6 Systems of Linear Equations

A system of equations of the form

\[ \begin{align*}
  x + 2y &= 7 \\
  2x - y &= 4
\end{align*} \]

or

\[ \begin{align*}
  5p - 6q + r &= 4 \\
  2p + 3q - 5r &= 7 \\
  6p - q + 4r &= -2
\end{align*} \]

is called a linear system of equations.

**Definition.** An equation involving certain variables is **linear** if each side of the equation is a sum of **constants** and **constant multiples of the variables**.

**Examples.** The equations \(2x = 3y + 4\) and \(3q = 0\) are linear.

**Non-examples.** The equations \(y = x^2\) and \(3xy + z + 7 = 2\) are not linear since they involve **products** of variables.

### 6.1 Linear Equations and Matrices

Any system of linear equation can be rearranged to put all the constant terms on the right, and all the terms involving variables on the left. This will yield a system something like . . .

\[ \begin{align*}
  4.238x - 1.297y &= 3.114 \\
  2.088x + 2.971y &= 0.277
\end{align*} \]

. . . so from now on we will assume all systems have this form. The system can then be written in matrix form:

\[ \begin{bmatrix}
  4.238 & -1.297 \\
  2.088 & 2.971
\end{bmatrix} \begin{bmatrix}
  x \\
  y
\end{bmatrix} = \begin{bmatrix}
  3.114 \\
  0.277
\end{bmatrix}. \]

You can check (by multiplying out the matrix equation using the definition of matrix multiplication) that it holds if and only if both of the original (scalar) equations hold. We can express the equations even more concisely as an augmented matrix:

\[ \begin{bmatrix}
  4.238 & -1.297 & 3.114 \\
  2.088 & 2.971 & 0.277
\end{bmatrix}. \]

### 6.2 Solutions to Systems of Equations

A **solution** to a system of equations is a way of assigning a value to each variable, so as to make all the equations true.

**Example.** The system of equations . . .

\[ \begin{align*}
  4x - 2y &= 22 \\
  9x + 3y &= 12
\end{align*} \]

. . . has a solution \(x = 3\) and \(y = -5\). (In fact in this case this is the only solution.)

In general, a system of linear equations could have no solutions, a unique solution, or infinitely many solutions. The set of all solutions is called the **solution set** of the system. Suppose we have \(k\) linear equations in \(n\) variables. What can the solution sets look like?

**Same Number of Equations and Variables.** If \(k = n\) then the “typical” behaviour is that there will be a unique solution. But there are also “special cases” with either no solution, or an infinite family of solutions.

**Fewer Equations than Variables.** If \(k < n\) then there will “typically” be an infinite family of solutions, but there could be no solution. (There can never be a unique solution in this case!)
More Equations than Variables. If \( k > n \) then there will “typically” be no solutions, but there could be a unique solution, or an infinite family of solutions.

Warning. When we talk about solving a system of equations, we mean describing all the solutions (if any), not just finding one possible solution!

6.3 Solving Systems of Equations

At school you probably learnt to solve systems of linear equations by simple algebraic manipulation. This \textit{ad hoc} approach is fine for small examples, but it is helpful to have a \textbf{systematic method} which allows us to solve even large systems of equations.

We start with the augmented matrix of our system (see Section 6.1), and we allow ourselves to modify the matrix by certain kinds of steps. Specifically, we allow ourselves to:

- \textbf{add} a multiple of one row to another (written \( r_i \rightarrow r_i + \lambda r_j \));
- \textbf{swap} two rows of the matrix (written \( r_i \leftrightarrow r_j \)).

These are called \textbf{elementary row operations}. It can be shown (exercise*) that applying them to an augmented matrix \textbf{does not change the set of solutions} of the corresponding equations.

Our strategy is to use these operations to transform our matrix, so as to obtain a system of equations which is easier to solve. For example, consider the system:

\[
\begin{align*}
10y - 3z &= -19 \\
-x + 4y - z &= -9 \\
10y - 2z &= -16
\end{align*}
\]

First we convert this into augmented matrix form (noting that the “missing” \( x \) terms are really \( 0x \) and become 0s in the matrix):

\[
\begin{pmatrix}
0 & 10 & -3 & | & -19 \\
-1 & 4 & -1 & | & -9 \\
0 & 10 & -2 & | & -16
\end{pmatrix}
\]

Now we swap rows 1 and 2 \((r_1 \leftrightarrow r_2)\), to get

\[
\begin{pmatrix}
-1 & 4 & -1 & | & -9 \\
0 & 10 & -3 & | & -19 \\
0 & 10 & -2 & | & -16
\end{pmatrix}
\]

Next we add \(-1\) times row 2 to row 3 \((r_3 \rightarrow r_3 - r_2)\), giving:

\[
\begin{pmatrix}
-1 & 4 & -1 & | & -9 \\
0 & 10 & -3 & | & -19 \\
0 & 0 & 1 & | & 3
\end{pmatrix}
\]

Now we convert the augmented matrix back to a system of equations:

\[
\begin{align*}
-x + 4y - z &= -9 \quad [1] \\
10y - 3z &= -19 \quad [2] \\
z &= 3 \quad [3]
\end{align*}
\]

It turns out that these equations are easier to solve than the original ones. Equation [3] tells us straight away that \( z = 3 \). Now substituting this into equation [2] we get \( 10y - 9 = -19 \), in other words, \( y = -1 \). Finally, substituting both of these into equation [2] gives \( -x - 4 - 3 = -9 \), that is, \( x = 2 \).

So we have found the solutions to equations [1], [2] and [3]. But these were obtained from our original system of equations by elementary row operations, so they are also the solutions to the original system of equations. (\textbf{Check this for yourself}!)  

Remark Notice how we converted the matrix into a roughly “triangular” form, with all the entries towards the “bottom left” being 0. It was this property which made the new equations easy to solve.
6.4 Gaussian Elimination

The approach we applied in the previous example forms the basis for a very general algorithm, called Gaussian Elimination. To describe this we shall need to define a few more terms.

Pivots. The **pivot** of a row in an augmented matrix is the position of the leftmost non-0 entry to the left of the bar. (If all entries left of the bar are 0, we call the row a *zero row* and it has no pivot.)

**Example.** In the matrix $A$ on the right

- the pivot of the first row is the 5;
- the pivot of the second row is the 3;
- the third row is a zero row (it has no pivot)

$A = \begin{pmatrix} 0 & 5 & 1 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

Row Echelon Form. An augmented matrix is in **row echelon form** if

(i) any zero rows come at the bottom, and

(ii) the pivot of each of the other rows is **strictly to the right of the pivot of the row above**.

**Example.** The matrix $A$ above is not in row echelon form, because the pivot of the second row is not to the right of the pivot in the first row. The matrix $B$ on the right is in row echelon form.

$B = \begin{pmatrix} 0 & 5 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}$

Gaussian Elimination is a simple recursive algorithm which uses elementary row operations to reduce the augmented matrix to row echelon form:

- By swapping rows if necessary, make sure that no non-0 entries (in any row) are to the left of the first row pivot.
- Add multiples of the first row to each of the rows below, so as to make all entries below the first row pivot become 0. (Since there were no non-0 entries left of the first row pivot, this this ensures all pivots below the first row are strictly to the right of the pivot in the first row.)
- Now ignore the first row, and repeat the entire process with the second row. (This will ensure that all pivots below the second row are to the right of the pivot in the second row, which in turn is to the right of the pivot in the first row.)
- Keep going like this (with the third row, and so on) until either we have done all the rows, or all the remaining rows are zero.
- At this point the matrix is in row echelon form.

**Exercise.** Go back to the example in Section 6.3, and compare what we did with the procedure above.

6.5 Another Example

**Example.** Solve the system of equations:

\[
\begin{align*}
2x & + y - z = -7 \\
6x & \quad - z = -10 \\
-4x & + y + 7z = 31
\end{align*}
\]

**Solution.** First we express the system as an augmented matrix:

\[
\begin{pmatrix}
2 & 1 & -1 & | & -7 \\
6 & 0 & -1 & | & -10 \\
-4 & 1 & 7 & | & 31
\end{pmatrix}
\]
Notice that there are no non-0 entries to the left of the pivot in the first row (there can’t be, since the pivot is in the first column!). So we do not need to swap rows to ensure this is the case.

Next we seek to remove the entry in row 2 which is below the row 1 pivot. We can do this with the operation \( r_2 \rightarrow r_2 - 3r_1 \), giving the matrix:

\[
\begin{pmatrix}
2 & 1 & -1 & -7 \\
0 & -3 & 2 & 11 \\
-4 & 1 & 7 & 31
\end{pmatrix}.
\]

Now we need to remove the entry in row 3 below the row 1 pivot. We can do this with \( r_3 \rightarrow r_3 + 2r_1 \):

\[
\begin{pmatrix}
2 & 1 & -1 & -7 \\
0 & -3 & 2 & 11 \\
0 & 3 & 5 & 28
\end{pmatrix}.
\]

Now we ignore the first row, and repeat the process with the remaining rows. Notice that there are no pivots to the left of the row 2 pivot. (The one in row 1 doesn’t count, since we are ignoring row 1!). Now we remove the entry in row 3 which is below the row 2 pivot, with \( r_3 \rightarrow r_3 + r_2 \), giving:

\[
\begin{pmatrix}
2 & 1 & -1 & -7 \\
0 & -3 & 2 & 11 \\
0 & 0 & 7 & 28
\end{pmatrix}.
\]

Our matrix is now in row echelon form, so we convert it back to a system of equations:

\[
\begin{align*}
2x + y - z &= -7 \quad [1] \\
-3y + 2z &= 11 \quad [2] \\
7z &= 28 \quad [3]
\end{align*}
\]

These matrices can easily be solved by “backtracking” through them. Specifically:

- equation [3] gives \( 7z = 28 \), so \( z = 4 \);
- substituting into equation [2] gives \( -3y + 8 = 11 \), so \( y = -1 \);
- substituting into equation [1] gives \( 2x + (-1) - 4 = -7 \), so \( x = -1 \).

