Chapter 1

The Euclidean Space

The objects of study in advanced calculus are differentiable functions of several variables. To set the stage for the study, the Euclidean space as a vector space endowed with the dot product is defined in Section 1.1. To aid visualizing points in the Euclidean space, the notion of a vector is introduced in Section 1.2. In Section 1.3 Euclidean motions, mappings preserving the Euclidean distance, are briefly discussed. The last Section 1.4 contains a discussion on the cross product which is only defined for vectors in the three dimensional Euclidean space, that is, our physical space.

1.1 The Dot Product

An n-tuple is given by

$$\mathbf{x} = (x_1, x_2, \cdots, x_n) \ , \quad x_j \in \mathbb{R}, \quad j = 1, \cdots, n \ .$$

It is called an ordered pair when $n = 2$. Denote by $\mathbb{R}^n$ the collection of all n-tuples. The zero n-tuple, $(0, 0, \cdots, 0)$, will be written as $\mathbf{0}$ from time to time. There are two algebraic operations defined on $\mathbb{R}^n$, namely, the addition

$$\mathbf{x} + \mathbf{y} \ = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) \ ,$$

and the scalar multiplication

$$\alpha\mathbf{x} \ = (\alpha x_1, \alpha x_2, \cdots, \alpha x_n) \ , \quad \alpha \in \mathbb{R} \ ,$$

where

$$\mathbf{x} = (x_1, x_2, \cdots, x_n), \quad \mathbf{y} = (y_1, y_2, \cdots, y_n) \ .$$

Recall that the ordinary multiplication assigns a number as the product of two numbers, so it can be regarded as a map from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$. One may expect a multiplication on $\mathbb{R}^n$

assigns an n-tuple to a given pair of n-tuples, that is, it is a map from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$. But here the scalar multiplication is not such a multiplication, instead it maps $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$.

From linear algebra we know that with these two operations, addition and scalar multiplication, $\mathbb{R}^n$ becomes a vector space of dimension n over the field of real numbers. The so-called canonical basis of $\mathbb{R}^n$ is given by

$$\mathbf{e}_1 = (1, 0, 0, \cdots, 0) \ , \ \mathbf{e}_2 = (0, 1, 0, \cdots, 0) \ , \cdots , \ \mathbf{e}_n = (0, 0, 0, \cdots, 1) \ .$$

Using this basis every n-tuple can be written as the linear combination of the basis elements in a very simple way,

$$\begin{aligned} \mathbf{x} &= (x_1, x_2, \cdots, x_n) \\ &= \sum_{j=1}^{n} x_j e_j \\ &= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n \ . \end{aligned}$$

For instance, the point $(2, -3, 6)$ in $\mathbb{R}^3$ is equal to

$$\begin{aligned} (2, -3, 6) &= 2(1, 0, 0) - 3(0, 1, 0) + 6(0, 0, 1) \\ &= 2\mathbf{e}_1 - 3\mathbf{e}_2 + 6\mathbf{e}_3 \ . \end{aligned}$$

In lower dimensions $n = 2, 3$, the notations **i, j, k** are used instead of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in some texts. We will not adopt them here though.

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, the **dot product** between $\mathbf{x}$ and $\mathbf{y}$ is defined by

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= \sum_{j=1}^{n} x_j y_j \\ &= x_1 y_1 + x_2 y_2 + \cdots x_n y_n \ . \end{aligned}$$

Recall that the axioms for an inner product on a vector space V over reals are: For $u, v, w, y, z \in V, \ \alpha, \beta \in \mathbb{R}$,

(a) $\langle u, u \rangle \geq 0$ and equals to 0 iff $u = 0$,

(b) $\langle u, v \rangle = \langle v, u \rangle$,

(c) $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$.

Note that (b) and (c) imply

(d) $\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$.

One has no difficulty in verifying the dot product satisfies these three axioms. Alternatively one may use $\langle \mathbf{x}, \mathbf{y} \rangle$ to denote $\mathbf{x}\cdot\mathbf{y}$. We will do this to avoid confusion occasionally. Note that $\mathbf{x}\cdot\mathbf{y} = \mathbf{y}\cdot\ \mathbf{x}$, so the dot product between $\mathbf{x}$ and $\mathbf{y}$ is the same as the dot product between $\mathbf{x}$ and $\mathbf{y}$. Things would be very different when we study the cross product later.

At this point let us make a digression to establish a fundamental inequality. Some of you may already learn this inequality, but its interesting proof is worth to go through once more.

Theorem 1.1 (**Cauchy-Schwarz Inequality**). *For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,*

$$\left|\sum_{j=1}^{n} x_j y_j\right| \leq \sqrt{\sum_{j=1}^{n} x_j^2}\sqrt{\sum_{j=1}^{n} y_j^2}\ .$$

Furthermore, equality sign holds if and only if either one of $\boldsymbol{x}, \boldsymbol{y}$ is zero n-tuple or there is some $\alpha \neq 0$ such that $\boldsymbol{y} = \alpha\boldsymbol{x}$.

