Linear Transformations

Objective: To understand linear transformations of 2- and 3-dimensional space and their representation as matrices.

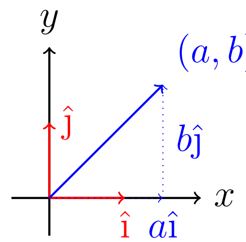
Warm-up:

Let's explore rotations of points about the origin.

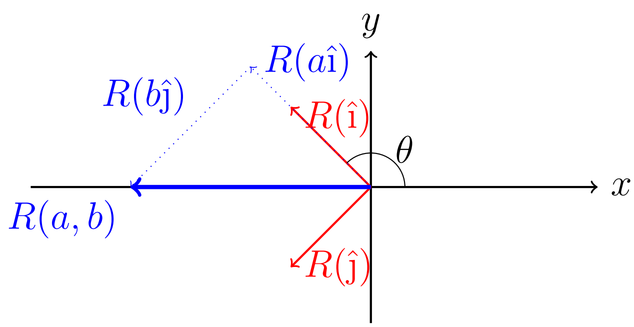
1. Suppose the point $(1,0)$ is rotated anticlockwise by an angle θ. What are its new coordinates?
2. Suppose the point $(0,1)$ is rotated anticlockwise by an angle θ. What are its new coordinates?
3. Suppose the points $(1,0)$, $(0,1)$, and (a,b) are rotated anticlockwise by $\frac{\pi}{2}$.
   (a) What are the coordinates of the rotations of $(1,0)$ and $(0,1)$?
   (b) The rotated coordinates of (a,b) are $(-b,a)$. We can relate the original coordinates to each other by the equation
   $$(a,b) = a(1,0) + b(0,1).$$
   Write a similar equation linking the coordinates of the rotated points.
4. Suppose the points $(1,0)$, $(0,1)$, and (a,b) are rotated anticlockwise by $\frac{3\pi}{4}$.
   (a) What are the coordinates of the rotations of $(1,0)$ and $(0,1)$?
   (b) The rotated coordinates of (a,b) are
   $$\left(\frac{-a}{\sqrt{2}} - \frac{b}{\sqrt{2}}, \frac{b}{\sqrt{2}} - \frac{a}{\sqrt{2}}\right).$$
   We can relate the original coordinates to each other by the equation
   $$(a,b) = a(1,0) + b(0,1).$$
   Write a similar equation linking the coordinates of the rotated points.

Theory—Rotations:

Think of the point $(1,0)$ as instructions "go one unit to the right." Call this instruction $\hat{\imath}$. Similarly, $(0,1)$ is the instruction "go one unit upwards." Call this $\hat{\jmath}$. Then a general point (a,b) may be thought of as $a\hat{\imath}+b\hat{\jmath}$—"go a to the right and b upwards." We call $\hat{\imath}$ and $\hat{\jmath}$ the **standard unit vectors**.



Let R denote rotation by an angle θ. Then $R(\hat{\imath})$ is the instruction "move one unit in the direction θ above the positive x-axis" and $R(\hat{\jmath})$ is the instruction "move one unit in the direction θ anticlockwise of the positive y-axis." Rotate the whole setup above by θ:



Rotating everything by θ doesn't affect the *relative angles*. So if R is any rotation, we have

$$R(a\hat{\imath}+b\hat{\jmath}) = aR(\hat{\imath}) + bR(\hat{\jmath}).$$

So to understand a rotation, it is enough to understand what it does to the **unit vectors** $\hat{\imath}$ and $\hat{\jmath}$. We saw in the warm-up that, for rotation by angle θ:

$$R(\hat{\imath}) = (\cos(\theta), \sin(\theta)) = \cos(\theta)\hat{\imath} + \sin(\theta)\hat{\jmath}$$

$$R(\hat{\jmath}) = (-\sin(\theta), \cos(\theta)) = -\sin(\theta)\hat{\imath} + \cos(\theta)\hat{\jmath}.$$

