Chapter 4

Vector Spaces

In this chapter you are introduced to the idea of vector spaces. The presentation of this topic so early in the book is to help you develop geometric intuition that will be useful in later chapters. The geometric insights can be hard to grasp at first, especially in higher-dimensional spaces, but the mathematical operations used throughout this chapter do not go beyond addition, multiplication, and taking square roots.

We begin by thinking of points in the 2-dimensional Cartesian plane as *vectors*, and consider ideas such as the *norm* of a vector, the *angle* between vectors, and the *linear combination* of vectors. This leads to the concepts of *vector spaces* and *subspaces*, and their *dimensions*. Pythagoras's Theorem is extended to the *Law of Cosines* for non-right-angled triangles, and the *Triangle Inequality* and the *Cauchy-Schwarz Inequality* are established. We then take these concepts to 3-dimensional space, and show how they extend to yet higher dimensional spaces. We also extend the Gaussian elimination method of Section 2.1.5 to solving systems of equations in 3 or more unknowns. A later chapter will explore the connection between vector spaces and the problem of solving systems of linear equations, although some remarks are made in this chapter regarding this connection.

We discuss Python *lists* and *tuples*, and *numpy arrays* in the programming section.

4.1 2-Dimensional Vector Spaces

Consider the set $\mathbb{R}^2$, comprising all points in the 2-dimensional Cartesian coordinate system. For any two points $u = (x_1, y_1)$ and $v = (x_2, y_2)$ in this set, define **scalar multiplication** to be the operation

$$\alpha u = (\alpha x_1, \alpha y_1) \tag{4.1}$$

where α is some real number, and define **vector addition** to be

$$u + v = (x_1 + x_2, y_1 + y_2)\,. \tag{4.2}$$

With these two definitions, the set of points in the Cartesian plane becomes a **vector space** (or a **linear space**) and the points are called **vectors**. We can use the scalar multiplication and addition operations to create **linear combinations** of vectors, such as

$$\alpha_1 u + \alpha_2 v\,.$$

These operations generally result in new vectors.

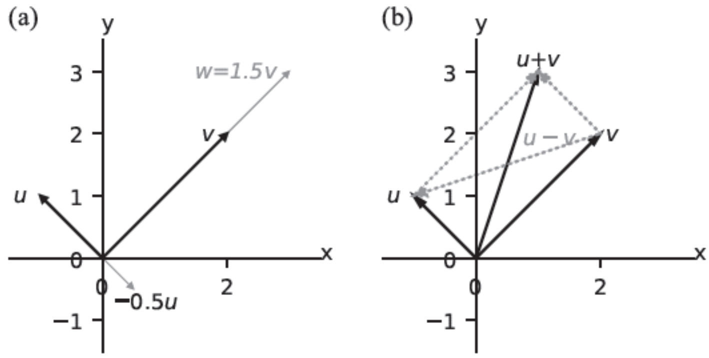

Fig. 4.1. 2-dimensional vectors.

The term "vector" usually refers to a quantity with direction and magnitude. When thought of in this way, a point such as $v = (2, 2)$ can be visualized or represented as an arrow from the origin to the point $(2, 2)$, and you can think of this as an action that moves an object located at the origin to the point $(2, 2)$, i.e., the object is displaced by a distance of $2\sqrt{2}$ in the northeast direction.[1] Scalar multiplication αv has the effect of reversing the direction of the vector v if α is negative, stretching it if $|\alpha| > 1$, and shrinking it if $|\alpha| < 1$. Fig. 4.1(a) shows the "arrow representation" of the vectors $u = (-1, 1)$ and $v = (2, 2)$ and the effect of multiplying the vectors by -0.5 and 1.5 respectively.[2]

We defined vectors as points in a space, and use the "arrow representation" to support the "magnitude with direction" *interpretation* of vectors. We will switch between these two perspectives freely. The arrow representation is very helpful in visualizing the outcome of vector operations. When we speak of the magnitude of vectors and the angle between two vectors we are thinking of the arrow representation. However, when we speak of the "distance between u and v" in Fig. 4.1(a) we are thinking of them as points. As another example, we say that "u and v in Fig. 4.1(a) lie on the line $y = \frac{4}{3} + \frac{1}{3}x$" because the points $(-1, 1)$ and $(2, 2)$ lie on this line, even though their "arrows" obviously do not.[3] Sometimes we use the term "points" instead of "vectors" when we want to emphasize the "points" perspective.

[1]In physics, the vector v might represent a force of $2\sqrt{2}$ Newtons in the northeast direction, and u might represent a force of $\sqrt{2}$ Newtons in the northwest direction.

[2]The zero vector $(0, 0)$ is a bit special, with no magnitude (or with magnitude 0) and no direction. Obviously, scalar multiplication of the zero vector returns the zero vector, and adding the zero vector to another vector does not change that vector.

[3]Of course, for points on lines that pass through the origin, both the points and their arrow representations lie on the line.

Vector addition is illustrated in Fig. 4.1(b) with $u+v$. You can view this sum as displacing an object at the origin by the distances and directions represented by u and v simultaneously. Alternatively, you can view $u+v$ as moving an object from the origin to u and then to $u+v$. In this sense the dotted arrow from u to $u+v$ also represents the vector v. Likewise, you can view $u+v$ as moving the object from the origin to v and then to $u+v$, so the dotted arrow from v to $u+v$ also represents the vector u. In the arrow representation, two arrows parallel to each other, with the same magnitude and pointing in the same direction, represent the same vector.

Notice that the arrow from the origin representing the vector $u+v$ is a diagonal of the parallelogram formed from the arrows from the origin to u and to v, and from u and v to $u+v$. Since $u = v + (u-v)$, $u-v$ can be represented by an arrow from v to u (the other diagonal).

Given the two vectors u and v in Fig. 4.1, we can express any vector in the space $\mathbb{R}^2$ by taking an appropriate linear combination of u and v. Suppose we want to express the vector $(6,4)$ as a linear combination of u and v. This means finding α and β such that

$$\alpha u + \beta v = \alpha(-1,1) + \beta(2,2) = (6,4)\,, \tag{4.3}$$

i.e.,

$$\begin{aligned} -\alpha + 2\beta &= 6 \\ \alpha + 2\beta &= 4\,. \end{aligned} \tag{4.4}$$

This system is easily solved to give $\alpha = -1$ and $\beta = 5/2$. Notice also that this is the only solution to (4.4), so there is only one linear combination of u and v that gives the vector $(6,4)$.

We can repeat this process for any arbitrary vector in the 2-dimensional space. For any vector (x^*, y^*), we can find α and β such that

$$\alpha u + \beta v = \alpha(-1,1) + \beta(2,2) = (x^*, y^*)\,.$$

The values $\alpha = (y^* - x^*)/2$ and $\beta = (y^* + x^*)/4$ will do the trick. Put another way, the set of ALL linear combinations of the form $\alpha u + \beta v$, when we consider all values of α and β, makes up the entire $\mathbb{R}^2$ space. We say that the vectors u and v **span** the entire 2-dimensional $\mathbb{R}^2$ vector space. We also say that the two vectors u and v form a **basis** for $\mathbb{R}^2$.

That we can express any vector in $\mathbb{R}^2$ as a linear combination of two fixed vectors is true for the pair $u = (-1,1)$ and $v = (2,2)$, and it is true for many other pairs as well, including $u = (0,1)$ and $v = (1,0)$, $u = (0,1)$ and $v = (1,1)$, and so on. But it is not true for all pairs of vectors. Suppose you start with some vector v and take the other vector to be $w = 1.5v$. Linear combinations of these two vectors will merely return another multiple of v:

$$\alpha v + \beta w = \alpha v + \beta 1.5 v = (\alpha + 1.5\beta)v\,.$$

That is, you will only be able to replicate vectors in the direction of v or its reverse, and not any other vector. The set of all linear combinations of u and w will form a strict subset of $\mathbb{R}^2$, making up a line passing through the origin. We say that the vectors v and w do not span the entire 2-dimensional vector space, but only span a 1-dimensional **subspace** of it. Two 2-dimensional vectors will span a 1-dimensional subspace (a line passing through the origin) if one of the vectors is a multiple of the other, or if one is zero and the other not zero. If both u and v are zero, then of course all linear combinations of u and v result in the zero vector. We view this as a "0-dimensional" subspace of $\mathbb{R}^2$.

Because u and v spans $\mathbb{R}^2$, for any vector b in $\mathbb{R}^2$ there is exactly one linear combination of u and v that gives b. In particular, the only linear combination that gives the zero vector $(0, 0)$ is

$$0\,u + 0\,v = (0, 0)\,.$$

In other words,

$$c_1 u + c_2 v = (0, 0) \quad \Rightarrow \quad c_1 = c_2 = 0\,. \tag{4.5}$$

We say that u and v are **linearly independent** if they satisfy (4.5). If two vectors do not satisfy this condition, we say they are **linearly dependent**. This will be the case if one of u and v is a multiple of the other, or one or both are zero vectors. If $u = \alpha v$, then $u - \alpha v = (0, 0)$ so we have non-zero c_1 and c_2 such that $c_1 u + c_2 v = 0$. If one (say u) is the zero vector, then $c_1 u + 0v = (0, 0)$ for any c_1, including non-zero c_1. If both are zero vectors, then $c_1 u + c_2 v = (0, 0)$ regardless of c_1 and c_2.

If two 2-dimensional vectors span the entire space $\mathbb{R}^2$, then any third vector can be written as a combination of the first two. This means that a set of n 2-dimensional vectors, $n \geq 3$, cannot satisfy the condition

$$c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = (0, 0) \quad \Rightarrow \quad c_1 = c_2 = \cdots = c_n = 0\,. \tag{4.6}$$

A set of three or more 2-dimensional vectors must be linearly dependent. Of course, such a set could well span only a lower dimensional subspace of $\mathbb{R}^2$, but it can span at most the entire space.[4]

What happens if we take *restricted* linear combinations of the form

$$(1 - \alpha)u + \alpha v\,? \tag{4.7}$$

Writing (4.7) as $u + \alpha(v - u)$, it should be clear that if we consider the set of all such linear combinations, with $\alpha \in \mathbb{R}$, then we get the entire line passing through the two points u and v. Fig. 4.2 shows $u = (-1, 1)$ and $v = (2, 2)$ with linear combinations of the form (4.7) with $\alpha = -0.3, 0.2, 0.5$ and 1.5 in grey. Setting $\alpha = 0.5$ gives the point that lies exactly midway between u and v. Can you tell which grey vector corresponds with which value of α?