The condition $\mathbf{x} = \alpha\mathbf{y}$ means $\mathbf{x}$ and $\mathbf{y}$ are proportional to each other. Using the language of linear algebra, the equality condition is simply $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

Proof. First assume not all x_j's are zero in $\mathbf{x}$. Consider the expression

$$\sum_{j=1}^{n}(x_j t - y_j)^2\ ,$$

which is a sum of squares and so must be non-negative for all $t \in \mathbb{R}$. We can express it as a quadratic polynomial in t as

$$p(t) \equiv at^2 - 2bt + c\ ,$$

where

$$a = \sum_{j=1}^{n} x_j^2\ , \quad b = \sum_{j=1}^{n} x_j y_j, \quad c = \sum_{j=1}^{n} y_j^2\ .$$

Since $a > 0$, $p(t)$ tends to ∞ as $t \to \pm\infty$. Therefore, it is non-negative if and only if its discriminant is non-positive, that is, $4b^2 - 4ac \leq 0$, which yields $|b| \leq \sqrt{ac}$ after taking square root. Our inequality follows. Moreover, the equality sign holds if and only if $4b^2 - 4ac = 0$. In this case the quadratic equation $at^2 - 2bt + c = 0$ has a (double) root, say, t_1. Going back to the original expression, we have

$$\sum_{j=1}^{n}(x_j t_1 - y_j)^2 = 0\ ,$$

which forces $t_1 x_j = y_j$ for all $j = 1, \cdots, n$. So we can take $\alpha = t_1$ in case $c = \sum_j y_j^2 > 0$.

When all x_j's vanish but not all y_j's, we exchange $\mathbf{x}$ and $\mathbf{y}$ to get the same conclusion.

Finally, when all x_j's and y_j's vanish, the inequality clearly holds.

$\square$

Accompanying with the notion of the inner product are those of the norm and the distance. Indeed, the **Euclidean norm** of an n-tuple is defined to be

$$\begin{aligned} |\mathbf{x}| &= (\mathbf{x}\cdot\mathbf{x})^{1/2} \\ &= \left(\sum_{j=1}^n x_j^2\right)^{1/2} \\ &= \sqrt{x_1^2 + x_2^2 + \cdots x_n^2}\ . \end{aligned}$$

The **Euclidean distance** between $\mathbf{x}$ and $\mathbf{y}$ is defined by

$$|\mathbf{x}-\mathbf{y}| = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2}\ .$$

In terms of these notions, Cauchy-Schwarz Inequality can be rewritten in a compact form

$$|\mathbf{x}\cdot\mathbf{y}| \le |\mathbf{x}||\mathbf{y}|\ .$$

In mathematics, a distance is a rule to assign a non-negative number to any pair of elements in a set under consideration. The rule consists of three "axioms": For a, b, c in this set,

(i) $d(a,b) \ge 0$, and equal to 0 iff $a = b$,

(ii) $d(a,b) = d(b,a)$, and

(iii) $d(a,b) \le d(a,c) + d(c,b)$.

Now, taking $d(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}|$, we see that it satisfies all these three axioms: $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$,

(a) $|\mathbf{x}-\mathbf{y}| \ge 0$ and equal to 0 if and only if $\mathbf{x} = \mathbf{y}$,

(b) $|\mathbf{x}-\mathbf{y}| = |\mathbf{y}-\mathbf{x}|$,

(c) $|\mathbf{x}-\mathbf{y}| \le |\mathbf{x}-\mathbf{z}| + |\mathbf{z}-\mathbf{y}|$.

Indeed, (a) and (b) are obvious. To prove (c), write $\mathbf{u} = \mathbf{x} - \mathbf{z}, \mathbf{v} = \mathbf{z} - \mathbf{y}$ to get

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}| \ ,$$

which holds after taking square of both sides and then applying the Cauchy-Schwarz Inequality. I let you work out the details in the exercise. Note that $|\mathbf{x}| = |\mathbf{x} - \mathbf{0}|$, so the norm of $\mathbf{x}$ is its distance to the zero n-tuple. In these notes, norm and distance are referred to Euclidean norm and Euclidean distance without further specification.

Note that the same notation $|x|$ stands for the absolute value of x when x is a real number. The norm of $\mathbf{x}$ is the same as its absolute value when $n = 1$. When $n \geq 2$, the notation $|\mathbf{x}|$ stands for the norm only as there is no such a definition of absolute value for an n-tuple. The notation $\|\mathbf{x}\|$ is also used to denote the norm of $\mathbf{x}$, but it will not be used here.

Recall that the cosine function $\cos t$ is strictly decreasing from 1 to -1 as t goes from 0 to π. Keeping this in mind, we are going to define the angle between two non-zero n-tuples. By Cauchy-Schwarz Inequality, the absolute value of the expression $\mathbf{x} \cdot \mathbf{y}/|\mathbf{x}||\mathbf{y}|$ lies in the interval $[-1, 1]$. Therefore, by what we just said, there exists a unique $\theta \in [0, \pi]$ satisfying $\cos\theta = \mathbf{x} \cdot \mathbf{y}/|\mathbf{x}||\mathbf{y}|$, that is,

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos\theta \ .$$

We define the **angle** between two non-zero n-tuples $\mathbf{x}$ and $\mathbf{y}$ to be θ where $\mathbf{x}, \mathbf{y}$ are arbitrary n-tuples. The angle between two n-tuples makes no sense when of them is zero. By definition this angle must belong to $[0, \pi]$. Moreover, it is symmetric, that is, the angle between $\mathbf{x}$ and $\mathbf{y}$ is the same as the angle between $\mathbf{y}$ and $\mathbf{x}$. At this stage, the notion of an angle is defined purely in an analytical manner and does not bear any geometric meaning. We will link it to geometry in the next section.

