldsymbol{P}_\mathcal{N}^2 = \boldsymbol{P}_\mathcal{N},$$

$$\det \boldsymbol{P}_\mathcal{N} = 0, \quad \mathrm{tr}\boldsymbol{P}_\mathcal{N} = n - m, \quad \|\boldsymbol{P}_\mathcal{N}\| = \sqrt{n-m}. \tag{2.51}$$

2.1.5 Orthogonal matrices and rotations

Matrix $\boldsymbol{R}$ is *orthogonal* if one of the following conditions holds (all are equivalent to each other):

$$\boxed{\boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{I}, \quad \boldsymbol{R}^\top\boldsymbol{R} = \boldsymbol{I}, \quad \boldsymbol{R}^{-1} = \boldsymbol{R}^\top.} \tag{2.52}$$

Equivalently, matrix $\boldsymbol{R} = (\boldsymbol{r}_1, ..., \boldsymbol{r}_n)$ is orthogonal if and only if its columns form an orthonormal set of vectors: $(\boldsymbol{r}_i, \boldsymbol{r}_j) = \delta_{ij}$.

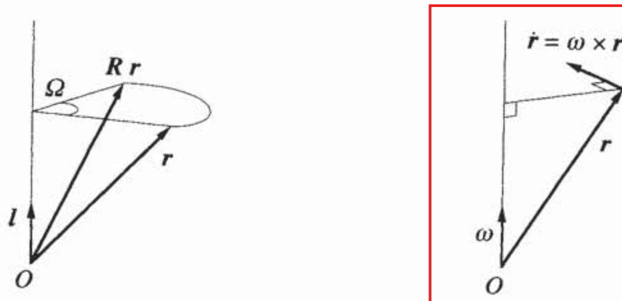
For an orthogonal matrix $\boldsymbol{R}$ and vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we have

$$(\boldsymbol{Ra}, \boldsymbol{Rb}) = (\boldsymbol{a}, \boldsymbol{b}), \qquad \|\boldsymbol{Ra}\| = \|\boldsymbol{a}\|. \tag{2.53}$$

The second equation implies that the length of a vector is unchanged after multiplication by an orthogonal matrix. The first one together with eqs. (2.31) implies that in three dimensions the angle that two vectors make is also unchanged.

Applying eq. (2.12) to eqs. (2.52), we see that $\det \boldsymbol{R} = \pm 1$ for an orthogonal matrix $\boldsymbol{R}$. If $\det \boldsymbol{R} = 1$, the orthogonal matrix $\boldsymbol{R}$ is said to be a *rotation*

[7]The notation given here is non-traditional: the projection onto subspace $\mathcal{S}$ is usually denoted by $\mathbf{P}_\mathcal{S}$. Our definition is in conformity to the notation $\mathbf{P_n}$ given by eq. (2.46).



(a) (b)

Fig. 2.3. (a) Axis and angle of rotation. (b) Instantaneous rotation.

matrix[8]. In three dimensions, the orthonormal Cartesian coordinate basis vectors are

$$\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.54}$$

The columns of a three-dimensional rotation matrix $\boldsymbol{R} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3)$ define a right-handed orthonormal system $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3\}$. The matrix $\boldsymbol{R}$ maps the coordinate basis $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ to $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3\}$. Such a map is realized as a rotation along an *axis* $\boldsymbol{l}$ by an *angle* Ω of rotation (*Euler's theorem*; Fig. 2.3a). The axis $\boldsymbol{l}$ (unit vector) and the angle Ω (measured in the screw sense) of rotation $\boldsymbol{R}$ are computed as follows:

$$\boldsymbol{l} = N[\begin{pmatrix} R_{32} - R_{23} \\ R_{13} - R_{31} \\ R_{21} - R_{12} \end{pmatrix}], \qquad \boxed{\Omega = \cos^{-1} \frac{\mathrm{tr}\boldsymbol{R} - 1}{2}.} \tag{2.55}$$

Conversely, an axis $\boldsymbol{l}$ and an angle Ω define a rotation $\boldsymbol{R}$ in the following form:

$$\boldsymbol{R} = \begin{pmatrix} \cos\Omega + {l_1}^2(1-\cos\Omega) & l_1 l_2(1-\cos\Omega) - l_3\sin\Omega & \cdot \\ l_2 l_1(1-\cos\Omega) + l_3\sin\Omega & \cos\Omega + {l_2}^2(1-\cos\Omega) & \cdot \\ \cdot & \cdot & \cos\Omega + {l_3}^2(1-\cos\Omega) \end{pmatrix}. \tag{2.56}$$

Rodrigues: R = I cosQ + (1-cosQ) nn^T + sinQ n^x
= I + sinQ n^x + (1-cosQ) $(n^x)^2$

From this equation, we see that a rotation around unit vector $\boldsymbol{l}$ by a small angle $\Delta\Omega$ is expressed in the form

$$\boldsymbol{R} = \boldsymbol{I} + \Delta\Omega \boldsymbol{l} \times \boldsymbol{I} + O(\Delta\Omega^2), \tag{2.57}$$

[8] The set of all n-dimensional rotation matrices forms a group, denoted by $SO(n)$, under matrix multiplication. It is a subgroup of $O(n)$, the group consisting of all n-dimensional orthogonal matrices. The group consisting of all nonsingular nn-matrices is denoted by $GL(n)$, and the group consisting of all nn-matrices of determinant 1 is denoted by $SL(n)$.