Therefore we have:

$$\begin{aligned} R(a,b) &= aR(\hat{\imath}) + bR(\hat{\jmath}) \\ &= a(\cos(\theta)\hat{\imath} + \sin(\theta)\hat{\jmath}) + b(-\sin(\theta)\hat{\imath} + \cos(\theta)\hat{\jmath}) \\ &= (a\cos(\theta) - b\sin(\theta))\hat{\imath} + (a\sin(\theta) + b\cos(\theta))\hat{\jmath} \\ &= (a\cos(\theta) - b\sin(\theta),\ a\sin(\theta) + b\cos(\theta)) \end{aligned}$$

Theory—Linear Transformations:

We saw on the last page that if R is rotation by an angle θ anticlockwise about the origin, then for any point (a, b):

$$R(a, b) = (a\cos(\theta) - b\sin(\theta),\ a\sin(\theta) + b\cos(\theta)).$$

Representing points by column vectors, we can write this as a matrix equation:

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

Therefore the matrix

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

gives a concise description of the rotation of any point about the origin.

The key to us showing this was that once you know what R does to the standard unit vectors $\hat{\imath}$ and $\hat{\jmath}$, then what it does to any other vector is determined by the equation

$$R(a, b) = aR(\hat{\imath}) + bR(\hat{\jmath}).$$

This is because rotation about the origin is a **linear transformation**. A function T from the plane to itself (or 3D space to itself, or in general any vector space to itself) is a **linear transformation** if it satisfies the two properties

$$T(u + v) = T(u) + T(v)$$
$$T(\lambda v) = \lambda T(v)$$

for any vectors u and v and scalar λ.

1. Let A be a 2×2 matrix and let $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the "multiply by A" function defined by $T_A(v) = Av$. Show that T_A is a linear transformation.
2. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation.
   (a) Show that for any point (x, y), we have $T(x, y) = xT(\hat{\imath}) + yT(\hat{\jmath})$.
   (b) Let $T(\hat{\imath}) = (a, c)$ and $T(\hat{\jmath}) = (b, d)$. Show that (writing points as column vectors) T is given by multiplication by the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

So multiplication by a matrix is a linear transformation, and any linear transformation can be written in this form! This is true in higher dimensions as well.

Practice:

Rotation by angle θ anticlockwise about the origin:

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

For any linear transformation T, if $T(1,0) = (a,c)$ and $T(0,1) = (b,d)$, then T is represented by the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

1. Calculate the inverse of the general rotation matrix above; compare with the rotation matrix by $-\theta$.

2. Write down the matrices representing rotations about the origin by
   (a) $\frac{\pi}{4}$ anticlockwise
   (b) $\frac{3\pi}{2}$ anticlockwise
   (c) $\frac{\pi}{2}$ clockwise
   (d) $\frac{2\pi}{3}$ clockwise

3. Identify the rotation represented by the matrix

$$\begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

4. Write down matrices representing the following linear transformations:
   (a) Reflection in the line $y = 0$
   (b) Reflection in the line $x = 0$
   (c) Reflection in the line $y = x$
   (d) Reflection in the line $y = -x$
   (e) Enlargement with scale factor 7 about the origin
   (f) Contraction by $\frac{1}{3}$ about the origin

   and calculate their determinants.

5. Describe the linear transformations represented by the following matrices:

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}, \quad \begin{pmatrix} -5 & 0 \\ 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & -7 \end{pmatrix}.$$

Successive Transformations:

Suppose S and T are two linear transformations $\mathbb{R}^2 \to \mathbb{R}^2$. Let M_S and M_T be the matrices representing them. Let $R : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation defined by $R(v) = S(T(v))$; so R is the result of doing first T and then S. Show that the matrix representing R is simply the matrix product $M_S M_T$.

1. Write down the matrices representing rotation by $45°$ anticlockwise, and reflection in the line $y = -x$. Hence compute the matrices representing the compound transformations "first rotate by $45°$ anticlockwise, then reflect in the line $y = -x$" and "first reflect in $y = -x$, then rotate by $45°$ anticlockwise."

2. Show that, for any angle θ,
$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(-\theta) & \cos(\theta) \end{pmatrix}.$$
Hence conclude that reflecting the plane in the line $y = x$, then rotating by an angle θ anticlockwise is the same as first rotating by θ clockwise, then reflecting in $y = x$.