Two n-tuples $\mathbf{x}$ and $\mathbf{y}$ are **perpendicular** or **orthogonal** to each other if $\mathbf{x} \cdot \mathbf{y} = 0$. In terms of the angle, they are perpendicular if and only if their angle is $\pi/2$. The zero n-tuple is perpendicular to all n-tuples. By Cauchy-Schwarz Inequality, we also know that two non-zero $\mathbf{x}$ and $\mathbf{y}$ satisfy $\mathbf{x} = c\mathbf{y}$ for some $c > 0$ when their angle $\theta = 0$ and satisfy $\mathbf{x} = c\mathbf{y}, c < 0$ when $\theta = \pi$.

Example 1.1. Find all n-tuples $\mathbf{x}$ that are perpendicular to $(1, -1, 2)$ and $(-1, 0, 3)$. These points satisfy

$$(1, -1, 2) \cdot \mathbf{x} = 0 \ , \quad (-1, 0, 3) \cdot \mathbf{x} = 0 \ ,$$

that is, the linear system

$$\begin{cases} x - y + 2z &= 0 \ , \\ -x \qquad + 3z &= 0 \ . \end{cases}$$

We solve this system (see Comments at the end of this chapter) to get $\mathbf{x} = (x, y, z) = a(3, 5, 1), a \in \mathbb{R}$. By varying a, we obtain infinitely many solutions.

Summing up, we have defined the Euclidean space $(\mathbb{R}^n, +, \cdot, \langle \cdot, \cdot \rangle)$ which is the set of all n-tuples with two algebraic operations— the addition and the scalar multiplication— as well as the dot product. From now on it will be abbreviated in a single symbol $\mathbb{R}^n$.

1.2 Vector Representation

Visualizing n-tuples for $n = 2, 3$ as vectors has been used widely in physics and engineering. In this section we discuss how to do it. However, despite its convenience and usefulness, it should be understood that the notion of a vector is only an auxiliary tool. Analysis on the Euclidean space can be carried out without referring to this notion.

To start off, imagine that the coordinates axes have been introduced in the plane. The x-axis consists of all ordered pairs of the form $(x, 0),\ x \in \mathbb{R}$, and the y-axis all ordered pairs of the form $(0, y),\ y \in \mathbb{R}$. So every ordered pair (x, y) can be written as $x(1, 0) + y(0, 1)$ and (x, y) is a point in the coordinate plane. Points in the plane are in one-to-one correspondence with ordered pairs. The same idea applies to all other dimensions. However, since the two dimensional case is easy to see and the three dimensional case can be seen with some imagination, in the following we will focus on these two spaces. It will be apparent that most of our discussions can be extended to all dimensions.

For a point (x, y) in the plane, it associates to a **vector** which is an arrow pointing from the base $(0, 0)$ to the tip (x, y). The vector degenerates into a point when it is the zero ordered pair $(0, 0)$. We call it the **zero vector** and denoted it by $(0, 0)$ or simply $\mathbf{0}$.

After the introduction of the vector for an ordered pair, we interpret the algebraic operations of $\mathbb{R}^2$ as follows. Indeed, by drawing pictures, it is not hard to convince oneself that the addition of two ordered pairs is accomplished by the parallelogram law. Specifically, first form the parallelogram using the two vectors as line segments corresponding to the two points. Then the diagonal of this resulting parallelogram, regarded as a vector pointing from the origin to the other end, is the sum of these two vectors. The same situation holds in $n \geq 3$, as one can always restrict to the plane containing these two vectors provided they are linearly independent. When they are linearly dependent, the geometric interpretation is apparent.

The scalar multiplication $(x, y) \mapsto \alpha(x, y), \alpha > 0$, means changing the vector by a scale of α along the same direction. It is a prolongation if $\alpha > 1$ and a shortening if $\alpha \in (0, 1)$.

On the other hand, when $\alpha < 0$, the resulting vector points in the opposite direction of the original vector with size changes equal to $|\alpha|$.

How about substraction of two vectors? Let $\mathbf{w} = \mathbf{v} - \mathbf{u}$. Then $\mathbf{w}$ can be obtained by first drawing the triangle with vertices at $(0,0), \mathbf{u}$ and $\mathbf{v}$ and then translate the side from $\mathbf{u}$ to $\mathbf{v}$ so that its base is located at the origin. The translated side is $\mathbf{w}$.

We may also find the midpoint of two ordered pairs. For $\mathbf{u}, \mathbf{v}$, its midpoint is given by $(\mathbf{u}+\mathbf{v})/2$. Regarding as a vector, this midpoint can be described in the following way. First draw the parallelogram formed by $\mathbf{u}$ and $\mathbf{v}$. Then the intersection point of the two diagonal lines of this parallelogram is the tip of the midpoint (vector).

When we regard an n-tuple $\mathbf{x}$ as a vector, it is more convenient to call its norm the **magnitude** of the vector. It is a **unit vector** if its magnitude is equal to 1. Likewise, the distance between two points may be called the **length** of the line segment connecting $\mathbf{x}$ and $\mathbf{y}$. It is consistent with the classical Pythagoras theorem. In fact, the definition of the Euclidean norm and distance were inspired by this classical theorem.