[4]Incidentally, the term "basis for $\mathbb{R}^2$ is reserved for sets of *exactly* two vectors that span the entire $\mathbb{R}^2$ space.

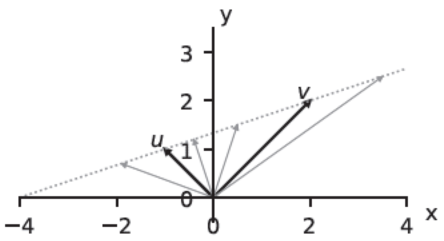

Fig. 4.2. Affine combinations of two vectors.

Linear combinations of the form (4.7) are called **affine combinations**. Notice that in general the line formed by affine combinations of two vectors will not pass through the origin, unless the two vectors already lie on a line passing through the origin.

We will not call the set of vectors obtained by affine combinations of vectors a "vector subspace", unless the line that it forms passes through the origin (in which case it will include the zero vector). The idea of a *vector space* or *vector subspace* is that we can take linear combinations of any vectors in the space/subspace and get back another vector in the same space/subspace. This will be the case if the set of vectors lie on a line that passes through the origin. It is not true of the set of vectors that form the line in Fig. 4.2.[5]

We know that a line can be represented by an equation of the form $ay + bx + c = 0$. The fact that linear combinations of the form (4.7) make up a line gives us another way to represent lines in the Cartesian plane. If the points u and v lie on the line, then any other point (x, y) on the line can be written as

$$(x, y) = (1 - \alpha)u + \alpha v = u + \alpha(v - u)$$

for some $\alpha \in \mathbb{R}$. We can describe this line as "the line passing through u in the direction of $v - u$". If $u = (x_1, y_1)$ and $v = (x_2, y_2)$, then we have

$$(x, y) = (x_1, y_1) + \alpha(x_2 - x_1, y_2 - y_1)\,,\ \alpha \in \mathbb{R}\,. \tag{4.8}$$

This is the **vector equation of a line**. We can write (4.8) as two separate equations:

$$\begin{aligned} x - x_1 &= \alpha(x_2 - x_1) \\ y - y_1 &= \alpha(y_2 - y_1)\,,\ \alpha \in \mathbb{R}\,. \end{aligned} \tag{4.9}$$

The equations in (4.9) are the **parametric equations** of the line, with α as the "parameter". The slope of this line, if the line is not vertical, is $(y_2 - y_1)/(x_2 - x_1)$.

[5]Sets of vectors such as in Fig. 4.2 are sometimes called "affine subspaces". The word "affine" means "related".

4.1.1 *The Dot Product*

Define the **dot product** of two vectors $u = (x_1, y_1)$ and $v = (x_2, y_2)$ to be

$$u \cdot v = x_1 x_2 + y_1 y_2$$

The dot product is also sometimes called the **inner product** or **scalar product**. It is straightforward to show (see Ex. 4.1) that

- $u \cdot v = v \cdot u$,
- $(u + v) \cdot z = u \cdot z + v \cdot z$ (where z is a third vector), and
- $(\alpha u) \cdot v = \alpha(u \cdot v)$ (where α is some number).

The second of these implies that for any four vectors u, v, w and z, we have

$$(u + v) \cdot (w + z) = u \cdot w + v \cdot w + u \cdot z + v \cdot z \ .$$

We refer to the distance of the point $u = (x_1, y_1)$ from the origin, i.e., $(x_1^2 + y_1^2)^{1/2}$, as the *magnitude*[6] or **norm** of the vector u, denoted as $\|u\|$. Since $u \cdot u = x_1 x_1 + y_1 y_1 = x_1^2 + y_1^2 = \|u\|^2$, we have

$$\|u\| = (u \cdot u)^{1/2} \, .$$

Obviously $u \cdot u \geq 0$, and equal to zero if (and only if) $u = (0, 0)$. Similarly, the distance between two points $u = (x_1, y_1)$ and $v = (x_2, y_2)$ is the square root of $\|v - u\|$, since

$$(v - u) \cdot (v - u) = (x_2 - x_1)^2 + (y_2 - y_1)^2 = \|v - u\|^2 \, .$$

Furthermore, we have $\|v - u\| = \|u - v\|$.

If a vector u has magnitude $\|u\|$, then $u/\|u\|$ will be a **unit vector**, i.e., a vector with magnitude 1. For example, $u = (1, -1)$ has norm $\sqrt{2}$, whereas the vector

$$\frac{u}{\|u\|} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

has norm $\sqrt{1/2 + 1/2} = 1$.

Suppose two vectors u and v are perpendicular to each other. Then $u - v$ is the hypotenuse of the right-angled triangle formed by the origin, u and v, and by Pythagoras's Theorem, we have

$$\|u\|^2 + \|v\|^2 = \|u - v\|^2 \, .$$

In terms of the dot product, this says

$$\begin{aligned}
u \cdot u + v \cdot v &= (u - v) \cdot (u - v) \\
x_1^2 + y_1^2 + x_2^2 + y_2^2 &= (x_1 - x_2)^2 + (y_1 - y_2)^2 \\
x_1^2 + y_1^2 + x_2^2 + y_2^2 &= x_1^2 + x_2^2 - 2x_1 x_2 + y_1^2 + y_2^2 - 2y_1 y_2 \\
x_1 x_2 + y_1 y_2 &= 0
\end{aligned}$$

[6]We refrain from referring to $\|u\|$ as the length of the vector, since in computer programming the "length" of a vector refers to the number of elements in the vector.

In other words, $u \cdot v = 0$ for all perpendicular vectors u and v. For example, the vectors $u = (-1, 1)$ and $v = (2, 2)$ in Fig. 4.1 are perpendicular, and we have $u \cdot v = (-1)(2) + (1)(2) = 0$. The vectors $u + v = (1, 3)$ and v are not perpendicular, and we have $(u + v) \cdot v = (1)(2) + (3)(2) \neq 0$.

4.1.2 *Law of Cosines*

The fact that perpendicular vectors have zero dot product is a special case of the following result:

Law of Cosines: If θ is the angle between two vectors u and v, then

$$\frac{u}{\|u\|} \cdot \frac{v}{\|v\|} = \cos\theta \quad \text{i.e.,} \quad u \cdot v = \|u\|\,\|v\|\,\cos\theta\,. \tag{4.10}$$

Proof: Given vectors u and v, the vectors

$$\frac{u}{\|u\|} \quad \text{and} \quad \frac{v}{\|v\|}$$

are unit vectors. This means they are located on the unit circle, and can be written as

$$\frac{u}{\|u\|} = (\cos\alpha, \sin\alpha) \quad \text{and} \quad \frac{v}{\|v\|} = (\cos\beta, \sin\beta)$$

where the angle between the two vectors is $\theta = \alpha - \beta$. This is illustrated in Fig. 4.3(a) for a particular pair u and v. Taking the dot product of $u/\|u\|$ and $v/\|v\|$ gives

$$\frac{u}{\|u\|} \cdot \frac{v}{\|v\|} = \cos\alpha\cos\beta + \sin\alpha\sin\beta\ . \tag{4.11}$$

A fundamental result in trigonometry called the "subtraction formula for cosines" (which we will prove momentarily) says that

$$\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta)\,. \tag{4.12}$$

Applying (4.12) to (4.11) gives

$$\frac{u}{\|u\|} \cdot \frac{v}{\|v\|} = \cos(\alpha - \beta)\ = \cos\theta\ , \quad \text{i.e.,} \quad u \cdot v = \|u\|\,\|v\|\cos\theta\,.$$

To show (4.12), rotate the vectors in Fig. 4.3(a) clockwise by an angle of β to get Fig. 4.3(b). Using the formula for the distance between two points, we have from Fig. 4.3(a) that

$$d^2 = (\cos\alpha - \cos\beta)^2 + (\sin\alpha - \sin\beta)^2\,, \tag{4.13}$$

and from Fig. 4.3(b) that

$$d^2 = (\cos(\alpha - \beta) - 1)^2 + \sin^2(\alpha - \beta)\,. \tag{4.14}$$

Equating these two expressions for d^2, simplifying, and using the fact that $\cos^2\gamma + \sin^2\gamma = 1$ for any angle γ gives (4.12). (See Ex. 4.2.)

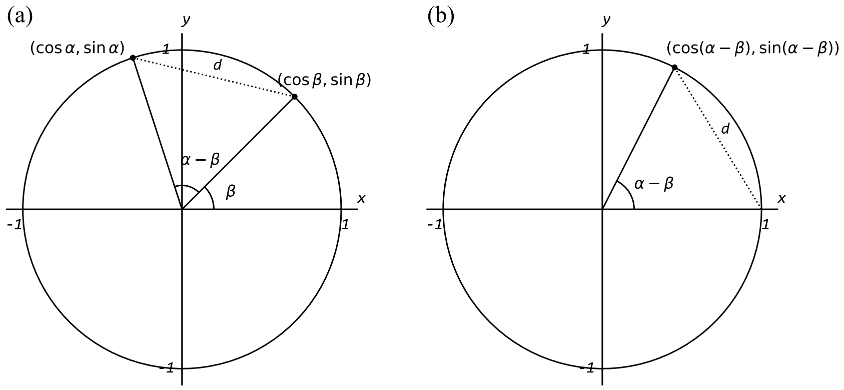

Fig. 4.3. The subtraction formula for cosines.

