A First Course in

Mathematics for Computing

Ludmila I Kuncheva

ISBN 978–1–291–78563–0

Preface

How do I teach mathematics to students who believe that minus two squared equals minus four? And when I challenge their stunning, world-changing discovery, they show me the result on their calculator... Students who think that the word “logarithm” is misspelt “algorithm”? Is it any surprise then that, in one of the lecture rooms, a “well-wisher” (a student) engraved on the desk ”Lecturer Destruct Button - Please Press”? Well... thank you. Luckily, the room was refurbished shortly after.

Joke aside, this book is a collection of lecture notes for the module Mathematics for Computing taught to the first year computer science degree at Bangor University, UK. The collection of topics might look a bit like cherry-picking. Indeed, the module was designed with a view to cover the basics (be it in the form of cooking recipes) needed for the rest of the degree scheme. And before you start sneering at us from your high Algebra horse, think about how much stuff you have forgotten since your high school or university.

The book follows the lectures with simple worked examples and problems at the end of each chapter.

Ludmila Kuncheva
Bangor, 17 December 2012

Contents

Chapter 1

Matrices

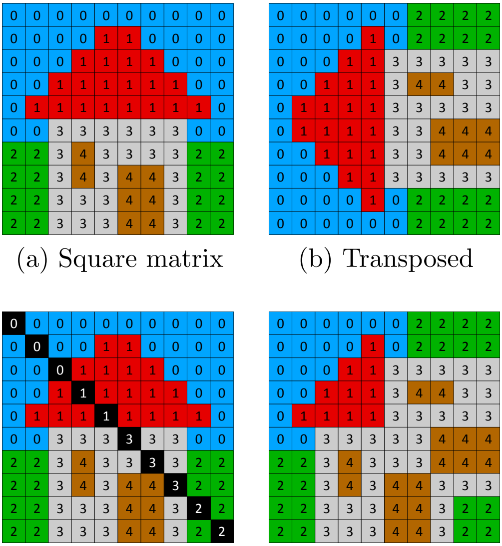
1.1 Definitions

A *matrix* is a rectangular table of numbers. Denote the number of rows of the table by m and the number of columns by n. A matrix with m rows and n columns is said to be of size m-by-n (denoted also $m \times n$). A matrix is called *square* when $m = n$, otherwise we shall call it a *non-square* matrix.

$\oplus\oplus\oplus$ **Example**

Figure 1.1 (a) shows a 10-by-10 (square) matrix, like the ones you may see in a "paint-by-numbers" children's book. $\ominus\ominus\ominus$

To make it easier to discuss matrices, we introduce the following



(a) Square matrix (b) Transposed

(c) Main diagonal (d) Symmetric

Figure 1.1: Examples of matrix-related concepts.

notations.

$$A = [a_{i,j}]_{m\times n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ \dots & \dots & a_{i,j} & \dots \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}.$$

The common element of the matrix, positioned in row i and column j, is denoted by $a_{i,j}$. Depending on the size of the matrix we may have the following special cases (also illustrated in Figure 1.2):

- a scalar (one value), $m = n = 1$,

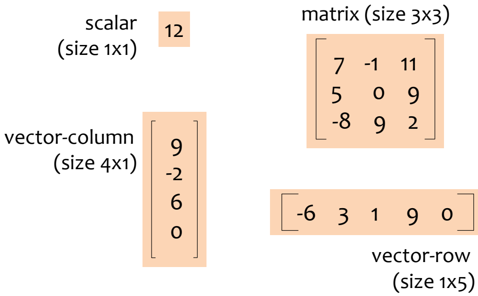


Figure 1.2: Special cases of matrix sizes.

- a vector-column, $m > 1, n = 1$,
- a vector-row, $m = 1, n > 1$,
- a (2d) matrix, $m > 1, n > 1$.

The term ‘matrix’ refers to the structure only. Matrices can be perceived as tables, and may contain entities other than signed integers, as in the examples thus far.

A matrix which we will use a lot, and will often return to, is the *rotation matrix*

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

where θ is the rotation angle.

A transpose matrix of a matrix A, denoted A^T or A' is a new matrix whose rows are the columns of A. First row of A^T is the first column of A, second row is the second column of A and so on. Figure 1.1 (b) shows the transpose of the matrix in (a). Obviously, if A is of size m-by-n, A^T is of size n-by-m. Note that the definition means that the elements of the transpose, $a^T_{i,j}$, are obtained by swapping the row and column indices, that is, $a^T_{i,j} = a_{j,i}$.

Elements $a_{i,i}$ of a square matrix of size $m \times m$, are called *diagonal elements.* The diagonal elements lie on the *main diagonal* of the matrix as shown with black tiles in Figure 1.1 (c).

The *trace* of a matrix is the sum of the elements on its main diagonal. The trace of an $n \times n$ matrix is

$$a_{1,1} + a_{2,2} + \cdots + a_{n,n} = \sum_{i=1}^{n} a_{i,i}.$$

A square matrix is *diagonal* if all its non-zero elements are on the main diagonal. For example,

$$A = \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

are diagonal matrices, while

$$C = \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 7 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are not. A very important diagonal matrix is the *identity matrix.* . The identity matrix (typically denoted by I) of size m is a diagonal matrix such that $a_{i,i} = 1$ for all $i = 1, \ldots, m$. For example, identity matrix I of size 3 is

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

A matrix is called *symmetric* if its elements are symmetric about the main diagonal. The element in row i, column j is equal to the element in row j, column i, that is $a_{i,j} = a_{j,i}$. In other words, a symmetric matrix is the same as its transpose. We can state formally that a matrix A is symmetric if and only if $A = A^T$. A visual example of a symmetric matrix is shown in Figure 1.1 (d). Note that the main diagonal may contain any elements.

$\oplus \oplus \oplus$ **Example**

The matrix below is symmetric

$$\begin{bmatrix} 28 & 18 & 2 & 4 \\ 18 & 29 & -14 & 11 \\ 2 & -14 & 18 & 0 \\ 4 & 11 & 0 & 18 \end{bmatrix}.$$

$\ominus \ominus \ominus$

A matrix cannot be symmetric if it is not square because the concept of 'main diagonal' is not defined for non-square matrices.

MATLAB

The word MATLAB is an acronym for MATrix LABoratory[1]. MATLAB's main advantage over many other programming languages and environments is the ease and elegance of matrix calculation. Here is how you enter a matrix in MATLAB. Type at the command prompt >>:

```
a = [1 2 3;4 5 6;7 8 9];
```

The matrix `a` entered into MATLAB memory is $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

[1] http://www.mathworks.co.uk/

The rows' delimiter is a semicolon. You can put commas or spaces between the elements of a row. The elements must be numerical. All rows must have the same number of elements. The semicolon at the end of the line suppresses the output in the command window. If you want to see the matrix that is being stored, omit the semicolon.

Matlab has a wide pre-set gallery of useful matrices. The most important for our class would be `eye(k)`, which returns the identity matrix of size k.

$\oplus \oplus \oplus$ **Example**

Try this `imagesc(rand(50)), axis square off`

MATLAB will generate a square matrix 50-by-50, filled with random numbers drawn from the interval [0,1]. It will then open up a figure window and display a colourful picture. $\ominus \ominus \ominus$

To obtain a transpose of a matrix entered into the MATLAB memory, type an apostrophe after the variable name. For example `a'` will return the transpose of `a`.

$\oplus \oplus \oplus$ **Example**

Here is how we create and make into an image (Figure 1.3) a random symmetric matrix

```
a = rand(9); imagesc(a*a'), axis square off
```

$\ominus \ominus \ominus$

1.2 Simple Operations with Matrices

1.2.1 Equivalence

Let A and B be two matrices. A is said to be *equivalent* to B if they are of the same size and all corresponding elements are equal

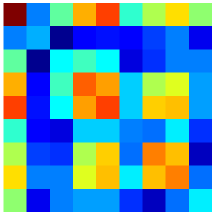
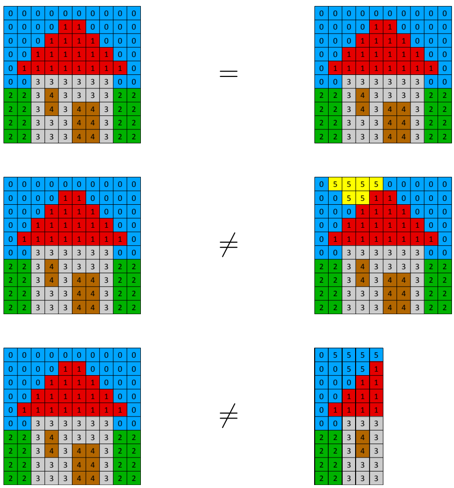


Figure 1.3: A graphical example of a random 9-by-9 symmetric matrix

($a_{i,j} = b_{i,j}$ for any a and j).

⊕ ⊕ ⊕ **Example**



⊖ ⊖ ⊖

1.2.2 Addition and multiplication by a scalar

To be able to add two matrices A and B, they must be of the same size. $A + B$ gives a new matrix C, whose elements are the sums of the corresponding elements from A and B.

$\oplus \oplus \oplus$ **Example**

If

$$A = \begin{bmatrix} 1 & 0 & -3 \\ 4 & 2 & -2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 4 & -3 & -1 \\ -3 & 1 & -2 \end{bmatrix}$$

then

$$C = A + B = \begin{bmatrix} 5 & -3 & -4 \\ 1 & 3 & -4 \end{bmatrix}.$$

$\ominus \ominus \ominus$

The addition of matrices is *commutative*, which means that $C = A + B = B + A$.

A matrix A is multiplied by a scalar λ by multiplying each element of A by λ.

$\oplus \oplus \oplus$ **Example**

Multiply matrix A above by $\lambda = -16$

$$\lambda A = (-16) \begin{bmatrix} 1 & 0 & -3 \\ 4 & 2 & -2 \end{bmatrix} = \begin{bmatrix} -16 & 0 & 48 \\ -64 & -32 & 32 \end{bmatrix}.$$

$\ominus \ominus \ominus$

1.2.3 Matrix expressions and equations

Equipped with just these two operations, we can construct linear equations of matrices. Start with the simplest: how do we subtract matrix B from matrix A? First multiply B by (-1) and then add the result to A. This will amount to subtracting from the elements of A the corresponding elements of B. If we denote $D = A - B$, the common element of D will be $d_{i,j} = a_{i,j} - b_{i,j}$.

$\oplus \oplus \oplus$ **Example**

Continuing the above example,

$$D = A - B = \begin{bmatrix} -3 & 3 & -2 \\ 7 & 1 & 0 \end{bmatrix}.$$

$\ominus \ominus \ominus$

The multiplication by a scalar is *distributive* with respect to addition and subtraction, just as in simple algebra, which means that

$$\lambda(A + B) = \lambda A + \lambda B.$$

$\oplus \oplus \oplus$ **Example**

How about spicing it a bit?[2] Find $M = (2x)A$, where $A = \begin{bmatrix} -x^3 & x^{-3} \\ 3x & \frac{1}{4x^2} \end{bmatrix}$.
The answer is

$$M = (2x) \begin{bmatrix} -x^3 & x^{-3} \\ 3x & \frac{1}{4x^2} \end{bmatrix} = \begin{bmatrix} -2x^4 & 2x^{-2} \\ 6x^2 & \frac{1}{2x} \end{bmatrix}.$$

$\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

Try this problem on your own (and then look at the solution below). Given below is matrix A. Find matrix B so that $A - 3B = I$

$$A = \begin{bmatrix} 0 & -2 \\ 3 & p \end{bmatrix}.$$

[2] Don't panic at the sight of "x"! Go back to your GCSE textbooks and brush up your knowledge of simple algebra. Your intelligent dog can consult you on that :)...

Solution: Recall that I was a notation for the identity matrix. Then you have the following matrix equation, where the elements of B are unknown

$$A - 3B = \begin{bmatrix} 0 & -2 \\ 3 & p \end{bmatrix} - 3 \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Knowing that each element in the matrix that will result from the sum on the left hand side *must* be identical to the corresponding element from I, we have to solve 4 little equations of one variable. The equation for the element in row 1, column 1 of B is $0-3b_{1,1} = 1$ which gives $b_{1,1} = -\frac{1}{3}$. Continuing in the same way for the remaining elements of B, we get

$$B = \begin{bmatrix} -\frac{1}{3} & -\frac{2}{3} \\ 1 & \frac{p-1}{3} \end{bmatrix}.$$

Note: To find $b_{2,2}$, you need to solve for $b_{2,2}$ the parametric equation

$$p - 3\, b_{2,2} = 1$$

with parameter p. ⊖⊖⊖

1.3 Matrix multiplication

While the matrix operations seen so far resemble algebraic operations with numbers, matrix multiplication does not. MATLAB will let you multiply two matrices of the same size *element-wise*, called the Hadamard product, but this is **not** matrix multiplication.

1.3.1 How is it done?

First, in multiplying two matrices A and B, we need to make sure that the sizes satisfy the following condition: the number of columns of A must be equal to the number of rows of B, that is,

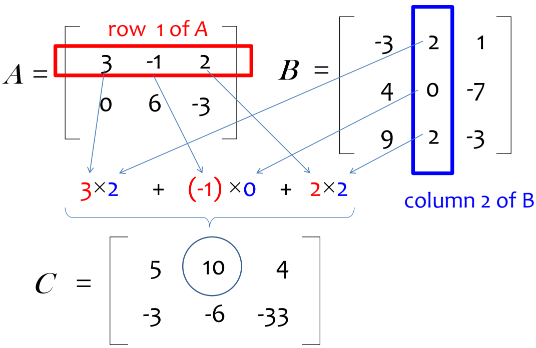
$$A_{m \times k} \qquad B_{k \times n}$$

the same

If the condition is not met, the matrix multiplication AB is impossible. When the condition is met, the result is a matrix $C = AB$ of size $m \times n$. The element of C in row i and column j is calculated from the ith row of A and the jth column of B. The corresponding elements are multiplied, and the results are summed up. The following equation shows the calculation in general notation

$$c_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots a_{i,k}b_{k,j} = \sum_{t=1}^{k} a_{i,t}b_{t,j}.$$

This equation is illustrated below in obtaining element $c_{1,2}$ of matrix $C = AB$.



You can see now why the middle dimension must be the same, because all k elements (columns) of the ith row of A must be paired with k elements (rows) of the jth column of B.

As an exercise, verify the remaining elements of C by performing the calculation for the respective rows and columns.

Try the following little problem. Dream up a matrix A of size (4×3). Find the products IA and AI where I denotes the identity matrix. Knowing that I must be a square matrix, what would be the appropriate size of I to be able to calculate IA? How about the size of I for AI to be possible?

In both cases you should arrive at A, that is

$$IA = AI = A, \quad \text{for any matrix } \ A,$$

where the identity matrix may have different sizes in the two parts of the equation.

Note that for any matrix A, the following two products are always possible:

$$AA^T \quad \text{and} \quad A^T A.$$

If A is of size $m \times n$, what would be the size of AA^T? And of $A^T A$?

Note also that both AA^T and $A^T A$ are symmetric matrices. Why does this happen? Consider $C = AA^T$ and elements $c_{i,j}$ and $c_{j,i}$ which lie symmetrically about the main diagonal. Recall that the elements of the transpose will have the row and the column index swapped. Then we can write

$$c_{i,j} = \sum_{t=1}^{k} a_{i,k} a_{j,k}.$$

Swapping the order of $a_{i,k}$ and $a_{j,k}$ in the product will not change the result, therefore

$$c_{i,j} = \sum_{t=1}^{k} a_{i,k} a_{j,k} = \sum_{t=1}^{k} a_{j,k} a_{i,k} = c_{j,i}.$$

Since the above equation is valid for all i and j, C is a symmetric matrix.

$\oplus \oplus \oplus$ **Example**
By way of illustrating the above proof, here is an example

$$\left[\begin{array}{rrrr} 1 & -3 & 6 & -2 \\ -5 & 6 & -4 & -4 \\ -5 & -3 & 7 & -3 \end{array}\right] \times \left[\begin{array}{rrr} 1 & -5 & -5 \\ -3 & 6 & -3 \\ 6 & -4 & 7 \\ -2 & -4 & -3 \end{array}\right] = \left[\begin{array}{rrr} 50 & -39 & 52 \\ -39 & 93 & -9 \\ 52 & -9 & 92 \end{array}\right].$$

$\ominus \ominus \ominus$

1.3.2 Properties of matrix multiplication

Commutativity (or the lack thereof). Unlike algebraic multiplication, matrix multiplication is not commutative. For scalars, we have $ab = ba$. For matrices, AB is not necessarily equal to BA.

First, the multiplication may not be possible because of the dimension rule. A matrix A of size $m \times k$ can be multiplied by matrix B of size $k \times n$. However, for BA to be possible, the number of columns of B, n must be equal to the number of rows of A, m.

Assume that $m = n$, hence both AB and BA are possible. Is matrix multiplication commutative for such matrices? The size of AB is $m \times n$. If $m = n \neq k$, $AB \neq BA$.

Finally, let $m = n = k$, that is both A and B are square matrices of the same size. Is matrix multiplication commutative for such matrices? Both AB and BA exist, and the size of both results is $m \times m$. One negative example is sufficient to disprove the positive statement that $AB = BA$ always holds for square matrices of the same size. Unless you are particularly lucky, a randomly picked pair of matrices may give us the desired counter-proof. For example

$$A = \left[\begin{array}{rr} 8 & 5 \\ 0 & -5 \end{array}\right] \qquad B = \left[\begin{array}{rr} -1 & 5 \\ 7 & 8 \end{array}\right]$$

$$AB = \left[\begin{array}{rr} 27 & 80 \\ -35 & -40 \end{array}\right] \quad \neq \quad BA = \left[\begin{array}{rr} -8 & -30 \\ 56 & -5 \end{array}\right].$$

Associativity and distributivity. Consider a chain of possible matrix multiplications $ABCD$. The associativity property states that the result will be the same regardless of the order in which we multiply the neighbouring matrices. For example

$$ABCD = ((AB)C)D = (A(BC))D = \cdots$$

The following distributivity property also holds

$$A(B+C) = AB + AC \quad \text{and} \quad (A+B)C = AC + BC.$$

⊕ ⊕ ⊕ **Example**

$$\begin{bmatrix} 1 & 0 \\ -8 & 2 \end{bmatrix} \left(\begin{bmatrix} 0 & -4 & 3 \\ 7 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 5 & 4 & 6 \\ -1 & 1 & 0 \end{bmatrix} \right)$$

Calculated as shown

$$\begin{bmatrix} 1 & 0 \\ -8 & 2 \end{bmatrix} \begin{bmatrix} 5 & 0 & 9 \\ 6 & 2 & 2 \end{bmatrix} = \boxed{\begin{bmatrix} 5 & 0 & 9 \\ -28 & 4 & -68 \end{bmatrix}}$$

Calculated using distributivity property

$$\begin{bmatrix} 0 & -4 & 3 \\ 14 & 34 & -20 \end{bmatrix} + \begin{bmatrix} 5 & 4 & 6 \\ -42 & -30 & -48 \end{bmatrix} = \boxed{\begin{bmatrix} 5 & 0 & 9 \\ -28 & 4 & -68 \end{bmatrix}}$$

⊖ ⊖ ⊖

Transpose of a product. The rule is as follows. To calculate the transpose of a product, reverse the order of the matrix chain and transpose each of the matrices. For example, for three matrices, we have $(ABC)^T = C^T B^T A^T$.

⊕ ⊕ ⊕ **Example**

An illustration for two matrices is given below

$$\left(\begin{bmatrix}1&0\\-8&2\end{bmatrix}\begin{bmatrix}0&-4&3\\7&1&2\end{bmatrix}\right)^T$$

Calculated as shown

$$\begin{bmatrix}0&-4&3\\14&34&-20\end{bmatrix}^T = \boxed{\begin{bmatrix}0&14\\-4&34\\3&-20\end{bmatrix}}$$

Calculated using transpose formula

$$\begin{bmatrix}0&7\\-4&1\\3&2\end{bmatrix}\begin{bmatrix}1&-8\\0&2\end{bmatrix} = \boxed{\begin{bmatrix}0&14\\-4&34\\3&-20\end{bmatrix}}$$

⊖ ⊖ ⊖

Multiplication by the identity matrix. It is important to note that, in matrix multiplication, the identity matrix I serves the same purpose as the number 1 in scalar multiplication. In other words, for any matrix A of size $(n \times m)$

$$IA = AI = A.$$

To make multiplication possible, the first I is of size $n \times n$ and the second I is of size $m \times m$.