Next we show that the angle defined in the last section, which purely depends on analytical terms, is the same as the "geometric angle". To see it, let $\mathbf{x} = (a, b)$ and $\mathbf{y} = (c, d)$ be two non-zero vectors in the plane. By the Law of Cosines in trigonometry (see Comments at the end of this chapter),

$$(c-a)^2 + (d-b)^2 = (a^2+b^2) + (c^2+d^2) - 2\sqrt{c^2+d^2}\sqrt{a^2+b^2}\cos\phi \ ,$$

where $\phi \in [0, \pi]$ is the "geometric angle" between $\mathbf{x}$ and $\mathbf{y}$. Simplifying, we have

$$-2(ac+db) = -2\sqrt{c^2+d^2}\sqrt{a^2+b^2}\cos\phi \ ,$$

which is equal to

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}|\cos\phi \ .$$

Comparing with the definition of θ, we have $\cos\phi = \cos\theta$ so that $\phi = \theta$. In other words, the geometric angle coincides with the analytical angle. The same argument works in higher dimensions as we can restrict to the plane containing any two given vectors.

A vector is uniquely determined by its magnitude and direction. To be more precise we fix them in definition. Any vector with unit length is called a **direction.** Each non-zero vector $\mathbf{x}$ can be written as

$$\mathbf{x} = |\mathbf{x}| \ \boldsymbol{\xi},$$

where $|\mathbf{x}|$ is its magnitude and

$$\boldsymbol{\xi} = \frac{\mathbf{x}}{|\mathbf{x}|}$$

its direction. Every direction $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$ can further be expressed as

$$\boldsymbol{\xi} = (\cos\alpha_1, \cos\alpha_2, \cdots, \cos\alpha_n) \ ,$$

where $\alpha_k \in [0, \pi]$ are called the **direction angles** of $\boldsymbol{\xi}$. From $\boldsymbol{\xi}\cdot\mathbf{e}_k = \cos\alpha_k$ we see that α_k is the angle between $\boldsymbol{\xi}$ and the e_k-axis. These $\cos\alpha_k$'s are called the **direction cosines** of $\mathbf{x}$.

Example 1.2. Find the magnitude and direction of $(1, 2, -7)$ and determine the vector $(2, a, 6)$ that is perpendicular to $(1, 2, -7)$. The magnitude of $(1, 2, -7)$ is

$$|(1, 2, -7)| = \sqrt{1^2 + 2^2 + (-7)^2} = \sqrt{54} \ ,$$

and its direction is $(1, 2, -7)/\sqrt{54}$. By orthogonality,

$$0 = (1, 2, -7) \cdot (2, a, 6) = 2 + 2a - 42 = 0 \ ,$$

which implies $a = 20$. The vector $(2, 20, 6)$ is perpendicular to $(1, 2, -7)$.

One may also consider the **vector from the initial point x to the terminal point y,** or **the vector based at some point**. Unlike a vector, a vector from $\mathbf{x}$ to $\mathbf{y}$ is an arrow whose base and tip are $\mathbf{x}$ and $\mathbf{y}$ respectively. Obviously such a "vector" is parallel to the position vector of $\mathbf{y} - x$ whose base is now at the origin. The length and direction of a vector from $\mathbf{x}$ to $\mathbf{y}$ are defined as the respective length and direction of $\mathbf{y}-\mathbf{x}$.

Example 1.3. Consider the triangle with vertices at $(1, 2), (3, 4), (0, -1)$. Find the direction of the vector pointing at the midpoint of the side connecting $(1, 2)$ and $(3, 4)$ from $(0, -1)$. Well, first we translate $(0, -1)$ to the origin so that the triangle is congruent to the one whose vertices are $((1, 2) - (0, -1), (3, 4) - (0, -1), (0, -1) - (0, -1)$, that is, $(1, 3), (3, 5), (0, 0)$. The midpoint of the side from $(1, 3)$ and $(3, 5)$ is given by

$$\frac{1}{2}\left((1, 3) + (3, 5)\right) = (2, 4) \ ,$$

and its direction is given by

$$\frac{(2, 4)}{\sqrt{2^2 + 4^2}} = \frac{(2, 4)}{\sqrt{20}} = \frac{(1, 2)}{\sqrt{5}} \ .$$

(No need to simplify further.)

Example 1.4. (a) Find the magnitude and direction of the vector from $(1, -1)$ to $(-2, 5)$.
(b) Find all directions that are perpendicular to the vector in (a).

The magnitude and direction of the vector from $(1,-1)$ to $(-2,5)$ are the same as those of the position vector $(-2,5)-(1,-1)=(-3,6)$. Its magnitude is given by

$$|(-3,6)| = \sqrt{(-3)^2+6^2} = 3\sqrt{5}\ ,$$

and the direction is given by

$$\frac{(-3,6)}{3\sqrt{5}} = \frac{(-1,2)}{\sqrt{5}}\ .$$

(No need to simplify further.)

A vector (a,b) perpendicular to $(-3,6)$ satisfies

$$(-3,6)\cdot(a,b) = -3a+6b = 0\ .$$

By varying a and b according to this relation, there are infinitely many vectors (a,b) satisfying this requirement. For instance, we may take $a=2, b=1$ so $(2,1)$ is one choice. However, to be a direction there is another condition, namely, the length of the vector has to equal to one. There are two such vectors:

$$\frac{(2,1)}{\sqrt{5}}\ , \quad -\frac{(2,1)}{\sqrt{5}}\ .$$

(Again no need to simplify.)