Thus, the solution is \( x = -1, y = -1 \) and \( z = 4 \). (You should check this by substituting the values back into the original equations!)

**Remark.** Notice how the row echelon form of the matrix facilitated the “backtracking” procedure for solving equations [1], [2] and [3].

### 6.6 Useful Tricks for Gaussian Elimination by Hand

The Gaussian elimination algorithm described above will always work. But if you are doing Gaussian elimination by hand, you may spot ways to speed up or otherwise simplify the process, by throwing in some extra elementary row operations.

**Extra Swaps.** Suppose you find yourself with the augmented matrix:

\[
\begin{pmatrix}
12 & 12 & -1 & -7 \\
4 & -3 & 2 & 11 \\
8 & 0 & 7 & 28
\end{pmatrix}.
\]

Gaussian elimination tells you to add \(-\frac{1}{3}\) row 1 to row 2, and \(\frac{2}{3}\) row 1 to row 3. This is fine in principle, but you will end up working with awkward fractions all over the place. A solution is to swap row 1 and row 2 first \((r_1 \leftrightarrow r_2)\), and then everything works much more nicely! This is perfectly okay, because \(r_1 \leftrightarrow r_2\) is an elementary row operation and doesn’t change the solution set.
Scaling Rows. Another useful operation is **multiplying a row by a non-zero scalar**. For example, if we have the augmented matrix

\[
\begin{pmatrix}
\frac{1}{2} & \frac{5}{2} & -1 \\
4 & -3 & 2 \\
\end{pmatrix}
\begin{pmatrix}
-\frac{7}{2} \\
12 \\
\end{pmatrix}
\]

then we can make life easier for ourselves by multiplying the top row by 6 (written \( r_1 \rightarrow 6r_1 \)):

\[
\begin{pmatrix}
3 & 15 & -6 \\
4 & -3 & 2 \\
\end{pmatrix}
\begin{pmatrix}
-14 \\
12 \\
\end{pmatrix}
\]

Of course multiplying a row by a non-zero scalar is **not** technically an elementary row operations (see Section 6.3) but, like them, it does not change the solution set, and so is fine. (On the other hand, multiplying a row by **zero** is definitely **not** permitted, as it can change the solution set!)

6.7 Elimination With No Solutions

What we have seen looks like a foolproof strategy for solving systems of linear equations. But on the other hand, we know (from Section 6.2) that some systems have **no** solutions. Clearly our strategy cannot find solutions in this case, so what goes wrong? Consider:

\[
\begin{align*}
-x + 2y + z &= 2 \\
3x + y - 2z &= 10 \\
x + 5y &= 7
\end{align*}
\]

The augmented matrix is

\[
\begin{pmatrix}
-1 & 2 & 1 & 2 \\
3 & 1 & -2 & 10 \\
1 & 5 & 0 & 7
\end{pmatrix}
\]

Removing entries below the first row pivot (\( r_2 \rightarrow r_2 + 3r_1, r_3 \rightarrow r_3 + r_1 \)) gives:

\[
\begin{pmatrix}
-1 & 2 & 1 & 2 \\
0 & 7 & 1 & 16 \\
0 & 7 & 1 & 9
\end{pmatrix}
\]

Now we remove the entry below the second row pivot (\( r_3 \rightarrow r_3 - r_2 \)):

\[
\begin{pmatrix}
-1 & 2 & 1 & 2 \\
0 & 7 & 1 & 16 \\
0 & 0 & 0 & -7
\end{pmatrix}
\]

Converting back to equations, we get \(-x + 2y + z = 2, 7y + z = 16\) and (most interestingly) \(0 = -7\). Since no values of the variables will make 0 equal to \(-7\), this system has no solutions. And since the row operations have preserved the solution set, this tells us that the **original system has no solutions**.

6.8 Elimination With Redundant Equations

Now consider the (very similar!) system

\[
\begin{align*}
-x + 2y + z &= 2 \\
3x + y - 2z &= 10 \\
x + 5y &= 14
\end{align*}
\]

The only thing changed from the previous section is the RHS of the third equation. This time the augmented matrix is:

\[
\begin{pmatrix}
-1 & 2 & 1 & 2 \\
3 & 1 & -2 & 10 \\
1 & 5 & 1 & 14
\end{pmatrix}
\]

and reducing to row echelon form we get

\[
\begin{pmatrix}
-1 & 2 & 1 & 2 \\
0 & 7 & 1 & 16 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

This time the equations are \(-x + 2y + z = 2, 7y + z = 16\) and \(0 = 0\). Instead of being never satisfied, the last equation is **always** satisfied! This means it is irrelevant. So we can discard it completely from the system, and just go on to solve the other equations.

(We now have fewer equations than variables: we will see how to handle this in Section 6.10 below.)
6.9 Elimination with More Equations Than Variables

So far we have looked at examples where there are the same number of equations and variables. If there are more equations than variables, we can still apply the elimination process, but we will always end up with a zero row in the matrix, either as in Section 6.7 or as in Section 6.8. So either we will be able to throw away equations until we have only as many equations as variables, or there will turn out to be no solutions.

6.10 Elimination with More Variables than Equations

What if there are more variables than equations? This can happen either because the system we started with was like this, or because we threw away some redundant “0 = 0” equations (see Section 6.8 above) after elimination.

Suppose there are \( k \) equations and \( n \) variables, where \( n > k \). We can still apply Gaussian elimination to find a row echelon form. But when we convert back to equations and “backtrack” to find solutions, we will sometimes still encounter an equation with more than one unknown variable.

If this happens, there are infinitely many solutions. We can describe all the solutions by introducing parameters to replace \( n - k \) of the variables.

**Example.** Let us continue to solve the system of equations from Section 6.8 above. Recall that we started with three equations, used Gaussian elimination to reduce the augmented matrix, and extracted the equations

\[
\begin{align*}
-x + 2y + z &= 2 \quad [1] \\
7y + z &= 16 \quad [2]
\end{align*}
\]

(plus the “always satisfied” equation \( 0 = 0 \), which we threw away).

The last equation has two unknown variables \( (y \text{ and } z) \) so we introduce a parameter \( \lambda \) to stand for one of them, say \( y \). Now solving we get \( z = 16 - 7\lambda \). Substituting into \([1]\) gives \(-x + 2\lambda + (16 - 7\lambda) = 2\), or \( x = 14 - 5\lambda \). So the solutions are:

\[
\begin{align*}
x &= 14 - 5\lambda, & y &= \lambda, & z &= 16 - 7\lambda.
\end{align*}
\]

**Remark.** When we say these are the solutions, we mean that substituting in different values of \( \lambda \) will give all the solutions to the equations. For example, \( \lambda = 1 \) would give \( x = 9, y = 1 \) and \( z = 9 \), so this is one solution to the system of equations (check this!). Or then again, \( \lambda = 0 \) gives \( x = 14, y = 0 \) and \( z = 16 \), so this is another possible solution.

**Example.** Find the solutions of the system

\[
\begin{align*}
2x + 4y + 16z &= 16 \\
x + 2y + 5z &= 5
\end{align*}
\]

**Solution.** The augmented matrix is

\[
\begin{pmatrix}
2 & 4 & 16 \\
1 & 2 & 5
\end{pmatrix}
\]

To simplify the calculations, we use a trick from Section 6.6 and swap the two rows \((r_1 \leftrightarrow r_2)\) to get

\[
\begin{pmatrix}
1 & 2 & 5 \\
2 & 4 & 16
\end{pmatrix}
\]

Now \( r_2 \rightarrow r_2 - 2r_1 \) gives

\[
\begin{pmatrix}
1 & 2 & 5 \\
0 & 0 & 6
\end{pmatrix}
\]

This is an row echelon form so we convert back to equations:

\[
\begin{align*}
x + 2y + 5z &= 5 \quad [1] \\
6z &= 6 \quad [2]
\end{align*}
\]

Now backtracking, the equation \([2]\) gives \( z = 1 \). Substituting into equation \([1]\) gives \( x + 2y + 5 = 5 \). This has two unknown variables, so we introduce a parameter \( \lambda \) to stand for one of them, say \( \lambda = y \), giving \( x + 2\lambda + 5 = 5 \), or \( x = -2\lambda \). So the solutions here are:

\[
x = -2\lambda, \ y = \lambda, \ z = 1.
\]
6.11 Further Exercises

- Use Gaussian elimination to solve the following systems of equations:

\[(i) \quad 4x + y = 9 \quad 2x - 3y = 1 \quad (ii) \quad 2x - 4y = 10 \quad -x + 2y = -5 \quad (iii) \quad 2x - 4y = 12 \quad -x + 2y = -5\]

6.12 Homogeneous Systems of Equations

A system of linear equations is called **homogeneous** if the constant terms are all 0. For example:

\[
\begin{align*}
x + 7y - 4z &= 0 \\
2x + 4y - z &= 0 \\
3x + y + 2z &= 0
\end{align*}
\]

An obvious observation about homogeneous systems is that they **always have at least one solution**, given by setting all the variables equal to 0 (in the above example, \(x = y = z = 0\)). Of course, there could also be other solutions.