1.4 Linear Transformations in 2d

An important application of matrix multiplication is for linear transformation from one space to another. Let us denote the coordinate axes in 2-dimensions x and y. A point in this space is defined by the pair of coordinates (x, y). We can organise the coordinates in a vector $[x, y]^T$ (which is the *position* vector of that point). A transformation will map (x, y) to a new point (x', y') obtained as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = T \begin{bmatrix} x \\ y \end{bmatrix}, \tag{1.1}$$

where T is a transformation matrix of size 2×2.

$\oplus \oplus \oplus$ **Example**

Let $A = (2, 3)$ be a point in the two-dimensional space, and let the transformation matrix be $T = \begin{bmatrix} -2 & 4 \\ 3 & 0 \end{bmatrix}$. The new coordinates of A are calculated as

$$\begin{bmatrix} x'_A \\ y'_A \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}.$$

$\ominus \ominus \ominus$

1.4.1 Rotation

The rotation transformation rotates the coordinate system at angle θ about the origin (0,0). The transformation matrix in this case is the Rotation matrix R

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

Figure 1.4 illustrates the rotation transformation. Note that the coordinate system rotates anti-clockwise. This means that, when plotted with their new coordinates, the points will appear to get rotated clock-wise.

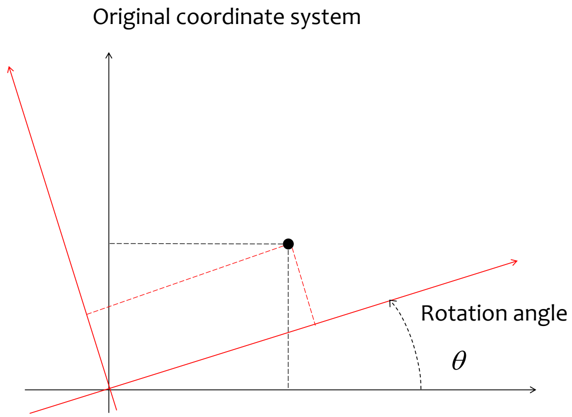


Figure 1.4: Rotation of the coordinate system.

$\oplus\oplus\oplus$ **Example**

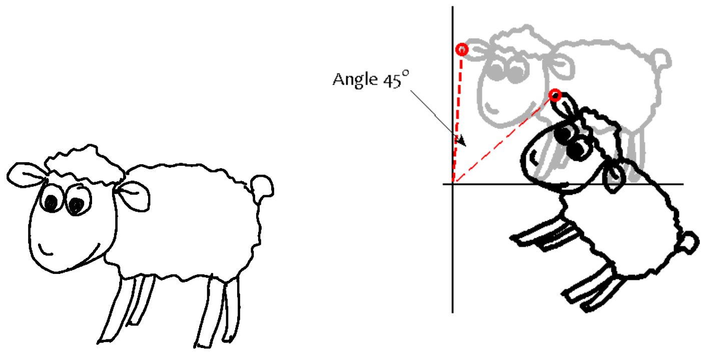
A lovely sheep cartoon (Figure 1.5 (a)) will be used as a graphical illustration of the transformations.[3] Figure 1.5 (b) shows the "rotated sheep". Each point on the original plot (grey) is mapped to a corresponding black point. One example of such mapping is highlighted with a red circle.

The rotation matrix for $\theta = 45^o$ is $\frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

The grey circled point has coordinates (17,265). The position vector (containing the coordinates) of the new point is

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 17 \\ 265 \end{bmatrix},$$

[3]Redrawn from `http://www.brentwindsor.co.uk/draw_cartoon_sheep.php`.



(a) (b)

Figure 1.5: Original and rotated sheep

$$x' = \frac{\sqrt{2}}{2} \times 1 \times 17 + \frac{\sqrt{2}}{2} \times 1 \times 265 = 199.40,$$
$$y' = \frac{\sqrt{2}}{2} \times (-1) \times 17 + \frac{\sqrt{2}}{2} \times 1 \times 265 = 175.36.$$

⊖ ⊖ ⊖

1.4.2 Reflections

The transformation matrices of reflections about various lines are shown in Figure 1.6. The position vector of each point on the graph is multiplied by the transformation matrix as in equation (1.1).

1.4.3 Expansions, contractions and shearing

Expansions and contractions on the x-axis are described by the following matrix

$$T = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}.$$

(a) about x	(b) about y	(c) about $x = y$	(d) about $x = -y$
$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

Figure 1.6: Example of reflections and the respective matrices. The grey sheep is the original image and the black one is the reflection.

For $0 < k < 1$, the transformation is a contraction because the x-coordinate will be smaller, while the y-coordinate will stay the same. And for $k > 1$, the transformation is an expansion. Similarly, for contractions and expansions on the y-axis, the transformation matrix is

$$T = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}.$$

⊕ ⊕ ⊕ **Example**

Consider the unit square with vertices $A = (0, 0)$, $B = (1, 0)$, $C = (1, 1)$ and $D = (0, 1)$. Let us apply expansion on the y-axis using matrix

$$T = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

The new coordinates of the vertices are

$$A' = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad B' = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$C' = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad D' = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

The calculations show that vertices C and D will be "lifted" from 1 to 3, while vertices A and B will not change.

$\ominus\ominus\ominus$

$\oplus\oplus\oplus$ Example

It is more fun to illustrate the transformations with the sheep, isn't it? Here is a snippet of MATLAB code which will shrink the sheep in proportion using transformation matrix

$$T = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}.$$

where the constant $0 < k < 1$ is the same for both axes. The code plots the Mama Sheep and the Baby Sheep, where the baby is translated on the x-axis. Note that in order to be able to run this code, you must have vectors x and y in MATLAB memory, containing the coordinates of the points of the sheep.

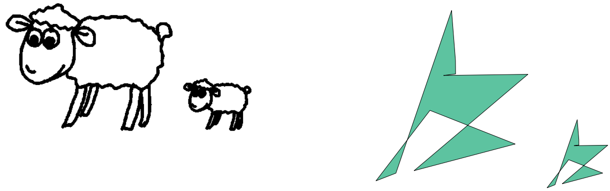
```
XY = [x(:)';y(:)']; % place all x and y coordinates
% in one matrix where the top row contains the x's
% and the bottom row contains the y's

figure, hold on
T = [0.4 0;0 0.4]; % transfromation matrix
plot(XY(1,:),XY(2,:),'k.') % mama sheep
XY = T * XY; % multiply by the transformation matrix
plot(XY(1,:)+450,XY(2,:),'k.') % baby sheep translated on x
axis equal off
```

Of course, you can try with an extravagant Tate Modern "sheep" like this:

```
XY = rand(2,10); co = rand(1,3);
figure, hold on
T = [0.4 0;0 0.4]; % transfromation matrix
fill(XY(1,:),XY(2,:),co) % mama sheep
XY = T * XY; % multiply by the transformation matrix
```

```
fill(XY(1,:)+1,XY(2,:),co) % baby sheep translated on x
axis equal off
```

This piece of code will work without loading up sheep coordinates. See the outcomes in Figure 1.7 (a) and (b).



(a) Mama Sheep and Baby Sheep (b) Tate Modern version

Figure 1.7: Examples of contraction transformation (MATLAB)

⊖ ⊖ ⊖

The *horizontal shearing* transformation matrix is

$$T = \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}, \quad \lambda \neq 0.$$

and the *vertical shearing* transformation matrix is

$$T = \begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix}, \quad \lambda \neq 0.$$

Figure 1.8 shows examples of horizontal shearing of the sheep along with the respective transformation matrices.

1.4.4 A sequence of linear transformations

A sequence of linear transformations can be implemented with a single transformation matrix. Suppose that we want to apply to

(a) $\begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & -0.5 \\ 0 & 1 \end{bmatrix}$

Figure 1.8: Example of horizontal shear and the respective matrices. The grey sheep is the original image and the black one is the transformed one.

$[x, y]^T$ a transformation with matrix T_1 first, and a transformation with matrix T_2 on the result. Then we have

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = T_2 \begin{bmatrix} x' \\ y' \end{bmatrix} = T_2 \left(T_1 \begin{bmatrix} x \\ y \end{bmatrix} \right)$$

We can open the brackets and regroup

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = T_2 T_1 \begin{bmatrix} x \\ y \end{bmatrix} = T \begin{bmatrix} x \\ y \end{bmatrix}$$

where $T = T_2 T_1$. Note the order of multiplying the transformation matrices. The latest transformation is first. Thus if we need to apply a sequence of m transformations in order T_1, T_2, ..., T_m, the resultant transformation matrix must be calculated as $T = T_m T_{m-1} \dots T_1$.

$\oplus \oplus \oplus$ **Example**

Let's reflect the sheep in the water of a lake (Figure 1.9). First, reflect across x, second contract the reflection on y, and finally sheer the reflection to simulate a flat horizontal surface. The transformation matrix should be

$$T = \underbrace{\begin{bmatrix} 1 & 1.3 \\ 0 & 1 \end{bmatrix}}_{\text{horizontal shear}} \times \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0.3 \end{bmatrix}}_{\text{contraction on } y} \times \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}}_{\text{reflection across } x}$$

$$= \begin{bmatrix} 1 & -0.39 \\ 0 & -0.30 \end{bmatrix}$$

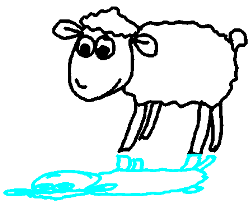


Figure 1.9: "Reflection in the lake" example. A sequence of three transformations were applied.

⊖⊖⊖

1.5 Matrix inverse for simultaneous equations

1.5.1 Inverse of a square matrix

Squared matrices can be invertible (called also *non-singular*) or not invertible (*singular*). A matrix A of size $n \times n$ is invertible if there exist a matrix B such that

$$AB = BA = I,$$

where I is the identity matrix. A common notation for matrix B is A^{-1}. Note that as A^{-1} is the inverse of A, A itself is the inverse of A^{-1}.

$\oplus \oplus \oplus$ **Example**

Show that matrix B below is the inverse of matrix A, also given below

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 2 \\ 2 & 3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 9 & -5 \\ 0 & -2 & 1 \\ -1 & -3 & 2 \end{bmatrix}.$$

All you need to do is to multiply A by B (or B by A) and obtain the identity matrix. $\ominus \ominus \ominus$

Is it enough to show this only one way? In other words, does $AB = I$ guarantee $BA = I$? The answer is 'yes', but the proof is a bit complicated, so we will not detail it here.

1.5.2 How do we find the inverse?

Every matrix has a number called *rank*, associated with it. The rank of a matrix filled with zeros is a zero. The rank of a non-zero matrix is the number of independent rows or columns in the matrix.[4] The rank of a non-zero square matrix of size $n \times n$ is an integer between 1 and n. If the rank of the matrix is n, then the matrix is invertible (non-singular).

We will confine our study to the petite 2×2 matrices. The rank of such a matrix can be 0, 1 or 2. If the rank is 2, the matrix is invertible.

[4]We haven't defined yet what independent rows and columns mean. A row (column) is independent of the others, if its entries cannot be obtained by a linear combination containing any of the other rows (columns).

$\oplus\oplus\oplus$ **Example**

$$\text{Invertible:} \quad A = \begin{bmatrix} -1 & 3 \\ 2 & 5 \end{bmatrix}$$

$$\text{Not invertible (singular):} \quad B = \begin{bmatrix} -1 & 3 \\ 2 & -6 \end{bmatrix}.$$

Matrix B is singular because the bottom row is obtained by multiplying the top row by -2. The rank of A is 2 but the rank of B is 1.

$\ominus\ominus\ominus$

Can we use a calculation to give us the answer of whether a matrix is singular? We can! Every square matrix has another number associated with it, called the *determinant*. If the determinant is non-zero, the matrix is non-singular and can be inverted. For a matrix of size 2×2, the determinant is calculated as

$$\det(A) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

Put in words, this is the product of the elements of the main diagonal minus the product of the elements of the second diagonal. Note that this calculation is valid *only* for 2×2 matrices. In the example above,

$$\det(A) = (-1)\times5-2\times3 = -11, \qquad \det(B) = (-1)\times(-6)-2\times3 = 0,$$

indicating that A is invertible but B is not.

Take this as a cooking recipe: If it exists, the inverse of matrix A is calculated as

$$A^{-1} = \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Note that $ad - bc = \det(A)$, hence

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

$\oplus \oplus \oplus$ **Example**

Let $A = \begin{bmatrix} 7 & 2 \\ 3 & 1 \end{bmatrix}$. Then $\det(A) = 7 - 6 = 1 \neq 0$, therefore

$$A^{-1} = \begin{bmatrix} 1 & -2 \\ -3 & 7 \end{bmatrix}.$$

Verify that $AA^{-1} = I$. $\ominus \ominus \ominus$

Note that the inverse of the transformation matrix (determinant $\neq$ zero) will restore the original data. Figure 1.10 shows our favourite sheep cartoon deformed through a random transformation and duly recovered by the inverse.

(a)
black = $T\times$ original

(b)
black = $T^{-1}\times$ grey
= $T^{-1}T\times$ original = original

Figure 1.10: Example of a random transformation and its reverse.

1.5.3 Solving simultaneous equations

Consider a system of simultaneous linear equations

$$\begin{aligned}
a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1\\
a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2\\
&\vdots\\
a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n &= b_n
\end{aligned}$$

where $a_{i,j}$ and b_i are constants. We can organise the constants a in an $n \times n$ matrix A

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n}\\ a_{2,1} & a_{2,2} & \ldots & a_{2,n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n,1} & a_{n,2} & \ldots & a_{n,n} \end{bmatrix}.$$

We can also arrange the unknowns x_i and the right-hand side constants in vector-columns

$$\mathbf{x} = \begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} b_1\\ b_2\\ \vdots\\ b_n \end{bmatrix}.$$

We can now re-write the system of linear equations as a *matrix equation*

$$A\,\mathbf{x} = \mathbf{b}.$$

The solution of this system can be found though calculating the matrix inverse of A. If A is non-singular, we can calculate A^{-1} such that $A^{-1}A = AA^{-1} = I$. Then we can multiply both sides of the above matrix equation by A^{-1} *on the left*

$$A^{-1}\,A\,\mathbf{x} = A^{-1}\,\mathbf{b}.$$

Since $A^{-1}A = I$ and $I\mathbf{x} = \mathbf{x}$, the above equation becomes

$$\mathbf{x} \;=\; A^{-1}\,\mathbf{b}.$$

This gives us a simple way to solve the system of linear equations. Taking our baby-steps into the scary world of linear algebra, we will keep the examples to two linear equations with two unknowns.

$\oplus \oplus \oplus$ **Example**

Use the 'express-and-substitute' method, and then the matrix equation method to solve the following system of simultaneous linear equations

$$\begin{aligned} -9x + 3y &= -6 \\ -11x - 9y &= 12 \end{aligned}$$

Answer:

(i) Express-and-substitute (and don't tell anybody you have seen this in a textbook for university students)

$$\begin{aligned} 3y &= 9x - 6 \\ y &= 3x - 2 \\ -11x - 9(3x - 2) &= 12 \\ -11x - 27x + 18 &= 12 \\ -38x &= -6 \\ x &= 0.1579 \approx 0.16 \\ y &= 3x - 2 = 3(0.1579) - 2 = -1.5263 \approx -1.53. \end{aligned}$$

(ii) Matrix equations

$$A = \begin{bmatrix} -9 & 3 \\ -11 & -9 \end{bmatrix}$$

$$det(A) = (-9)(-9) - 3(-11) = 81 + 33 = 114 \neq 0.$$

The inverse of A is

$$A^{-1} = \frac{1}{114}\begin{bmatrix} -9 & -3 \\ 11 & -9 \end{bmatrix}.$$

Then the solution is

$$\mathbf{w} = \begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}\begin{bmatrix} -6 \\ 12 \end{bmatrix}$$

$$= \frac{1}{114}\begin{bmatrix} -9 & -3 \\ 11 & -9 \end{bmatrix}\begin{bmatrix} -6 \\ 12 \end{bmatrix}$$

$$= \frac{1}{114}\begin{bmatrix} 18 \\ -174 \end{bmatrix}$$

$$= \begin{bmatrix} 0.1579 \\ -1.5263 \end{bmatrix}$$

Ideally, you should make a check with the original equations here even though the solutions from the two methods match.

⊖ ⊖ ⊖

1.6 Exercises

1. If matrix A has 5 rows and 4 columns, and matrix B has 10 rows and two columns, are the two matrices of the same size? Explain your answer.

2. Calculate matrix A knowing that

   (a) A is of size 3-by-6, and the common element is

   $$a_{i,j} = 3 \times i \times (j - 2)$$

(b) A is symmetric; all elements are positive integers; the number of rows is two, the trace of the matrix is 2, and the sum of all elements is 26.

3. Fill in the following table by entering 'Y' or 'N' in the cells.

If a matrix is:	then it necessarily is			
	Square	Diagonal	Identity	Symmetric
Square	Y			
Diagonal		Y		
Identity			Y	
Symmetric				Y

4. Each of the matrices below can be characterised with one word that describes it best. Find the respective word for each matrix.

(a)
$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

(b)
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(c)
$$\begin{bmatrix} 5 & 4 & 1 \\ 4 & 6 & -3 \\ 1 & -3 & 6 \end{bmatrix}$$

(d)
$$\begin{bmatrix} -4 & 2 & 0 \\ 3 & 6 & 3 \\ 0 & 2 & -4 \end{bmatrix}$$

(e)
$$\begin{bmatrix} 6 & 1 & -2 & -1 \\ 1 & 1 & -1 & -1 \\ -2 & -1 & 11 & -1 \end{bmatrix}$$

(f)
$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

(g)
$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

(h)
$$\begin{bmatrix} -2 & -2 & -2 \\ 4 & 4 & 4 \\ 7 & 7 & 7 \end{bmatrix}$$

5. What is the transpose of

   (a) a diagonal matrix? (b) the identity matrix?

6. How will the transpose of a matrix affect its trace?

7. Matrices A, B and C are given below

$$A = \begin{bmatrix} 4 & 0 & 4 \\ 2 & -2 & 4 \\ -1 & -2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -3 & -4 & 4 \\ 2 & 3 & 4 \\ -3 & 5 & -3 \end{bmatrix}$$

$$C = \begin{bmatrix} -1 & -4 & -2 \\ 0 & 1 & -3 \\ 6 & 2 & -1 \end{bmatrix}.$$

   Calculate the following sums

   (a) $-7A - 7B - 7C$ (b) $-A + 4B + 3C$

8. Evaluate the following matrix expression

$$-\frac{x}{y}\begin{bmatrix} 4 & x^2y^3 \\ -2\sqrt{y} & y \end{bmatrix} + 4\begin{bmatrix} \frac{x}{y} & 1 \\ \frac{x}{\sqrt{y}} & \frac{x}{4} \end{bmatrix}.$$

9. Find matrix A so that the following matrix equation holds

$$4I - 2A + \begin{bmatrix} -2 & -1 & 2 \\ 3 & 1 & 1 \\ 3 & 3 & -5 \end{bmatrix} = \begin{bmatrix} 6 & -1 & 2 \\ 13 & -1 & 15 \\ -11 & 11 & -1 \end{bmatrix}.$$

10. Find matrix A so that the following matrix equation holds

$$-3A^T + \begin{bmatrix} \sin\theta & \cos\theta \\ 2\theta & \theta^2 \end{bmatrix} = 5I.$$

11. If multiplication is possible, find AB and BA

$$A = \begin{bmatrix} 3 & 7 & 4 & -5 \\ -7 & 3 & -1 & 4 \\ 6 & 5 & 3 & -7 \end{bmatrix} \qquad B = \begin{bmatrix} -3 & 4 & 0 \\ -7 & -2 & -1 \\ -6 & 8 & 5 \\ 6 & -7 & 5 \end{bmatrix}$$

12. Calculate the following sum

$$\left(3\begin{bmatrix} 7 & -1 \\ 2 & 1 \end{bmatrix} - (-2)\begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix}\right)\begin{bmatrix} 4 & 1 \\ 6 & 3 \end{bmatrix}.$$

13. Propose three matrices A, B and C, of different sizes, such that the product ABC is possible. Illustrate the associativity property using these matrices.