1.3 Euclidean Motions

A **Euclidean motion** is a map from $\mathbb{R}^n$ to itself of the form

$$T\mathbf{x} = A\mathbf{x} + \mathbf{b}\ ,$$

where $\mathbf{b} \in \mathbb{R}^n$ and A is an $n \times n$-matrix, that preserves the distance between two points, that is, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$|T\mathbf{x} - T\mathbf{y}| = |\mathbf{x} - \mathbf{y}|\ .$$

Here in $A\mathbf{x}$ the vector $\mathbf{x}$ should be understood as a column vector.

Recall that a square matrix R is called an orthogonal matrix if $R'R = RR' = I$, where R' is the transpose of R and I is the identity matrix.

Proposition 1.2. *A map $T\boldsymbol{x} = A\boldsymbol{x} + \boldsymbol{b}$ is a Euclidean motion if and only if A is an orthogonal matrix.*

Proof. In the following we use $\langle \mathbf{x}, \mathbf{y} \rangle$ instead $\mathbf{x} \cdot \mathbf{y}$ to denote the dot product. First of all, let T be a Euclidean motion. Then it follows from the definition that

$$|\mathbf{x} - \mathbf{y}| = |T\mathbf{x} - T\mathbf{y}| = |A\mathbf{x} - A\mathbf{y}| = |A(\mathbf{x} - \mathbf{y})|,$$

which yields immediately that

$$\begin{aligned}|A(\mathbf{x}-\mathbf{y})|^2 &= \langle A(\mathbf{x}-\mathbf{y}), A(\mathbf{x}-\mathbf{y})\rangle \\ &= |\mathbf{x}-\mathbf{y}|^2 \\ &= |\mathbf{x}|^2 - 2\langle \mathbf{x},\mathbf{y}\rangle + |\mathbf{y}|^2 \ .\end{aligned}$$

On the other hand, a direct calculation shows that

$$\begin{aligned}\langle A(\mathbf{x}-\mathbf{y}), A(\mathbf{x}-\mathbf{y})\rangle &= \langle A\mathbf{x}-A\mathbf{y}, A\mathbf{x}-A\mathbf{y}\rangle \\ &= |A\mathbf{x}|^2 - 2\langle A\mathbf{x}, A\mathbf{y}\rangle + |A\mathbf{y}|^2.\end{aligned}$$

By comparing, we see that for all $\mathbf{x}, \mathbf{y}$,

$$\langle A'A\mathbf{x}, \mathbf{y}\rangle = \langle A\mathbf{x}, A\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle \ ,$$

which implies that $A'A = I$. Thus A is a orthogonal matrix. Finally, this relation also shows that T is a Euclidean motion whenever A is orthogonal. □ □

Here we have used the following derivation in linear algebra: For a matric $(B\mathbf{x})_j = \sum_k b_{kj}x_k$,

$$\langle \mathbf{x}, B\mathbf{y}\rangle = \langle B'\mathbf{x}, \mathbf{y}\rangle \ .$$

Indeed,

$$\begin{aligned}\langle \mathbf{x}, B\mathbf{y}\rangle &= \sum_j x_j \sum_k b_{kj}y_k \\ &= \sum_k \sum_j b_{kj}x_jy_k \\ &= \sum_k y_k \sum_j b'_{jk}x_j = \langle B'\mathbf{x}, \mathbf{y}\rangle \ .\end{aligned}$$

Here are some examples of Euclidean motions.

(1) Take A to be the identity and $\mathbf{b}$ a nonzero vector. Then $T\mathbf{x} = \mathbf{x} + \mathbf{b}$ is a translation. The origin is moved to $\mathbf{b}$ after the motion.

(2) When $n = 2$, the Euclidean motion

$$T\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is the reflection with respect to the x-axis and

$$T\mathbf{x} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is the reflection with respect to the y-axis. (In matrix form the vector $\mathbf{x}$ is understood as a column vector.) On the other hand, given any plane in $\mathbb{R}^3$, one may consider the reflection with respect to this plane. For instance,

$$T\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is the reflection with respect to the xy-plane in $\mathbb{R}^3$. The reflection with respect to any straight line in $\mathbb{R}^2$ or with respect to any plane in $\mathbb{R}^3$ can be defined similarly.

(3) The (counterclockwise) rotation of angle θ in $\mathbb{R}^2$ is given by the Euclidean motion

$$T\mathbf{x} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} , \quad \theta \in (0, 2\pi) .$$

In $\mathbb{R}^3$, one can perform a rotation around a fixed axis. For instance, the rotation

$$T\mathbf{x} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

takes the z-axis as its axis of rotation.

Let us verify that Euclidean motions are closed under compositions. Let $T\mathbf{x} = A\mathbf{x}+\mathbf{b}$ and $S\mathbf{x} = B\mathbf{x}+\mathbf{c}$ be two Euclidean motions. Its composition is given by

$$ST\mathbf{x} = B(A\mathbf{x}+\mathbf{b}) + \mathbf{c} = C\mathbf{x} + \mathbf{d} , \quad C \equiv BA, \quad \mathbf{d} = B\mathbf{b} + \mathbf{c} .$$

As

$$\begin{aligned} C'C &= (BA)'BA \\ &= A'B'BA \\ &= A'IA \\ &= I, \end{aligned}$$

we conclude that ST is again a Euclidean motion. Furthermore, we claim that each Euclidean motion admits an inverse. Indeed, letting $U\mathbf{x} = A^{-1}\mathbf{x} - A^{-1}\mathbf{b}$ which is obviously an Euclidean motion, we have

$$UT\mathbf{x} = A^{-1}(A\mathbf{x}+\mathbf{b}) - A^{-1}\mathbf{b} = \mathbf{x} .$$

Summing up, the collection of all Euclidean motions forms a group called the Euclidean group of $\mathbb{R}^n$. (It is alright if you have not learned what a group is. You will learn it in MATH2070.)