When it comes to solving them, homogeneous systems of equations are treated exactly like non-homogenous ones. The only difference is that nothing interesting ever happens on the RHS of the augmented matrix!

**Example.** Find solutions to the system of equations above.

**Solution.** The augmented matrix is:

\[
\begin{pmatrix}
1 & 7 & -4 & 0 \\
2 & 4 & -1 & 0 \\
3 & 1 & 2 & 0
\end{pmatrix}
\]

Using row 1 to remove the entries below its pivot gives:

\[
\begin{pmatrix}
1 & 7 & -4 & 0 \\
0 & -10 & 7 & 0 \\
0 & -20 & 14 & 0
\end{pmatrix}
\]

Now using row 2 to remove entries below its pivot gives:

\[
\begin{pmatrix}
1 & 7 & -4 & 0 \\
0 & 1 & -\frac{7}{10} & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

This is in row echelon form. Converting back to equations we have:

\[
\begin{align*}
x + 7y - 4z &= 0 & [1] \\
-10y + 7z &= 0 & [2]
\end{align*}
\]

plus a “0 = 0” type equation which can be discarded.

We now have more variables than equations (see Section 6.10 above). Introducing a parameter \(\lambda\) in place of \(z\), equation [2] gives \(y = \frac{7}{10}\lambda\), and then equation [1] gives \(x = -\frac{9}{10}\lambda\). So, the solutions here are: \(x = -\frac{9}{10}\lambda, y = \frac{7}{10}\lambda\) and \(z = \lambda\).

**Remark.** Notice how everything stayed “homogeneous” throughout. All the augmented matrices in the elimination process, and all the resulting equations, had only 0’s on the RHS.

6.13 2-dimensional Geometry of Linear Equations

Consider a linear equation in two variables \(x\) and \(y\), say:

\[3x + 4y = 4\]

Each solution consists of an \(x\)-value and a \(y\)-value. We can think of these as the coordinates of a point in 2-dimensional space. For example, a solution to the above equation is \(x = -4\) and \(y = 4\). This solution gives the point \((-4, 4)\).

In general, the set of solutions to a linear equation with variables \(x\) and \(y\) forms a line in 2-dimensional space. So each equation corresponds to a line — we say that the equation defines the line.

**Systems of Equations.** Now suppose we have a system of such equations. A solution to the system means an \(x\)-value and a \(y\)-value which solve **all the equations at once**. In other words, a solution corresponds
to a point which lies on all the lines. So in geometric terms, the solution set to a system of linear equations with variables \(x\) and \(y\) is an **intersection of lines**.

**Question.** What can an intersection of two lines in 2-space look like?

- Usually it is a **single point.** This is why 2 equations in 2 variables usually has a unique solution!
- Alternatively, the lines could be different but parallel. Then the intersection is the **empty set.** This is why 2 equations in 2 variables can sometimes have no solutions!
- Or then again, the two lines could actually be the same. In this case the intersection is the **whole line.** This is why 2 equations in 2 variables can sometimes have an infinite family of solutions.

### 6.14 Higher Dimensional Geometry of Linear Equations

Now suppose we have an equation in the variables \(x, y\) and \(z\), say

\[
3x + 4y + 7z = 4.
\]

This time we can consider a solution as a point in 3-dimensional space. The set of all solutions (assuming there are some) forms a **plane** in space.

**Exercise.** What can the intersection of 2 planes in 3-space look like? Try to imagine all the possibilities, as we did for lines in the previous section. How do they correspond to the possible forms of solution sets of 2 equations with 3 variables?

**Exercise.** Now try to do the same for intersections of 3 planes. (Recall that the intersection of three sets is the set of points which lie in all three). How do the possible geometric things you get correspond to the possible forms of solution sets for 3 equations with 3 variables?

**Fact.** In yet higher dimensions, the solution set to a linear equation in \(n\) variables defines an \((n - 1)\)-dimensional subspace inside \(n\)-dimensional space. Such a subspace is called a **hyperplane.** So geometrically, the solution set of \(k\) equations with \(n\) variables will be an **intersection of \(k\) hyperplanes** in \(n\)-dimensional space.

### 7 Lines, Planes and Hyperplanes

One of the most important geometric ideas of that of a **line.** In this section we will see how to use vectors to describe and work with lines.

#### 7.1 The Vector Equation of a Line

A line (in any number of dimensions) can be described by two pieces of information:

- a point on the line (which can be described by its position vector); and
- the direction\(^2\) of the line (which can be described by a vector in the appropriate direction).

Suppose \(\mathbf{a}\) is the position vector of a point on the line and \(\mathbf{b}\) is a vector in the direction of the line. It follows from the definition of vector addition that every point on the line is the sum of \(\mathbf{a}\) and a scaling of \(\mathbf{b}\). In the other words, the line is the set of points \(\mathbf{r}\) satisfying

\[
\mathbf{r} = \mathbf{a} + \lambda \mathbf{b}
\]

for some real number \(\lambda\). This equation is called the**\(^3\) vector equation of the line**.

Each value of the parameter \(\lambda\) gives a different point on the line. The following picture gives some examples:

---

\(^2\)Of course, being strictly accurate, there are two (opposite) directions of the line. It doesn’t matter here which we use.

\(^3\)The use of the word “the” here is standard but a bit misleading. Different choices of \(\mathbf{a}\) and \(\mathbf{b}\) would have led to different vector equations describing the same line.
Example. Find the vector equation of the \(x\)-axis in 3-space.

Solution. The \(x\)-axis is the line through the origin in the direction of the unit vector \(\mathbf{i}\). The position vector of the origin is \(\mathbf{0}\). So the vector equation of the \(x\)-axis is:

\[
\mathbf{r} = \mathbf{0} + \lambda \mathbf{i}
\]
or alternatively just \(\mathbf{r} = \lambda \mathbf{i}\).

### 7.2 The Line Through Two Points

Suppose \(P\) and \(Q\) are points in space (of any dimension). Provided \(P \neq Q\), there will always be exactly one line going through both \(P\) and \(Q\). In fact, it is easy to find the vector equation of this line.

Let \(\mathbf{p}\) be the position vector of \(P\), and \(\mathbf{q}\) be the position vector of \(Q\). Then \(\mathbf{p}\) is the vector of a position on the line. And the displacement vector \(\overrightarrow{PQ} = \mathbf{q} - \mathbf{p}\) is a vector in the direction of the line. So the vector equation of the line is

\[
\mathbf{r} = \mathbf{p} + \lambda (\mathbf{q} - \mathbf{p})
\]

Example. Find the vector equation of the line through the points \(P = (2, 5)\) and \(Q = (4, 1)\) in 2-space.

Solution. The position vector of point \(P\) is \(2\mathbf{i} + 5\mathbf{j}\). The displacement vector from \(P\) to \(Q\) is \(2\mathbf{i} - 4\mathbf{j}\). So, the line has equation

\[
\mathbf{r} = 2\mathbf{i} + 5\mathbf{j} + \lambda (2\mathbf{i} - 4\mathbf{j}) = \left( \begin{array}{c} 2 \\ 5 \\ \lambda (2) \\ -4 \end{array} \right).
\]

Example. Find the vector equation of the line through the points \(R = (1, 3, 8)\) and \(S = (2, -4, 7)\) in 3-space.

Solution. The position vector of point \(R\) is \(\mathbf{i} + 3\mathbf{j} + 8\mathbf{k}\). The displacement vector from \(R\) to \(S\) is \(\mathbf{i} - 7\mathbf{j} - \mathbf{k}\). So, the line has equation

\[
\mathbf{r} = \mathbf{i} + 3\mathbf{j} + 8\mathbf{k} + \lambda (\mathbf{i} - 7\mathbf{j} - \mathbf{k}) = \left( \begin{array}{c} 1 \\ 3 \\ 8 \\ \lambda (1) \\ -7 \\ -1 \end{array} \right).
\]

### 7.3 Checking if a Point Lies on a Line

Consider the vector equation of a line:

\[
\mathbf{r} = \mathbf{a} + \lambda \mathbf{b}.
\]

A point \(P\) lies on the line if there is a choice of the parameter \(\lambda\) which makes \(\mathbf{a} + \lambda \mathbf{b}\) equal to the position vector of \(P\).

Example. Which of the points \((4, 7, 7), (0, -1, 2), (6, 3, -5)\) lie on the line

\[
\mathbf{r} = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k} + \lambda (\mathbf{i} + 2\mathbf{j} + 4\mathbf{k})?
\]
**Solution.** A point \((x, y, z)\) lies on this line if there is a \(\lambda\) such that

\[
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.
\]

Consider the point \((4, 7, 7)\). Comparing the \(x\)-coordinates gives \(4 = 2 + \lambda\). So the only choice of \(\lambda\) with any chance of working is \(\lambda = 2\). Checking whether this does indeed work, we see that

\[
\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 7 \end{pmatrix}
\]

so the point \((4, 7, 7)\) does indeed lie on the line.