14. Considering the sizes of the matrices indicated in the expressions below, is the matrix chain product possible? If not, explain why. If yes, what would be the size of the resultant matrix?

    (a) $A_{(3\times 12)}B_{(12\times 6)}C_{(6\times 4)}D_{(4\times 10)}E_{(10\times 12)}F_{(12\times 7)}$

    (b) $A_{(8\times 1)}B_{(1\times 9)}C_{(9\times 2)}D_{(3\times 4)}E_{(4\times 10)}F_{(10\times 6)}$

15. (M)[5] Write a MATLAB script that will rotate and plot a simple shape. You can invent a shape; could be something random such as

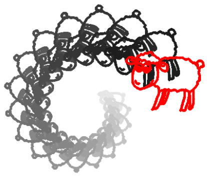
```
h = fill(rand(6,1),rand(6,2),rand(1,3));
```

Hint: Use the handle `h` to change the XData and YData of the object.

[5] Problems labelled (M) require some knowledge of MATLAB.

16. (M) Using MATLAB, can you plot what you see below? Use any shape you like instead of the sheep.

(a)



(b)

17. The vertices of a triangle have coordinates: $A\ (3, -1)$, $B\ (2, 8)$, and $C\ (4, -2)$. Use the respective reflection matrix to calculate the new positions of the vertices reflected about $x = -y$. Show your calculations. Plot the original and the reflected triangles.

18. What reflection will the following matrix perform?

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

Illustrate your answer by an example.

19. Consider the following chain product of matrices

$$A_{(3,5)}B_{(5,11)}C_{(11,7)}D_{(7,2)}.$$

(a) What would be the size of the resultant matrix?

(b) The computational load in chain matrix multiplication will depend on the order of multiplications. We can choose which product of two matrices to calculate first and put a pair of brackets around that. Then, again, we can choose the next pair of neighbouring matrices, and so on. For the sake of the argument, assume that a summation costs 1 pound (of any number of terms) and a scalar multiplication costs 3 pounds.

Evaluate the cost of the calculation of the matrix product for the following cases

$$((AB)C)D, \quad A((BC)D), \quad (AB)(CD).$$

(c) What is the total number of ways in which you can group the calculations of a product of 4 matrices?

20. Give an example of a singular 2×2 matrix.

21. Give an example of a singular 3×3 matrix.

22. For what values of x is the following matrix invertible?

$$\begin{bmatrix} 5 & -2 \\ -1 & x \end{bmatrix}$$

Find the inverse as an expression of x.

23. Use the 'express-and-substitute' method, and then the matrix equation method to solve the following system of simultaneous linear equations

$$\begin{aligned} 2x - 3y &= -5 \\ -6x + y &= 7 \end{aligned}$$

24. Use the matrix equation method to solve the following system of simultaneous linear equations, where p, q, r, w, a and b are constant parameters.

$$\begin{aligned} px + qy &= a \\ rx + wy &= b \end{aligned}$$

Chapter 2

Vectors

2.1 Definitions and notations

The term *vector* is often used to mean a list of numbers. Vectors are usually denoted with lower-case letters: $\mathbf{a}$ (boldface), $\underline{a}$ (underlined) or $\vec{a}$ (with an arrow on top). The preferred notation in most scientific texts, which we will also adopt, is $\mathbf{a}$.

A vector with n entries "lives" in the n-dimensional space $\mathbb{R}^n$. By default, all vectors are *column* vectors, i.e., the components are arranged vertically. As an example, consider $\mathbf{a} \in \mathbb{R}^5$ expressed as $\mathbf{a} = [-1, 3, 5, 0, -4]^T$.

A vector is characterised by a direction and magnitude. In other words, it is not fixed in the space. As an example consider the bunch of pencils to the left. They are all of the same length and orientation, parallel to one another. Hence this is the same vector at different locations in $\mathbb{R}^2$.

Examples of vectors are wind velocity and force. Both can be applied at different locations in the space, having the same specific direction and magnitude.

2.2 Operations with vectors

As vectors are special cases of matrices, all that is valid for matrices is valid for vectors too. For example, two vectors are equal if they have the same number of components and all corresponding entries are equal. Like with matrices, we can write matrix equations containing vectors.

$\oplus \oplus \oplus$ **Example**

Let $\mathbf{x} = [3, -2, 5]^T$ and $\mathbf{y} = [8, 3, -3]^T$. Then

$$4\mathbf{x} - 6\mathbf{y} = 4 \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} - 6 \begin{bmatrix} 8 \\ 3 \\ -3 \end{bmatrix} = \begin{bmatrix} -36 \\ -26 \\ 38 \end{bmatrix}.$$

$\ominus \ominus \ominus$

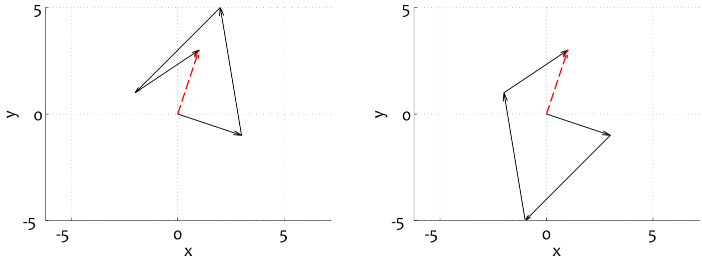
There is an interesting geometrical interpretation of summation of vectors called the "principle of superposition". It states that the order of summation of the vector does not matter; the same final vector will be obtained for any permutation of the summation terms.

$\oplus \oplus \oplus$ **Example**

While this principle is clear from the algebraic representation of vector summation, Figure 2.1 demonstrates its 2-d geometrical interpretation.

$\ominus \ominus \ominus$

The multiplication by a constant scales the vector (stretches or shrinks it). If the sign of the constant is negative, the scaled vector points in the opposite direction.



The vectors in this example (in order as in the left sub-plot) are as follows

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} \begin{bmatrix} -1 \\ 6 \end{bmatrix} \begin{bmatrix} -4 \\ -4 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Figure 2.1: Principle of superposition. The order of summation of vectors does not matter; the end vector (dashed red line) is the same in both subplots.

2.3 Magnitude of a vector

2.3.1 Calculating the magnitude

The *magnitude* of a vector $\mathbf{a} = [a_1, \ldots, a_n]^T$, also called the *length* of the vector, is calculated as follows

$$|\mathbf{a}| = \sqrt{a_1^2 + a_1^2 + \cdots + a_n^2}.$$

It is easy to verify that, in 2-d, the magnitude of the vector is calculated through the Pythagoras theorem. As displayed in Figure 2.2, vector $\mathbf{a} = [a_x, a_y]^T$ sits as the hypotenuse of a right-angle triangle with sides a_x and a_y.

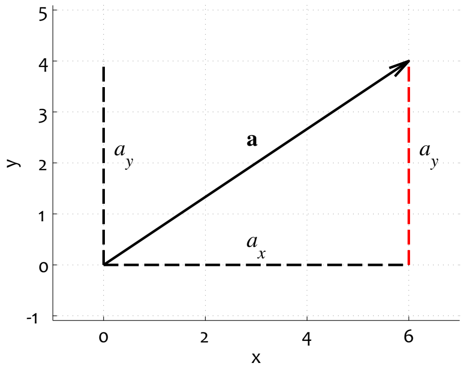


Figure 2.2: Magnitude of vector **a** calculated using the Pythagoras theorem.

2.3.2 Scaling a vector to a unit length

A *unit vector* is a vector of magnitude 1. How can we construct a unit vector which will have the same direction as a given vector **a**? Easy! Divide each component by the magnitude of the vector. Here is why this will work.

Given is vector $\mathbf{a} = [a_1, \ldots, a_n]^T$. Let $m = |\mathbf{a}|$ be the magnitude of **a**. Construct vector

$$\mathbf{u} = \frac{1}{m}\mathbf{a} = \left[\frac{a_1}{m}, \frac{a_2}{m}, \cdots, \frac{a_n}{m}\right]^T.$$

Since **u** is obtained by multiplying **a** by a positive constant, its direction will be the same as that of **a**. Its magnitude, on the other hand, becomes

$$|\mathbf{u}| = \sqrt{\frac{a_1^2}{m^2} + \frac{a_2^2}{m^2} + \cdots + \frac{a_n^2}{m^2}}$$

$$= \sqrt{\frac{1}{m^2}}\ \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

$$= \frac{m}{m} = 1.$$

$\oplus\oplus\oplus$ **Example**

Design a unit vector which has the opposite direction of vector $\mathbf{a} = [-1, 2, 4, -3]^T$. First, calculate the magnitude of $\mathbf{a}$

$$|\mathbf{a}| = \sqrt{(-1)^2 + 2^2 + 4^2 + (-3)^2} = \sqrt{30} = 5.4772.$$

Then

$$\mathbf{u} = -\frac{1}{5.4772}\ [-1, 2, 4, -3]^T = [0.18,\ \ -0.37,\ \ -0.73,\ \ 0.55]^T.$$

$\ominus\ominus\ominus$

2.4 Coordinate decomposition

Three important *unit* vectors in $\mathbb{R}^3$ are shown in Figure 2.3. They are called *basis* vectors and can be thought of as the columns of the identity matrix $I_{3\times 3}$. They are

$$\mathbf{i} = [1, 0, 0]^T, \quad \mathbf{j} = [0, 1, 0]^T, \quad \text{and} \quad \mathbf{k} = [0, 0, 1]^T.$$

Every vector in $\mathbb{R}^3$ can be represented as a linear combination of the basis vectors.

$\oplus\oplus\oplus$ **Example**

Consider the example shown in Figure 2.4. Vector $\mathbf{a} = [2, 2, 4]^T$ is obtained by adding two of vector $\mathbf{i}$, two of vector $\mathbf{j}$ and 4 of vector $\mathbf{k}$. [1] Written as a vector equation,

$$\mathbf{a} = 2\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}.$$

[1] The illustration is in this order but, according to the principle of superposition, any permutation of the terms will result in $\mathbf{a}$.

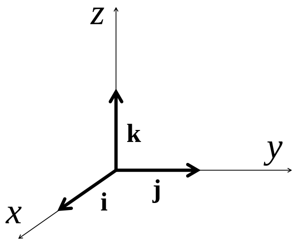


Figure 2.3: Basis vectors in $\mathbb{R}^3$.

$\ominus\ominus\ominus$

Look how simple it is! The coordinate decomposition of a vector is obtained by multiplying each coordinate by the respective basis vector and then summing up.

Note that the n basis vectors of the n-dimensional real space $\mathbb{R}^n$ are the columns of the $n \times n$ identity matrix. And, again, any vector in $\mathbb{R}^n$ can be represented as a linear combination of these basis vectors.

2.5 Scalar product of vectors

2.5.1 Definition

The *scalar product* of two vectors is a number. The vectors must be of the same size. For $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, the scalar product is calculated as

$$\mathbf{a}^T\mathbf{b} = [a_1, a_2, \ldots, a_n] \times \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

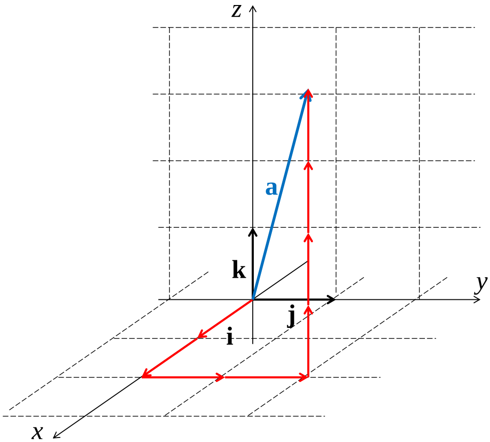
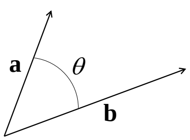


Figure 2.4: Coordinate decomposition of vector $\mathbf{a} = [2, 2, 4]^T$

This is nothing more than the matrix multiplication that we learned in the previous chapter. The matrices are of sizes $(1 \times n)$ and $(n \times 1)$, respectively, leading to a scalar (size (1×1)).

Alternative notations to $\mathbf{a}^T\mathbf{b}$ for scalar product are $\langle \mathbf{a}, \mathbf{b} \rangle$ and $\mathbf{a} \cdot \mathbf{b}$, the latter giving the name *dot product* to the scalar product of vectors.



A useful expression of the scalar product is based upon the vectors' magnitudes and the angle θ between the vectors (illustrated on the left)

$$\mathbf{a}^T\mathbf{b} = |\mathbf{a}|\ |\mathbf{b}|\ \cos(\theta).$$

2.5.2 Orthogonality of vectors

Two vectors are orthogonal if and only if their scalar product is zero. Remember that. Note that if the magnitude of any of the two vectors is zero, the scalar product is zero too. However in this we cannot talk about orthogonality or an angle because the vector with magnitude zero degenerates into a point.

$\oplus \oplus \oplus$ **Example**

Find x for which the following two vectors are orthogonal: $\mathbf{a} = [-3, 2, 4]^T$ and $\mathbf{b} = [2, -3x, 6]^T$?

Answer: To find x, we must solve $\mathbf{a}^T\mathbf{b} = 0$ for x, that is,

$$(-3) \times 2 + 2 \times (-3x) + 4 \times 6 = 0$$

$$-6 - 6x + 24 = 0,$$

therefore $x = 3$. $\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

Construct a unit vector orthogonal to vector $\mathbf{a} = [-1, 3]^T$.

Answer: We can first choose one of the components, calculate the second one so that the scalar product is zero, and then scale the vector to unit magnitude. For example, take vector $\mathbf{w}$ to be $[3, x]^T$. Then $-3 + 3x = 0$, hence $x = 1$. Then $\mathbf{w} = [3, 1]^T$, and $|\mathbf{w}| = \sqrt{9 + 1} = \sqrt{10} = 3.1623$ Rescaling $\mathbf{w}$, we get

$$\mathbf{u} = [0.95,\ 0.32]^T .$$

$\ominus \ominus \ominus$

Try the following problem on your own before looking at the solution. Let $\mathbf{a} = [a_1, a_2]^T$ be a unit vector. Construct a unit vector $\mathbf{b}$ which is orthogonal to $\mathbf{a}$.

Solution: Let us first construct a vector $\mathbf{v} = [1, x]^T$ that is orthogonal to $\mathbf{a}$. The first component is $v_1 = 1$ because we only have one equation coming from the orthogonality, from which to identify x.

$$a_1 + a_2 x = 0, \qquad \text{hence} \qquad x = -\frac{a_1}{a_2}$$

and

$$\mathbf{v} = \left[1, -\frac{a_1}{a_2}\right]^T$$

To scale $\mathbf{v}$ to a unit length, find its magnitude and divide both components by it.

$$|\mathbf{v}| = \sqrt{1^2 + \left(-\frac{a_1}{a_2}\right)^2} = \frac{1}{a_2} \underbrace{\sqrt{a_1^2 + a_2^2}}_{\text{magnitude of } \mathbf{a}} = \frac{1}{a_2}.$$

Then

$$\mathbf{b} = \frac{\mathbf{v}}{\frac{1}{a_2}} = a_2\mathbf{v} = \left[a_2, -\frac{a_1 a_2}{a_2}\right]^T = [a_2, -a_1]^T.$$

⊕ ⊕ ⊕ Example

And here is a small MATLAB script which plots 200 random pairs of unit orthogonal vectors, positioned at random locations and displayed with random colours, as shown in Figure 2.5

```
figure, hold on

% format the axes
axis([-5 5 -5 5])
axis square off

for i = 1:200
    a = rand(2,1)-0.5; % generate a random 2d vector 'a'
    maga = sqrt(a'*a); % magnitude of 'a'
    a = a/maga; % rescale 'a' to magnitude 1
    b = [a(2) -a(1)]; % construct vector 'b'
    p = 8*rand(2,1)-4; % choose a random point in 2d to
    % place the pair of vectors there
    plot([p(1)+a(1),p(1),p(1)+b(1)],...
```

```
        [p(2)+a(2),p(2),p(2)+b(2)],'k-','color',...
        rand(1,3),'linewidth',3)
    drawnow % show the incremental build-up of the picture
end
```

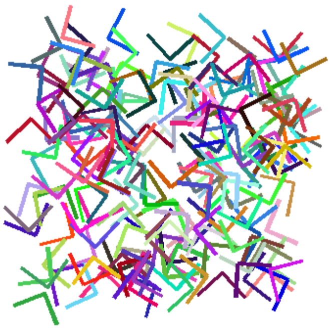


Figure 2.5: MATLAB output from the script for plotting 200 random pairs of orthogonal unit vectors.

⊖⊖⊖

2.6 Euclidean distance

The Euclidean distance between points $A(a_1, a_2, \ldots, a_n)$ and $B(b_1, b_2, \ldots, b_n)$, $A, B \in \mathbb{R}^n$, is calculated as

$$\begin{aligned} d(A,B) &= \sqrt{(a_1-b_1)^2+\cdots+(a_n-b_n)^2} \\ &= \sqrt{\sum_{i=1}^{n}(a_i-b_i)^2} \end{aligned} \tag{2.1}$$

$\oplus\oplus\oplus$ **Example**

Given are points $A(2, -1, 10, 4)$ and $B(1, -3, -9, 5)$. Calculate the Euclidean distance between them.

Answer:

$$d(A, B) = \sqrt{(2-1)^2 + (-1-(-3))^2 + (10-(-9))^2 + (4-5)^2}$$
$$= \sqrt{1+4+361+1} = \sqrt{367} \approx 19.16$$

$\ominus\ominus\ominus$

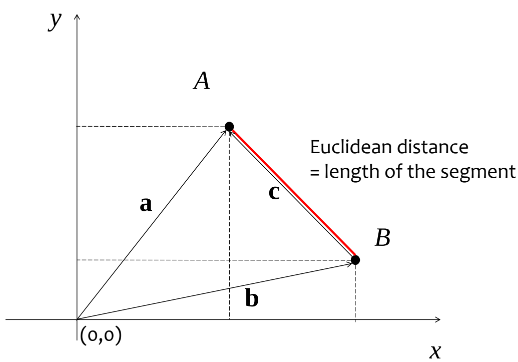


Figure 2.6: Illustration of the calculation of Euclidean distance using position vectors.

Figure 2.6 illustrates the calculation of the Euclidean distance in $\mathbb{R}^2$. Let $A(a_1, a_2)$ and $B(b_1, b_2)$ be two-dimensional points as shown in the figure. The Euclidean distance between the points is the length of the segment AB. Recall the *position vectors* of points introduced in the previous chapter (a vector from the origin of the coordinate system to the point). The position vectors of A and

B are plotted in the figure. These vectors are $\mathbf{a} = [a_1, a_2]^T$ and $\mathbf{b} = [b_1, b_2]^T$. Let $\mathbf{c}$ be the vector from B to A. According to the principle of superposition, when added to $\mathbf{b}$, this vector gives vector $\mathbf{a}$, which means that the following vector equation holds

$$\mathbf{b} + \mathbf{c} = \mathbf{a},$$

therefore

$$\mathbf{c} = \mathbf{a} - \mathbf{b} = [(a_1 - b_1), (a_2 - b_2)]^T.$$

Notice that the Euclidean distance between points A and B is, in fact, the magnitude of $\mathbf{c}$

$$|\mathbf{c}| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2},$$

which is exactly equation (2.1)

2.7 Exercises

1. Calculate the sums

   (i). $2\mathbf{x} - \mathbf{y}$ (ii). $\mathbf{x} + 7\mathbf{y}$ (iii). $3\mathbf{x} - 5\mathbf{y}$

   where $\mathbf{x} = [-4, 0, 1, -5, 7]^T$ and $\mathbf{y} = [1, 5, 5, -2, -3]^T$.