In the following we study the structure of Euclidean motions for $n = 2, 3$. Apparently it suffices to look at the orthogonal matrix A.

Theorem 1.3. *In $\mathbb{R}^2$, every orthogonal matrix can be written as*

(a)

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} ,$$

or

(b)

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} , \quad \theta \in [0, 2\pi) \ .$$

Case (a) is a genuine rotation for $\theta \in (0, 2\pi)$ and reduces to the identity at $\theta = 0$. Case (b) is the reflection with respect to the x-axis and then followed by a rotation.

Proof. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \ .$$

By orthogonality $A'A = I$ we have

$$a^2 + c^2 = 1 \ , \quad b^2 + d^2 = 1 \ , \quad ab + cd = 0 \ .$$

Since a is a number between -1 and 1, we can find a unique $\theta \in [0, 2\pi)$ such that $a = \cos\theta, \ c = \sin\theta$. Then either $b = -\sin\theta, \ d = \cos\theta$ or $b = \sin\theta, d = -\cos\theta$, so (a) or (b) must hold. □

In the following we consider the three dimensional case. Denote by $R_z(\theta)$ the rotation around the z-axis by an angle θ:

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \ .$$

Similarly we define

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} ,$$

and

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \ .$$

Also denote by L_z the reflection with respect to the xy-plane:

$$L_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \ .$$

Similarly

$$L_x = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ,$$

and

$$L_y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Theorem 1.4. ** In $\mathbb{R}^3$, every orthogonal matrix can be written as*

(a) $R_z(\alpha)R_x(\beta)R_z(\gamma)$, *or*

(b) $R_z(\alpha)R_x(\beta)R_z(\gamma)L_z$,

for some α, β, and γ.

Proof. * Let $A = (a_{ij})\,, i, j = 1, 2, 3$, be orthogonal. We have

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta a_{11} - \sin\theta a_{21} & \cos\theta a_{12} - \sin\theta a_{22} & \cos\theta a_{13} - \sin\theta a_{23} \\ \sin\theta a_{11} + \cos\theta a_{21} & \sin\theta a_{12} + \cos\theta a_{22} & \sin\theta a_{13} + \cos\theta a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} .$$

Choose θ so that $\cos\theta a_{13} - \sin\theta a_{23} = 0$ and write the resulting matrix as

$$\begin{bmatrix} b_{11} & b_{12} & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} .$$

We further have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11} & b_{12} & 0 \\ \cos\varphi b_{21} - \sin\varphi b_{31} & \cos\varphi b_{22} - \sin\varphi b_{32} & \cos\varphi b_{23} - \sin\varphi b_{33} \\ \sin\varphi b_{21} + \cos\varphi b_{31} & \sin\varphi b_{22} + \cos\varphi b_{32} & \sin\varphi b_{23} + \cos\varphi b_{33} \end{bmatrix} .$$

Choose φ so that $\cos\varphi b_{23} - \sin\varphi b_{33} = 0$ and write the resulting matrix as

$$\begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33} \end{bmatrix} .$$

This matrix is the product of three orthogonal matrices, again it is orthogonal. Therefore, $c_{33} = \pm 1$. Moreover, from

$$c_{11} \times 0 + c_{21} \times 0 + c_{31} \times c_{33} = 0 \ ,$$

we deduce $c_{31} = 0$. Similarly, $c_{32} = 0$. The matrix is in fact of the form

$$\begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & \pm 1 \end{bmatrix} \ ,$$

where the 2×2-matrix is orthogonal. It can be written as $R_z(\gamma)$ or $R_z(\gamma)L_z$ for some γ according to Proposition 1.3. We conclude that

$$R_x(\varphi)R_z(\theta)A = R_z(\gamma) \ ,$$

or

$$R_x(\varphi)R_z(\theta)A = R_z(\gamma)L_z \ ,$$

that is,

$$A = R_z(-\theta)R_x(-\varphi)R_z(\gamma) \ ,$$

or

$$A = R_z(-\theta)R_x(-\varphi)R_z(\gamma)L_z \ .$$

The desired result follows by taking $\alpha = -\theta$ and $\beta = -\varphi$. □

1.4 The Cross Product in $\mathbb{R}^3$

The cross product assigns a 3-vector to two given 3-vectors. There is no such product in the general dimension. Somehow it shows how special our physical space is. The cross product is important due to its relevance in physics and engineering.

Notations like $\mathbf{x}, \mathbf{y}, \mathbf{u}$ and $\mathbf{v}$ are common for vectors. We have used $\mathbf{x}, \mathbf{y}$ in the previous sections. Now we use $\mathbf{u}, \mathbf{v}$ in this one.