Now consider the point \((0, -1, 2)\). This time comparing the \(x\)-coordinates gives \(0 = 2 + \lambda\), so the only possible \(\lambda\) which might give our point in the equation of the line is \(\lambda = -2\). Checking the other coordinates we have

\[
\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -9 \end{pmatrix}
\]

Since this is not the point we were looking for, we conclude that our point does not lie on the line.

Finally consider \((6, 3, -5)\). Comparing the \(x\)-coordinates gives \(6 = 2 + \lambda\), so the candidate is \(\lambda = 4\). Setting \(\lambda = 4\), we have

\[
\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 11 \\ 15 \end{pmatrix}
\]

which is not the point we wanted, so our point does not lie on the line.

Thus \((4, 7, 7)\) lies on the line but \((0, -1, 2)\) and \((6, 3, -5)\) do not.

### 7.4 Line Segments

In mathematics, when we talk about a line we mean the whole of the line, stretching away to infinity in both directions. But sometimes it is helpful to consider a finite portion, or **segment**, of a line. These can be described by using the vector equation of the line, and placing some restrictions on the values of the parameter \(\lambda\).

For example, the equation

\[
\vec{r} = 2\hat{i} - 3\hat{j} + \lambda(3\hat{i} - \hat{j}) \quad -1 \leq \lambda \leq 4
\]

describes the line segment with end-points at \((-1, -2)\) (the point obtained when \(\lambda = -1\)) and \((14, -7)\) (the point obtained when \(\lambda = 4\)).

![Diagram showing line segments with \(\lambda\) values](image)

If we impose a lower bound but not an upper bound (or an upper bound but not a lower bound!) on \(\lambda\), then we will get a **half-line**: a line segment which stretches off to infinity in one direction but ends at a point in the other. For example, the equation

\[
\vec{r} = \vec{0} + \lambda\hat{i} \quad \lambda \leq 0
\]

describes the half of the \(x\)-axis to the left of the origin.
From Vector to Cartesian Equations for Lines

In this section we have seen how to express a line using vectors. In Section 6.13 we saw that a line in 2-space can also be described as the solution set to a single linear equation in the variables \(x\) and \(y\). It is straightforward to calculate one description from the other.

Consider the line \( \mathbf{r} = \mathbf{a} + \lambda \mathbf{b} \). If \( \mathbf{r} = (x, y) \), \( \mathbf{a} = (a_x, a_y) \) and \( \mathbf{b} = (b_x, b_y) \)
then this gives
\[
\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \lambda \begin{pmatrix} b_x \\ b_y \end{pmatrix}.
\]
Considering the components we obtain a system of two linear equations in the three variables \(x\), \(y\) and \(\lambda\):
\[
\begin{align*}
x &= a_x + \lambda b_x, \\
y &= a_y + \lambda b_y
\end{align*}
\]
The equation of our line can be obtained by eliminating \(\lambda\) from these equations. Providing \(b_x \neq 0\) and \(b_y \neq 0\) the resulting equation can be written as a simple formula:
\[
\frac{x - a_x}{b_x} = \frac{y - a_y}{b_y}.
\]

Example. Find the equation describing the line \( \mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \).

Solution. Note that the formula above is not applicable here, as here \(b_y = 0\). Here we have
\[
\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 0 \end{pmatrix},
\]
or as a system of equations:
\[
\begin{align*}
x &= 1 + 3\lambda, \\
y &= 2.
\end{align*}
\]
Since \(\lambda\) is already missing from one of the equations, there is no work to do. The equation of the line is just \(y = 2\).

Remark. You might wonder why we can ignore the remaining equation with a \(\lambda\) in it – doesn’t that need to hold as well? The point is that this equation can be made to hold for any \(x\) and \(y\) by choosing \(\lambda\) appropriately, so it does not place any restriction on which points lie on the line. Thus, we can safely ignore it.

Terminology. The equation whose solution set is the line is called the cartesian equation of the line, to distinguish it from the vector equation.

Three Dimensions. We know that a single linear equation in three variables \(x\), \(y\) and \(z\) describes a plane. A line can be described as the intersection two planes, that is, the solution to a system of two equations. These can be obtained in a similar way: turn the vector equation into three equations in the variables \(x\), \(y\), \(z\) and \(\lambda\). Then eliminate \(\lambda\) from two of them to get the two equations defining the line.

Again, if \( \mathbf{r} = \mathbf{a} + \lambda \mathbf{b} \) where \( \mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j} + a_z \mathbf{k} \) and \( \mathbf{b} = b_x \mathbf{i} + b_y \mathbf{j} + b_z \mathbf{k} \) and \(b_x\), \(b_y\) and \(b_z\) are all non-zero there is a formula:
\[
\begin{align*}
\frac{x - a_x}{b_x} &= \frac{y - a_y}{b_y} = \frac{z - a_z}{b_z}.
\end{align*}
\]
Notice that the formula is actually two linear equations, expressed in a concise way. (One says that the first bit equals the second, and the other says that the second bit equals the third.)

Example. Find the cartesian system of equations for the line
\[
\mathbf{r} = \begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}.
\]
Solution. Here we can use the formula to get
\[
\frac{x - 1}{2} = \frac{y - 2}{3} = \frac{z + 4}{2}
\]
or rearranged in a more conventional form:
\[
3x - 2y = -1, \quad 2y - 3z = 16.
\]

7.6 Intersection of Lines

Imagine two lines (in any number of dimensions) with equations \( \mathbf{r}_1 = \mathbf{a} + \lambda \mathbf{b} \) and \( \mathbf{r}_2 = \mathbf{c} + \mu \mathbf{d} \). They will meet (or intersect) if there are values of \( \lambda \) and \( \mu \) which make \( \mathbf{r}_1 = \mathbf{r}_2 \).

Example. Do the lines \( \mathbf{r}_1 = \begin{pmatrix} 2 \\ 3 \\ 7 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} \) and \( \mathbf{r}_2 = \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \) meet? If so, at what point(s)?

Solution. We are looking for values of \( \lambda \) and \( \mu \) which make \( \mathbf{r}_1 = \mathbf{r}_2 \), that is, solutions to the equation:
\[
\begin{pmatrix} 2 \\ 3 \\ 7 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.
\]

Considering each coordinate gives a system of linear equations in the variables \( \lambda \) and \( \mu \):
\[
\begin{align*}
2 - \lambda &= 1 + \mu \\
3 + \lambda &= -5 + 2\mu \\
7 + 3\lambda &= 1
\end{align*}
\]
which rearrange to:
\[
\begin{align*}
-\lambda - \mu &= -1 \\
\lambda - 2\mu &= -8 \\
3\lambda &= -6.
\end{align*}
\]

Applying Gaussian elimination (or just solving by ad hoc manipulation) tells us that the unique solution to these equations is \( \lambda = -2 \) and \( \mu = 3 \). So the intersection of the lines is the point given by \( \lambda = -2 \) in the equation for the first line (which is the same as the point given by \( \mu = 3 \) in the equation for the second line - check this!), namely
\[
\begin{pmatrix} 2 \\ 3 \\ 7 \end{pmatrix} - 2 \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 + 2 \\ 3 - 2 \\ 7 - 6 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}.
\]
So the point of intersection is \( (4, 1, 1) \).

Remark. We have seen that two lines in 2-space will intersect at a single point unless they are parallel. In more than 2 dimensions, two lines can fail to meet without being parallel. Such a pair of lines is called skew.

7.7 Vector Equation of a Plane

A plane in 3-space can be efficiently described by giving

- a point in the plane (which can be described by its position vector); and
- a vector perpendicular or normal to the plane.
Suppose \( A \) is a point on the plane (with position vector \( \mathbf{a} \)) and \( \mathbf{n} \) is a vector normal to the plane. Suppose \( R \) is another point, with position vector \( \mathbf{r} \). Then \( R \) lies on the plane if and only if the displacement vector \( \overrightarrow{AR} = \mathbf{r} - \mathbf{a} \) is perpendicular to \( \mathbf{n} \). We know (see Section 4.2) that two vectors are perpendicular if and only if their scalar product is 0. So \( \overrightarrow{AR} \) lies on the plane precisely if:

\[
(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0.
\]

Using laws of scalar products (Section 4.5) this can be rewritten:

\[
\mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}.
\]

This is called the\(^4\) vector equation of the plane. Notice that the right hand side of the equation does not involve \( \mathbf{r} \). Once we have fixed on our \( \mathbf{a} \) and \( \mathbf{n} \), it can be (and usually is) evaluated to a constant.

**Example.** Find the vector equation of the plane passing through the point \((2, 1, 5)\) which is normal to the vector \(2\mathbf{i} - 3\mathbf{j} + 7\mathbf{k}\).