2. Design a unit vector which has the opposite direction of vector $\mathbf{a} = [-8, 2, 1, 0, 4, 1, -7]^T$.

3. Design a unit vector which has the opposite direction of vector $\mathbf{a} = [4, w, -3]^T$, where w is a parameter.

4. For what value of x is vector $\mathbf{z} = [-0.1, x, 0.4]^T$ a unit vector?

5. Show the coordinate decomposition of vector $\mathbf{a} = [-2, 1, -1]^T$.

6. Check whether any pair of these vectors are orthogonal vectors

   $\mathbf{a} = [-2, -6, 4]^T$ $\mathbf{b} = [1, -1, -1]^T$ $\mathbf{c} = [2, -2, 4]^T$

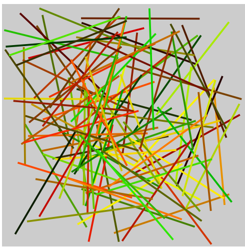
7. For what value of x are the following vectors orthogonal?

$$\mathbf{a} = [1, 0, -3x+1, -2, 4]^T \qquad \mathbf{b} = [-x, 3, -5, 6x, -1]^T$$

8. The scalar product of two unit vectors is 0.5. What is the angle between these vectors?

9. Find the Euclidean distance between points
   $A(3, 1, -6, -1, 2)$ and B(4, -5, 3, 0, -4) (2dp).

10. (M) Write a MATLAB script that will plot 100 randomly positioned "straws" in the unit square. The length of each straw must be between 0.49 and 0.51. An example is shown below.



Hint: Use a 'while' loop; generate a pair of points in the unit square using the command 'rand'. Then check the Euclidean distance between the points and accept the straw if the distance is within the limits. Only then draw the straw and increment the straw counter.

You can add an artistic touch to that by colouring the straws with random colours in the green and red palette (this will include yellow, oranges and browns). Just suppress the blue colour to achieve this effect.

Chapter 3

Lines, Planes, Circles

3.1 Lines in 2d

The generic equation of a line in $\mathbb{R}^2$ is

$$ax + by + c = 0.$$

3.1.1 Equation of a line between two points

There are many ways to derive the equation of a line between two points. Here we will use a formula that is easily generalisable for three dimensions. Let the two points be $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$, where $x_1 \neq x_2$ and $y_1 \neq y_2$. The equation of the line passing through the two points is

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}. \tag{3.1}$$

What happens if we swap the order of the points? Nothing! The equation that you derive will describe the same line.

$\oplus\oplus\oplus$ **Example**

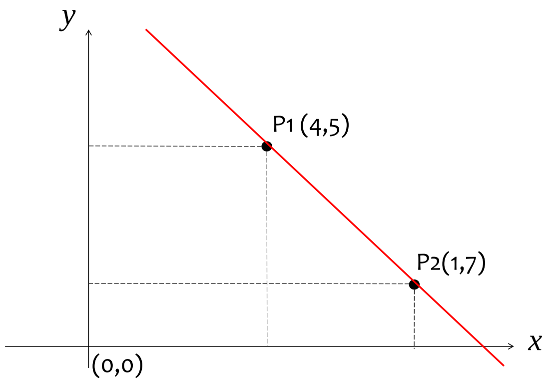


Figure 3.1: Line through two points $2x + 3y - 23 = 0$.

Figure 3.1 shows an example. The equation of the line through points $P_1(4, 5)$ and $P_2(1, 7)$ is

$$\frac{x-4}{1-4} = \frac{y-5}{7-5},$$

leading to

$$2(x-4) = (-3)(y-5)$$

$$2x - 8 = -3y + 15$$

$$2x + 3y - 23 = 0.$$

If we took the points in a reverse order,

$$\frac{x-1}{4-1} = \frac{y-7}{5-7}$$

$$(-2)(x-1) = 3(y-7)$$
$$-2x + 2 = 3y - 21$$
$$-2x - 3y + 23 = 0.$$

Since we can multiply both sides by any non-zero constant without changing the equation, multiplying by (-1), we arrive at

$$2x + 3y - 23 = 0.$$

⊖ ⊖ ⊖

3.1.2 Special cases of lines in 2d

Recall the generic equation of a line in 2d, $ax + by + c = 0$. When $c = 0$, the line passes through the origin of the coordinate system. Indeed, point $(0,0)$ satisfies the equation $ax + by = 0$. Figure 3.2 Shows an example of the three special cases.

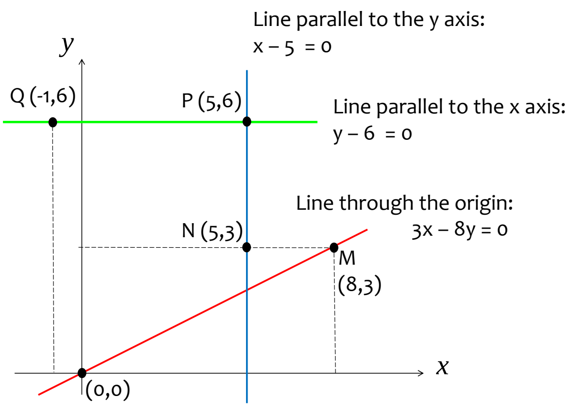


Figure 3.2: Examples of special cases of lines between two points.

The line through the origin, $3x - 8y = 0$, is calculated from points $(0, 0)$ and $M(8, 3)$. However, the formula cannot be applied for obtaining lines through $P(5, 6)$ and $N(5, 3)$ or through $P(5, 6)$ and $Q(-1, 6)$. The reason is that the denominator of one of the fractions will turn into zero. Instead of using the formula, we can inspect the coordinates of the two points and apply the following logic

- if $x_1 = x_2 = k$ and $y_1 \neq y_2$, the equation of the line is $x - k = 0$.
- if $y_1 = y_2 = k$ and $x_1 \neq x_2$, the equation of the line is $y - k = 0$.
- if $y_1 = y_2 = k$ and $x_1 = x_2 = m$, the two points are identical and do not define a line.

Notice that the coefficient and the respective variable which are missing from the generic equation of the line tell you which axis this line is parallel to. For example, line $5x - 7 = 0$ is parallel to the y-axis.

3.1.3 Points on the side of a line

For all point on a line, the equation of the line will hold. If we substitute the coordinates of a point that is not on the line in the left-hand side of the equation, the sum will not be zero. All points in the half-space on one side of the line will have sums with the same sign. Thus we can define a positive and a negative side of a line. These sides are relative because if both sides of the equation are multiplied by (-1), the $+$ and $-$ sides will swap. This property of the line equation is useful when we want to know whether the line cuts through a segment or a figure.

$\oplus \oplus \oplus$ **Example**

Find out whether the line with equation $3x - y + 4 = 0$ intersects the segment between points $A(2, 3)$ and $B(-1, -6)$. All we need to do is substitute the coordinates of the points in the left-hand side (LHS) of the line equation and then compare the signs of the sums.

Point A:	$3 \times 2 - 3 + 4 = 7 > 0$
Point B:	$3 \times (-1) - (-6) + 4 = 7 > 0$

Both sums are positive, therefore A and B are on the same side of the line and the line does not cut through the segment. $\ominus \ominus \ominus$

$\oplus \oplus \oplus$ Example

Let's play! The MATLAB code below generates a random shape in the top right corner of the unit square. Then it generates a ray from the origin that 'scans' the first quadrant. Only the rays that don't intersect the shape are displayed. We just check whether the points of the shape have LHS sums with the same sign.

At each step, we need the equation of the line passing through the origin, and sitting at angle θ from the x axis. Suppose that the ray has length 1. Then the end-point of the ray is at $(\cos(\theta), \sin(\theta))$. The equation of the line is

$$\sin(\theta)x - \cos(\theta)y = 0.$$

Notice the calculation of the signs (line 7). The point coordinates are generated as an array 'sha' with 10 rows and 2 columns. The coefficients of the line are organised in a column vector $[\sin(\theta), -\cos(\theta)]^T$, size 2×1. Then the product `sha * [sin(theta);-cos(theta)]` gives a 10×1 array with the LHS sums. Taking the sign using the MATLAB function of the same name, we check whether the signs are all 1s or all are -1s.

```
1 figure, hold on
2 axis([0 1 0 1]), axis square, grid on
3 sha = rand(10,2)*.35 + 0.5; % generate the shape coordinates
4 fill(sha(:,1),sha(:,2),rand(1,3)) % plot the form
5 % with parts filled with a random colour
```

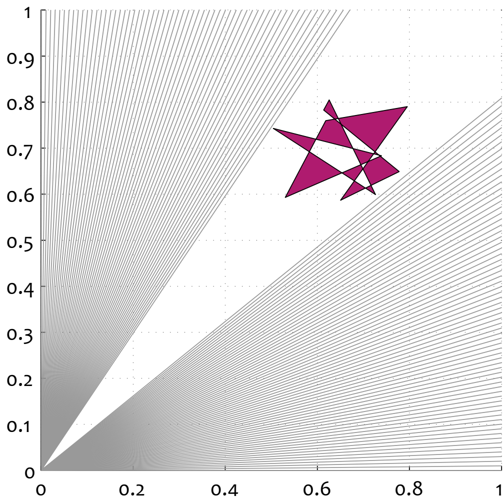


Figure 3.3: Line through two points $2x + 3y - 23 = 0$.

```
6  for theta = 0:0.01:pi/2
7      point_signs = sign(sha * [sin(theta);-cos(theta)]);
8      if all(point_signs == 1) || all(point_signs =  -1)
9          plot([0,1.41*cos(theta)],[0,1.41*sin(theta)],'k-',...
10             'color',[0.6 0.6 0.6]) % plot a grey ray
11     end
12     pause(0.02)
13 end
```

The ‘pause’ command ensures that you see the rays as they are plotted on the figure.

You may also have spotted the constant 1.41 that multiplies the $\cos(\theta)$ and $\sin(\theta)$ (line 9). Why is it there? Because the maximum length that we want to plot is the diagonal of the unit square. The excess of the lines (outside the unit square) is cut through specifying the axes limits (line 2).

$\ominus \ominus \ominus$

3.2 Planes

3.2.1 Equation of a plane and a hyperplane

The generic equation of a plane in $\mathbb{R}^3$ is

$$ax + by + cz + d = 0. \tag{3.2}$$

In larger-dimensional spaces, the linear equation defines a *hyperplane* as follows

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n + b = 0, \tag{3.3}$$

where $\mathbf{x} = [x_1, x_2 \ldots, x_n]^T \in \mathbb{R}^n$.

3.2.2 Plane defined by three points

Every three 3d points that do not lie on the same line define a plane. The equation of the plane can be derived from the coordinates of the points. Let the three points be $P(p_1, p_2, p_3)$, $Q(q_1, q_2, q_3)$ and $R(r_1, r_2, r_3)$. The coordinates of these points must satisfy the equation of the plane, which makes three equations with four unknowns, a, b, c and d. Suppose that the three x-values are not identical. Then the coefficient a, in front of x, is non-zero. We can divide both sides of the equation of the plane by a, and rewrite the equation of the plane as

$$x + b'y + c'z + d' = 0$$

where $b' = b/a$, $c' = c/a$ and $d' = d/a$. The system of three linear equations with three unknowns is

$$\begin{aligned} p_1 + b'p_2 + c'p_3 + d' &= 0 \\ q_1 + b'q_2 + c'q_3 + d' &= 0 \\ r_1 + b'r_2 + c'r_3 + d' &= 0 \end{aligned}$$

Rearranging the equations so that the constants are moved on the right-hand side (RHS),

$$\begin{aligned} b'p_2 + c'p_3 + d' &= -p_1 \\ b'q_2 + c'q_3 + d' &= -q_1 \\ b'r_2 + c'r_3 + d' &= -r_1 \end{aligned}$$

The vector with the unknowns is $\mathbf{u} = [b', c', d']^T$. In matrix equation terms, the system of the three equations is $A\mathbf{u} = \mathbf{v}$ where

$$A = \begin{bmatrix} p_2 & p_3 & 1 \\ q_2 & q_3 & 1 \\ r_2 & r_3 & 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{v} = \begin{bmatrix} -p_1 \\ -q_1 \\ -r_1 \end{bmatrix}$$

If $det(A) \neq 0$, A is invertible and the solution is

$$\mathbf{u} = A^{-1}\,\mathbf{v}$$

As we haven't studied how to calculate the determinant or the inverse of a matrix of sizes larger than 2×2, we will delegate these operations to existing software. In MATLAB, commands `rank(A)` and `det(A)` can be used to find out if A is invertible, followed by `inv(A)` to obtain A^{-1}.

⊕ ⊕ ⊕ **Example**

Figure 3.4 is produced by the MATLAB code below. It illustrates the calculations of the equation of the plane through points $A(12, -4, 5)$, $B(3, -1, 11)$ and $C(4, 2, 10)$. Triangle ABC is shaded in red. This time, we assumed $d = 1$ in the plane equation, and calculated a, b and c (in variable 'solution'). Three hundred points are generated so that the x and y coordinates are random, and the z coordinate is obtained from the derived equation. Thus all 300 points lie on the plane. Finally, the 'fill' command shades random spots on the plane. The last line in the code allows the plot to be rotated with the mouse to see that the calculated plane is indeed the one containing triangle ABC.

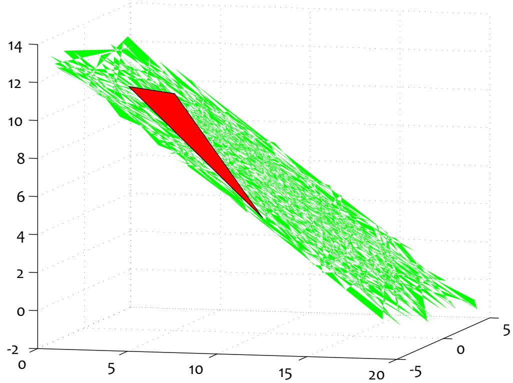


Figure 3.4: Plane through three 3d points.

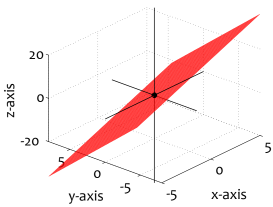
```
figure, hold on
A = [12 -4 5];B = [3 -1 11];C = [4 2 10]; % the 3 points
D = [A;B;C]; solution = inv(D) * [-1;-1;-1];
% plane coefficients
x = rand(300,1)*20; y = rand(300,1)*10-5;
% random points on the plane
z = (-1 - solution(1)*x -solution(2)*y) / solution(3);
fill3(x,y,z,'g','marker','none','edgecolor','none')
% plot triangle ABC
fill3(D(:,1),D(:,2),D(:,3),'r'), grid on
% shade random spots on plane
rotate3d % enable rotation of the plot with the mouse
```

3.2.3 Special cases of planes

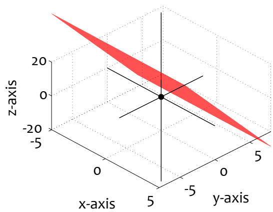
Just like the special cases of lines in 2d, we distinguish the following special cases of planes

$$
\begin{array}{rl}
ax + by + cz = 0 & \text{passing through the origin } (0,0,0) \\
by + cz + d = 0 & \text{parallel to the } x \text{ axis} \\
ax + cz + d = 0 & \text{parallel to the } y \text{ axis} \\
ax + by + d = 0 & \text{parallel to the } z \text{ axis} \\
ax + d = 0 & \text{parallel to the } (y,z) \text{ plane} \\
by + d = 0 & \text{parallel to the } (x,z) \text{ plane} \\
cz + d = 0 & \text{parallel to the } (x,y) \text{ plane}
\end{array}
$$

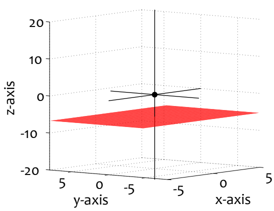
Again, the missing variable or variables tell you which axis or plane defined by axes the plane is parallel to. Three examples of special cases are shown in Figure 3.5.



$2x - 3y - z = 0$
passing through $(0,0,0)$



$5y + z - 10 = 0$
parallel to x axis



$z + 6 = 0$
parallel to (x,y) plane

Figure 3.5: Special cases of planes.

$\oplus\oplus\oplus$ **Example**

Let us construct the equation of a plane parallel to the y axis and passing through point $A(3, -2, 6)$. The equation must be of the form $ax+cz+d = 0$. Also, the coordinates of A must satisfy the equation, hence

$$3a + 6c + d = 0.$$

Any set of values (a, c, d) that satisfy the equation will define a plane parallel to the y axis. Thus we may choose non-zero values for two of the parameters and calculate the third one. Let $a = 1$ and $c = 1$. Then $d = -9$, and the desired equation of the plane parallel to the y axis and containing point A is

$$x + z - 9 = 0.$$

$\ominus\ominus\ominus$

$\oplus\oplus\oplus$ **Example**

Construct the equation of a plane parallel to the (x, z) plane and containing point $B(-8, 9, 6)$. In this case, the plane equation is $by + d = 0$. Substituting B in the equation, we get $9b + d = 0$. Again, we may choose a value for one of the unknown parameters, say $b = 1$.

Then the equation of the plane is

$$y - 9 = 0.$$

$\ominus\ominus\ominus$

3.2.4 Points on the side of a plane

A plane splits $\mathbb{R}^3$ into two half-spaces. We can think of these, again, as a positive and a negative sides of the plane. Substituting the coordinates of points on the same side of the plane in the LHS of the plane equation will result in sums with the same sign.

$\oplus \oplus \oplus$ **Example**

Let's find whether the plane with equation

$$4x - y + 9z - 12 = 0$$

intersects the *unit cube*. (The unit cube has its bottom left corner at (0,0,0), and side of length 1.)

For convenience, the coordinates of the cube vertices can be organised in a 3×8 matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Multiplying on the left by the transposed vector of coefficients $[4, -1, 9]$, we get

$$[4, -1, 9] \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = [0, 9, -1, 8, 4, 13, 3, 12].$$

Adding the free term -12 to all elements, the LHS scores for the 8 vertices of the unit cube are

$$[-12, -3, -13, -4, -8, 1, -9, 0]$$

The signs are different, therefore the plane cuts into the unit cube. What is more, vertex (1,1,1) lies *on* the plane. Figure 3.6 shows the cube and the plane.

$\ominus \ominus \ominus$

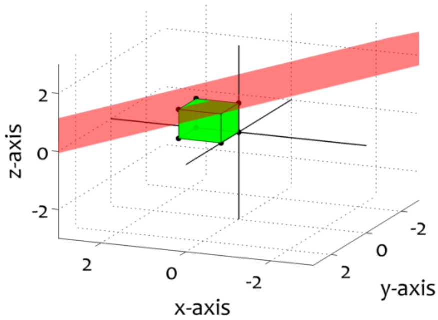


Figure 3.6: Plane $4x - y + 9z - 12 = 0$ intersecting the unit cube.

3.3 Lines in 3d

A line in $\mathbb{R}^3$ is obtained by the intersection of two planes. This means that we need two equations to describe the line

$$ax + by + cz + d = 0$$

$$ex + fy + gz + h = 0$$

To derive the equations of the line between two points $P(x_1, y_1, z_1)$ and $Q(x_2, y_2, z_2)$, we can use any two of the following equations

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

$\oplus \oplus \oplus$ **Example**

Find the line defined by points $A(3, 1, -8)$ and $B(4, 2, -5)$. We can use the following two equations

$$\begin{aligned} \frac{x-3}{4-3} &= \frac{y-1}{2-1} \\ \frac{x-3}{4-3} &= \frac{z-(-8)}{-5-(-8)}. \end{aligned}$$

which reduce to

$$x - y - 2 = 0$$
$$3x - z - 17 = 0$$

$\ominus \ominus \ominus$

3.4 Intersections

3.4.1 Lines in 2d

The coordinates of the intersection point between two lines in 2d must satisfy both line equations. Hence we must solve the two simultaneous equations.