Digression: Trigonometric Identities. The identity (4.12) is one of a set of identities called the "addition and subtraction formulas". Together with $\sin^2\theta + \cos^2\theta = 1$, the addition and subtraction formulas are worth committing to memory. The **addition formulas** are:

$$\begin{aligned}\sin(\theta+\gamma) &= \sin\theta\cos\gamma + \cos\theta\sin\gamma \\ \cos(\theta+\gamma) &= \cos\theta\cos\gamma - \sin\theta\sin\gamma\end{aligned} \tag{4.15}$$

and the **subtraction formulas** are

$$\begin{aligned}\sin(\theta-\gamma) &= \sin\theta\cos\gamma - \cos\theta\sin\gamma \\ \cos(\theta-\gamma) &= \cos\theta\cos\gamma + \sin\theta\sin\gamma\end{aligned} \tag{4.16}$$

We have just proven the cosine subtraction formula. Applying the cosine subtraction formula and the identities in Table 2.3 to

$$\sin(\theta-\gamma) = \cos(\theta - (\gamma + \pi/2))$$

gives the sine subtraction formula. Applying the subtraction formulas to

$$\sin(\theta+\gamma) = \sin(\theta - (-\gamma)) \ \text{ and } \ \cos(\theta+\gamma) = \cos(\theta - (-\gamma))\,,$$

gives the addition formulas.

There are yet other trigonometric identities. Setting $\gamma = \theta$ in $\sin(\theta+\gamma)$ and $\cos(\theta+\gamma)$ gives the **Double-Angle Formulas**:

$$\begin{aligned}\sin 2\theta &= 2\sin\theta\cos\theta \\ \cos 2\theta &= \cos^2\theta - \sin^2\theta\,.\end{aligned}$$

From the double-angle formula for $\cos 2\theta$ and $\sin^2\theta + \cos^2\theta = 1$, we get the **Half-Angle Formulas**:

$$\cos^2\theta = \frac{1+\cos 2\theta}{2} \qquad \text{and} \qquad \sin^2\theta = \frac{1-\cos 2\theta}{2}\,.$$

The addition and subtraction formulas also gives the **Product Formulas**:

$$\sin x \cos y = \tfrac{1}{2}[\sin(x+y) + \sin(x-y)]$$
$$\cos x \cos y = \tfrac{1}{2}[\cos(x+y) + \cos(x-y)]$$
$$\sin x \sin y = \tfrac{1}{2}[\cos(x-y) - \cos(x+y)]\ .$$

The Law of Cosines gives a nice geometric interpretation to the dot product: the dot product of two unit vectors is the cosine of the angle between them. But why do we refer to it as the Law of Cosines? First note that for any vectors u and v, we have

$$\|u-v\|^2 = \|u\|^2 - 2u \cdot v + \|v\|^2 = \|u\|^2 - 2\|u\|\,\|v\| \cos\theta + \|v\|^2\,. \qquad (4.17)$$

The first equality comes from the dot product definition (see Ex. 4.3). The second equality comes from (4.10).

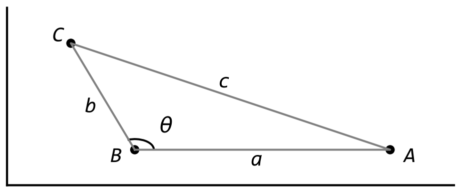

Fig. 4.4. The Law of Cosines and the Triangle Inequality.

Now suppose A, B and C are three arbitrary points, as illustrated in Fig. 4.4. Let $c = \|A-C\|$, $a = \|A-B\|$, $b = \|C-B\|$, and let θ be the angle made by the vectors $A-B$ and $C-B$. Substituting $u = A-B$ and $v = C-B$ into (4.17) gives

$$\|A-C\|^2 = \|A-B\|^2 + \|C-B\|^2 - 2\|A-B\|\,\|C-B\| \cos\theta\,. \qquad (4.18)$$

Writing in terms of a, b and c reveals the form of the Law of Cosines that you may be more familiar with:

$$c^2 = a^2 + b^2 - 2ab\cos\theta\,. \qquad (4.19)$$

The Law of Cosines is a generalization of Pythagoras's Theorem. Eq. (4.19) simplifies to $c^2 = a^2 + b^2$ when the triangle is right-angled, i.e., when $\cos\theta = 0$.

4.1.3 *The Triangle Inequality*

The Law of Cosines also gives us another important fact about geometric distances, that is, that the straight line distance from point A to B, and then onto to point C, is no less than the straight line distance from A directly to C. This is the **Triangle inequality**, which is implied by the Law of Cosines. Because $-1 \leq \cos\theta \leq 1$, the maximum value of the RHS of (4.18) occurs when $\cos\theta = -1$, that is,

$$\begin{aligned} \|A-C\|^2 &\leq \|A-B\|^2 + \|C-B\|^2 + 2\|A-B\|\,\|C-B\| \\ &= (\|A-B\| + \|C-B\|)^2. \end{aligned} \tag{4.20}$$

Taking square roots on both sides gives

$$\|A-C\| \leq \|A-B\| + \|C-B\|\,. \tag{4.21}$$

4.1.4 *The Cauchy-Schwarz Inequality*

The fact that $u \cdot v = \|u\|\,\|v\|\cos\theta$ and $-1 \leq \cos\theta \leq 1$ means that

$$-\|u\|\,\|v\| \leq u \cdot v \leq \|u\|\,\|v\|\,, \text{ i.e., } |u \cdot v| \leq \|u\|\|v\|\,. \tag{4.22}$$

This is a version of the **Cauchy-Schwarz inequality**. The importance of this inequality will not be obvious at this point, but it is a key result that appears in many different areas of mathematics.

4.1.5 *Connection with Complex Numbers*

How do 2-dimensional vector spaces relate to complex numbers? We have seen that complex numbers are points $(a,b) \in \mathbb{R}^2$ that can be added in exactly the same way as the 2-dimensional vectors we have discussed in this section. The main difference is that for complex numbers we also define *complex multiplication*, in a way that (i) allows us to work with complex numbers in the same manner that we work with real numbers, (ii) encompasses the real number system as a special case of complex number system (real numbers are just complex numbers of the form $(a,0)$), and (iii) allows us to interpret the second coordinate as the square root of negative numbers, e.g., $(0,1)(0,1) = -1$. We define scalar multiplication for 2-dimensional vector spaces but we do not consider multiplication of vectors the way we multiply complex numbers. In the following sections, we extend our discussion of 2-dimensional vector spaces to 3- and higher-dimensional vector spaces. The complex number system can be extended to 4-dimensional systems (quaternions) and 8-dimensional systems (octonions) but we will not cover these in this book.

4.1.6 *Exercises*

Ex. 4.1 Let $u = (x_1, y_1)$, $v = (x_2, y_2)$, $z = (x_3, y_3)$ and $z = (x_4, y_4)$. Show that

(a) $u \cdot v = v \cdot u$,
(b) $(u+v) \cdot z = u \cdot z + v \cdot z$, and
(c) $(\alpha u) \cdot v = \alpha(u \cdot v)$ (where α is some number).

Use result (b) to show that,

(d) $(u + v) \cdot (w + z) = u \cdot w + v \cdot w + u \cdot z + v \cdot z$.

Ex. 4.2 Equate (4.13) and (4.14) to derive the subtraction formula for cosines.

Ex. 4.3 Let $u = (x_1, y_1)$, $v = (x_2, y_2)$. Show that $\|u - v\|^2 = \|u\|^2 - 2u \cdot v + \|v\|^2$.

Ex. 4.4 Referring to the Triangle Inequality as expressed in (4.21), when will equality hold?

Ex. 4.5 Consider the line described by the equation $y = 2 + 3x$. Find a vector equation that describes this line.

4.2 3-Dimensional Vector Spaces

The 3-dimensional Cartesian coordinate system "assigns addresses" to points in 3-dimensional space.[7] We have already used this to visualize functions of two variables. Fig. 4.5(a) shows a 3-dimensional Cartesian coordinate system with axes labelled x, y and z. We include the point $(x, y, z) = (3, 5, 5)$ as an example, together with a few lines to help situate the point. We draw an arrow from the origin to the point to reflect its interpretation as a vector, in the "magnitude with direction" sense.

(a) (b)

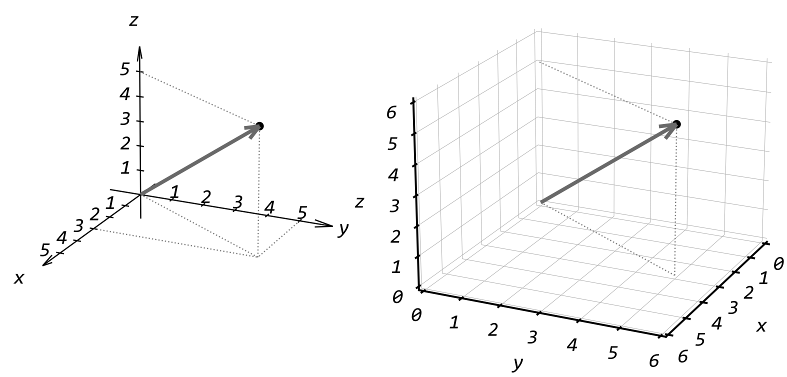

Fig. 4.5. The 3-dimensional Cartesian coordinate system.

Fig. 4.5(a) is the "classic" style of drawing a 3-dimensional diagram, with the axes passing through the origin. By convention, the z-axis points upwards, and if the x-axis points toward you, then the y-axis points to the right. This is the "right-hand rule": if you use your right thumb, index and

[7]There is also a polar coordinate system for 3-dimensional spaces, but we will not need it in this book.

middle fingers to make three perpendicular axes, the z-axis is your thumb, the x-axis is your index finger, and the y-axis is your middle finger. We plot the same diagram in Fig. 4.5(b) following a style preferred in modern computer visualizations. This style also follows the right-hand rule, but the axes "spines" and ticks are placed on the outside of the plot region, and a grid is included on the floor and on the back panels, like a holodeck.