First the definition. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, define the cross product of $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ to be

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2, -(u_1v_3 - u_3v_1), u_1v_2 - u_2v_1) \ .$$

In particular, we have

$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \quad \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \quad \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2 \ .$$

To aid memorizing, we can formally express it as the determinant

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} .$$

Expand the determinant by the first row yields the formula above. Here it is formal because it does not make sense to put the unit vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ as entries in a matrix.

The cross product is in many aspects in sharp contrast with an ordinary product. Some of its properties are list below:

Theorem 1.5. *For $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$,*

(a) $(\alpha \boldsymbol{u} + \beta \boldsymbol{v}) \times \boldsymbol{w} = \alpha \boldsymbol{u} \times \boldsymbol{w} + \beta \boldsymbol{v} \times \boldsymbol{w}$, $\quad \forall \alpha, \beta \in \mathbb{R}$.

(b) $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$. *In particular,* $\boldsymbol{u} \times \boldsymbol{u} = 0$.

(c) $(\boldsymbol{u} \times \boldsymbol{v}) \times \boldsymbol{w} = \boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w})$ *is not always true.*

The proofs of Theorem (a) and (b) are straightforward from definition. As for (c), which asserts that the associative law does not hold, some examples suffice:

$$(\mathbf{e}_1 \times \mathbf{e}_2) \times \mathbf{e}_2 = -\mathbf{e}_1 \ , \quad \mathbf{e}_1 \times (\mathbf{e}_2 \times \mathbf{e}_2) = \mathbf{0} \ ,$$

and

$$(1,1,1)\times\big((1,0,-1)\times(2,1,1)\big) = (3,-1,-5) \ , \quad \big((1,1,1)\times(1,0,-1)\big)\times(2,1,1) = (4,0,-4).$$

As a vector is completely determined by its magnitude and direction, let us consider the magnitude and direction of the cross product. First of all, using the definition of the cross product, one can verify directly that

$$\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0 \ , \quad \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0 \ ,$$

so

$$(\alpha \mathbf{u} + \beta \mathbf{v}) \cdot \mathbf{u} \times \mathbf{v} = 0 \ ,$$

that is, it is perpendicular to the two dimensional subspace spanned by the vectors $\mathbf{u}$ and $\mathbf{v}$. After the definition of a plane is introduced in the next chapter, we can say that the cross product of two linearly independent vectors points in the normal direction of the plane spanned by $\mathbf{u}$ and $\mathbf{v}$. When they are linearly dependent, their cross product is the zero vector which does not form a normal direction. There are two normal directions, pointing upward or downward, so to speak. The choice of the direction of the cross product is determined by the right hand rule. That is, with the thumb making a right angle with the other four fingers of your right hand, first point the four fingers along the direction of $\mathbf{u}$ and then move them to $\mathbf{v}$ in an angle less than π. The direction of $\mathbf{u} \times \mathbf{v}$

is where your thumb points to. To see this, one identifies the direction of $\mathbf{u}$ with $\mathbf{e}_1$. If $\mathbf{v}$ lies in the first or second quadrants, $\mathbf{v} = \alpha\mathbf{e}_1 + \beta\mathbf{e}_2$, $\beta > 0$, and $\mathbf{u} \times \mathbf{v}$ points to $\mathbf{e}_3$. If $\mathbf{v}$ lies in the third or fourth quadrants, $\mathbf{v} = \alpha\mathbf{e}_1 + \beta\mathbf{e}_2$, $\beta < 0$, and $\mathbf{u} \times \mathbf{v}$ points to $-\mathbf{e}_3$. This is the right hand rule.

We have described the direction of the cross product. How about its magnitude ? We have

Theorem 1.6. *For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$,*

$$|\boldsymbol{u} \times \boldsymbol{v}| = |\boldsymbol{u}||\boldsymbol{v}| \sin\theta, \quad \theta \in [0, \pi] \ ,$$

where θ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$.

Proof. The proof depends on the identity

$$|\mathbf{u} \times \mathbf{v}|^2 = |\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \ .$$

Indeed, by brute force

$$\begin{aligned} |\mathbf{u} \times \mathbf{v}|^2 &= (u_2v_3 - u_3v_2)^2 + (u_1v_3 - u_3v_1)^2 + (u_1v_2 - u_2v_1)^2 \\ &= u_2^2v_3^2 + u_3^2v_2^2 + u_1^2v_3^2 + u_3^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2v_3u_3v_2 - 2u_1v_3u_3v_1 - 2u_1v_2u_2v_1 \ . \end{aligned}$$

On the other hand,

$$\begin{aligned} & |\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \\ = \ & (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2 \\ = \ & u_2^2v_3^2 + u_3^2v_2^2 + u_1^2v_3^2 + u_3^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2v_3u_3v_2 - 2u_1v_3u_3v_1 - 2u_1v_2u_2v_1 \ , \end{aligned}$$

whence the identity holds. Now, by the Cosine Law,

$$\begin{aligned} |\mathbf{u} \times \mathbf{v}| &= \sqrt{|\mathbf{u}|^2|\mathbf{v}|^2 - |\mathbf{u}|^2|\mathbf{v}|^2\cos^2\theta} \\ &= |\mathbf{u}||\mathbf{v}||\sin\theta| \\ &= |\mathbf{u}||\mathbf{v}|\sin\theta \ , \end{aligned}$$

as $\sin\theta \geq 0$ on $[0, \pi]$. □

In conclusion the magnitude and direction of the decomposition of the cross product is given by

$$\mathbf{u} \times \mathbf{v} = |\mathbf{u}||\mathbf{v}|\sin\theta \ \mathbf{n} \ ,$$

where $\mathbf{n}$ is the unit vector determined by the right hand rule (when $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, that is, when $\sin\theta \neq 0$).