$\oplus \oplus \oplus$ **Example**

Find the intersection point between the lines defined by the following equations

$$5x - 10y - 11 = 0 \qquad \text{and} \qquad 2x + 5y + 6 = 0.$$

This system of equations can be solved through expressing one unknown from one of the equations and substituting it in the other equation (express-and-substitute method). However, we will put to practice our matrix multiplication knowledge. The matrix with coefficients is

$$A = \begin{bmatrix} 5 & -10 \\ 2 & 5 \end{bmatrix},$$

and the right-hand side vector is $[11, -6]^T$. The determinant of A is

$$5 \times 5 - (-10) \times 2 = 45 \neq 0,$$

therefore A is invertible. The solution is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{45} \begin{bmatrix} 5 & 10 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 11 \\ -6 \end{bmatrix} = \begin{bmatrix} -0.1111 \\ -1.1556 \end{bmatrix}$$

The intersection point is $P(-0.1111, -1.1556)$. An illustration is given in Figure 3.7.

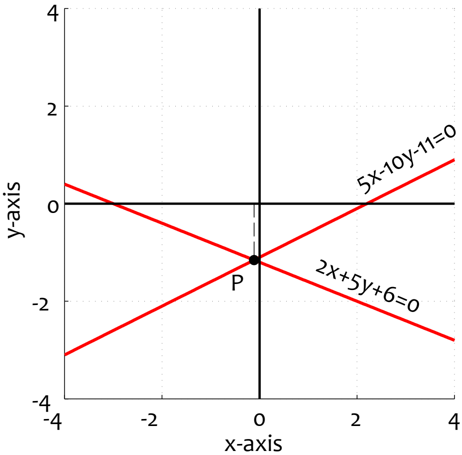


Figure 3.7: Intersection point between lines $5x - 10y - 11 = 0$ and $2x + 5y + 6 = 0$.

⊖⊖⊖

The system of linear equations may have no solution or infinitely many solutions. In both cases the determinant of A is zero. There wil be no solution if the lines are parallel and infinitely many solutions if the lines are identical.

3.4.2 Plane and line in 3d

To find the intersection point between a plane and a line in 3d, we must solve three simultaneous linear equations: two of them defining the line and one defining the plane.

$\oplus\oplus\oplus$ **Example**

Find the intersection point of a line that passes through points $P(4,-1,3)$ and $Q(6,-5,1)$ and a plane defined by the equation

$$x - y + 5z - 10 = 0.$$

First derive the equations of the line

$$\frac{x-4}{6-4} = \frac{y-(-1)}{-5-(-1)} = \frac{z-3}{1-3}$$

$$\frac{x-4}{2} = \frac{y+1}{-4} = \frac{z-3}{-2}.$$

The first equation is

$$(-4)(x-4) = 2(y+1)$$

$$-4x + 16 - 2y - 2 = 0$$

$$2x + y - 7 = 0.$$

The second equation is

$$(-2)(x-4) = 2(z-3)$$

$$x + z - 7 = 0.$$

The system of three equations becomes

$$\begin{aligned} 2x + y &= 7 \\ x + z &= 7 \\ x - y + 5z &= 10 \end{aligned}$$

The coefficient matrix and the RHS vector are respectively

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 5 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 7 \\ 7 \\ 10 \end{bmatrix}.$$

Getting the inverse of A from MATLAB, the solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -0.5 & 2.5 & -0.5 \\ 2.0 & -5.0 & 1.0 \\ 0.5 & -1.5 & 0.5 \end{bmatrix} \begin{bmatrix} 7 \\ 7 \\ 10 \end{bmatrix} = \begin{bmatrix} 9 \\ -11 \\ -2 \end{bmatrix}.$$

The intersection point between the line containing P and Q, and the plane is $(9, -11, -2)$. The solution is illustrated in Figure 3.8.

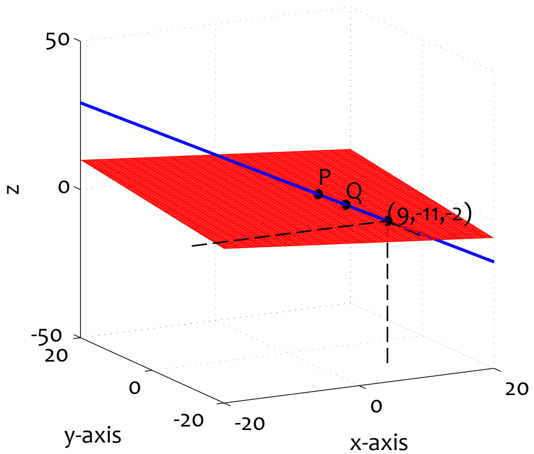


Figure 3.8: Intersection of a 3d line through points P and Q and plane $x - y + 5z - 10 = 0$.

3.5 Circles and spheres

3.5.1 Circles

A circle in $\mathbb{R}^2$ centred at (c_x, c_y) with radius r is described by the following equation

$$(x - c_x)^2 + (y - c_y)^2 = r^2.$$

The left-hand side is the squared Euclidean distance between point (x, y) and the centre of the circle. For all points inside the circle,

$$(x - c_x)^2 + (y - c_y)^2 - r^2 < 0$$

and for all points outside the circle,

$$(x - c_x)^2 + (y - c_y)^2 - r^2 > 0$$

⊕ ⊕ ⊕ **Example**

Find out whether triangle with vertices $A(-2, 5)$, $B(-10, 9)$ and $C(-4, 3)$ is completely enclosed in the circle centred at $(-6, 11)$ and with radius 8.

The equation of the circle is

$$(x + 6)^2 + (y - 11)^2 - 64 = 0.$$

Substitute the coordinates of the three vertices

A: $(-2 + 6)^2 + (5 - 11)^2 - 64 = 16 + 36 - 64 = -12$ inside the circle.
B: $(-10 + 6)^2 + (9 - 11)^2 - 64 = 16 + 4 - 64 = -44$ inside the circle.
C: $(-4 + 6)^2 + (3 - 11)^2 - 64 = 4 + 64 - 64 = 4$ outside the circle.

The circle intersects the triangle. ⊖ ⊖ ⊖

⊕ ⊕ ⊕ Example

The MATLAB code below generates 5000 points and checks which ones fall in the circles with centres $(0.3, 0.7)$ and $(0.6, 0.5)$, and respective radii 0.25 and 0.35. The points in the different regions are plotted with different colours as shown in Figure 3.9.

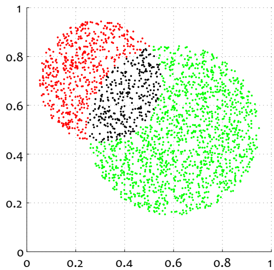


Figure 3.9: Examples of random points inside circles.

```
figure, hold on
axis([0 1 0 1]), axis square, grid on
x = rand(5000,1);y = rand(5000,1); % 5000 random 2d points
index1 = (x-0.3).^2 + (y-0.7).^2 - 0.25^2 < 0; % in c 1
index2 = (x-0.6).^2 + (y-0.5).^2 - 0.35^2 < 0; % in c 2
plot(x(index1&¬index2),y(index1&¬index2),'r.')
plot(x(¬index1&index2),y(¬index1&index2),'g.')
plot(x(index1&index2),y(index1&index2),'k.')
```

⊖ ⊖ ⊖

3.5.2 Intersection between a circle and a line

There are three possible cases as show in Figure 3.10: (1) No intersection, (2) One intersection point and (3) Two intersection points

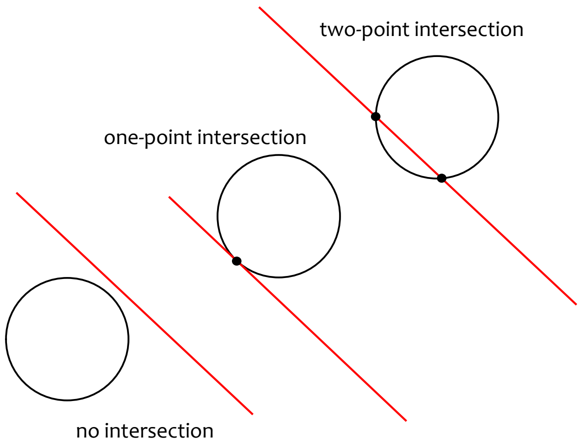


Figure 3.10: Intersection between a line and a circle: the three possible cases

To find which case we have at hand, we form the system of two simultaneous equations: the equation of the line and the equation of the circle. This time the matrix calculations will not help because the equation of the circle is not linear. The way to solve the problem is to express one of the unknowns from the linear equation and substitute in the quadratic equation. This will lead to a quadratic equation of one unknown. If the discriminant is positive, there are two intersection points. Otherwise, if the discriminant is zero, there is one intersection point. Finally, if the discriminant is negative, the line does not intersect the circle.

$\oplus \oplus \oplus$ Example

Find the intersection point(s) of a line passing through points $A(4,-1)$ and $B(-6,9)$ and a circle centred at $(5,3)$ with radius 6.

The equation of the line is

$$\frac{x-4}{-6-4} = \frac{y-(-1)}{9-(-1)}$$

which gives $10x - 40 = -10y - 10$, and finally

$$x + y - 3 = 0.$$

Expressing x from this equation, we have $x = 3 - y$. The equation of the circle is

$$(x-5)^2 + (y-3)^2 - 36 = 0.$$

Then

$$\begin{aligned}
(3-y-5)^2 + (y-3)^2 - 36 &= 0 \\
4 + 4y + y^2 + y^2 - 6y + 9 - 36 &= 0 \\
2y^2 - 2y - 23 &= 0
\end{aligned}$$

The discriminant $(-2)^2 - 4 \times 2 \times (-23) = 4 + 184 = 188$ is positive, therefore there are two intersection points

$$\begin{aligned}
y_1 &= \frac{2+\sqrt{188}}{4} \approx 3.93 \\
x_1 &= 3 - 3.93 = -0.93 \\
y_2 &= \frac{2-\sqrt{188}}{4} \approx -2.93 \\
x_2 &= 3 - (-2.93) = 5.93
\end{aligned}$$

The two intersection points are $P(-0.93, 3.93)$ and $Q(5.93, -2.93)$ as illustrated in Figure 3.11.

⊖ ⊖ ⊖

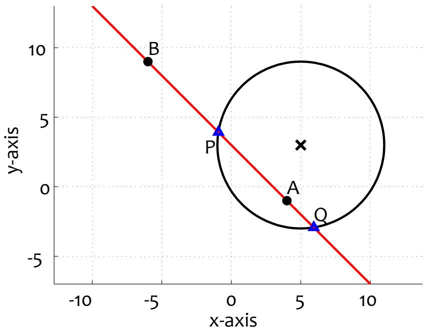


Figure 3.11: Illustration of intersection of a circle and a line.

3.5.3 Spheres

The equation of a sphere in $\mathbb{R}^3$ with centre (c_x, c_y, c_z) and radius r is

$$(x - c_x)^2 + (y - c_y)^2 + (z - c_z)^2 = r^2.$$

Similarly, a *hypersphere* in $\mathbb{R}^n$ with centre $(c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$ and radius r has the following equation

$$(x_1 - c_1)^2 + (x_2 - c_2)^2 + \cdots + (x_n - c_n)^2 = r^2.$$

Again we have the left-hand side of the equation equal to the squared Euclidean distance between the point $(x_1, x_2, \ldots, x_n)$ and the centre. Thus we have the concept of points inside and outside the sphere (hypersphere).

To find the intersection point of a line in 3d with a sphere, we need to solve a system with three equations: the two equations of the line and the equation of the sphere. Again, the system cannot be solved using the matrix inverse approach because the equation of the sphere is not linear.

3.6 Exercises

1. Find the equation of the line passing through points $P(-2,4)$ and $Q(-1,-2)$.

2. Find the equation of the line passing through points $A(t,-1)$ and $B(-2,-3t)$, where t is a parameter.

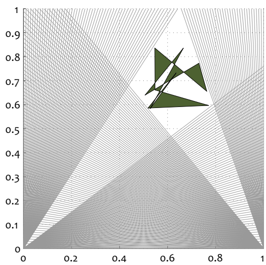
3. Find the equation of the line passing through points $M(12,-1)$ and $N(12,6)$.

4. What is the equation of the line passing through point $A(3,-2)$ and parallel to the y axis?

5. Find out whether the line with equation $-x+4y-6=0$ intersects the triangle with vertices $A(-1,2)$, $B(4,-3)$ and $C(5,2)$.

6. Does the line defined by points $A(4,-1)$ and $B(-9,2)$ intersect the square with lower left corner at point $C(-1,-3)$ and side 4.

7. (M) Use the code that produces Figure 3.3 and add a second ray. It should be fixed at $(1,0)$, and scan the same quadrant starting at the x axis and rotating clockwise. Again, plot only the rays which do not intersect the shape. The final figure should look like the one below.



Hint: The line of the second ray source does not pass through (0,0). Its equation should be derived from two points. One of these can be the source $(1,0)$, and the other can be the end point of a ray of length 1, at angle θ.

8. Construct an equation of a plane containing point $A(-1, 3, 8)$, parallel to the x axis.

9. Construct an equation of a plane containing point $B(4, -1, -7)$, parallel to the (y, z) plane.

10. Construct an equation of a plane parallel to the z axis and containing points $P(9, -1, 0)$ and $Q(6, 9, -6)$.

11. Check whether the plane with equation $x - 3y + 2z + 9 = 0$ intersects the line segment between points $A(-2, 6, 8)$ and $B(6, 0, 5)$.

12. Find the equations of the line defined by points $A(2, 11, -7)$ and $B(-4, -2, 1)$.

13. Using matrix equations, find the intersection point of the lines defined by the following equations:

$$x - 5y + 8 = 0 \qquad \text{and} \qquad 4x - 12y - 9 = 0.$$

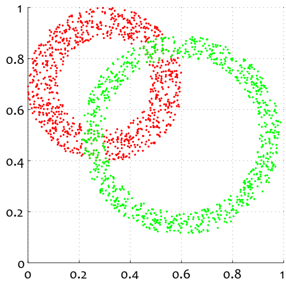
14. Using matrix equations, find the intersection point of the lines defined by the following pairs of points

$$\begin{array}{lll} \text{line 1:} & A(9, -2) \text{ and } B(6, 1) \\ \text{line 2:} & C(-6, 1) \text{ and } D(4, 5) \end{array}$$

15. Write down the equation of a circle centred at $(6, -1)$ and with radius 5.

16. Will a circle centred at $(4, -1)$ and radius 10 completely enclose the segment between points $A(-4, 4)$ and $B(12, -5)$?

17. (M) Use the MATLAB code for Figure 3.9 as a template to write a piece of code that will generate a figure similar to the one shown below.



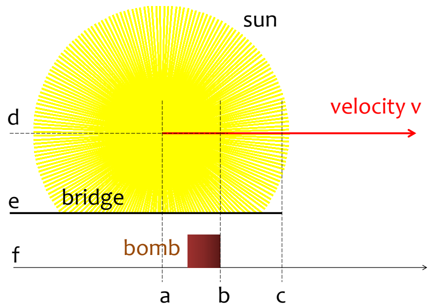
18. If it exists, find the intersection point (or points) between the line passing through points $A(3, -1)$ and $B(-1, -4)$, and a circle centred at $(1, -5)$ with radius 6.

19. Your space ship is visiting 10 planets with space coordinates (in light years) as shown below. Find the length of the path.

PLANET	X	Y	Z
Arit	18	17	-15
Zenit	9	7	15
Dipuzoz	3	0	18
Damet	16	7	-11
Azoben	15	17	19
Ezet	14	-1	4
Putep	-19	14	8
Xuka	17	4	13
Axamu	-1	15	4
Xetata	11	-7	-15

20. The enemy have planted a bomb under a bridge. The bomb is in the shade at the moment, but will go off as soon as a direct sun ray touches it. The sun moves horizontally as shown below. (Note that the figure is not drawn to scale.) The measurements

are as follows. The x-values are $a = 20$, $b = 28$ and $c = 92$, and the y-values are $d = 78$, $e = 68$, and $f = 58$. If the sun moves with velocity $v = 4$ units per hour, how much time (in seconds) do you have to disarm the bomb?



Chapter 4

Orthogonality and collinearity

4.1 Normal vector for a line

4.1.1 Vector orthogonal to a line through two points

A vector is orthogonal to a line if it is orthogonal to any vector on that line. Such a vector is called *a normal vector* for this line.

Consider a line passing through points $A(a_x, a_y)$ and $B(b_x, b_y)$. A vector $\mathbf{v} = [v_1, v_2]^T$ is orthogonal to the line if it is orthogonal to the vector $\mathbf{w}$ beginning at A and ending at B. To understand the calculation of $\mathbf{w}$, refer to the diagram in Figure 4.1.

The position vectors of the two points are $\mathbf{a} = [a_x, a_y]^T$ for A and $\mathbf{b} = [b_x, b_y]^T$ for B. Since $\mathbf{w}$ is the vector between the end-points of these vectors, we can write the following equation

$$\mathbf{a} + \mathbf{w} = \mathbf{b},$$

hence $\mathbf{w} = \mathbf{b} - \mathbf{a} = [b_x - a_x, b_y - a_y]^T$. Then a vector $\mathbf{v}$ will be

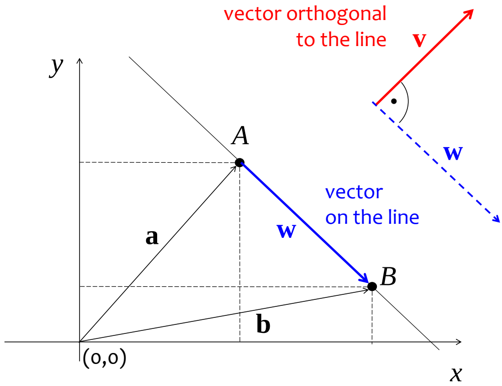


Figure 4.1: Vector $\mathbf{v}$ orthogonal to the line passing through A and B ($\mathbf{v}^T\mathbf{w} = 0$).

orthogonal to the line through A and B if the scalar product of $\mathbf{v}$ and $\mathbf{w}$ is zero

$$\mathbf{v}^T\mathbf{w} = [v_1, v_2] \begin{bmatrix} b_x - a_x \\ b_y - a_y \end{bmatrix} = 0. \tag{4.1}$$

$\oplus \oplus \oplus$ **Example**

Try it with numbers. Find out that whether vector $\mathbf{p} = [4, -6]^T$ is orthogonal to the line through points $A(5, -3)$ and $B(-2, -8)$.

$$[4, -6] \begin{bmatrix} -2 - 5 \\ -8 - (-3) \end{bmatrix} = [4, -6] \begin{bmatrix} -7 \\ -5 \end{bmatrix} = -2 \neq 0.$$

The scalar product is not zero, therefore $\mathbf{p}$ in not orthogonal to the line.

$\ominus \ominus \ominus$

4.1.2 Constructing a unit normal vector for a line through two points

A *unit normal vector* for a line is a vector with magnitude one, orthogonal to the line.

$\oplus \oplus \oplus$ **Example**

Let us construct a unit normal vector $\mathbf{u}$ orthogonal to the line passing through points $A(-1, 9)$ and $B(4, 2)$. To start with, we will construct a normal vector $\mathbf{v}$ and then will scale it to unit length. Vector $\mathbf{v}$ must satisfy (4.1), hence

$$[v_1, v_2] \begin{bmatrix} 4-(-1) \\ 2-9 \end{bmatrix} = [v_1, v_2] \begin{bmatrix} 3 \\ -7 \end{bmatrix} = 3v_1 - 7v_2 = 0.$$

Since we have only one equation to determine the two components of $\mathbf{v}$, we can assign a non-zero value to one of the components and determine the other one. This will ensure that $\mathbf{v}$ satisfies (4.1). Let $v_1 = 7$. Then $3 \times 7 - 7v_2 = 0$, and $v_2 = 3$. (Check $[7, 3][3, -7]^T = 21 - 21 = 0$.)

The next step is to scale $\mathbf{v}$ to unit length. The magnitude of $\mathbf{v}$ is

$$|\mathbf{v}| = \sqrt{7^2 + (-3)^2} = \sqrt{49 + 9} = \sqrt{58} \approx 7.62.$$

Then the unit normal vector for the line through A and B is

$$\mathbf{u} = \left[\frac{7}{7.62}, \frac{3}{7.62}\right]^T = \begin{bmatrix} 0.92 \\ 0.39 \end{bmatrix}$$

$\ominus \ominus \ominus$

The steps can be summarised as follows: (1) Pick a value for v_1; (2) Solve equation (4.1) for v_2; (3) Scale $\mathbf{v}$ to unit length.