The plots in Fig. 4.5 are viewed at an **elevation** of 20^o and an **azimuth** of 25^o. In computer graphics, azimuth refers to the rotation about the vertical axis. We will take azimuth 0 to be the viewpoint where the x-axis points directly towards you. Positive azimuth is when you move counter clockwise around the figure (or if you stay still and rotate the figure clockwise). An elevation of 90^o means you are looking down from above the z-axis. If you set elevation and azimuth to 0 you will only see a vertical z-axis, and a horizontal y-axis. If you set elevation at 90 and azimuth at -90, you will see only a vertical y axis, and a horizontal x-axis, with the z-axis pointing right at you. Choose elevation and azimuth to maximize readability of your plot.

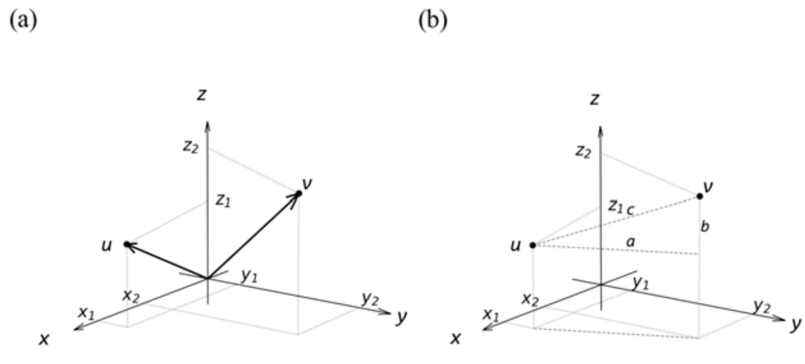

Fig. 4.6. Two 3-dimensional vectors and the distance between them.

We denote the set of all 3-dimensional points in the Cartesian coordinate space by $\mathbb{R}^3$ and refer to the points as vectors. In Fig. 4.6(a) we plot two points $u = (x_1, y_1, z_1)$ and $v = (x_2, y_2, z_2)$ together with the vector arrows. In Fig. 4.6(b), we removed the arrows, added in a few additional lines, and marked the lengths of some of those lines by a, b, and c, with c indicating the distance between u and v. One application of Pythagoras's Theorem shows that $a^2 = (x_2-x_1)^2+(y_2-y_1)^2$. Another application of Pythagoras's Theorem shows that

$$c^2 = a^2 + b^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 .$$

In other words, the distance (or norm) between two points $u = (x_1, y_1, z_1)$

and $v = (x_2, y_2, z_2)$ in 3-dimensional space is

$$\text{distance}(u, v) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}\,.$$

Likewise, the magnitude of a vector $u = (x_1, y_1, z_1)$ is

$$\|u\| = \sqrt{x_1^2 + y_1^2 + z_1^2}\,.$$

Define scalar multiplication and addition in the usual way, i.e., for any vectors $u = (x_1, y_1, z_1)$ and $v = (x_2, y_2, z_2)$ in $\mathbb{R}^3$, define

$$\alpha u = (\alpha x_1, \alpha y_1, \alpha z_1) \quad \text{and} \quad u + v = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

This makes $\mathbb{R}^3$ a *vector space.* We can take linear combinations of vectors in the usual way. We can also define the dot product $u \cdot v = x_1x_2 + y_1y_2 + z_1z_2$. As before, we have

$$\|u\|^2 = u \cdot u\,.$$

You probably can see (in your mind's eye) that:

(a) Scalar multiplication of a non-zero vector by α stretches or shrinks the vector, reversing its direction if α is negative.

(b) If one vector is a scalar multiple of another, they lie on the same line. Taking all linear combinations of the two vectors gives you that line, which passes through the origin. It is a 1-dimensional vector subspace of the 3-dimensional space $\mathbb{R}^3$.

(c) Two non-zero vectors that do not lie on the same line will lie on a plane. A linear combination of the two vectors results in a new vector lying on that same plane. The set of all linear combinations of the two vectors gives you the entire plane, which passes through the origin. It is a 2-dimensional vector subspace of $\mathbb{R}^3$.

(d) If you have two vectors that are not scalar multiples of each other, and you add a third vector that does not lie on the plane spanned by the first two, then you have three **linearly independent** vectors; you cannot express any one of these vectors as a multiple or linear combination of the others. Taking the set of all linear combinations of three linearly independent vectors in $\mathbb{R}^3$ gives you the entire space.

As in the case of 2-dimensional vectors, we say that a set of three 3-dimensional vectors u_1, u_2, u_3 are **linearly independent** if

$$c_1u_1 + c_2u_2 + c_3u_3 = 0 \;\; \Rightarrow c_1 = c_2 = c_3 = 0\,.$$

If this condition holds, then the three vectors will span the entire space $\mathbb{R}^3$. If any of the vectors are zero, or if any of the vectors are multiples of another, or if any of the vectors can be constructed as a linear combination of the others, then this condition will not hold, and the vectors are **linearly dependent**. We can extend this condition to any number of 3-dimensional vectors. Two 3-dimensional vectors can be linearly dependent or linearly independent. Four or more 3-dimensional vectors must be linearly dependent.

Recall that if you have two non-identical 2-dimensional vectors (points), their affine combinations (linear combinations where the coefficients add to one) will form a straight line passing through the two points. The same is true of 3-dimensional vectors:

(e) The affine combinations $(1-\alpha)u + \alpha v$ of two non-identical points u and v will lie on a line passing through the two points.

This gives us a way to describe lines in $\mathbb{R}^3$. Suppose a line passes through the points u and v. Then this line satisfies

$$(x, y, z) = (1-\alpha)u + \alpha v = u + \alpha(v-u)\,.$$

Writing $u = (x_1, y_1, z_1)$ and $v = (x_2, y_2, z_2)$, we can write

$$(x, y, z) = (x_1, y_1, z_1) + \alpha(x_2 - x_1, y_2 - y_1, z_2 - z_1)$$

which is the **vector equation of a line** in $\mathbb{R}^3$. Alternatively, we can express this in the form of the parametric equations

$$x - x_1 = \alpha(x_2 - x_1)\,, \quad y - y_1 = \alpha(y_2 - y_1)\,, \quad z - z_1 = \alpha(z_2 - z_1)\,,\ \alpha \in \mathbb{R}$$

These are straightforward extensions of the vector and parametric equations of a line in the 2-dimensional space $\mathbb{R}^2$.

We reiterate the idea of *vector spaces* and *vector subspaces*. We call $\mathbb{R}^3$, the set of all 3-dimensional vectors, a vector space because if u and v are in $\mathbb{R}^3$, then $c_1 u + c_2 v$ is also in $\mathbb{R}^3$, regardless of what c_1 and c_2 are. A vector subspace is a subset of a vector space, but with the property that if u and v are in the subset, then $c_1 u + c_2 v$ is also in that subset, regardless of the values of c_1 and c_2. For instance, the set of all vectors that make up a line or a plane in $\mathbb{R}^3$ that passes through the origin is a vector subspace of $\mathbb{R}^3$. The set comprising only the point $(0, 0, 0)$ is also a vector subspace of $\mathbb{R}^3$, called the "trivial vector subspace". However, the set of 3-dimensional vectors that make up a line or a plane in $\mathbb{R}^3$ that does *not* pass through the origin is *not* a vector subspace of $\mathbb{R}^3$.

If we have three non-identical 3-dimensional vectors u_1, u_2 and u_3, these vectors form a triangle lying on a plane (see Fig. 4.7).

(f) The set of all affine combinations of these three vectors

$$(1 - \alpha - \beta)u_1 + \alpha u_2 + \beta u_3 \tag{4.23}$$

will give you the plane containing the triangle. If we further restrict the coefficients of the affine combinations to be positive (but still

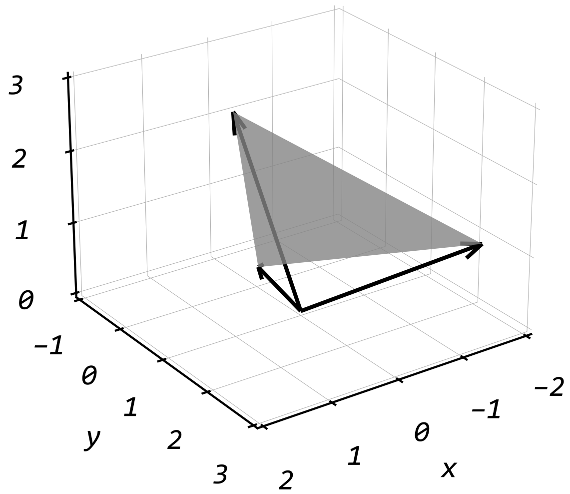

Fig. 4.7. A triangle in 3-dimensional space.

summing to one), then the affine combinations will give you the filled triangle.

Geometric arguments similar to those made in the previous section will reveal that the Law of Cosines will continue to hold (a triangle, after all, is a 2-dimensional object, even in 3-dimensional space). The fact that two vectors are perpendicular if their dot product is zero follows as a result, as does the Triangle Inequality and the Cauchy-Schwarz Inequality.