**Solution.** Here, \( \mathbf{a} \), the position vector of a point in the plane is given by \(2\mathbf{i} + \mathbf{j} + 5\mathbf{k}\) and the normal vector is \( \mathbf{n} = 2\mathbf{i} - 3\mathbf{j} + 7\mathbf{k}\). So, the equation of the plane is given by:

\[
\mathbf{r} \cdot (2\mathbf{i} - 3\mathbf{j} + 7\mathbf{k}) = (2\mathbf{i} + \mathbf{j} + 5\mathbf{k}) \cdot (2\mathbf{i} - 3\mathbf{j} + 7\mathbf{k}).
\]

However, the scalar product \((2\mathbf{i} + \mathbf{j} + 5\mathbf{k}) \cdot (2\mathbf{i} - 3\mathbf{j} + 7\mathbf{k})\) evaluates to 36, so the equation is

\[
\mathbf{r} \cdot (2\mathbf{i} - 3\mathbf{j} + 7\mathbf{k}) = 36.
\]

**Example.** Do the points \((8, 8, 7)\) and \((4, 6, 1)\) lie in the \( \mathbf{r} \cdot (-2\mathbf{i} + 8\mathbf{j} - 1\mathbf{k}) = 41 \)?

**Solution.** Let \( \mathbf{r} = 8\mathbf{i} + 8\mathbf{j} + 7\mathbf{k} \) be the position vector of \((8, 8, 7)\). Here

\[
\mathbf{r} \cdot (-2\mathbf{i} + 8\mathbf{j} - 1\mathbf{k}) = (8\mathbf{i} + 8\mathbf{j} + 7\mathbf{k}) \cdot (-2\mathbf{i} + 8\mathbf{j} - 1\mathbf{k}) = -16 + 64 = 41.
\]

As \( \mathbf{r} \cdot (-2\mathbf{i} + 8\mathbf{j} - 1\mathbf{k}) \) is indeed equal to 41, the point \((8, 8, 7)\) lies on the plane.

Now let \( \mathbf{r} = 4\mathbf{i} + 6\mathbf{j} + 1\mathbf{k} \) be the position vector of \((4, 6, 1)\). Here

\[
\mathbf{r} \cdot (-2\mathbf{i} + 8\mathbf{j} - 1\mathbf{k}) = (4\mathbf{i} + 6\mathbf{j} + 1\mathbf{k}) \cdot (-2\mathbf{i} + 8\mathbf{j} - 1\mathbf{k}) = -8 + 48 = 39.
\]

So in this case the vector equation is not satisfied and the point \((4, 6, 1)\) does not lie in the given plane.

### 7.8 Plane Through Three Points

Suppose \( A, B \) and \( C \) are three points in 3-space. There will usually be a unique plane which passes through all three points. To find the vector equation for the plane, we need:

1. The position vector of a point in the plane. For this we can pick any of the three points, so suppose \( \mathbf{a} \) is the position vector of \( A \).
2. A vector normal to the plane. For this we can take the vector/cross product of any two vectors in different directions which lie in the plane (see Section 4.8). For example, \( \overrightarrow{AB} \times \overrightarrow{AC} \).

**Example.** Find the vector equation of the plane through the points \((2, 6, 3)\), \((0, 6, 7)\) and \((5, 7, 5)\).

**Solution.** Let \( A = (2, 6, 3) \), \( B = (0, 6, 7) \) and \( C = (5, 7, 5) \). First we need the position vector of a point in the plane. We can use the position vector of \( A \) which is \( \mathbf{a} = 2\mathbf{i} + 6\mathbf{j} + 3\mathbf{k} \). Also, \( \overrightarrow{AB} = -2\mathbf{i} + 0\mathbf{j} + 4\mathbf{k} = -2\mathbf{i} + 4\mathbf{k} \) and \( \overrightarrow{AC} = 3\mathbf{i} + \mathbf{j} + 2\mathbf{k} \) so for our normal vector we may take

\[
\mathbf{n} = (-2\mathbf{i} + 0\mathbf{j} + 4\mathbf{k}) \times (3\mathbf{i} + \mathbf{j} + 2\mathbf{k})
= |0 \times 2 - 4 \times 1|\mathbf{i} + |4 \times 3 - (-2) \times 1|\mathbf{j} + |(-2) \times 1 - 0 \times 3|\mathbf{k}
= -4\mathbf{i} + 16\mathbf{j} - 2\mathbf{k}.
\]

\(^4\)Again, the use of the word “the” is not strictly accurate, as other choices of \( \mathbf{a} \) and \( \mathbf{n} \) could have given different equations for the same plane.
The equation of the plane is thus
\[ \mathbf{r} \cdot (-4\mathbf{i} + 16\mathbf{j} - 2\mathbf{k}) = (2\mathbf{i} + 6\mathbf{j} + 3\mathbf{k}) \cdot (-4\mathbf{i} + 16\mathbf{j} - 2\mathbf{k}). \]

But \((2\mathbf{i} + 6\mathbf{j} + 3\mathbf{k}) \cdot (-4\mathbf{i} + 16\mathbf{j} - 2\mathbf{k}) = 82\) so the equation is:
\[ \mathbf{r} \cdot (-4\mathbf{i} + 16\mathbf{j} - 2\mathbf{k}) = 82. \]

**Warning.** The above method will not work if the three points are **collinear**, that is, if they lie in a straight line. In this case there are lots of planes containing all three points, but the above method won’t actually find any of them.

**Exercise*.** Follow the method through in the case of three collinear points, and work out what goes wrong!

### 7.9 From Vector to Cartesian Equations of Planes

In Section 6.14, we saw that a plane in 3-dimensions can also be described by a single linear equation in the variables \(x, y\) and \(z\). In Section 7.7 we saw how to describe a plane using its vector equation. These two descriptions are closely related.

Consider the plane with vector equation
\[ \mathbf{r} \cdot \mathbf{n} = c \]
where \(\mathbf{n} = n_x\mathbf{i} + n_y\mathbf{j} + n_z\mathbf{k}\). A point \((x, y, z)\) will lie on the plane if and only if \(\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}\) satisfies the equation, which means that
\[ xn_x + yn_y + zn_z = c. \]

This is the cartesian equation of our plane.

**Remark.** It may seem a bit odd that our vector description of a plane works only in 3 dimensions. What happens if we look at an equation like \(\mathbf{r} \cdot \mathbf{n} = c\) where \(\mathbf{r}\) and \(\mathbf{n}\) are \(n\)-dimensional vectors with \(n \neq 3\)? The argument above allows us to transform this equation into a single linear equation, but this time with one variable for each of the \(n\) coordinates. As we saw in Section 6.14, such an equation defines a **hyperplane** in \(n\)-space, that is, a subspace of dimension \(n - 1\). So when \(n = 3\) such an equation defines a plane, but when \(n = 4\) it defines a 3-dimensional space within 4-space, and so on.

**Exercise*.** Investigate what happens when \(n = 2\). What is a hyperplane in 2 dimensions? How does the description of it relate to things you have already seen in the course?

### 7.10 Further Exercises

- Find both the vector and the cartesian equations of the line through the points \((2, 6)\) and \((5, -1)\).

- Does the point \((7, 5)\) lie on the line \(\mathbf{r} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -1 \end{pmatrix}\)?

- Find the vector equation of the plane containing the point \((2, 0, -3)\) and perpendicular to the vector \(\mathbf{i} + \mathbf{j} - 2\mathbf{k}\).

### 7.11 Intersection of a Line and a Plane

The line
\[ \mathbf{r} = \mathbf{a} + \lambda\mathbf{b} \quad (4) \]
and the plane
\[ \mathbf{r} \cdot \mathbf{n} = c \cdot \mathbf{n} \quad (5) \]
will intersect if a value of \(\lambda\) can be found which (from equation (4)) gives an \(\mathbf{r}\) that satisfies equation (5).

**Example.** Do the line \(\mathbf{r} = \begin{pmatrix} 8 \\ -1 \\ 22 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 7 \end{pmatrix}\) and the plane \(\mathbf{r} \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = 3\) meet? If so, where?
Solution. Substitute
\[
\mathbf{r} = \begin{pmatrix} 8 \\ -1 \\ 22 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 7 \end{pmatrix} = \begin{pmatrix} 8 + 2\lambda \\ -1 - \lambda \\ 22 + 7\lambda \end{pmatrix}
\]
into the equation for the plane. This gives:
\[
\begin{pmatrix} 8 + 2\lambda \\ -1 - \lambda \\ 22 + 7\lambda \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = 3.
\]
Expanding out using the component definition of the scalar product gives
\[
(8 + 2\lambda) \times 1 + (-1 - \lambda) \times (-1) + (22 + 7\lambda) \times 3 = 3.
\]
or \(\lambda = -3\). A unique solution is found so the line and the plane intersect at a single point. This is found by putting \(\lambda = -3\) into the equation of the line:
\[
\mathbf{r} = \begin{pmatrix} 8 \\ -1 \\ 22 \end{pmatrix} - 3 \begin{pmatrix} 2 \\ -1 \\ 7 \end{pmatrix} = \begin{pmatrix} 8 - 6 \\ -1 + 3 \\ 22 - 21 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.
\]
So the line and plane intersect at the point \((2, 2, 1)\).

Example. Do the line \(\mathbf{r} = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ -8 \end{pmatrix}\) and the plane \(\mathbf{r} \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = 5\) meet? If so, where?