$\oplus \oplus \oplus$ **Example**

Construct a unit normal vector $\mathbf{u}$ orthogonal to the line passing through points $A(5,-1)$ and $B(5,9)$. Notice the special case – points A and B have the same x coordinate. Following the three steps, let $v_1 = 5$. Then

$$[5, v_2] \begin{bmatrix} 5-5 \\ 9-(-1) \end{bmatrix} = 10v_2 = 0.$$

Then $v_2 = 0$, therefore $\mathbf{v} = [5,0]^T$ and $\mathbf{u} = [1,0]^T$. Do you recognise this vector? This is the basis vector $\mathbf{i} = [1,0]^T$ in $\mathbb{R}^2$. Indeed, the line has equation $x = 5$, and is parallel to the y axis. Therefore the unit normal vector will be a unit vector along the x axis, which is exactly $\mathbf{i}$.

If the special case happens to be such that the y coordinates of the two points are the same, the line is parallel to the x axis and the unit normal vector is the basis vector $\mathbf{j} = [0,1]^T$.

⊖ ⊖ ⊖

4.1.3 Constructing a unit normal vector for a line $ax + by + c = 0$

How do we construct a unit normal vector for a line given by the equation $ax + by + c = 0$?

First, we will prove that vector $[a,b]^T$ is orthogonal to the line. To do this, we must show that if we pick any two points on the line, say $P(p_1,p_2)$ and $Q(q_1,q_2)$, their coordinates will satisfy equation (4.1)

$$[a, b] \begin{bmatrix} q_1 - p_1 \\ q_2 - p_2 \end{bmatrix} = a(q_1 - p_1) + b(q_2 - p_2) = 0.$$

The coordinates of the points satisfy the equation of the line

$$ap_1 + bp_2 + c = 0 \qquad (4.2)$$

$$aq_1 + bq_2 + c = 0 \qquad (4.3)$$

Subtract (4.2) from (4.3) to get

$$aq_1 + bq_2 + c - ap_1 - bp_2 - c = 0.$$

Cancelling the free terms and rearranging, we arrive at

$$a(q_1 - p_1) + b(q_2 - p_2) = 0,$$

which proves that vector $[a, b]^T$ is orthogonal to the line. The magnitude of the vector is $\sqrt{a^2 + b^2}$. Then the unit normal vector for the line $ax + by + c = 0$ is

$$\mathbf{u} = \left[\frac{a}{\sqrt{a^2 + b^2}} \,, \quad \frac{b}{\sqrt{a^2 + b^2}} \right]^T .$$

$\oplus \oplus \oplus$ **Example**

Find a unit normal vector for the line

$$-4x - 7y + 2 = 0.$$

The unit normal vector is

$$\mathbf{u} = \left[\frac{-4}{\sqrt{(-4)^2 + (-7)^2}} \,, \quad \frac{-7}{\sqrt{(-4)^2 + (-7)^2}} \right]^T = [-0.4961, \ -0.8682]^T.$$

$\ominus \ominus \ominus$

How do we check if a vector $\mathbf{v} = [v_1, v_2]$ is orthogonal to a line with equation $ax + by + c = 0$? For this to be the case, $\mathbf{v}$ must be a re-scaled version of $[a, b]^T$. Then the following will hold

$$\frac{v_1}{a} = \frac{v_2}{b}, \qquad a \neq 0, \ b \neq 0.$$

If the above does not hold, $\mathbf{v}$ is not orthogonal to the line. It is even easier to check when either $a = 0$ or $b = 0$. For example, a vector $[0, k]$ is orthogonal to line $by + c = 0$ for any b, c and k. Similarly, a vector $[k, 0]$ is orthogonal to line $ax + c = 0$ for any a, c and k.

4.1.4 Line through a point and orthogonal to a vector

How can we find the equation of a line which passes through a given point $A(a_1, a_2)$ and is orthogonal to vector $\mathbf{v} = [v_1, v_2]^T$?

Start by constructing an equation of a line using the components of $\mathbf{v}$ as the coefficients

$$v_1 x + v_2 y + c = 0.$$

This way we guarantee that the line is orthogonal to $\mathbf{v}$. Second, point A must satisfy the equation because A lies on the line. Therefore

$$v_1 a_1 + v_2 a_2 + c = 0,$$

and

$$c = -v_1 a_1 - v_2 a_2.$$

The equation of the line is

$$v_1 x + v_2 y - v_1 a_1 - v_2 a_2 = 0.$$

$\oplus \oplus \oplus$ **Example**

Find the equation of a line which passes through a given point $P(6, -4)$ and is orthogonal to vector $\mathbf{v} = [-2, 10]^T$?

$$-2x + 10y + 12 + 40 = 0$$

$$x - 5y - 26 = 0.$$

Check with P: $6 - 5(-4) - 26 = 6 + 20 - 26 = 0$.

$\ominus \ominus \ominus$

4.2 Normal vector for a plane

4.2.1 Normal vector for a plane through three points

A vector is orthogonal to a plane if it is orthogonal to any two vectors on this plane. Such a vector is called a *normal vector* for this plane. Figure 4.2 illustrates the concept.

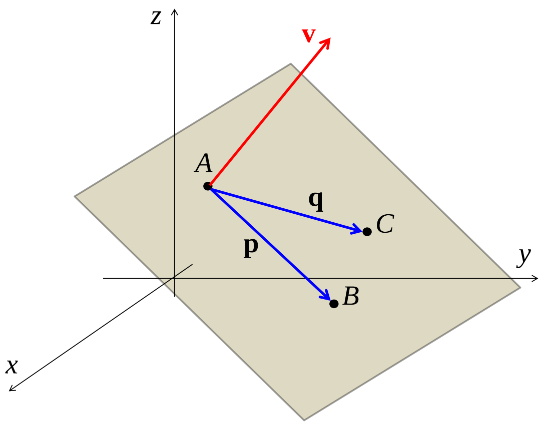


Figure 4.2: Vector $\mathbf{v}$ orthogonal to the plane passing through A, B and C ($\mathbf{v}^T\mathbf{p} = 0$ and $\mathbf{v}^T\mathbf{q} = 0$).

To check whether a vector $\mathbf{v}$ is orthogonal to a plane containing three points, $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$ and $C(c_1, c_2, c_3)$, we should calculate the scalar products of $\mathbf{v}$ and two vectors between any pairs of points, as shown in the figure. Let $\mathbf{p} = [b_1 - a_1, b_2 - a_2, b_3 - a_3]^T$ be the vector from A to B, and $\mathbf{q} = [c_1 - a_1, c_2 - a_2, c_3 - a_3]^T$, the vectors from A to C.

$\oplus \oplus \oplus$ **Example**

Let the three points be $A(2, 8, -3)$, $B(0, 9, -3)$ and $C(3, 2, -2)$, and $\mathbf{v} = [1, 2, 3]^T$. Check whether $\mathbf{v}$ is orthogonal to the plane defined by A, B and C. Construct the two vectors contained in the plane.

$$\mathbf{p} = [0 - 2, 9 - 8, -3 - (-3)]^T = [-2, 1, 0]^T$$

$$\mathbf{q} = [3 - 2, 2 - 8, -2 - (-3)]^T = [1, -6, 1]^T$$

Calculate the scalar products

$$\mathbf{v}^T\mathbf{p} = [1, 2, 3][-210]^T = -2 + 2 + 0 = 0$$

$$\mathbf{v}^T\mathbf{q} = [1, 2, 3][1, -6, 1] = 1 - 12 + 3 = -8 \neq 0.$$

Since the two scalar products are not *both* zero, $\mathbf{v}$ is not orthogonal to the plane defined by A, B and C.

$\ominus \ominus \ominus$

4.2.2 Constructing a unit normal vector for a plane through three points

If we have the coordinates of three points on a plane, $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$ and $C(c_1, c_2, c_3)$, we can define two vectors: $\mathbf{p} = [b_1 - a_1, b_2 - a_2, b_3 - a_3]^T$ and $\mathbf{q} = [c_1 - a_1, c_2 - a_2, c_3 - a_3]^T$. For a vector $\mathbf{v} = [v_1, v_2, v_3]^T$ to be orthogonal to the plane, the following two equations must hold

$$\mathbf{v}^T\mathbf{p} = (b_1 - a_1)\, v_1 + (b_2 - a_2)\, v_2 + (b_3 - a_3)\, v_3 = 0$$

and

$$\mathbf{v}^T\mathbf{q} = (c_1 - a_1)\, v_1 + (c_2 - a_2)\, v_2 + (c_3 - a_3)\, v_3 = 0$$

Since we have two equations with three unknowns, we can choose one of the components of $\mathbf{v}$, and solve the system of two simultaneous linear equations to determine the other two components. In the general case, any non-zero value can be assigned to any of the components. After calculating **v**, we should calculate its magnitude and divide each component by that value. This will scale the normal vector **v** for the plane to unit length.

$\oplus\oplus\oplus$ **Example**

Let $A(-1,3,5)$, $B(-3,-2,6)$ and $C(5,-4,1)$ be three points in a plane. Construct a unit normal vector for this plane.

Form vectors **p** and **q**

$$\mathbf{p} = [(-3-(-1)),(-2-3),(6-5)]^T = [-2,-5,1]^T$$

$$\mathbf{q} = [(5-(-1)),(-4-3),(1-5)]^T = [6,-7,-4]^T.$$

Choose $v_1 = 10$. Then the two equations for v_2 and v_3 are

$$\begin{aligned} 10(-2) - 5v_2 + v_3 &= 0 \\ 10 \times 6 - 7v_2 - 4v_3 &= 0 \end{aligned}$$

The system can be rewritten as

$$\begin{aligned} -5v_2 + v_3 &= 20 \\ 7v_2 + 4v_3 &= 60 \end{aligned}$$

The coefficient matrix is

$$A = \begin{bmatrix} -5 & 1 \\ 7 & 4 \end{bmatrix}.$$

$\det(A) = -20 - 7 = -27 \neq 0$, therefore A is non-singular, and the solution of the system is

$$\begin{bmatrix} v_2 \\ v_3 \end{bmatrix} = -\frac{1}{27} \begin{bmatrix} 4 & -1 \\ -7 & -5 \end{bmatrix} \begin{bmatrix} 20 \\ 60 \end{bmatrix}$$

$$= -\frac{1}{27} \begin{bmatrix} 20 \\ -440 \end{bmatrix} = \begin{bmatrix} -0.7407 \\ 16.2963 \end{bmatrix}$$

Then $\mathbf{v} = [10, -0.7407, 16.2963]^T$. The magnitude of $\mathbf{v}$ is

$$|\mathbf{v}| = \sqrt{10^2 + (-0.7407)^2 + (16.2963)^2} = \sqrt{366.1180} \approx 19.13$$

The unit normal vector for the plane containing A, B and C is therefore

$$\mathbf{u} = \left[\frac{10}{19.13}, \frac{-0.7407}{19.13}, \frac{16.2963}{19.13}\right]^T = [0.52,\ -0.04,\ 0.85]^T$$

⊖⊖⊖

4.2.3 Constructing a unit normal vector for a plane $ax + by + cz + d = 0$

The argument here is very similar to that about the normal vector of a line given with its generic equation. It can be shown that vector $\mathbf{v} = [a, b, c]^T$ is a normal vector for the plane with equation $ax + by + cz + d = 0$. To derive the unit normal vector, scale $\mathbf{v}$ to magnitude 1

$$\mathbf{u} = \left[\frac{a}{\sqrt{a^2 + b^2 + c^2}}, \frac{b}{\sqrt{a^2 + b^2 + c^2}}, \frac{c}{\sqrt{a^2 + b^2 + c^2}}\right]^T.$$

4.2.4 Plane with a given normal vector passing through a point

Let us find the equation of a plane with a given normal vector $\mathbf{v} = [v_1, v_2, v_3]^T$, knowing that point $P(p_1, p_2, p_3)$ lies on that plane.

Constructing the equation

$$v_1 x + v_2 y + v_3 z + d = 0,$$

we guarantee that, for any d, the plane with this equation has $\mathbf{v}$ as a normal vector. To find d, we can use the equation of the plane for the given point. Since P is on the plane, its coordinates must satisfy the above equation

$$v_1 p_1 + v_2 p_2 + v_3 p_3 + d = 0.$$

The only unknown in this equation is d, therefore

$$d = -v_1 p_1 - v_2 p_2 - v_3 p_3.$$

$\oplus \oplus \oplus$ **Example**

Find the equation of a plane with normal vector $\mathbf{v} = [7, -2, 4]^T$, knowing that point $P(2, -1, 3)$ lies on that plane.

Find the free term

$$d = -7 \times 2 - (-2)(-1) - 4 \times 3 = -14 + 2 - 12 = -24$$

and write the equation of the plane

$$7x - 2y + 4z - 24 = 0.$$

$\ominus \ominus \ominus$

4.3 Vector parallel to a line

A vector is parallel to a line if it is parallel to any vector on this line.

4.3.1 Vector parallel to a line through two points

A line through points A and B contains vector from A to B, $\mathbf{w} = \mathbf{b} - \mathbf{a}$, where $\mathbf{a}$ and $\mathbf{b}$ are the position vectors of A and B, respectively (Figure 4.3).

Vector $\mathbf{w}$ and any vector obtained as a multiple of $\mathbf{w}$, say $\mathbf{p} = k\mathbf{w}$, where $k \neq 0$ are parallel to the line containing points A and B.

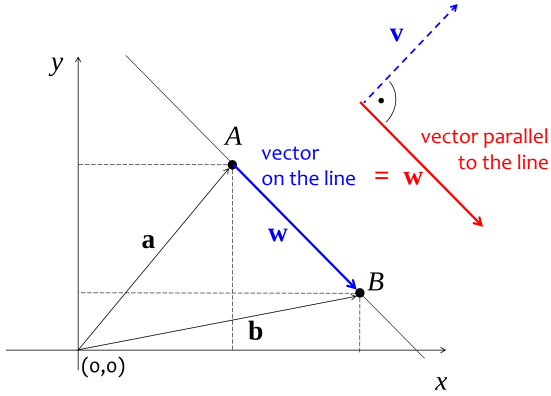


Figure 4.3: Vector $\mathbf{v}$ parallel to the line passing through A and B.

⊕ ⊕ ⊕ **Example**

Find a vector $\mathbf{q}$ of magnitude 10, parallel to the line passing through points $A(-3, 5)$ and $B(-6, 1)$.

Vector $\mathbf{w} = [(-3 - (-6)), (5 - 1)]^T = [3, 4]^T$ is parallel to the line. Scaling to unit magnitude, we have $|\mathbf{w}| = \sqrt{9 + 16} = 5$, and

$$\mathbf{u} = \left[\frac{3}{5}, \frac{4}{5}\right]^T = [0.6, 0.8]^T,$$

which is a unit vector parallel to the line. Finally, to arrive at the desired vector, we expand $\mathbf{u}$ to magnitude 10 by just multiplying it by 10

$$\mathbf{q} = [6, 8]^T.$$

⊖ ⊖ ⊖

4.3.2 Constructing a unit vector parallel to line $ax + by + c = 0$

Consider the generic equation of the line $ax + by + c = 0$. How do we construct a vector $\mathbf{w} = [w_1, w_2]^T$, parallel to the line? Knowing that vector $\mathbf{v} = [a, b]^T$ is orthogonal to the line, it must also be orthogonal to $\mathbf{w}$. Then $\mathbf{v}^T\mathbf{w} = 0$, therefore

$$aw_1 + bw_2 = 0.$$

Any w_1 and w_2 that satisfy the equation will make a vector parallel to the line. Then we can choose one of the values and solve for the other. Let $w_1 = b$ ($b \neq 0$). Then $w_2 = -a$, and the vector parallel to line $ax + by + c = 0$ is

$$\mathbf{w} = \begin{bmatrix} b \\ -a \end{bmatrix}.$$

To transform $\mathbf{w}$ into a unit vector $\mathbf{u}$, find its magnitude and divide each element by it

$$\mathbf{w} = \left[\frac{b}{\sqrt{a^2 + b^2}}, \frac{-a}{\sqrt{a^2 + b^2}}\right]^T.$$

4.3.3 Line through a point and parallel to a vector

Let us find the equation of a line passing through point $A(a_x, a_y)$ and parallel to vector $\mathbf{w} = [w_1, w_2]^T$.

Knowing that vector $\mathbf{w} = [w_2, -w_1]^T$ is orthogonal to the line, we can derive the following equation for the line

$$w_2x - w_1y + c = 0.$$

To find c, we substitute the coordinates of point A in the equation

$$a_xw_2 - a_yw_1 + c = 0$$

$$c = a_y w_1 - a_x w_2.$$

Then the equation of the line is

$$w_2 x - w_1 y + a_y w_1 - a_x w_2 = 0.$$

$\oplus\oplus\oplus$ **Example**

Find the equation of the line through point $A(9,-1)$ parallel to vector $\mathbf{w} = [8,3]^T$. Using the above equation,

$$3x - 8y + (-1) \times 8 - 9 \times 3 = 0$$

$$3x - 8y - 35 = 0.$$

$\ominus\ominus\ominus$

4.4 A little MATLAB

Hello MATLAB! Long time no see. The function below takes a point $A(a_1, a_2)$ and a vector $\mathbf{v} = [v_1, v_2]^T$ and returns a, b and c of a line passing through A. The third input parameter, 'mode' should be set to 1 for 'line orthogonal to $\mathbf{v}$', and to 2 for 'line parallel to $\mathbf{v}$'. Notice how we handle the special cases

- $\mathbf{v} = [0,0]^T$. Return an error.
- $\mathbf{v} = [0, v_2]^T$. For mode = 1 (orthogonal), the line has equation $x - a_1 = 0$. For mode = 2 (parallel), the equation is $y - a_2 = 0$.
- $\mathbf{v} = [v_1, 0]^T$. For mode = 1 (orthogonal), the line has equation $y - a_2 = 0$. For mode = 2 (parallel), the equation is $x - a_1 = 0$.

```
function [a,b,c] = line_through_point(A,v,mode)
% A        point coordinates
```

```
% v        vector with 2 components
% mode     1 for "orthogonal", 2 for "parallel"

if ¬any(v), error('Vector must have a non-zero entry.'), end
switch mode
    case 1
        if v(1) == 0, [a,b,c] = deal(1,0,-A(1));
        elseif v(2) == 0, [a,b,c] = deal(0,1,-A(2));
        else  [a,b,c] = deal(v(1),v(2),-v(:)'*A(:));
        end
    case 2
        if v(1) == 0, [a,b,c] = deal(0,1,-A(2));
        elseif v(2) == 0, [a,b,c] = deal(1,0,-A(1));
        else  [a,b,c] = deal(v(2),-v(1),A(2)*v(1)-A(1)*v(2));
        end
    otherwise
        error('Mode must be 1 (orthogonal) or 2 (parallel)')
end
```

⊕ ⊕ ⊕ **Example**

Let's try the function with the MATLAB script below. The function must be saved as `line_through_point.m`. Put as input parameters $A(2,-10)$, $\mathbf{v} = [-6, 9]^T$ and mode 1 (orthogonal). Then choose two points on the line on either side of $x = 2$. Calculate the respective y values on the line. Plot the segment between the two points, then plot point A, and finally, plot vector v with its tail at A. See the result in Figure 4.4.

```
clear, clc, close all
[a,b,c] = line_through_point([2,-10],[-6,9],1);
% line: ax + by + c = 0
y = @(x) (-c-a*x)/b;
% calculate y on the line for a given x
% (assuming b is not = 0)
[x1,x2] = deal(-7,16); % set x for the segment
[y1,y2] = deal(y(x1),y(x2)); % calculate y for the segment
figure('color','w'), hold on, grid on % set up the figure
plot([x1,x2],[y1,y2],'k-') % plot the segment
plot(2,-10,'ro','linewidth',3) % plot point A
quiver(2,-10,-6,9,'k--') % plot v
```

```
axis equal
```

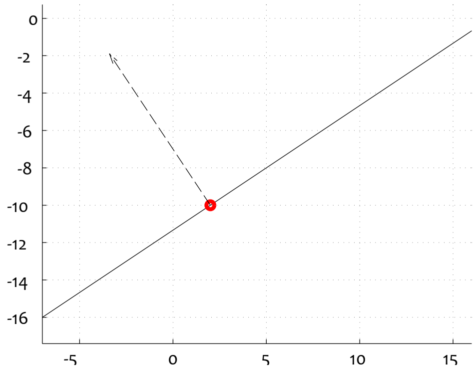


Figure 4.4: MATLAB output from the script calling function

$\ominus\ominus\ominus$

Well, that's the end of me giving you certain assignment questions. But I can always stick them in the exam script... ha-ha-ha

4.5 Exercises

1. Check whether vector $\mathbf{a} = [-2, 3]^T$ is orthogonal to the line passing through points $P(4, -1)$ and $Q(5, 8)$.