What is the equation that describes a plane in 3-dimensional space? Imagine a plane (not necessarily passing through the origin), and picture a vector that is perpendicular to this plane (this vector is called the **normal vector** of the plane). Suppose the point (x_0, y_0, z_0) lies on the plane. Then for any point (x, y, z) on the plane, the vector $(x - x_0, y - y_0, z - z_0)$ is parallel to this plane. If (a, b, c) is the normal vector of the plane, then

$$(a, b, c) \cdot (x - x_0, y - y_0, z - z_0) = 0\,.$$

Expanding this, we get $ax + by + cz - (ax_0 + by_0 + cz_0) = 0$. Collecting the constants in the parenthesis into the constant $-d$, we see that the equation of a plane can be written

$$ax + by + cz + d = 0\,.$$

Another way is to view the plane as formed by the affine combinations (4.24), without restrictions on α and β. For any point (x, y, z) on the plane containing the points $u_1 = (x_1, y_1, z_1)$, $u_2 = (x_2, y_2, z_2)$ and $u_3 = (x_3, y_3, z_3)$, we have

$$\begin{aligned}(x, y, z) &= (1 - \alpha - \beta)u_1 + \alpha u_2 + \beta u_3 \\ &= u_1 + \alpha(u_2 - u_1) + \beta(u_3 - u_1) \\ &= (x_1, y_1, z_1) + \alpha(x_2 - x_1, y_2 - y_1, z_2 - z_1) \\ &\qquad\qquad + \beta(x_3 - x_1, y_3 - y_1, z_3 - z_1)\,.\end{aligned} \tag{4.24}$$

We can write this as three separate equations, which gives the parametric equations of a plane:

$$\begin{aligned}x &= x_1 + \alpha(x_2 - x_1) + \beta(x_3 - x_1) \\ y &= y_1 + \alpha(y_2 - y_1) + \beta(y_3 - y_1) \\ z &= z_1 + \alpha(z_2 - z_1) + \beta(z_3 - z_1)\,.\end{aligned}$$

The difference between the parametric equations for a plane in $\mathbb{R}^3$ and a line in $\mathbb{R}^3$ is that the equations for the plane have two parameters α and β whereas the equation for the line has just one parameter. This corresponds to the notion that a line is "1-dimensional" whereas a plane is "2-dimensional".

4.2.1 *Exercises*

Ex. 4.6 Let $u = (1, 1, 1)$, $v = (1, -2, 1)$ and $w = (2, 1, 0)$. Which vector has the largest magnitude? Which pairs of vectors are perpendicular? Find the angle between any pair of non-perpendicular vectors.

Ex. 4.7 Consider a plane in $\mathbb{R}^3$ passing through the points $A = (2, 2, 4)$, $B = (1, 0, 3)$ and $C = (-1, 3, 2)$. Show that the vector $(5, 1, -7)$ is a normal vector to this plane. Find the equation of this plane.

4.3 n-Dimensional Vector Spaces

An n-dimensional vector is an ordered collection of n numbers

$$(x_1, x_2, \dots, x_n)\,.$$

If all of the entries in a vector are zero, we call it the zero vector. Given any two such vectors

$$u = (u_1, u_2, \dots, u_n) \quad \text{and} \quad v = (v_1, v_2, \dots, v_n)\,,$$

define **scalar multiplication** and **vector addition** in the usual way:

$$\begin{aligned}\alpha u &= (\alpha u_1, \alpha u_2, \dots, \alpha u_n)\,, \\ u + v &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)\,.\end{aligned}$$

We can take linear combinations in the usual way. The set of all n-dimensional vectors together with the scalar multiplication and vector addition definitions make up the n-dimensional vector space, denoted $\mathbb{R}^n$.

We can also define the dot product

$$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \,.$$

The dot product continues to have the following properties, presented earlier for 2- and 3-dimensional vectors.

Theorem 4.1 (The Dot Product) *Let u, v and w be arbitrary n-dimensional vectors. Then*

(a) $u \cdot u \geq 0$ *with* $u \cdot u = 0$ *if and only if* u *is the zero vector,*

(b) $u \cdot v = v \cdot u$,

(c) $(u + v) \cdot w = u \cdot w + v \cdot w$,

(d) $(\alpha u) \cdot v = \alpha(u \cdot v)$.

The proofs are straightforward extensions of the proofs in the 2-dimensional case.

We relied heavily on geometric arguments when discussing 2- and 3-dimensional vector spaces. For instance, the proof that the dot product of two 2-dimensional unit vectors is the cosine of the angle between them made direct use of geometric arguments. But what geometric interpretations can we give to ideas such as the distance between two n-dimensional points and the angle between n-dimensional vectors when $n > 3$? We are not going to try to give *literal* geometric interpretations to such concepts in higher-dimensional spaces, but it turns out that results like the Triangle Inequality and the Law of Cosines still hold, and that we can still usefully make use of geometric *intuition* gained from 2- and 3-dimensional spaces.

As an example, suppose we extend the formulas for the geometric distance between two points in $\mathbb{R}^3$ to $\mathbb{R}^n$ in the "obvious way". Does

$$\|u - v\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2} \tag{4.25}$$

have a meaningful interpretation as the distance between two n-dimensional vectors u and v when $n > 3$? Can we interpret

$$\|u\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} \tag{4.26}$$

as the "magnitude" of an n-dimensional vector when $n > 3$?

The answer is yes, in the sense that the "distance" concept given by (4.25) can be shown to share all of the essential properties of geometric distance in 2- and 3-dimensions, namely: (i) it is never negative, (ii) the distance from u to v is the same as the distance from v to u, and (iii) it satisfies the Triangle Inequality. This means that we can treat (4.25) very much like a "distance" between n-dimensional points, even if we cannot literally measure this distance with a ruler. Likewise, $\|u\|$ can be interpreted as the "magnitude" or "norm" of the vector u.

The non-negativity and symmetry properties (i) and (ii) obviously hold. The key to showing that the Triangle Inequality (iii) also holds in $\mathbb{R}^n$ lies in the Cauchy-Schwarz Inequality. Earlier we derived this inequality for 2-dimensional vectors using geometric arguments. It turns out that there is a purely algebraic proof that does not rely on geometric arguments at all, and which holds for general n-dimensional vector spaces.

Theorem 4.2 (The Cauchy-Schwarz Inequality) *For any two n-dimensional vectors u and v, we have*

$$|u \cdot v| \leq \|u\| \, \|v\| \,. \tag{4.27}$$

Equality holds if and only if $u = \alpha v$ or one of the vectors is a zero vector.

Proof: If one of the vectors is a zero vector, then (4.27) obviously holds trivially as an equality. If $u = \alpha v$, then $|u \cdot v| = |\alpha v \cdot v| = |\alpha| \|v\|^2$ and $\|u\| \, \|v\| = |\alpha| \|v\|^2$, so again (4.27) holds with equality. Now suppose neither u or v is a zero vector, and that $u \neq \alpha v$ for any α. Then we have

$$0 < (u - \alpha v) \cdot (u - \alpha v) = u \cdot u - 2\alpha u \cdot v + \alpha^2 v \cdot v \,. \tag{4.28}$$

The inequality (4.28) holds for all α. Evaluating it at the particular value

$$\alpha = \frac{u \cdot v}{v \cdot v} \,,$$

we get $0 < u \cdot u - 2\dfrac{u \cdot v}{v \cdot v} u \cdot v + \dfrac{(u \cdot v)^2}{(v \cdot v)^2} v \cdot v = u \cdot u - \dfrac{(u \cdot v)^2}{v \cdot v}$. Rearranging this inequality gives

$$(u \cdot v)^2 < (u \cdot u)(v \cdot v) \,.$$

Taking square root gives us our desired result.

We use the Cauchy-Schwarz inequality to show that the distance measure (4.25) satisfies the Triangle Inequality.

Theorem 4.3 (The Triangle Inequality) *For any two n-dimensional vectors u and v, we have*

$$\|u - v\| \leq \|u\| + \|v\| \,.$$

Proof: We first note that

$$\|u - v\|^2 = (u - v) \cdot (u - v) = u \cdot u - 2u \cdot v + v \cdot v = \|u\|^2 - 2u \cdot v + \|v\|^2 \,.$$

Since $u \cdot v \leq |u \cdot v| \leq \|u\| \, \|v\|$, the largest value on the RHS occurs when $u \cdot v = -\|u\| \, \|v\|$, therefore

$$\|u - v\|^2 \leq \|u\|^2 + 2\|u\| \, \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \,.$$

Taking square roots gives us the desired result.

Let x, y and z be any three points in n-dimensional space. Setting $u = x - y$ and $v = z - y$, we get

$$\|x - z\| \leq \|x - y\| + \|z - y\| \,.$$

The "distance" between x and z is no greater than the sum of the distances between x and y, and y and z.

For $\mathbb{R}^2$ and $\mathbb{R}^3$ we derived the Law of Cosines $u \cdot v = \|u\| \, \|v\| \cos\theta$ geometrically, and then obtained the Cauchy-Schwarz inequality from it. Here we derived the Cauchy-Schwarz inequality $u \cdot v \leq \|u\| \, \|v\|$ directly, without appealing to geometric arguments. Is there a Law of Cosines in n-dimensions? Can we even talk about the angle between two n-dimensional vectors when $n > 3$? If we take advantage of the fact that

$$-1 \leq \frac{u \cdot v}{\|u\| \, \|v\|} \leq 1 \,,$$

we can take the bold step of *defining* the angle between two n-dimensional vectors to be value $\theta \in [0, \pi]$ such that

$$\cos\theta = \frac{u \cdot v}{\|u\| \, \|v\|} \quad \text{or} \quad \theta = \cos^{-1}\left(\frac{u \cdot v}{\|u\| \, \|v\|}\right) .$$

This provides a notion of perpendicularity for n-dimensional vectors. We say that two n-dimensional vectors u and v are **orthogonal** if $u \cdot v = 0$ (we reserve the word 'perpendicular' for the 2- and 3-dimensional cases).

Given a set of n-dimensional vectors, we can ask if the vectors span the entire $\mathbb{R}^n$ space, or a lower dimensional vector subspace of $\mathbb{R}^n$. For instance, the vectors $u_1 = (1, 2, 4, 3)$ and $v_1 = (2, 4, 8, 6)$ span only a one-dimensional "line" in $\mathbb{R}^4$, since v_1 is a multiple of u_1, and so every linear combination of u_1 and v_1 is also a multiple of u_1. As another example, the vectors $u_2 = (1, 0, 0, 0)$ and $v_2 = (0, 1, 0, 0)$ span a two-dimensional "plane" in $\mathbb{R}^4$. In order to span the entire $\mathbb{R}^4$ space, we need at least four linearly independent vectors, i.e., four vectors u_1, u_2, u_3 and u_4 such that

$$c_1 u_1 + c_2 u_2 + c_3 u_3 + c_4 u_4 = 0 \quad \Rightarrow \quad c_1 = c_2 = c_3 = c_4 = 0 \,.$$

One question is how to determine the dimension of the subspace spanned by a set of vectors. We discuss this briefly in the next section.