3. Let $T \colon \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation and let $\lambda \in \mathbb{R}$ be a scalar. Show that first applying T and then enlarging about the origin by a factor of λ is the same as enlarging by λ and then applying T. Note: this can be done using matrices, or the abstract definition of a linear transformation. Can you see both ways to do it?

4. Let θ and ϕ be two angles. Let R_θ, R_ϕ, and $R_{\theta+\phi}$ be the rotations anticlockwise about the origin by the angles θ, ϕ, and $\theta + \phi$ respectively. Geometrically, we must have $R_{\theta+\phi}(v) = R_\theta(R_\phi(v)) = R_\phi(R_\theta(v))$ for any vector v. Prove this using matrices.

5. The matrix
$$M = \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}$$
represents a rotation by θ followed by an enlargement by λ. Compute the determinant of M and hence find the value of λ (hint: the determinant is the *area* scale factor, λ is the *length* scale factor). Hence find the value of θ.

Linear Transformations of $\mathbb{R}^3$:

In $\mathbb{R}^2$, every point (x, y) can be written as $x\hat{\imath}+y\hat{\jmath}$, where $\hat{\imath}$ and $\hat{\jmath}$ are the standard unit vectors. Similarly, in 3-dimensional space $\mathbb{R}^3$, we can define

$$\hat{\imath} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{\jmath} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \hat{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and write any point (x, y, z) as $x\hat{\imath} + y\hat{\jmath} + z\hat{k}$. Then, given a linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^3$, we have

$$T(x, y, z) = xT(\hat{\imath}) + yT(\hat{\jmath}) + zT(\hat{k}),$$

so T is entirely determined by what it does to the three standard unit vectors. Given any 3×3 matrix A, multiplication by A (the function T_A defined by $T_A(v) = Av$) is a linear transformation, and conversely any linear transformation T can be written as T_A where A is the matrix

$$\begin{pmatrix} T(\hat{\imath}) & T(\hat{\jmath}) & T(\hat{k}) \end{pmatrix},$$

whose columns are given by applying T to the standard unit vectors. So just as in 2 dimensions, there is a correspondence between matrices and linear transformations.

1. Describe the linear transformation represented by the matrix

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

   where θ is any angle. Compute the determinant; does this make sense geometrically (in terms of volume scaling).

2. Describe the linear transformation represented by the matrix

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

   where θ is any angle.

3. Write down the matrix representing a rotation about the z-axis by an angle θ, anticlockwise in a right-handed coordinate system.

4. Describe the linear transformations represented by the matrices
$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
and compute their determinants.

5. Compute the matrix which represents a rotation of 60° anticlockwise about the z-axis, followed by reflection in the plane $y = 0$.

6. Let S be the linear transformation "reflect in the plane $x = 0$, then rotate by π about the y-axis" and let T be the linear transformation "reflect in the plane $z = 0$." Compute the matrices M_S and M_T representing these transformations. What can you conclude about the transformations S and T? What is the transformation "do S, then do T?"

7. Compute the matrix which represents reflection in the plane $z = 0$, followed by rotation of $\frac{\pi}{2}$ about the y-axis, followed by rotation of $-\frac{5\pi}{3}$ about the z-axis, followed by reflection in the plane $x = 0$.

8. Let R be the rotation about the x-axis by an angle $\frac{2\pi}{3}$ anticlockwise, and S the reflection in the plane $y = 0$. Let M_R and M_S be the matrices representing R and S respectively.
   (a) Compute M_R^{-1} and M_S^{-1} (hint: if you think about the geometry, you **don't** need to do any matrix calculations!).
   (b) Compute the matrix representing the transformation "do R, then do S."
   (c) Compute the inverse of the matrix from part (b).
   (d) Compute the matrix products $M_R^{-1}M_S^{-1}$ and $M_S^{-1}M_R^{-1}$. Compare with your answer to part (c), and explain what you observe.

9. Describe the linear transformation represented by the matrix
$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$
Compute the determinant of this matrix and compare with what you would expect this transformation to do to volumes.

Further Practice:

These questions go beyond the A-level syllabus.

1. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ be the function defined by $f(x, y) = (x, y, x + y)$.
   (a) Viewing points as column vectors, write down $f(\hat{\imath})$ and $f(\hat{\jmath})$ as column vectors, and create a 3×2 matrix A with these as columns.
   (b) Compute
   $$A \begin{pmatrix} x \\ y \end{pmatrix}$$
   and compare with the definition of f.
   (c) Show that f is linear—*i.e.*, that
   $$f(u + v) = f(u) + f(v) \quad \text{and} \quad f(\lambda v) = \lambda f(v),$$
   for any scalar λ and vectors u and v.

2. Let $g\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the function defined by $g(x, y, z) = (x, y)$.
   (a) Describe geometrically the effect of the function g.
   (b) Show that g is linear.
   (c) Write down a 2×3 matrix representing g.
   (d) Let f be the function from the previous question. Let $fg\colon \mathbb{R}^3 \to \mathbb{R}^3$ denote the transformation "first do g, then do f," so $fg(v) = f(g(v))$. Similarly, let $gf\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation "do f, then do g." Write down formulae expressing fg and gf.
   (e) Let M_f be the 3×2 matrix found in question 1, representing f, and let M_g be the 2×3 matrix representing g, found in part (c). Compute $M_f M_g$ and $M_g M_f$, and compare with your answers to part (d).

3. Let P be the set of all polynomials in the variable x with real coefficients. Let $D\colon P \to P$ be the function "differentiate with respect to x:" so if f is a polynomial, $D(f)$ is the derivative $f'(x)$.
   (a) Show that D is linear—*i.e.*, that for any two polynomials $f(x)$ and $g(x)$, and any scalar λ, we have
   $$D(f + g) = D(f) + D(g) \quad \text{and} \quad D(\lambda f) = \lambda D(f).$$
   (b) There are "standard unit vectors" in P: $1, x, x^2, x^3, \ldots, x^n, \ldots$, and every polynomial can be written in terms of finitely many of these "standard unit

vectors." Describe the effect of D on a general one of these "standard unit vectors." Hence write down the start of an "infinite matrix" representing the action of D.

4. Recall that the **transpose** of a matrix is the matrix made by swapping its rows and columns (equivalently, reflecting in the leading diagonal). We say that an $n \times n$ matrix M is **orthogonal** if $MM^T = I_n = M^T M$—*i.e.*, if the transpose of M is equal to the inverse of M.

   (a) Let θ be an angle; show that the matrix representing the rotation of $\mathbb{R}^2$ by an angle θ anticlockwise about the origin is an orthogonal matrix.

   (b) Let M be a general 2×2 matrix; show that $|M| = |M^T|$.

   (c) Let M be a 2×2 orthogonal matrix; show that $|M| = \pm 1$. You may assume for this question that $|AB| = |A| \times |B|$ for any $n \times n$ matrices A and B; this is true, but hard to prove.

   (d) Let M and N be orthogonal matrices. Show that MN is also orthogonal (hint: $(MN)^T = N^T M^T$).

   (e) Let M be an orthogonal, 2×2 matrix and let N be the matrix representing reflection in the y-axis. Show that MN is orthogonal and $|MN| = -|M|$.

   By part (c), every orthogonal matrix has determinant ± 1. If M is orthogonal and $|M| = +1$, we call M **special orthogonal**. It can be shown that the special orthogonal matrices are precisely the rotation matrices! Moreover, part (e) shows that if M is orthogonal but not special orthogonal (*i.e.*, $|M| = -1$), then multiplying M by reflection in the y-axis makes M special orthogonal (any other reflection would also work for this). It follows that any orthogonal matrix is either special orthogonal (and so is a rotation matrix) or else represents reflection followed by a rotation. A reflection followed by a rotation is actually itself a reflection; so any orthogonal matrix is either a rotation or a reflection!

   This remains true in higher dimensions; the special orthogonal matrices are rotations, and the non-special orthogonal matrices are reflections, and can each be written as a single, fixed reflection, followed by a rotation. Once you see dot products, you will be able to prove that orthogonal matrices are precisely the linear transformations which preserve all lengths and angles. Putting that together with what we've looked at here: a linear transformation which preserves lengths and angles is either a rotation (if its determinant is $+1$) or a reflection (if its determinant is -1), and any reflection can be written as one fixed reflection followed by a rotation.