2. For what value of z is vector $\mathbf{a} = [-7, (1 - z)]^T$ orthogonal to a line through points $M(-9, 2)$ and $N(4, -4)$?

3. Construct a unit normal vector $\mathbf{u}$ orthogonal to the line passing through points $P(6, -8)$ and $Q(3, -1)$.

4. Construct a unit normal vector $\mathbf{u}$ orthogonal to the line passing through points $A(-1, 4)$ and $B(-5, 4)$.

5. Find a unit normal vector for the line $5x - 3y + 6 = 0$. Plot the line and the unit normal vector.

6. Is vector $[6, -8]^T$ orthogonal to the line with equation $-3x + 4y - 5 = 0$?

7. For what value of w is vector $[-10,\ 5w - 6]^T$ orthogonal to the line with equation $5x - 2y - 9 = 0$?

8. Find the equation of a line passing through point $M(-4, 8)$ and orthogonal to vector $[9, -5]^T$.

9. Find the equation of a line passing through point $M(-6, 8)$ and orthogonal to vector $[9, 0]^T$.

10. Find the equation of a line passing through point $A(3t - 4, 4t)$ and orthogonal to vector $[3, -5]^T$, where t is a parameter.

11. Check whether vector $[-1, 4, 6]^T$ is orthogonal to the plane containing points $A(3, -1, 8)$, $B(9, -4, 11)$ and $C(7, 3, 6)$.

12. For what values of the parameters t and k, is vector $\mathbf{v} = [3, t, k]^T$ orthogonal to the plane containing points $A(3, -1, 8)$, $B(9, -4, 11)$ and $C(7, 3, 6)$?

13. Find a unit normal vector for the plane passing through points $A(6, 2, -4)$, $B(1, 9, -6)$ and $C(-3, 1, 5)$.

14. Find a unit normal vector for the plane with equation $6x - 3y - 12z + 8 = 0$.

15. Find the equation of a plane with normal vector $\mathbf{v} = [1, 8, -6]^T$, knowing that point $P(4, -10, -1)$ lies on that plane.

16. Find a unit vector parallel to the line passing through points $A(5, 1)$ and $B(-4, -2)$.

17. Find a vector of magnitude k, parallel to the line passing through points $A(m, -2)$ and $B(6, t)$, where k, m and t are parameters.

18. Construct a unit vector parallel to line with equation $5x-3y+6=0$.

19. Find the equation of the line through point $A(7,4)$ parallel to vector $\mathbf{w}=[-2,8]^T$.

20. (M) Write a MATLAB function that will return a unit normal vector to a line or a plane. The first line should be

```
function v = unit_normal_vector(z)
```

where z is either $[a,b,c]^T$ (coefficients of a line equation) or $[a,b,c,d]^T$ (coefficients of a plane equation). The output v should have 2 or 3 components depending on the size of z. Write a MATLAB script to illustrate the calculations by a plot for the 2-dimensional case.

Chapter 5

Probability

Alea iacta est

Here we will only glide over the fabulously complex and intriguing world of probability theory. This chapter will introduce some basic concepts and ideas, and hopefully lure you into exploring the subject further.

5.1 Counting rule, permutations, combinations

5.1.1 Counting rule

The counting rule says that, if there are two events, A and B, and A can give n different outputs, while B can give m different outputs, the total number of possible outputs is $m \times n$.

$\oplus \oplus \oplus$ **Example**

Figure 5.1 illustrates the counting rule. If you have 4 shirts and 3 neck ties, how many possible different outfits are there?

Figure 5.1: Illustration of the counting rule

As any shirt can be worn with any tie (however hideous the colour combination) there are $4 \times 3 = 12$ possible outfits. ⊖ ⊖ ⊖

The counting generalises for k events $A_1, A_2, \ldots, A_k$. Let m_i be the number of possible outcomes for event A_i. Then the total number of possible different outcomes is $m_1 \times m_2 \times \ldots \times m_k$.

⊕ ⊕ ⊕ **Example**

Suppose that there are 20 question types for the exam on Maths for Computing. For each question type, I have 10 sample questions. How many different exam scripts can I generate from this collection?

Since I can combine any of the 10 questions of type 1 with any question of type 2, should there be only 2 questions on the exam (type 1 and type 2), I will have $10 \times 10 = 100$ choices. Each of these can be combined with any of the 10 questions of type 3, giving 1000 possibilities, and so on. Therefore, I have

$$\underbrace{10 \times 10 \times \cdots \times 10}_{20 \text{ times}} = 10^{20}$$

possible scripts. Million billions and more, so, prepare yourselves! :)

⊖⊖⊖

5.1.2 Permutations

Arrangement of n different elements

In how many ways can we arrange 3 books on a shelf? Let's call them A, B and C. Then we have the following 6 possible arrangements

ABC ACB BAC BCA CAB CBA

Each such arrangement is called *a permutation*. The number of permutations can be calculated using the counting rule as illustrated in Figure 5.2. Suppose that you have n different objects and n fixed positions to place them into. For position number 1, there are n possible candidates. So, according to the counting rule, there are n outcomes for position 1. Once the element is placed in position 1, there are $n-1$ elements left, from which we can choose the resident in position number 2. Thus the number of outcomes for position 2 is $n-1$. By the same logic, position number 3 will have $n-2$ possible outcomes, and so on. Position number n will have no alternatives because there will be only one element left, unassigned to a position thus far. According to the counting rule, the number of possible arrangements is

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1.$$

This quantity is known as n factorial (pronounced [fak'toreeul]). n factorial ($n!$) is the product of all integers from n down to 1. Table 5.1 shows the factorials for the integers from 1 to 12. By convention, $0! = 1$.

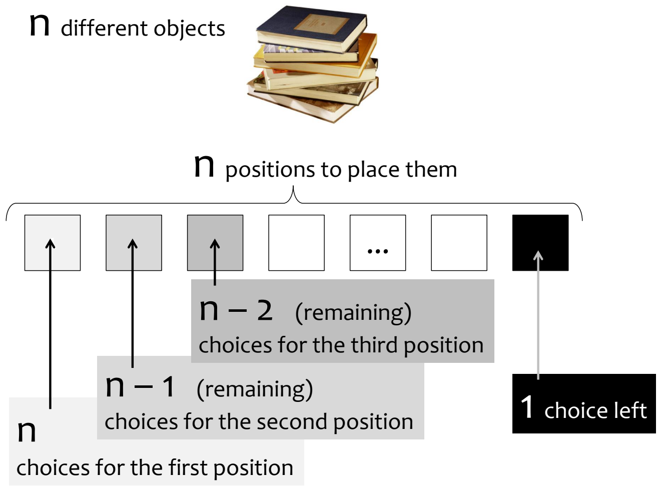


Figure 5.2: Illustration of permutation count.

$\oplus \oplus \oplus$ **Example**

In how many ways can you arrange 6 swimmers in 6 corridors? And the answer is $6! = 6.5.4.3.2.1 = 720$ ways. $\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

The MATLAB command `randperm(n)` will return a random permutation of the integers from 1 to n. The code below plays a game where the player has to guess the name of a European country from an anagram displayed with capital letters.

```
clear, close, clc
```

```
2  C = {'Albania';'Andorra';'Austria';'Belarus';'Belgium';...
3      'Bosnia and Herzegovina';'Bulgaria';'Croatia';...
4      'Cyprus';'Czech Republic';'Denmark';'Estonia';...
5      'Finland';'France';'Germany';'Greece';'Hungary';...
6      'Iceland';'Ireland';'Italy';'Kosovo';'Latvia';...
7      'Liechtenstein';'Lithuania';'Luxembourg';'Malta';...
8      'Moldova';'Monaco';'Montenegro';'Netherlands';...
9      'Norway';'Poland';'Portugal';...
10     'Republic of Macedonia';'Romania';'Russia';...
11     'San Marino';'Serbia';'Slovakia';'Slovenia';...
12     'Spain';'Sweden';'Switzerland';'Turkey';'Ukraine';...
13     'United Kingdom';'Vatican City'};
14 S = upper(C{randi(numel(C))}); % choose a country
15 % and turn upper case
16 SS = S(randperm(length(S))); % permute the letters
17 fprintf('Guess the country: %s\n',SS) % print the anagram
18 UserC = input('Your guess?  --> ','s'); % get the user guess
19 if strcmpi(UserC,S) % compare the string ignoring the case
20     fprintf('Congratulations! %s is correct!\n',S)
21 else disp(['Bad luck!!! The country was ' S '.']), end
```

Note the **randperm** command on line 13. It is used to scramble the letters of the name of the randomly chosen country. The **randi(k)** command generates an integer between 1 and k. It is used to select at random one of the elements of the cell array C.

⊖ ⊖ ⊖

Arrangement of r out of n different elements

Suppose that we want to place only r of the n elements in r fixed places. Then the number of possible outcomes is

$$\underbrace{n \times (n-1) \times \cdots \times (n-r+1)}_{r \text{ positions}} = \frac{n!}{(n-r)!}.$$

⊕ ⊕ ⊕ Example

A club has 50 members. The club needs to select a governing body consisting of a President, Deputy-president, Secretary and Treasurer. Each member of the club is qualified for each of these positions. How many possible different governing bodies can be formed from the members of this club?

Since there are 4 named positions, we have

$$50 \times 49 \times 48 \times 47 = 5527200 \text{ possible governing bodies.}$$

⊖ ⊖ ⊖

Table 5.1: Table of factorials

n	$n!$
1	1
2	2
3	6
4	24
5	120
6	720
7	5040
8	40320
9	362880
10	3628800
11	39916800
12	479001600

Arrangement of n elements where r are the same

Up to now, the elements we were arranging were all different. Assume now that r of the n elements are not distinguishable. For example, how many permutations of the letters of the word MUM are there? According to the argument above, if all letters were different, we would have $3! = 6$ permutations. However, the different permutations are actually 3: MUM, UMM and MMU. The reason for this discrepancy is that there are repeated letters. Suppose that the Ms are tagged. For any permutation, the two Ms may be swapped, and the permutation will not change. If there were three Ms in the word, any "mini"-permutation of the three Ms would leave to the same permutation. Therefore, the total number of permutations must be divided by the number of repeated permutations. Then the number of permutations of n objects where r are the same is

$$\frac{n!}{r!}.$$

Let's check with the MUM example. There are 2 Ms, therefore the number of repeated permutations is $2! = 2$. The number of possible permutations is therefore $3!/2! = 6/2 = 3$, which confirms our finding.

It is not difficult to show, using the same argument, that if we have k repeated elements, where element i is repeated r_i times, the total number of permutations is

$$\frac{n!}{r_1!r_2!\cdots r_k!}.$$

⊕ ⊕ ⊕ **Example**

How many different anagrams of the word MISSISSIPPI are there?

There are one 'M', 4 'I's, 4 'S's and 2 'P's. Then the number of different permutations is

$$\frac{11!}{4!4!2!} = \frac{11.10.9.8.7.6.5.4.3.2.1}{4.3.2.1.4.3.2.1.2.1} = 11.10.9.7.5 = 34650.$$

⊖ ⊖ ⊖

5.1.3 Combinations

Suppose that the places of the elements are not fixed, their arrangement does not matter, and we are only interested to know in how many ways we can select k out of n different objects. Each such selection of objects is called a *combination*.

Let's start with selecting k objects with *fixed places*. We found before that the number of possible arrangements is

$$M = \frac{n!}{(n-k)!}$$

Each such arrangement is a permutation of the k selected elements. The same combination (a set of k elements) will have $k!$ copies counted within the total number of M arrangements. Therefore, the number of combinations (possible selections) of k elements out of n is

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

The notation $\binom{n}{k}$ is the *binomial coefficient* 'n-choose-k'.

$\oplus \oplus \oplus$ **Example**

Let's find out in how many ways we can select two out of four elements A, B, C and D. The possible choices are (AB), (AC), (AD), (BC), (BD) and (CD), a total of 6. Indeed, the binomial coefficient is

$$\binom{4}{2} = \frac{4!}{2!\ 2!} = \frac{4.3.2.1}{2.1.2.1} = 6.$$

$\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

Consider the lottery, 6 out of 49. How many possible different lottery combinations are there?

In a standard draw, 6 numbers are selected out of 49. Therefore the number of different lottery combinations is

$$\binom{49}{6} = \frac{49!}{43!\ 6!} = 13\,983\,816 \text{ combinations.}$$

$\ominus\ominus\ominus$

Here are some interesting properties of the binomial coefficient

- $\binom{n}{n} = \binom{n}{1} = 1$;
- $\binom{n}{k} = \binom{n}{n-k}$;
- $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

Finally, there is a short-cut formula for calculating the binomial coefficient $\binom{n}{k}$. Take the senior k terms from $n!$ and divide this product by $k!$.

$\oplus\oplus\oplus$ **Example**

Calculate $\binom{49}{6}$. Take $\underbrace{49.48.47.46.45.44}_{6\ \text{terms}}$ and divide by 6.5.4.3.2.1

$$\binom{49}{6} = \frac{49.48.47.46.45.44}{6.5.4.3.2.1} = 49.47.46.3.44 = 13\,983\,816.$$

$\ominus\ominus\ominus$

5.2 Events and Probability

5.2.1 Sample space

Random experiments are experiments with uncertain outcome. Let's call the outcome "an event". An event may or may not happen if the random experiment is repeated. For example, in tossing a coin, the event "Tail" may or may not occur.

A *sample space* is the set of all possible outcomes of a random experiment.

$\oplus \oplus \oplus$ **Example**

The sample space for the experiment 'Toss a coin' is {Head,Tail}. Similarly, the sample space for the experiment 'Toss a coin twice' is

$$\{HH, HT, TH, TT\}.$$

$\ominus \ominus \ominus$

The definition of a sample space depends on what outcomes we are interested in. For example, the experiment 'Toss a coin twice' may have the following different sample spaces:

- $\{HH, HT, TH, TT\}$
- {same, different}
- {0 Heads, 1 Head, 2 Heads}
- {0 Tails, 1 Tail, 2 Tails}

More often we will be interested in sample spaces where the outcomes have the same chance to appear, called *equiprobable* sample spaces. The first two items in the above list are such sample spaces. Note that the order of the elements in the sample space does not matter because the sample space is a *set*.

A question that will often appear is "What is the *cardinality* of the sample space?" Or, in other words, how many elements does the set contain?

$\oplus \oplus \oplus$ **Example**

For example, what is the cardinality of the following sample space for the experiment 'Toss a coin 4 times'?

$$S = \{HHHH,\ HHHT,\ \ldots,\ TTTT\} \tag{5.1}$$

Knowing that there are 4 tosses and each can result in 2 possible outcomes, according to the counting rule, there are $2 \times 2 \times 2 \times 2 = 2^4 = 16$ possible outcomes, therefore the cardinality of S is

$$|S| = 16.$$

$\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

Consider a random experiment where 4 students are randomly selected from a class of 42 to share their opinions of the degree course they are following with an accreditation committee. What is the sample space for this experiment and what is its cardinality?

The sample space is the collection of all possible sets of 4 students, and its cardinality is

$$|S| = \binom{42}{4} = \frac{42.41.40.39}{4.3.2.1} = 42.41.5.13 = 111\,930.$$

$\ominus \ominus \ominus$

5.2.2 Elementary and compound events

Every outcome of a random experiment is an event. When the event is a member of the sample space, it is called an *elementary event*. For example, the outcome $HHTH$ is an elementary event of the experiment 'Toss a coin 4 times', where the sample space is as in equation (5.1).

When an event includes more than one outcome from the sample space, the event is called a *compound* event.

$\oplus \oplus \oplus$ **Example**

We can form a compound event 'Fewer than three Heads' for the same experiment. This event happens with every outcome containing one or two Heads. Therefore we can create a set of elementary events which the compound event consists of, and which are also said to be 'favourable outcomes' for the compound event

$$\begin{array}{cccc} HTTT & THTT & TTHT & TTTH \\ HHTT & HTHT & HTTH & TTHH \\ THTH & THHT & TTTT & \end{array}$$

The compound event 'Fewer than three Heads' comprises 11 elementary events. The occurrence of any of these outcomes means that the event has occurred.

$\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

In a triathlon competition, there are 12 Russian, 8 German, 10 UK and 9 French athletes. The random experiment is the running of the competition to determine the winning athlete. Describe the sample space of this experiment. What is the cardinality of this sample space. Consider the event 'The competition is not won by a French athlete'. How many elementary events form this compound event?

The sample space for the experiment is the set of all athletes, as each one could be the winner of the competition, however unlikely this may be. Thus the cardinality of the sample space is $12 + 8 + 10 + 9 = 39$. The compound event 'The competition is not won by a French athlete' contains all the elementary events excluding the French athletes. Its cardinality is $12 + 8 + 10 = 30$.

$\ominus \ominus \ominus$

5.2.3 Probability

Probability is a measure of uncertainly about the outcomes of random experiments. By convention, we have the following interpretation of probability

1. Probability of an event E, denoted $P(E)$, is a number between 0 and 1, $0 \leq P(E) \leq 1$. It signifies the likelihood of E happening as an outcome of the random experiment.

2. Probability $P(E) = 0$ defines the *impossible* event. E will never occur, however many times we carry out the random experiment.

3. Probability $P(E) = 1$ defines the *sure* event. E will always occur, however many times we carry out the random experiment.

Think of the probability as a fixed amount equal to 1, which you have to spread across the elementary events in the sample space. If the sample space consists of a single elementary event, then all probability will fall onto it, and that will be the sure event. If the sample space consists of two elementary events $S = \{E_1, E_2\}$, they will share the probability "quantity" so that

$$P(E_1) + P(E_2) = 1.$$

Events E_1 and E_2 are called *complementary* events. If one does not happen in a given run of the experiment, the other one will happen. Thus E_2 is the opposite event to E_1, and E_1 is the opposite event to E_2.

Note that *only one* elementary event can happen at each run of the experiment. Such events are called *mutually exclusive.*

We can generalise the above argument for sample spaces with n elements.

Let $S = \{E_1, E_2, \dots, E_n\}$ be a sample space of cardinality n. Then the following holds

$$P(E_1) + P(E_2) + \cdots + P(E_n) = \sum_{i=1}^{n} P(E_i) = 1.$$

If all elementary events have the same likelihood to happen, they are called *equiprobable*, and so is called the sample space. In this case

$$P(E_1) = P(E_2) = \cdots = P(E_n) = \frac{1}{n}.$$

$\oplus \oplus \oplus$ **Example**

The random experiment is rolling a fair die. The sample space consists of the 6 faces $S = \{1, 2, 3, 4, 5, 6\}$. Since the die is fair, all elementary events have the same probability of occurring, which is $\frac{1}{6}$.

$\ominus \ominus \ominus$

The probability of a compound event is equal to the sum of the probabilities of the elementary events it consists of.

$\oplus \oplus \oplus$ **Example**

Consider a random experiment which determines the winning horse in a race. There are 5 horses: h_1, h_2, h_3 from stables A and h_4, h_5 from stables B. Their probabilities for winning the race are respectively

Horse	h_1	h_2	h_3	h_4	h_5
Probability of winning	0.04	0.34	0.11	0.28	0.23

The probabilities sum up to 1. Horse h_2 from stables A has the highest likelihood to win the race. However, consider the compound events E_A = 'Stables A wins the race' and E_B = 'Stables B wins the race'. The sets of elementary events for E_A and E_B are

$$E_A = \{h_1, h_2, h_3\} \qquad \text{and} \qquad E_B = \{h_4, h_5\}.$$

The respective probabilities are

$$P(E_A) = 0.04 + 0.34 + 0.11 = 0.49$$

$$P(E_B) = 0.28 + 0.23 = 0.51.$$

This result means that even though the favourite is from stables A, and, in spite of the fact that stables A has more horses in the race, stables B is more likely to be the home of the winner.