Finally, we note that up to now we have treated vectors as simply an ordered collection of n numbers, with no "shape" to it. Sometimes it is useful to put a shape to the vector. For instance, we call u in (4.29) a column vector, and v in (4.29) a row vector

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \,, \quad v = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \,. \tag{4.29}$$

It does not matter for our discussion if the vector is shaped or not. In the next section we will use row and column vectors explicitly, but the shape does not change any of the concepts we have discussed so far.

4.3.1 *Exercises*

Ex. 4.8 Let $\{x_i, y_i\}_{i=1}^n$ be a sample of n observations on two variables. The sample correlation between x_i and y_i is

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^n (x_i - \overline{x})^2}\sqrt{\sum_{i=1}^n (y_i - \overline{y})^2}}$$

where $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$ and $\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$.

(a) Explain why it must be that $-1 \leq r_{xy} \leq 1$. *Hint: Let*

$$u = (x_1 - \overline{x}, x_2 - \overline{x}, \dots, x_n - \overline{x}) \text{ and } v = (y_1 - \overline{y}, y_2 - \overline{y}, \dots, y_n - \overline{y})\,.$$

(b) What is the angle between the vectors u and v as defined in the hint for part (a) if (i) $r_{xy} = 1$, (ii) $r_{xy} = 0$, (iii) $r_{xy} = -1$?

Ex. 4.9 Let u and v be n-dimensional vectors. Find a scalar α and a vector w orthogonal to v such that

$$u = \alpha\frac{v}{\|v\|} + w\,.$$

Find an expression for $\alpha v/\|v\|$. Illustrate this result for the special case $n = 2$. *Remark: the vector $\alpha v/\|v\|$ is called the* **projection** *of u onto v (or more accurately, the subspace spanned by v).*

Ex. 4.10 Let $u = (u_1, u_2)$ and $v = (v_1, v_2)$. Show that the area of the parallelogram formed by these two vectors is $|u_1v_2 - u_2v_1|$.

4.4 Linear Systems of Equations in Three or more Unknowns

In the previous chapter we discussed solving systems of equations in two unknowns. In this section we extend the discussion to systems of equations in three or more unknowns, and connect the problem of solving systems of linear equations with the vector space concept.

4.4.1 *Linear Systems in Three Unknowns*

Suppose we have two linear equations in three unknowns x, y and z:

$$\begin{array}{ccccccc} a_{11}x & + & a_{12}y & + & a_{13}z & = & b_1 \\ a_{21}x & + & a_{22}y & + & a_{23}z & = & b_2 \end{array} \tag{4.30}$$

and we wish to solve this system, i.e., we wish to find all points (x, y, z) that satisfy the two equations simultaneously. We know that the graph of each of the two equations in (4.30) is a plane in 3-dimensional Cartesian space. With two planes, there are three possibilities: the planes are parallel in which case there are no solutions; the planes coincide in which case every point on the (common plane) is a solution; the planes intersect in which case every point on the *line* of intersection is a solution. So there are either no solutions, or an infinite number of solutions. In the latter case, the infinite solutions either make up a plane, or a line in 3-dimensional space.

Example 4.1 The following system has infinitely many solutions represented by a single line:

$$\begin{array}{rcrcrcl} 2x & - & 3y & + & z & = & 0 \\ x & + & y & + & z & = & 1\,. \end{array} \tag{4.31}$$

We solve this using Gaussian elimination:

$$\left[\begin{array}{ccc|c} \boxed{2} & -3 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{array}\right] \xrightarrow[{[2]=[2]-\frac{1}{2}[1]}]{} \left[\begin{array}{ccc|c} \boxed{2} & -3 & 1 & 0 \\ 0 & \boxed{5/2} & \frac{1}{2} & 1 \end{array}\right].$$

In each rectangular array, each column on the left of the vertical line represents the coefficients on x, y, and z respectively. The entries on the right of the vertical line are the constants on the right-hand-side of the system. The top-left entry is the *pivot* for x, and must be non-zero; if it is zero, re-order the equations so the top-left entry is not zero. Use the x-pivot and row operation to create zeros underneath it; in this case, subtract $1/2$ of the first row from the second row. If the $(2,2)$th element is non-zero, it becomes the y-pivot. There is no z-pivot; it is a "free" variable.

Let z be some "parameter" s. Then we have

$$\frac{5}{2}y + \frac{1}{2}s = 1 \quad \text{or} \quad y = \frac{2}{5} - \frac{1}{5}s\,.$$

The first row represents the equation $2x-3y+z=0$. Substituting in $z=s$ and $y=\frac{2}{5}-\frac{1}{5}s$ gives

$$x = \frac{3}{5} - \frac{4}{5}s\,.$$

The solution is the line comprising the points $(x,y,z) = (\frac{3}{5}-\frac{4}{5}s, \frac{2}{5}-\frac{1}{5}s, s)$.

The next example is similar, except that the pivots are associated with the first (x) and third (z) variables.

Example 4.2 Consider the system

$$\begin{array}{rcrcrcl} 2x & + & 2y & + & z & = & 2 \\ x & + & y & + & z & = & 0\,. \end{array} \tag{4.32}$$

Solving by Gaussian elimination gives

$$\left[\begin{array}{ccc|c} \boxed{2} & 2 & 1 & 2 \\ 1 & 1 & 1 & 0 \end{array}\right] \xrightarrow[{[2]=[2]-\frac{1}{2}[1]}]{} \left[\begin{array}{ccc|c} \boxed{2} & 2 & 1 & 2 \\ 0 & 0 & \boxed{1/2} & -1 \end{array}\right].$$

The $(2,2)$th element in the array on the right is zero, so the $(2,3)$th element becomes the second pivot. There is no y-pivot, y is the "free" variable. The second row says

$$\frac{z}{2} = -1 \quad \text{or} \quad z = -2\,.$$

Let y be some "parameter" s. The first row represents the equation $2x + 2y + z = 2$. Substituting in $y = s$ and $z = -2$ gives

$$x = 1 - s + 1 = 2 - s\,.$$

The solutions are the points $(x, y, z) = (2-s, s, -2)$, which makes up a line in 3-dimensional Cartesian space.

In Example 4.1 and Example 4.2 there are three variables and two pivots, leading to one free variable, so the solutions are the points that comprise a line. If the two planes representing the equations coincide, then every point on the common plane will be a solution. There will only be one pivot, and the second row will reduce to $0 = 0$. If the two planes are parallel, there will again only be one pivot, but the second equation reduces to a contradiction .

With three equations in three unknowns

$$\begin{aligned} a_{11}\,x + a_{12}\,y + a_{13}\,z &= b_1 \\ a_{21}\,x + a_{22}\,y + a_{23}\,z &= b_2 \\ a_{31}\,x + a_{32}\,y + a_{33}\,z &= b_3 \end{aligned} \tag{4.33}$$

we have three planes, and there are several possibilities: (a) the three equations intersect at a single point (imagine two of the planes intersecting at a line, and the third plane cutting this line). Then there is a single point that satisfies all three equations, that point being the solution; (b) the three equations could all intersect at a line — this can happen if two of the planes intersect at a line and the third plane is either coincident with one of the previous two planes, or is not coincident with either, yet cuts the first two planes exactly along their line of intersection. In this case the solution is the common line of intersection; (c) all three planes coincide, in which case the solution are all the points on the common plane, (d) the three planes do not all coincide, and do not all intersect at a common line or point.

For (a), Gaussian elimination will lead to three pivots. Three unknowns and three pivots lead to a unique solution. In the case of (b), there will be two pivots and no contradiction; the last row will be $0 = 0$. Three unknowns, two pivots, no contradiction implies the solutions make up a line. For (c), there will be one pivot an no contradictions, and the last two rows will be $0 = 0$. For (d), there will be fewer than three pivots, and a contradiction.

Example 4.3 Consider the system

$$\begin{aligned} 2x + 2y + 4z &= 4 \\ 3x + y + 2z &= 2 \\ 5x + 2y + z &= 7\,. \end{aligned} \tag{4.34}$$

Solving by Gaussian elimination gives

$$\left[\begin{array}{ccc|c} \boxed{2} & 2 & 4 & 4 \\ 3 & 1 & 2 & 2 \\ 5 & 2 & 1 & 7 \end{array}\right] \xrightarrow[{[2]=[2]-\frac{3}{2}[1]}]{} \left[\begin{array}{ccc|c} \boxed{2} & 2 & 4 & 4 \\ 0 & -2 & -4 & -4 \\ 5 & 2 & 1 & 7 \end{array}\right]$$

$$\xrightarrow[{[3]=[3]-\frac{5}{2}[1]}]{} \left[\begin{array}{ccc|c} \boxed{2} & 2 & 4 & 4 \\ 0 & \boxed{-2} & -4 & -4 \\ 0 & -3 & -9 & -3 \end{array}\right] \xrightarrow[{[3]=[3]+\frac{3}{2}[1]}]{} \left[\begin{array}{ccc|c} \boxed{2} & 2 & 4 & 4 \\ 0 & \boxed{-2} & -4 & -4 \\ 0 & 0 & \boxed{-3} & 3 \end{array}\right].$$

Substituting backwards, you can easily show that the unique solution is $(x, y, z) = (0, 4, -1)$.