Solution. This time substituting the equation of the line into the equation of the plane gives:
\[
\begin{pmatrix} 4 + \lambda \\ 2\lambda \\ 2 - 8\lambda \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = 5.
\]
Using the component definition of the scalar product, this means
\[
8 + 2\lambda + 6\lambda + 2 - 8\lambda = 5,
\]
in other words, \(10 = 5\). This is a contradiction, so there can be no values of \(\lambda\) for which the corresponding a point on the line also lies on the plane. In other words, the line is parallel to the plane, and does not meet it.

Remark. In three dimensions there are three possibilities for the intersection of a line and a plane:

* The line is not parallel to the plane. In this case the line and plane will intersect at a single point.
* The line lies in the plane. In this case every point on the line is in the plane, so the intersection is the whole of the line.
* The line is parallel to the plane and not in it. In this case they never meet so the intersection is empty.

In four or more dimensions, a line and a plane can also fail to meet without being parallel. But a line and a hyperplane will always behave as above.

7.12 Further Exercises

* Which of the points \((-3, 7, 2)\) and \((4, 7, -1)\) lie on the plane \(\mathbf{r} \cdot \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = 5\)?
• Find the point of intersection (if any) of the two lines

\[ \mathbf{r}_1 = \begin{pmatrix} 6 \\ 4 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{r}_2 = \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}. \]

• Find the intersection (if any) of the line

\[ \mathbf{r} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and the plane} \quad \mathbf{r} \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = 7. \]

• Find the vector equation of the plane through \((1,0,0), (1,1,0)\) and \((1,2,2)\). Now find another vector equation for the same plane.

8 Affine Transformations

A transformation is a way of moving things around in space (of any dimension). It maps every point to a point, and every vector to a vector. We are particularly interested in affine transformations, which satisfy the following two properties:

• Lines map to lines. In other words, three points which are colinear before the transformation will remain colinear after the transformation.

• Proportions in lines are conserved. For example, suppose \(A, B, C\) are colinear points, and the distance from \(A\) to \(B\) is twice the distance from \(B\) to \(C\). If the transformation takes \(A, B, C\) to \(A', B', C'\) respectively, then the distance from \(A'\) to \(B'\) will be twice the distance from \(B'\) to \(C'\).

Simple examples of affine transformations include rotations, reflections, scalings and translations. Here is a figure in the plane, and some examples of what can happen to it under some affine transformations:

![Original](original.png) ![Rotated](rotated.png) ![Scaled (y-axis)](scaled_y_axis.png) ![Scaled (both axes)](scaled_both_axes.png) ![Reflected](reflected.png)

Here are some things which cannot happen to it:

![Original](original.png) ![Rotated](rotated.png)

**Terminology.** The thing that a point (or line or figure etc.) gets taken to is called the image of the point (or line or figure) under the transformation. For example, if \((1,2)\) is taken to \((3,4)\), then \((3,4)\) is the image of \((1,2)\).

**Remark.** Note that when thinking about affine transformations, no special role is played by the origin. We can rotate through 90° about the point \((3,1)\) just as well as we can about the origin \((0,0)\).
8.1 Affine Transformations and Matrices

An affine transformation can be applied to a point or vector by multiplying the homogeneous coordinates by a matrix. Matrices for affine coordinates have the form

\[
\begin{bmatrix}
  a_{11} & a_{12} & a_{13} & b_1 \\
  a_{21} & a_{22} & a_{23} & b_2 \\
  a_{31} & a_{32} & a_{33} & b_3 \\
  0 & 0 & 0 & 1
\end{bmatrix}
\text{ or }
\begin{bmatrix}
  a_{11} & a_{12} & b_1 \\
  a_{21} & a_{22} & b_2 \\
  0 & 0 & 1
\end{bmatrix}
\]

in three and two dimensions respectively. Each possible transformation corresponds to such a matrix.

**Example.** Suppose \( A \) is the point \((2,4,1)\) and \( \mathbf{b} \) is the vector \(2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k} \). Find the images of \( A \) and \( \mathbf{b} \) under the transformation given by the matrix

\[
\begin{bmatrix}
  -2 & 4 & 3 & 1 \\
  -3 & 2 & 0 & 5 \\
  -5 & 0 & -4 & -4 \\
  0 & 0 & 0 & 1
\end{bmatrix}
\].

**Solution.** The homogeneous coordinates of \( A \) are \[ \begin{bmatrix} 2 \\ 4 \\ 1 \\ 1 \end{bmatrix} \] and the homogeneous coordinates of its image are

\[
\begin{bmatrix}
  -2 & 4 & 3 & 1 \\
  -3 & 2 & 0 & 5 \\
  -5 & 0 & -4 & -4 \\
  0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  2 \\
  4 \\
  1 \\
  1
\end{bmatrix} = \begin{bmatrix}
  16 \\
  7 \\
  -18 \\
  1
\end{bmatrix}, \text{ that is, the point (16, 7, -18).}
\]

The homogeneous coordinates of \( \mathbf{b} \) are \[ \begin{bmatrix} 2 \\ -3 \\ 4 \\ 0 \end{bmatrix} \] and the homogeneous coordinates of its image are

\[
\begin{bmatrix}
  -2 & 4 & 3 & 1 \\
  -3 & 2 & 0 & 5 \\
  -5 & 0 & -4 & -4 \\
  0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  2 \\
  -3 \\
  4 \\
  0
\end{bmatrix} = \begin{bmatrix}
  -4 \\
  -12 \\
  -26 \\
  0
\end{bmatrix}, \text{ that is, the vector } -4\mathbf{i} - 12\mathbf{j} - 26\mathbf{k}.
\]

**Example.** Consider the three points \( A = (0, 0), B = (0, 2) \) and \( C = (0, 4) \). These are colinear, with \( B \) lying half way between \( A \) and \( C \). Find the images of these points under the affine transformation described by

\[
\begin{bmatrix}
  1 & 2 & 5 \\
  -1 & 1 & 2 \\
  0 & 0 & 1
\end{bmatrix}.
\]

Show that the images are colinear with the image of \( B \) lying half-way between the images of \( A \) and \( C \).

**Solution.** The image of \( A \) is \[ \begin{bmatrix} 1 \\ 2 \\ 5 \\ 0 \end{bmatrix} \] \[ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \] \[ \begin{bmatrix} 0 \end{bmatrix} \] \[ \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix} \], that is, the point \((5, 2)\).

The image of \( B \) is \[ \begin{bmatrix} -1 \\ 2 \\ 0 \\ 0 \end{bmatrix} \] \[ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \] \[ \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} \], that is, the point \((9, 4)\).

The image of \( C \) is \[ \begin{bmatrix} -1 \\ 2 \\ 0 \\ 0 \end{bmatrix} \] \[ \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \] \[ \begin{bmatrix} 4 \\ 6 \\ 1 \end{bmatrix} \], that is, the point \((13, 6)\).

Call the images \( A' \), \( B' \) and \( C' \) respectively. To check that these three points are colinear, we find the unique line through two of them (as in Section 7.2) and check if the third point lies on this line (as in Section 7.3).

The position vector of \( A' \) is \( 5\mathbf{i} + 2\mathbf{j} \). The displacement vector \( \overrightarrow{A'C'} \) is \( 8\mathbf{i} + 4\mathbf{j} \). So the line through \( A' \) and \( C' \) has vector equation \( \mathbf{r} = 5\mathbf{i} + 2\mathbf{j} + \lambda(8\mathbf{i} + 4\mathbf{j}) \).
To check if $B'$ lies on this line we equate $x$-components to get $9 = 5 + 8\lambda$ so $\lambda = \frac{1}{2}$ is the only possible choice. Setting $\lambda = \frac{1}{2}$ gives the point $(9,4)$ which is $B'$ so $B'$ does indeed lie on the line containing $A'$ and $C'$ and the three points are colinear.

Also, the distance $A'B'$ is $\sqrt{(9 - 5)^2 + (4 - 2)^2} = \sqrt{20}$; the distance $B'C'$ is $\sqrt{(13 - 9)^2 + (6 - 4)^2} = \sqrt{20}$. So, $B'$ lies half way between $A'$ and $C'$, just as $B$ lies half way between $A$ and $C$.

### 8.2 Further Exercises

- Which of the following matrices represent affine transformations in two dimensions?

  \[
  \begin{bmatrix}
  2 & 6 & -1 \\
  -1 & 2 & 4 \\
  2 & -1 & 2
  \end{bmatrix}, \quad
  \begin{bmatrix}
  2 & 4 & -1 \\
  2 & 5 & -2 \\
  1 & 0 & 1
  \end{bmatrix}, \quad
  \begin{bmatrix}
  2 & -1 & 2 \\
  -1 & 6 & -3 \\
  0 & 0 & 1
  \end{bmatrix}, \quad
  \begin{bmatrix}
  1 & 2 & 3 \\
  -1 & 2 & 3 \\
  0 & 0 & 0
  \end{bmatrix}
  \]

- Apply the transformation matrix

  \[
  \begin{bmatrix}
  9 & 2 & 3 & 5 \\
  -3 & 2 & 0 & 4 \\
  -1 & 2 & 3 & -2 \\
  0 & 0 & 0 & 1
  \end{bmatrix}
  \]

  to the point $(4,6,-1)$ and the vector $2\mathbf{i} - \mathbf{j} - 6\mathbf{k}$.