Corollary 1.7. *(a) The area of the parallelogram spanned by $\boldsymbol{u}$ and $\boldsymbol{v}$ is equal to $|\boldsymbol{u} \times \boldsymbol{v}|$.*

(b) The area of the triangle with two sides given by $\boldsymbol{u}$ and $\boldsymbol{v}$ is equal to $1/2|\boldsymbol{u}\times\boldsymbol{v}|$.

(c) The volume of the parallelepiped spanned by $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ is equal to

$$V = |\boldsymbol{w}\cdot(\boldsymbol{u}\times\boldsymbol{v})| \ .$$

Proof. (a) follows immediately from Theorem 1.6 and (b) from (a). To prove (c), we may assume $\mathbf{u}$ and $\mathbf{v}$ lie on the xy-plane after a rotation. The volume of the parallelepiped is given by the product of the area of the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$ with its height. Now $|\mathbf{u}\times\mathbf{v}|$ is equal to the area of this parallelogram. On the other hand, its height is given by $|\mathbf{w}\cdot\mathbf{e}_3|$. Therefore, letting α be the angle between $\mathbf{w}$ and the z-axis,

$$\begin{aligned}
|\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})| &= |\mathbf{w}||\mathbf{u}\times\mathbf{v}||\cos\alpha| \\
&= |\mathbf{u}\times\mathbf{v}|\ |\mathbf{w}\cdot\mathbf{e}_3| \\
&= V\ .
\end{aligned}$$

□

Example 1.5. Determine if the four points

$$(1,0,1),\ (2,4,-6),\ (3,-1,5),\ (1,-9,19)\ ,$$

lie on the same plane in $\mathbb{R}^3$. Well, they lie on the same plane if and only if the parallelepiped formed by these vectors has zero volume. We compute the volume using this corollary after subtracting the first vector from the last three vectors (to make sure that the vectors are based at the origin):

$$\begin{aligned}
(1,4,-7)\cdot\big((2,-1,4)\times(0,-9,18)\big) &= (1,4,-7)\cdot(18,-36,-18) \\
&= 0\ ,
\end{aligned}$$

so they lie on the same plane.

Comments on Chapter 1.

1.1. A helping hand from linear algebra. In several occasions we need to solve homogeneous systems of linear equations. Let us review it by looking at some examples. First, consider the single equation

$$x-2y+6z=0, \quad (x,y,z)\in\mathbb{R}^3\ .$$

To solve this equation means to find all possible (x,y,z) satisfying this relation. Obviously $(0,0,0)$ is a solution, but there are many others, for instance, $(-6,0,1)$ and $(2,1,0)$ are also solutions. To find all solutions, we set $y=a$ and $z=b$. Then $x=2a-6b$, so $(2a-6b,a,b)=a(2,1,0)+b(-6,0,1)$ gives all solutions. A solution is obtained whenever values to a and b are assigned. We may say the solution is described by two parameters. Next, consider the system

$$\begin{cases} x - y + 5z &= 0 \ , \\ x + y - 3z &= 0 \ . \end{cases}$$

Setting $z = a$, the system becomes

$$\begin{cases} x - y &= -5a \ , \\ x + y &= 3a \ , \end{cases}$$

which is readily solved to yield $x = -a$ and $y = 4a$, so the general solution is given by $(x, y, z) = a(-1, 4, 1)$ where a is the only parameter. From these two examples, we see there are three principles governing homogeneous linear systems.

- $\mathbf{0}$ is always a solution (the trivial solution).
- The general solution consists of several free parameters. In most cases, the number of parameters is equal to $n - m$ where n is the number of unknowns and m is the number of equations.
- In some exceptional cases, the number of parameters is less than $n - m$.

Exceptional cases come up when the linear system is kind of cheating us. For instance, look at

$$\begin{cases} x - y + 5z &= 0 \ , \\ 2x - 2y + 10z &= 0 \ . \end{cases}$$

The second equation in this system comes from multiplying the first equation by 2, so essentially there is only one equation in this system. Its general solution contains two instead of one parameters. This situation occurs in the study of standard forms in Section 2.4.

1.2. The Law of Cosines states that, let ΔABC be a triangle and $a = \overline{CB}$, $b = \overline{AC}$ and $c = \overline{BA}$, and $\phi = \angle ACB$. Then

$$c^2 = a^2 + b^2 - 2ab\cos\phi \ .$$

To prove it let $h = \overline{AH}$ be the height of the triangle from A and s be $\overline{CH}$. By Pythagoras Theorem,

$$c^2 = h^2 + (a - s)^2 = b^2 - s^2 + (a - s)^2 = b^2 - 2as + a^2 \ ,$$

implies $s = (c^2 - a^2 - b^2)/2a$ and the Cosine Law follows after noting $s = b\cos\phi$.

1.3. You may look up Wikipedia under "Euclidean motion" and "orthogonal matrix" to find further information on these interesting topics. How the cross product is used in physics can be also be found in Wikipedia under "cross product". From the same source you will see how the cross product arises from the Lie algebra of the orthogonal group.

Supplementary Reading

1.1 and 1.2 in chapter 1, [Au].