Example 4.4 The system

$$\begin{aligned} x + y + 4z &= 4 \\ 3x + y + 2z &= 2 \\ 2x + y + 3z &= 3 \end{aligned} \tag{4.35}$$

has an infinite number of solutions, forming a (1-dimensional) line. Gaussian elimination, using the row operations $[2] = [2] - 3[1]$, $[3] = [3] - 2[1]$ and $[3] = [3] - \frac{1}{2}[2]$ yields

$$\left[\begin{array}{ccc|c} \boxed{1} & 1 & 4 & 4 \\ 3 & 1 & 2 & 2 \\ 2 & 1 & 3 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} \boxed{1} & 1 & 4 & 4 \\ 0 & \boxed{-2} & -10 & -10 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

There is no pivot associated with variable z. Assign a free parameter s to z. Then second row is $-2y - 10z = -10$, which gives $y = 5 - 5s$. The first row is the equation $x + y + 4z = 4$. Substituting for y and z gives $x = s - 1$. The solutions lie on the line $(x, y, z) = (s - 1, 5 - 5s, s)$.

Example 4.5 Consider the system

$$\begin{aligned} x + y + 4z &= 4 \\ 3x + y + 2z &= 2 \\ 2x + y + 3z &= 4\,. \end{aligned} \tag{4.36}$$

Using exactly the same elimination steps as in Example 4.4, we get

$$\left[\begin{array}{ccc|c} \boxed{1} & 1 & 4 & 4 \\ 3 & 1 & 2 & 2 \\ 2 & 1 & 3 & 4 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} \boxed{1} & 1 & 4 & 4 \\ 0 & \boxed{-2} & -10 & -10 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

There is therefore no solution.

4.4.2 *Larger Systems*

The general ideas carry over to larger systems (four equations in three unknowns, systems with four or more unknowns, ...) even though it is harder to describe the solutions geometrically. Suppose you have a system of m equations in n unknowns, as in:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m\,. \end{aligned} \tag{4.37}$$

If $b_1 = b_2 = \cdots = b_m = 0$, we call this a **homogeneous** system.

To use Gaussian elimination to solve this system, collect all of the constants into a rectangular array as in

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right].$$

The rectangular block on the left of the vertical line is called the “coefficient matrix”. The entire rectangular array is called the “augmented matrix”. Then carry out Gaussian elimination on the augmented matrix so as to reduce the coefficient matrix (the left part of the augmented matrix) into **row echelon form** (REF), meaning that

- all zero rows are at the bottom,
- the first non-zero item in each row (the “leading term”) is to the right of the leading terms of all of the rows above.

This is what we have done in all our earlier Gaussian elimination examples. The leading terms in the REF of the coefficient matrix are the pivots. Columns without pivots correspond to “free variables”. The final solution is obtained by backward substitution, assigning a “parameter” to each of the “free” variables. There will be a unique solution (which may be the trivial solution if the system is homogeneous) if there are as many pivots as there are variables (columns) in the augmented matrix and no contradictions. If there are fewer pivots than number of variables, then there will be an infinite number of solutions, the dimension of the solution being equal to the difference between the number of variables and the number of pivots (i.e., the number of "free variables").

Of course, any fully zero rows in the REF of the coefficient matrix on the left must be matched with zeros on the right, otherwise there are no solutions (these would be the "contradictions" found in earlier examples where there were no solutions).

The following is an example of a system of four linear equations in four unknowns, solved by Gaussian elimination. The solution also differs from previous examples in that we require swapping equations at some stage.

Example 4.6 Consider the system

$$\begin{aligned} w + \ \ x + 2y + 2z &= 1 \\ w + \ \ x + 2y + \ \ z &= 2 \\ 2w + 3x + 3y + \ \ z &= 1 \\ w + 3x + \ \ y + 2z &= 1\,. \end{aligned}$$

Solving by Gaussian Elimination:

$$\left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 1 & 1 & 2 & 1 & 2 \\ 2 & 3 & 3 & 1 & 1 \\ 1 & 3 & 1 & 2 & 1 \end{array}\right] \xrightarrow[\substack{[2]=[2]-\ [1] \\ [3]=[3]-2[1] \\ [4]=[4]-\ [1]}]{} \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & -3 & -1 \\ 0 & 2 & -1 & 0 & 0 \end{array}\right]$$

The next step would be to use the $(2,2)$th element to eliminate the $(3,2)$th and $(4,2)$th elements, but we can't do this, since the $(2,2)$th element is zero. There is no pivot there. This situation can be dealt with by swapping the order of the equations. We'll swap the second and fourth rows:

$$\left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & -3 & -1 \\ 0 & 2 & -1 & 0 & 0 \end{array}\right] \xrightarrow[{[2]\leftrightarrow[4]}]{} \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & \boxed{2} & -1 & 0 & 0 \\ 0 & 1 & -1 & -3 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{array}\right].$$

Now continuing with elimination gives

$$\left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & \boxed{2} & -1 & 0 & 0 \\ 0 & 1 & -1 & -3 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{array}\right] \xrightarrow[{[3]=[3]-\frac{1}{2}[2]}]{} \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & \boxed{2} & -1 & 0 & 0 \\ 0 & 0 & \boxed{-1/2} & -3 & -1 \\ 0 & 0 & 0 & \boxed{-1} & 1 \end{array}\right].$$

There are four pivots, so there is a unique solution. You should be able to find by substitution that the solution is $(w, x, y, z) = (-17, 4, 8, -1)$.

Example 4.7 Consider the system of three equations in four unknowns

$$\begin{aligned} w + x + 2y + 2z &= 1 \\ w + x + 2y + z &= 2 \\ 2w + 2x + 3y + z &= 1\,. \end{aligned}$$

Solving by Gaussian Elimination:

$$\left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 1 & 1 & 2 & 1 & 2 \\ 2 & 2 & 3 & 1 & 1 \end{array}\right] \xrightarrow[\substack{[2]=[2]-[1] \\ [3]=[3]-2[1]}]{} \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & -3 & -1 \end{array}\right].$$

We cannot find a pivot for the second column. Moving on to the third column, we find that we have to swap the second and third rows in order to get a pivot. Doing so we find we get a REF without further elimination:

$$\left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & -3 & -1 \end{array}\right] \xrightarrow[{[3] \leftrightarrow [2]}]{} \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 2 & 1 \\ 0 & 0 & \boxed{-1} & -3 & -1 \\ 0 & 0 & 0 & \boxed{-1} & 1 \end{array}\right].$$

The last row is $-z = 1$, so $z = -1$. Substituting into $-y - 3z = -1$ gives $y = 4$. Assign a parameter s to x. The first row says $w + x + 2y + 2z = 1$. Substituting our solutions (and parameters) obtained to this point gives $w = 1 - s - 2(4) - 2(-1) = -5 - s$. The solutions are $(w, x, y, z) = (-5 - s, s, 4, -1)$ which forms a "line" in $\mathbb{R}^4$.

Note that the row echelon form for a system of equations is not unique. If you start with the same equations, but arranged in a different order, you will get a different row echelon form. The ordering of the equations do not change the system, of course, and the solution that you get at the end will be the same.

Notice finally that in our examples, the number of free variables in our solution is the number of variables n minus the number of pivots r found when the coefficient matrix is reduced to row echelon form.

4.4.3 *Systems of Linear Equations and Vector Spaces*

There is a strong and important connection between systems of linear equations and the vector space concept. Consider the system of two equations in two unknowns

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2\,. \end{aligned} \tag{4.38}$$

Earlier we described the solution of such a system as the intersection of two lines. If the two lines intersect, there is a unique solution. If the two lines are parallel, there are no solutions. If the lines coincide, there are an infinite number of solutions.

An alternative view is to think of a system of equations in terms of vectors. In view of what is to come in later chapters, re-write vectors into "column vector" form, i.e., write a vector

$$a_1 = (a_{11}, a_{21}) \quad \text{as} \quad a_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} .$$

Then write (4.38) as

$$\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} x + \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} y = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} , \tag{4.39}$$

or

$$a_1 x + a_2 y = b$$

Then solving the system means finding the values x and y so that the linear combination $a_1 x + a_2 y$ results in the vector b. If a_1 and a_2 span the whole $\mathbb{R}^2$ space, then we must be able to obtain b as a linear combination of a_1 and a_2. If there are no solutions or an infinite number of solutions, then it must be that a_1 and a_2 does not span the entire space. If this is the case, then there are no solutions if b lies outside of the subspace spanned by a_1 and a_2, and there are infinitely many solutions if b lies in the subspace spanned by a_1 and a_2.

We can think of the general system with m equations in n unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{4.40}$$

in a similar way, and write (4.40) as

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{4.41}$$

or

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$$

where a_1, a_2, ..., a_n and b are the m-dimensional column vectors in (4.41). Then we see that finding a solution to the system (4.40) is equivalent to finding a n-dimensional vector $x = (x_1, x_2, ..., x_n)$ such that (4.41) is satisfied. The $a_1, a_2, \ldots, a_n$ span a certain subspace of $\mathbb{R}^m$. If b is in this subspace, then there will be at least one solution. If not, there are no solutions.

There is much much more to the connection between vector spaces and solving systems of equations. We will explore these connections and their implications in greater detail in Chapter 8 and Chapter 10.

4.4.4 *Exercises*

Ex. 4.11 We learnt in Section 2.1.5 that system (i) below has an infinite number of solutions, whereas system (ii) has no solution:

$$\text{(i)} \quad \begin{aligned} 2x + y &= 3 \\ 4x + 2y &= 6 \end{aligned} \qquad \text{(ii)} \quad \begin{aligned} 2x + y &= 3 \\ 4x + 2y &= 4\,. \end{aligned}$$

Write these two systems in vector form, where the vector containing the constants on the RHS of the system is expressed as a linear combination of two vectors, with weights x and y. Explain using the language of vectors why (i) has an infinite number of solutions whereas (ii) has no solution.

Ex. 4.12 Describe the graphs of the two equations in the system

$$\begin{array}{ccccccc} 2x & + & 2y & + & z & = & 2 \\ 4x & + & 4y & + & 2z & = & 0\,. \end{array}$$

Show that Gaussian elimination leads to a contradiction.

Ex. 4.13 Solve the following system by Gaussian elimination:

$$\begin{aligned} w + x + 2y + 2z &= 1 \\ w + x + 2y + z &= 2 \\ 2w + 2x + 4y + 3z &= 3 \\ w + 3x + y + 2z &= 1\,. \end{aligned}$$

[redacted]

4.7 Solutions to Exercises

Ex. 4.1: (a) $u \cdot v = (x_1x_2, y_1y_2) = (x_2x_1, y_2y_1) = v \cdot u$.