### 8.3 Translation

A translation is a transformation that moves every point by a fixed distance in a given direction. Intuitively, translations “slide things around” without changing their sizes or orientations:

A translation is described by a vector, indicating the direction in which things are moved and the distance they move. For example, in three dimensions, the translation through the vector $\alpha_x\mathbf{i} + \alpha_y\mathbf{j} + \alpha_z\mathbf{k}$ moves every point a distance of $\alpha_x$ in the $x$-direction, $\alpha_y$ in the $y$-direction and $\alpha_z$ in the $z$-direction. This translation is performed by applying the translation matrix

\[
T = \begin{bmatrix}
1 & 0 & 0 & \alpha_x \\
0 & 1 & 0 & \alpha_y \\
0 & 0 & 1 & \alpha_z \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

to homogeneous coordinates. The homogeneous coordinates of the new point $\mathbf{x}_{new}$ will be related to the homogeneous coordinates of the old point $\mathbf{x}_{old}$ by $\mathbf{x}_{new} = T\mathbf{x}_{old}$

**Example.** Find the matrix for a translation through the vector $2\mathbf{i} + 1\mathbf{j} + 3\mathbf{k}$. Hence find the images of the points $A = (1,1,1)$ and $B = (-1,2,4)$ and the vector $\overline{AB}$ under this transformation.

**Solution.** Here, $\alpha_x = 2$, $\alpha_y = 1$ and $\alpha_z = 3$, so the matrix is:

\[
\begin{bmatrix}
1 & 0 & 0 & 2 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

In homogeneous coordinates, $A = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $B = \begin{bmatrix} -1 \\ 2 \\ 4 \\ 1 \end{bmatrix}$ and $\overline{AB} = \begin{bmatrix} -2 \\ 1 \\ 3 \\ 0 \end{bmatrix}$.

The image of $A$ is

\[
\begin{bmatrix}
1 & 0 & 0 & 2 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+0+0+2 \\
0+1+0+1 \\
0+0+1+3 \\
0+0+0+1
\end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 4 \\ 1 \end{bmatrix}
\]

, that is, the point $(3,2,4)$. 39
The image of $B$ is given by the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 2 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 7 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 7 \\ 1 \end{bmatrix}$$

that is, the point $(1, 3, 7)$.

The image of $\overrightarrow{AB}$ is given by the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 3 \\ 0 \end{bmatrix}$$

that is, $\overrightarrow{AB}$ again!

**Remark.** As this example illustrates, vectors are not affected by translation! This is because, as we have already seen, the position of a vector is unimportant — only its magnitude and direction matter, and these do not change. Notice that, for the same reason, the displacement vector from the image of $A$ to the image of $B$ is the same as that from $A$ to $B$.

**Two Dimensions.** In two dimensions, the translation through $\alpha_1 x + \alpha_2 y$ is given by the matrix

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

**Example.** Find the transformation matrix $T$ in two dimensions for the translation which slides every point right 2 units and down 5 units.

**Solution.** This is the translation through $2\hat{i} - 5\hat{j}$, so here $\alpha_x = 2$, $\alpha_y = -5$ and the matrix is

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix}$$

8.4 Reversing a Translation

A translation through a vector $\mathbf{a}$ can be “reversed” or “undone” by translating through $-\mathbf{a}$.

**Example.** Let $\mathbf{v} = 1\hat{i} + 3\hat{j} - 4\hat{k}$. Find the matrices for a translation through $\mathbf{v}$ and a translation through $-\mathbf{v}$. Apply the first transformation to the point $(6, -1, 3)$. Now apply the second to the result.

**Solution.** The matrices are

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

respectively.

The first takes $(6, -1, 3)$ to

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 7 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \\ -1 \\ 1 \end{bmatrix}$$

that is, the point $(7, 2, -1)$.

The second takes this to

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ -1 \\ 3 \\ 1 \end{bmatrix}$$

which is back to the original point $(6, -1, 3)$!

This is what we should expect, since we have “done” a translation and then “undone” it.

8.5 Scaling

A scaling is a transformation that enlarges or contracts an object. A uniform scaling expands or contracts an object equally along all axes. A more general (non-uniform) scaling may stretch or squash the object by different factors in the directions of the different axes. Here are some examples of how scaling can affect a figure:
Figure (a) shows the result of a uniform scaling: the figure is scaled by a factor of $\frac{1}{2}$ in all directions, in other words, halved in size. Figure (b) shows a non-uniform scaling: the figure has been scaled up (expanded) along the $x$-axis and scaled down (contracted) along the $y$-axis. Figure (c) shows another non-uniform scaling: this time the figure has been scaled up (expanded) along both axes, but by differing amounts.

A scaling is always centred around some point, which it leaves fixed. We shall consider first scalings about the origin; later we shall see how to construct scalings about other points.

Scaling Matrices. In three dimensions, a scaling about the origin is performed by a matrix of the form

$$
\begin{bmatrix}
\beta_x & 0 & 0 & 0 \\
0 & \beta_y & 0 & 0 \\
0 & 0 & \beta_z & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

This matrix scales by a factor of $\beta_x$ along the $x$-axis, a factor of $\beta_y$ along the $y$-axis and a factor of $\beta_z$ along the $z$-axis.

Remark. For a uniform scaling by a factor of $\beta$, we simply set $\beta_x = \beta_y = \beta_z = \beta$ in the above matrix.

Example. Consider the point $(2, 4, 1)$ in three dimensions and apply the uniform scaling about the origin where each coordinate is expanded by a factor of 2. Give the new coordinates of the point.

Solution. Here, the transformation matrix is

$$
\begin{bmatrix}
2 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

This transforms the point $(2, 4, 1)$ to

$$
\begin{bmatrix}
2 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
2 \\
4 \\
1 \\
1
\end{bmatrix} =
\begin{bmatrix}
4 \\
8 \\
2 \\
1
\end{bmatrix},
$$

so the image of the point is $(4, 8, 2)$. (You could probably have worked that out without thinking about matrices, but the matrix approach will prove useful later for more complex transformations.)

Warning. To get a scaling, the factors $\beta_x$, $\beta_y$ and $\beta_z$ must all be positive. Replacing them with negative values can give other kinds of transformations (such as for example reflections, which we will see later).

Remark. Scaling by a factor bigger than 1 will make points get further apart; this is called an expansion. Scaling by a factor between 0 and 1 will cause points to get closer together; such a transformation is called a contraction or reduction or compression.

Warning/Terminology. Think carefully about terms such as “reduction by a factor of 2”. This means “scale by a factor of $\frac{1}{2}$”!

Two dimensions. In two dimensions, scaling about the origin by a factor $\beta_x$ in the $x$-direction and $\beta_y$ in the $y$-direction is carried out by the matrix

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

Example. Consider the square with vertices at $(0, 0)$, $(2, 0)$, $(2, 2)$ and $(0, 2)$. It has area $2 \times 2 = 4$. Carry out the transformation of uniform reduction by a factor of 2 about the origin and find the area of the image of the square under this scaling.
Solution. The transformation matrix is \( S = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \).

The point \((0,0)\) transforms to \( P = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \), that is, the point \((0,0)\).

The point \((2,0)\) transforms to \( P = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \), that is, \((1,0)\).

The point \((2,2)\) transforms to \( P = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \), that is, \((1,1)\).

The point \((0,2)\) transforms to \( P = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \), that is, \((0,1)\).

The image of the square now has area 1 \(\times\) 1 = 1 (as opposed to 4 before).

Remark. In general, a uniform scaling by a factor of \(\beta\) will scale lengths by \(\beta\), areas by \(\beta^2\) and volumes (if working in 3 dimensions) by \(\beta^3\).

Example. Consider a scaling which fixes the origin, stretches the \(x\)-axis by a factor of 2, compresses the \(y\)-axis by a factor of 3 and leaves the \(z\)-axis unchanged. Find what happens to the points \(A = (2,3,4)\) and \(B = (1,6,5)\) and the vector \(\overrightarrow{AB}\).

Solution. The transformation matrix for this scaling is \( S = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \).

The points \((2,3,4)\) and \((2,2,5)\) transform to

\[
\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}
\]

respectively, that is, the points \((4,1,4)\) and \((2,2,5)\) respectively. The vector \(\overrightarrow{AB} = -\mathbf{i} + 3\mathbf{j} + \mathbf{k}\) transforms to

\[
\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}
\]

which is the vector \(-2\mathbf{i} + \mathbf{j} + \mathbf{k}\).

Remark. Again, notice how the image of the displacement vector \(\overrightarrow{AB}\) is the displacement vector between the images. This holds for all affine transformations: if \(P\) and \(Q\) are points and an affine transformation takes \(P\) to \(P'\) and \(Q\) to \(Q'\) then it will transform \(\overrightarrow{PQ}\) to \(\overrightarrow{P'Q'}\).

Exercise*. Use matrices to prove this. (You will need to use the fact that matrix multiplication is distributive, as we saw in Section 5.11.)

8.6 Further Exercises

* Give the transformation matrix for the two-dimensional translation which moves everything \(-2\) units in the \(x\)-direction and \(+3\) units in the \(y\)-direction.
• Apply the translation matrix \[
\begin{bmatrix}
1 & 0 & -5 \\
0 & 1 & -3 \\
0 & 0 & 1
\end{bmatrix}
\] to the points \(A = (2, 5)\) and \(B = (-1, 7)\) and also the vector \(\overrightarrow{AB}\). What vector does this matrix translate through?