$\ominus \ominus \ominus$

For equiprobable sample spaces, the probability of a compound event E is easy to calculate

$$P(E) = \frac{\text{Cardinality of the set of favourable oucomes}}{\text{Cardinality of the sample space}}.$$

$\oplus \oplus \oplus$ **Example**

What is the probability to win the lottery with a single ticket?

Each possible selection of 6 out of 49 numbers is a candidate winner. The sample space of winning combination has cardinality $\binom{49}{6} = 13\,983\,816$. Of these, only one combination is a favourable outcome – the combination on your ticket. Therefore the probablity to win the lottery with one ticket is

$$\frac{1}{13\,983\,816} = 7.1511 \times 10^{-8}.$$

Hm, not particularly hopeful then...

$\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

If you draw at random 4 cards from a standard deck of 52 cards, what is the probability that the four cards will have the same value? (For example, $\{\clubsuit 8, \diamondsuit 8, \heartsuit 8, \spadesuit 8\}$)

The sample space consists of all possible selections of 4 cards out of 52. The cardinality of the space is therefore $\binom{52}{4} = 270725$. Favourable to our compound event 'Same value' are only 13 of these selections: four Aces, four 2s, etc, to four Queens and four Kings. Therefore the probability to select four cards of the same value is

$$\frac{13}{270725} = 4.8019 \times 10^{-5}.$$

$\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

On your shelf, there are 5 CDs with classical music, 4 CDs with Jazz and 11 CDs with pop music. If you grab, without looking, 6 CDs, what is the probability that you will have 2 CDs with classical music, 2 CDs with Jazz and 2 CDs with pop music?

The sample space for this experiment consists of all combinations of 6 CDs out of your collection of 20 CDs. The cardinality of the sample space is

$$\binom{20}{6} = \frac{20.19.18.17.16.15}{6.5.4.3.2.1} = 19.17.8.15 = 38760.$$

Out of these, favourable combinations are only those which contain two of each genre. Using the counting rule, the total number of such combinations is

$$\binom{5}{2}(\text{classical}) \times \binom{4}{2}(\text{jazz}) \times \binom{11}{2}(\text{pop}) = 10 \times 6 \times 55 = 3300.$$

Then the probability of having picked 2 CDs of each genre is

$$\frac{3300}{38760} = 0.0851 \ (4dp)$$

$\ominus \ominus \ominus$

$\oplus \oplus \oplus$ **Example**

The book will not be complete without mentioning the famous birthday problem. In a class of 45 students, what is the probability that at least two students share the same birthday (date and month)? Intuition may mislead you here into thinking that this probability is quite small.

Usually the most important part of solving a probability-related question is figuring out the random experiment and the sample space. We can approach this question from the side of the opposite event. What is the probability that all 45 students were born on different dates? Subtracting this probability from 1, will give us the answer.

As there are 45 students, and each student has one choice from 365 for their birthday, the total number of possible birthdays is 365^{45} (remember the counting rule). These birthdays are 'assigned' to the students. We are only interested in the 'bag' of birthdays, and not in how these dates are assigned to the students. Therefore we should divide the total by the number of permutations 45!. These combinations form the equiprobable sample space of possible combinations of 45 birthdays. How many of these are on different dates? To answer this question, we need to find out the number of combinations of 45 out of 365. This is given by the binomial coefficient $\binom{365}{45}$. Neither of the two quantities can be reliably calculated. However, using the short-cut formula for the binomial coefficient, we can arrive at a feasible way of calculating the probability of different birthdays

$$P(\text{different birthdays}) = \frac{\binom{365}{45}}{\frac{365^{45}}{45!}} = \frac{\overbrace{365.364.363...321}^{45}\ 45!}{\underbrace{365.365...365}_{45}\ 45!}$$

$$= \frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \cdots \times \frac{321}{365} = 0.0590.$$

Therefore, the probability of having matching birthdays in a class of 45 students is

$$1 - 0.0590 = 0.9410.$$

Surprisingly high, is it not? In fact, for a class of 23 students, the probability of matching birthdays is greater than 50%.

$\ominus \ominus \ominus$

5.3 Joint, marginal and conditional probabilities

5.3.1 The Monty Hall problem

Every year, at the start of the probability part of the module, I play a little game with the students. This is a famous probability puzzle, so counter-intuitive that it sparkled a lively debate in the media and even in the scientific literature. Our version of the puzzle is attributed to an American TV show 'Let's Make a Deal', and is named after its host, Monty Hall.[1] Ths problem is a perfect illustration of conditional probability, therefore we reproduce it here.

The rules of the game are illustrated in Figure 5.3. Suppose you are chosen as the player in the game show. You are given the choice of three doors. Behind one of the doors there is a car, and behind the other two doors, there are goats. You will be allowed to take home whatever you find behind your chosen door. Say, you choose door #3. Then the host, who knows where the car is, opens another door to reveal a goat. In our example in the figure, this is door #2. Then you are offered a chance to switch to the other unopened door. Will you switch? In our example in the figure, you will do well if you do switch, but is it always better to switch?

Intuition says that, after one door is open, there is 50% chance that the car is behind the door you already selected and it does

[1] `http://en.wikipedia.org/wiki/Monty_Hall_problem`

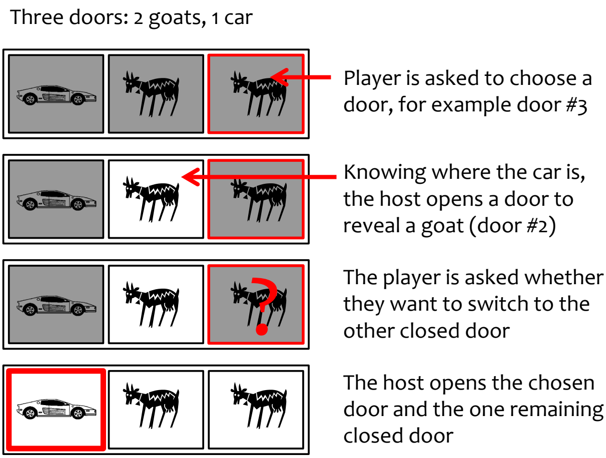


Figure 5.3: The Monty Hall game.

not matter whether you switch or not. Wrong! When you choose a door at random, the probability to hit the door with the car is $\frac{1}{3}$. Therefore the probability that the car is behind one of the other two doors is $\frac{2}{3}$. Note that the decision of the host of which door to open is not random! It depends on where the car is. The remaining closed door will 'absorb' all the remaining $\frac{2}{3}$ probability, and if you switch, your chances to go home in a car is twice the one of first choice. So, yes, switch!

Well, I was not giving out to students live stock or cars; our teaching budget wouldn't stretch to that. We would play the game a few times at the beginning of the lecture with a chocolate bar in the place of the car :).

5.3.2 Contingency tables

When we analyse data, is it convenient to organise the information in contingency tables. They represent the data distribution when we have two characteristics of interest.

$\oplus\oplus\oplus$ **Example**

A bookshop in a university town has 49 books on the Maths & Computer Science shelf. The books can be categorised by topic and level of difficulty. The table below gives number of books in each joint category.

	Level		
	Introductory	Medium	Specialist
Computer science	9	11	7
Maths	12	8	2

Suppose that the random experiment is to select one book from the shelf. The sample space for this experiment consists of 49 elementary events. The contingency table makes calculating probabilities of compound events straightforward. For example, what is the probability that a randomly picked book from this shelf will be a medium-level book on computer science? The total number of outcomes is 49; the number of favourable outcomes is 11 (the number in row 1, column 2 of the contingency table). Therefore the probability of selecting the desired book is $\frac{11}{49} \approx 0.2245$. $\ominus\ominus\ominus$

5.3.3 Joint, marginal and conditional probabilities

Joint probability is the probability of an event happening where all variables are taken into account. In the example above, we calculated the joint probability of the randomly picked book to have difficulty = medium and topic = computer science. The sample space is the whole set of possible outcomes. Its cardinality is the sum of all the entries in the contingency table. The number of favourable outcomes is found in the corresponding entry in the table.

Denote by X the first variable and by Y, the second variable. The lower-case x and y will denote specific values of the variables. The joint probability

$$P(X = x, Y = y)$$

is the probability that $X = x$ AND $Y = y$.

Marginal probability is the probability of one variable taking a specific value irrespective of the values of the other variables.

$\oplus \oplus \oplus$ **Example**

In the example above, consider the following question. What is the probability that a randomly picked book from this shelf is an introductory-level book? The answer is a marginal probability because the variable 'topic' is not taken into account. The sample space for this experiment is again the whole set of objects but the number of favourable outcomes is the sum of elements in the respective row or column of the contingency table. The favourable outcomes for the event 'introductory-level book' are in the first column of the table, in both rows. Therefore the probability for this event is $\frac{9+12}{49} \approx 0.4286$.

$\ominus \ominus \ominus$

The marginal probability is

$$P(X = x) \qquad \text{or} \qquad P(Y = y).$$

Conditional probability is the probability of an event happening given that another event has happened.

$\oplus \oplus \oplus$ **Example**

Find the probability that a randomly chosen book from this shelf is on computer science, given that its difficulty level is medium. The 'given' clause specifies the sample space this time. In our case, the sample space is all medium level books, with cardinality $11+8=19$. Favourable outcomes out of these 19 are the 11 computer science books, therefore the probability is $\frac{11}{19} \approx 0.5789$.

$\ominus\ominus\ominus$

The conditional probability is

$$P(X = x|Y = y)$$

and is the probability of $X = x$ GIVEN that $Y = y$. Notice the vertical line in the notation which separates the event of interest from the 'given' clause.

The three types of probability are illustrated in Figure 5.4.

$\oplus\oplus\oplus$ **Example**

There are 32 girls in a class of 126 students. The total number of left-handed students in the class is 21, of which 15 are boys. What is the probability that a randomly chosen student from this class is left-handed, given that this student is a boy?

Based on the information in the question, we can construct the following contingency table

	girls	boys
left-handed	6	15
right-handed	26	79

The probability we are interested in is conditional, therefore the sample space is defined by the condition. In this case the sample space is the set of boys in the class. The cardinality of this set is $15 + 79 = 94$. Favourable outcomes are the 15 left-handed boys, therefore

$$P(\text{left-handed}|\text{boy}) = \frac{15}{79} = 0.1899.$$

$\ominus\ominus\ominus$

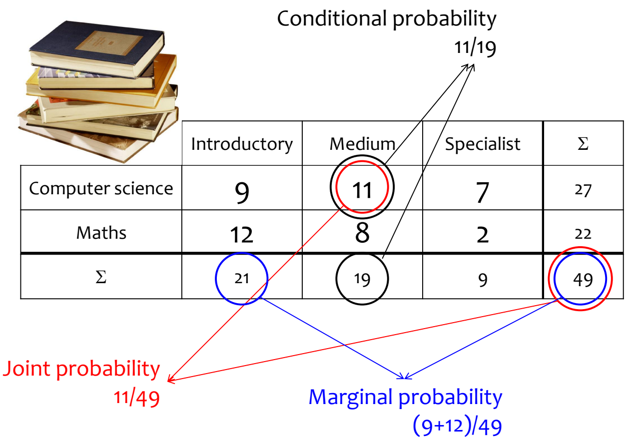


Figure 5.4: Types of probabilities calculated from a contingency table.

5.4 Exercises

1. How many possible outcomes are there from an experiments where a fair coin is tossed 12 times?

2. How many possible outcomes are there from an experiments where a fair die is rolled 8 times?

3. A restaurant has 5 types of starters, 6 types of main courses, and 4 types of deserts. How many possible different menus can this restaurant serve?

4. Your space ship is travelling from planet K to planet Q. However, you have to stop midway on planet M. There are 5 ways

to reach M from K, and 8 ways to reach Q from M. How many possible different paths are there from K to Q?

5. How many 4×3 matrices with binary entries (0 or 1) are there?

6. In how many ways can you arrange 11 people in a queue?

7. A veterinary surgery has 7 kittens to find homes for. Seven children from the neighbourhood have requested a kitten each. In how many ways can the kittens be paired with their new young human-servants?

8. How many possible 4-letter candidate-words can you form from the Latin alphabet (26 letters)?

   (a) with non-repeated letters

   (b) with repeated letters

9. A 6-symbol password must contain non-repeated characters which could be letters and numbers. A legitimate password must begin and end with a letter. How many different passwords are there?

10. In how many ways can you seat 10 people around a round table. Note that arrangement ABCDEFHGIJ is the same as JABCDEFHGI,
IJABCDEFHG, and so on.

11. How many different permutations of the letters of 'Republic of Macedonia' are there?

12. A biased die has a probability of a face appearing proportional to the number (1,2,3,4,5 or 6) occurring on that face.

    (a) Find these probabilities.

    (b) Calculate the probability that an even number or a number greater than 4 is thrown.

(c) Calculate the probability that an odd number which is not 5 is thrown.

13. The Hawaiian language has only 12 letters: the vowels a, e, i, o, and u, and the consonants h, k, l, m, n, p, and w.

    (a) How many "words" with 4 letters are there which begin with a consonant, end with a different consonant, and have two distinct vowels in between?

    (b) How many "words" with 4 letters are there which begin with a consonant, end with a different consonant, and have two not necessarily distinct vowels in between?

14. A class of 18 students must be split into teams of 3 students for their Software Hut projects. In how many ways can you split the class?

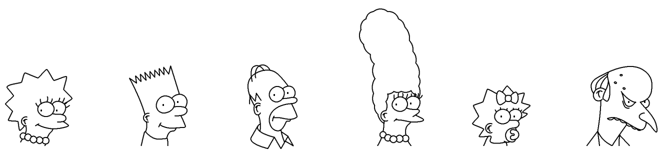
15. Calculate the binomial coefficient $\binom{10}{6}$.

16. Use the short-cut formula to calculate the binomial coefficient $\binom{150}{5}$.

17. In how many ways can we select four DVD movies from a collection of 84 DVD movies?

18. A poker hand consists of 5 cards randomly dealt from 52 cards. How many possible poker hands are there?

19. You are assigned the task to create a dialogue scene for an episode of 'The Simpsons' involving the following characters



Lisa Bart Homer Marge Maggie Burns

(a) In how many ways can you select a pair of characters for the scene?

(b) How many possible pairs of characters are there if one of them is either Lisa or Marge?

(c) Suppose that the scene includes a third characters who walks in during the dialogue. Assuming that this third character is neither Homer nor Burns, how many possible groups of three characters are there?

20. Write down the sample space for the experiment 'Rolling a fair die', where the outcomes are the faces. What will be the cardinality of the sample space for 'Rolling two fair dice', where the outcomes are (Face1,Face2)?

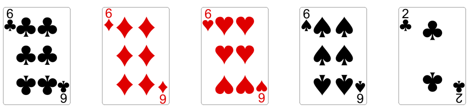
21. A play in a kinder-garden requires two children, a boy and a girl, to greet the parents. The group of children in consists of 7 boys and 11 girls. Describe the sample space for this experiment and calculate its cardinality.

22. Consider again problem 21, but this time suppose that two boys and two girls are needed to greet the parents. Describe the sample space for this experiment and calculate its cardinality.

23. What is the probability to roll two sixes if you roll two fair dice?

24. Knowing that a poker hand is a random subset of 5 elements from a deck of 52 cards, what is the probability of a hand containing

    (a) 4 four of a kind, as in the example below?



(b) a flush (5 cards of the same suit but not in a sequence) as in the example below?

25. Remember the "infinite monkey theorem"? That a monkey sitting at a typewriter and randomly hitting the keys, has a non-zero chance of typing the exact content of Encyclopaedia Britannica. Eh, well, the time to do that would exceed many times over the existence of the universe, but the chance is still non-zero! Let's modernise and simplify the game a little bit. Suppose that the monkey has grabbed your iPhone and is just about to type four numbers in. What is the probability that the monkey will punch in your 4-digit password and unlock your iPhone?

26. A random sample of 80 lawyers was taken, and they were asked if they were in favour of or against capital punishment. The following table gives the two-way classification of these lawyers.

	Favours capital punishment	Opposes capital punishment
Male	32	24
Female	13	11

If one lawyer is randomly selected from this group, find the probability that this lawyer

(a) favours capital punishment

(b) is a female

(c) opposes capital punishment given that the lawyer is a female

(d) a male given that he favours capital punishment

(e) a female and favours capital punishment

27. In a class of 49 students, 11 are female. Seven of the male students and six of the female students speak French. If a student is picked randomly from the class, what is the probability that this student

    (a) does not speak French?

    (b) is a female who does not speak French?

    (c) is a male?

    (d) speaks French, if it is known that the student is male?

28. In rolling a pair of fair dice

    (a) What is the probability that the sum of the two faces is less than 10?

    (b) What is the probability that the sum of the two faces is less than 10 given that exactly one of the faces is an odd number?

    (c) What is the probability that both faces are odd numbers?

    (d) What is the probability that both faces are odd numbers given that the sum is greater than 6?

    (e) What is the probability that the product of the two numbers divides by 3 given that both faces are even?

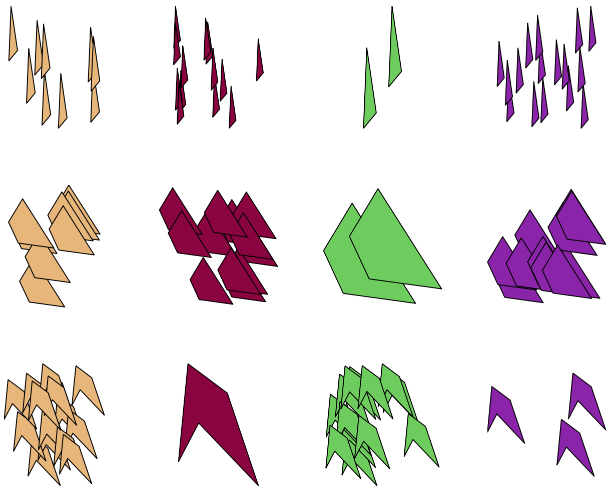
29. (M) Write a MATLAB function which will display an illustration of a contingency table. The inputs to the function should be an $m \times n$ matrix containing the contingency table. The function should open up a figure and place the respective number of shapes in each table cell. The rows will differ by the number of vertices (starting with 3), and the columns will

differ by colour. The shapes and the colours should be random. (Hint: use the `subplot` command.) Two examples are shown below.

Example 1. Contingency table:

$$\begin{bmatrix} 9 & 13 & 2 & 16 \\ 8 & 12 & 2 & 11 \\ 15 & 1 & 13 & 3 \end{bmatrix}$$

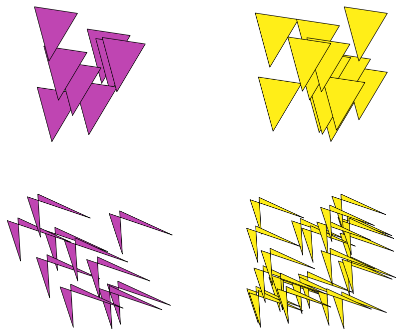
MATLAB output:



Example 2. Contingency table:

$$\begin{bmatrix} 8 & 14 \\ 10 & 26 \end{bmatrix}$$

MATLAB output:



Chapter 6

Further reading

You can consult any textbooks on linear algebra and probability.

1. Robert A Beezer. *A First Course in Linear Algebra*, 2004, `http://linear.ups.edu`, 3rd edition, 2012, ISBN-10: 0984417559 (ISBN-13: 978-0984417551)

2. Saeed Ghahramani. *Fundamentals of Probability*, Prentice Hall, 2nd edition, 1999, ISBN-10: 0130113298 (ISBN-13: 9780130113290)

3. David C. Lay. *Linear Algebra and Its Applications*, Addison Wesley, 3rd edition 2002, ISBN-10: 0201709708 (ISBN-13: 978-0201709704)

4. Prem S. Mann. *Introductory Statistics*, 4th edition, 2001, John Wiley & Sons, N.Y., ISBN-10: 0471440124 (ISBN-13: 978-0471440123)

Index