(b) $(x_1{+}x_2, y_1{+}y_2){\cdot}(x_3, x_3) = ((x_1{+}x_2)x_3, (y_1{+}y_2)y_3) = (x_1x_3{+}x_2x_3, y_1y_3{+}y_2y_3) = (x_1x_3, y_1y_3) + (x_2x_3, y_2y_3)$.

(c) $(\alpha x_1, \alpha y_1) \cdot (x_2, y_2) = (\alpha x_1x_2, \alpha y_1y_2) = \alpha(x_1x_2, y_1y_2)$.

(d) $(u + v) \cdot (w + z) = u \cdot (w + z) + v \cdot (w + z) = u \cdot w + v \cdot w + u \cdot z + v \cdot z$.

Ex. 4.2: We have

$$(\cos\alpha - \cos\beta)^2 + (\sin\alpha - \sin\beta)^2 = (\cos(\alpha-\beta) - 1)^2 + \sin^2(\alpha - \beta) \qquad (4.42)$$

Expanding the LHS of (4.42) and using the fact that $\cos^2\gamma + \sin^2\gamma = 1$ gives

$$\begin{aligned} & \cos^2\alpha - 2\cos\alpha\cos\beta + \cos^2\beta + \sin^2\alpha - 2\sin\alpha\sin\beta + \sin^2\beta \\ = \;& 2 - 2\cos\alpha\cos\beta - 2\sin\alpha\sin\beta\,. \end{aligned}$$

Likewise, the RHS of (4.42) gives

$$\cos^2(\alpha-\beta) - 2\cos(\alpha+\beta) + 1 + \sin^2(\alpha-\beta) = 2 - 2\cos(\alpha-\beta)\,.$$

Equating and simplifying the expanded and simplified versions of the LHS and RHS of (4.42) gives the subtraction formula for cosines.

Ex. 4.3: $\|u + v\|^2 = (u + v) \cdot (u + v) = u \cdot u + 2u \cdot v + v \cdot v = \|u\|^2 + 2u \cdot v + \|v\|^2$.

Ex. 4.4: When point B lies on the line segment joining points A and C.

Ex. 4.5: To find the vector equation

$$(x, y) = (u_x, u_y) + \alpha(w_x, w_y)\,,\ \ \alpha \in \mathbb{R}\,.$$

of a line given by $y = 2 + 3x$, we need a point (u_x, u_y) that lies on the line, and a vector (w_x, w_y) that is parallel to the line. We can pick any point on the line; the y-intercept $(0, 2)$ seems convenient. The slope of the line is w_y/w_x is 3, so we can choose $(w_x, w_y) = (1, 3)$. A vector equation of the line is then

$$(x, y) = (0, 2) + \alpha(1, 3)\,,\ \ \alpha \in \mathbb{R}\,.$$

Ex. 4.6: Given $u = (1, 1, 1)$, $v = (1, -2, 1)$ and $w = (2, 1, 0)$, we have $\|u\| = \sqrt{3}$, $\|v\| = \sqrt{6}$ and $\|w\| = \sqrt{5}$, so v has the largest norm, or magnitude. Since $u \cdot v = 1 - 2 + 1 = 0$, $u \cdot w = 2 + 1 + 0 = 3$ and $v \cdot w = 2 - 2 + 0 = 0$, the pairs $\{u, v\}$ and $\{v, w\}$ are perpendicular, whereas the vectors $\{u, w\}$ are not perpendicular. If θ is the angle between u and w, then $\cos\theta = 3/(\sqrt{3}\sqrt{5}) = \sqrt{3}/\sqrt{5}$, so $\theta \approx 0.685$.

Ex. 4.7: Since A, B and C are points on the plane, the vectors $A - B = (1, 2, 1)$ and $A - C = (3, -1, 2)$ lie on the plane. Since these two vectors are not linear dependent (one is not a multiple of the other), they completely identify the plane. To show the $(5, 1, -7)$ is the normal vector to the plane, we only need to show that it is orthogonal to $A - B$ and $A - C$, which it is, since $(1, 2, 1) \cdot (5, 1, -7) = 5 + 2 - 7 = 0$ and $(3, -1, 2) \cdot (5, 1, -7) = 15 - 1 - 14 = 0$. Now let (x, y, z) be any point on the plane. Then $(x, y, z) - (1, 0, 3)$ is orthogonal to $(5, 1, -7)$. (We could have picked any of the three points; we chose B). That is,

$$(5, 1, -7) \cdot (x - 1, y, z - 3) = 5x + y - 7z + 16 = 0\,.$$

Ex. 4.8: Let u and v be the n-dimensional vectors as given in the hint. Then

$$r_{xy} = \frac{u \cdot v}{\|u\| \, \|v\|} \, .$$

(a) $-1 \leq r_{xy} \leq 1$ follows immediately from the Cauchy-Schwarz Inequality.

(b) We have

$$\cos\theta = \frac{u \cdot v}{\|u\| \, \|v\|} = r_{xy} \;\text{ for }\; 0 \leq \theta \leq \pi \, .$$

If $r_{xy} = 1$, then $\theta = 0$ (the vectors u and v have the same direction). If $r_{xy} = 0$, then $\theta = \pi/2$ (the vectors u and v are orthogonal). If $r_{xy} = -1$, then $\theta = \pi$ (the vectors point in opposite directions).

Ex. 4.9: We want scalar α and vector w such that $u = \alpha \dfrac{v}{\|v\|} + w$ and $v \cdot w = 0$. Taking the dot product of both sides of the first equality with v gives

$$u \cdot v = \alpha \frac{v \cdot v}{\|v\|} + w \cdot v.$$

Setting $w \cdot v = 0$ gives $\alpha = \dfrac{u \cdot v}{\|v\|} = \|u\| \cos\theta$ where θ is the angle between the two vectors. The expression for the projection is then

$$\alpha \frac{v}{\|v\|} = \frac{u \cdot v}{\|v\|} \, \frac{v}{\|v\|} = \frac{u \cdot v}{v \cdot v} \, v \, .$$

Finally, we have $w = u - \alpha v$. Illustration for $n = 2$ case below (the completion of the parallelogram is for the next question).

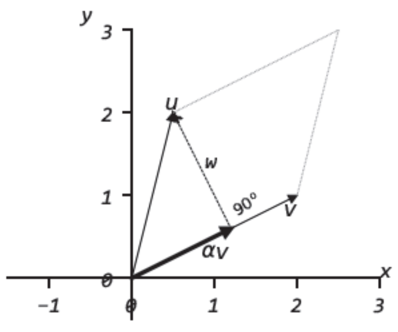

Fig. 4.8. Projecting u onto v

Ex. 4.10: The area of a parallelogram is base times height. See Fig. 4.8. If we taken the length of vector v as the base, then the height is the length of vector

w. Let A be the area. Then

$$\begin{aligned}
A^2 = \|v\|\|w\| = (v\cdot v)(w\cdot w) &= (v\cdot v)\left[\left(u - \frac{u\cdot v}{v\cdot v}v\right)\cdot\left(u-\frac{u\cdot v}{v\cdot v}v\right)\right]\\
&= (v\cdot v)\left(u\cdot u - 2\frac{u\cdot v}{v\cdot v}u\cdot v + \frac{(u\cdot v)^2}{(v\cdot v)^2}v\cdot v\right)\\
&= (v\cdot v)(u\cdot u) - (u\cdot v)^2\\
&= (v_1^2+v_2^2)(u_1^2+u_2^2) - (u_1v_1+u_2v_2)^2\\
&= u_1^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 + u_2^2v_2^2 - u_1^2v_1^2 - 2u_1u_2v_1v_2 - u_2^2v_2^2\\
&= u_1^2v_2^2 + u_2^2v_1^2 - 2u_1u_2v_1v_2\\
&= (u_1v_2 - u_2v_1)^2
\end{aligned}$$

Therefore $A = |u_1v_2 - u_2v_1|$. The expression $u_1v_2 - u_2v_1$ is the **determinant** of the matrix

$$\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}.$$

We study determinants in more detail in Chapter 8.

Ex. 4.11: The two systems can be written as

$$\text{(i)}\ \ (2,4)x + (1,2)y = (3,6) \qquad \text{and} \qquad \text{(ii)}\ \ (2,4)x + (1,2)y = (3,4)\,.$$

The vectors $(2,4)$ and $(1,2)$ lie on the same line, and span only a one-dimension subspace. The vector $(3,6)$ lies on this subspace, therefore we can find infinite number of weights (x,y) such that the equation in (i) holds. The vector $(3,4)$ does not lie on the subspace spanned by the vectors $(2,4)$ and $(1,2)$, therefore we cannot find weights (x,y) such that the equation in (ii) holds.

Ex. 4.12: The graphs of the two equations in the system

$$\begin{array}{ccccccc}
2x & + & 2y & + & z & = & 2\\
4x & + & 4y & + & 2z & = & 0\,.
\end{array}$$

are parallel planes, one passes through the origin, and the vertical distance between the two plans is 2 units. Gaussian elimination leads to

$$\begin{array}{ccccccc}
2x & + & 2y & + & z & = & 2\\
0 & + & 0 & + & 0 & = & -4\,.
\end{array}$$

The second equation is a "contradiction".

Ex. 4.13: You should find only three pivots (associated with variables w, x and z) and no contradiction. A possible row echelon form is

$$\left[\begin{array}{cccc|c}
\boxed{1} & 0 & \frac{5}{2} & 0 & 3\\
0 & \boxed{2} & -1 & 0 & 0\\
0 & 0 & 0 & \boxed{-1} & 1\\
0 & 0 & 0 & 0 & 0
\end{array}\right].$$

The solution is $(w,x,y,z) = (3 - \frac{5}{2}s, \frac{s}{2}, s, -1)$