• Consider the scaling in two dimensions where \(x\) is expanded by a factor of 2 and \(y\) is contracted by a factor of three. Consider the square with corners at \((0, 0), (6, 0), (6, 6)\) and \((0, 6)\). Apply the scaling to the square and draw the original and revised figures.

• Apply the scaling matrix \[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 4 & 0 & 0 \\
0 & 0 & 2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] to the points \((1, 1, 1)\) and \((1, 5, 4)\).

9 Combining Transformations and Matrices

9.1 Composition and Multiplication

Suppose \(A\) and \(B\) are the matrices representing two affine transformations. What is the combined effect of applying the transformation \(A\) followed by the transformation \(B\)?

Suppose \(\mathbf{x}\) is the homogeneous coordinate vector of a point in space. Applying transformation \(A\) takes it to the point \(A\mathbf{x}\). Now applying the transformation \(B\) takes it onward to \(B(A\mathbf{x})\). But we know that matrix multiplication is associative (see Section 5.11) so \(B(A\mathbf{x}) = (BA)\mathbf{x}\).

Since this argument works for every point, this means the combined transformation is the one performed by the product matrix \(BA\).

Warning. Notice carefully the order of composition. Because we apply our matrices on the left of a point or vector, the rightmost matrix in the product is the nearest to the point or vector, and hence the first to get applied to it. So the product \(BA\) corresponds to doing \(A\) first and then \(B\). It is important to get this right, because matrix multiplication is not commutative (see Section 5.10 above).

Example. Find a matrix for the combined transformation of a scaling by a uniform factor of 4 around the origin, followed by a translation through \(1\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}\).

Solution. The matrices for the two transformations are

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

respectively. The combined transformation is given by the matrix product.

\[
BA = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 & 1 \\ 0 & 4 & 0 & 3 \\ 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
\]

Again, notice carefully the order in which the matrices are multiplied.

Exercise. Find a matrix for the combined transformation of a translation through \(1\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}\) followed by a scaling by a uniform factor of 4 around the origin. Verify that it differs from the matrix \(BA\) above.

Remark. More generally, if \(A_1, A_2, \ldots, A_{k-1}, A_k\) are the matrices for a sequence of transformations, then the combined effect of applying the transformations in sequence \((A_1, \text{then} A_2, \text{and so on up to} A_k)\) is described by the matrix \(A_k A_{k-1} \ldots A_2 A_1\). (Yet again, notice the order of multiplication!)
9.2 Matrix Powers

A power of a matrix is a product of some number of copies of the matrix. Just as with numbers, we write $A^n$ for the $n$th power of $A$. For example,

$$A^7 = \underbrace{AAAAAAA}_{\text{is the 7th power of } A}$$

is the 7th power of $A$. If $A$ is the matrix of a transformation, it follows from the previous section that $A^n$ is the matrix which performs the transformation $n$ times in sequence.

**Example.** If $A$ is a matrix which translates through 3 units upwards, then $A^4$ will be the matrix which translates through 3 units upwards 4 times over. In other others, $A^4$ will translate through 12 units upwards.

9.3 The Identity Transformation

A surprisingly useful transformation is the one which does absolutely nothing! In other words, the transformation which takes every point $x$ to itself. This is called the identity transformation and is “performed” by the identity matrix (see Section 5.7 above) of the appropriate size (which because of the homogeneous coordinates is $I_3$ when working in 2 dimensions, $I_4$ in three dimensions!).

**Remark/Exercise.** The identity transformation can be described as a translation through $0$, or alternatively a uniform scaling by a factor of $1$! Use the methods in Section 8 to produce matrices for these two transformations, and check that these are in fact the identity matrix.

9.4 Reversal and Inverses

Many affine transformations, including those we have looked at so far, are reversible. In other words, there is another affine transformation which “undoes” them.

**Example.** The transformation which translates everything 5 units to the right can be undone by the transformation which translates everything 5 units to the left. (See Section 8.4 above.)

**Example.** The uniform scaling by a factor of 5 can be undone by a uniform scaling by a factor of $1/5$.

What does this mean in terms of matrices? When we say that $B$ “undoes” $A$, we mean that the combined effect of transformation $B$ followed by $A$ is to do nothing. In other words, the combined transformation is the identity transformation. It can be shown (challenging exercise*) that this means that $B = A^{-1}$.

To summarise, if $A$ corresponds to a transformation, then $A^{-1}$ (if it exists) corresponds to the “undoing” transformation.

**Example.** Find the inverse of the matrix $A = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

**Solution.** We recognise $A$ as the matrix describing a translation through $2\hat{i} + 3\hat{j} + 1\hat{k}$. The “undoing” transformation is clearly a translation through $-2\hat{i} - 3\hat{j} - 1\hat{k}$, which has matrix

$$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$ 

So this matrix is the inverse of $A$.

**Exercise.** Check this by multiplying the two matrices together.
9.5 Building Complex Transformations

The correspondence between products of matrices and sequences of transformations can be used to build complex affine transformations from sequences of simpler ones.

Example. Find a matrix to perform a uniform scaling by a factor of 2 around the point \( (2, 1, 3) \). Find the images of the points \( (2, 1, 3) \) and \( (4, 4, 7) \) under this transformation.

Solution. The desired transformation can be performed by

- moving the point \( (2, 1, 3) \) to the origin (by a translation through \( -2\mathbf{i} - \mathbf{j} - 3\mathbf{k} \));
- then scaling by a uniform factor of 2 about the origin;
- then finally moving the origin back to \( (2, 1, 3) \) (by a translation through \( 2\mathbf{i} + 1\mathbf{j} + 3\mathbf{k} \)).

The matrices associated with these steps are

\[
A = \begin{bmatrix}
1 & 0 & 0 & -2 \\
0 & 1 & 0 & -1 \\
0 & 0 & 1 & -3 \\
0 & 0 & 0 & 1
\end{bmatrix},
B = \begin{bmatrix}
2 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

and

\[
A^{-1} = \begin{bmatrix}
1 & 0 & 0 & 2 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

respectively. (Notice that the third operation is the reverse of the first, so the third matrix will be the inverse of the first). So the combined transformation will be performed by the product matrix:

\[
A^{-1}BA = \begin{bmatrix}
1 & 0 & 0 & 2 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
2 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & -2 \\
0 & 1 & 0 & -1 \\
0 & 0 & 1 & -3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
= \begin{bmatrix}
2 & 0 & 0 & -2 \\
0 & 2 & 0 & -1 \\
0 & 0 & 2 & -3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

As ever, notice the right-to-left order of multiplication. We wish to perform \( A \) first, then \( B \), then \( A^{-1} \) so we must consider \( A^{-1}BA \). So, the complete transformation is governed by the matrix

\[
\begin{bmatrix}
2 & 0 & 0 & -2 \\
0 & 2 & 0 & -1 \\
0 & 0 & 2 & -3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

The point \( (2, 1, 3) \) transforms to

\[
\begin{bmatrix}
2 & 0 & 0 & -2 \\
0 & 2 & 0 & -1 \\
0 & 0 & 2 & -3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
2 \\
1 \\
3 \\
1
\end{bmatrix}
= \begin{bmatrix}
4 - 2 \\
2 - 1 \\
6 - 3 \\
1
\end{bmatrix}
= \begin{bmatrix}
2 \\
1 \\
3 \\
1
\end{bmatrix},
\]

in other words, to itself. This should not come as a surprise, since this was the point about which we are rotating! The point \( (4, 4, 7) \) transforms to

\[
\begin{bmatrix}
2 & 0 & 0 & -2 \\
0 & 2 & 0 & -1 \\
0 & 0 & 2 & -3 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
4 \\
4 \\
7 \\
1
\end{bmatrix}
= \begin{bmatrix}
8 - 2 \\
8 - 1 \\
14 - 3 \\
1
\end{bmatrix}
= \begin{bmatrix}
6 \\
7 \\
11 \\
1
\end{bmatrix},
\]

45
that is, the point $(6, 7, 11)$.

**Mathematical Aside.** In the previous example we modified the action of a matrix $(B)$ by sandwiching it between another matrix $(A)$ and its inverse $(A^{-1})$. This basic idea is fundamental in many areas of mathematics (especially group theory) and is called **conjugation** of $B$ by $A$. We shall see more examples of this later.

### 9.6 Scaling Around a Point

In fact it is possible to give a general formula for a matrix to scale about a particular point. In three dimensions, the matrix to scale by $\beta_x$ in the $x$-direction, $\beta_y$ in the $y$-direction and $\beta_z$ in the $z$-direction around the point $(p, q, r)$ is:

$$
\begin{bmatrix}
\beta_x & 0 & 0 & p(1 - \beta_x) \\
0 & \beta_y & 0 & q(1 - \beta_y) \\
0 & 0 & \beta_z & r(1 - \beta_z) \\
0 & 0 & 0 & 1
\end{bmatrix}.
$$

In two dimensions, the matrix to scale by $\beta_x$ in the $x$-direction and $\beta_y$ in the $y$-direction around the point $(p, q)$ is

$$
\begin{bmatrix}
\beta_x & 0 & p(1 - \beta_x) \\
0 & \beta_y & q(1 - \beta_y) \\
0 & 0 & 1
\end{bmatrix}.
$$

**Exercise.** Use the approach of the previous example to derive this formula.