6. Linear transformations

Consider the function

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \qquad \text{which sends} \qquad (x,y) \longrightarrow (-y,x)$$

This is an example of a linear transformation. Before we get into the definition of a linear transformation, let's investigate the properties of this map.

What happens to the point $(1,0)$? It gets sent to $(0,1)$. What about $(2,0)$? It gets sent to $(0,2)$. In fact any point on the x-axis gets sent to a point on the y-axis. How about points on the y-axis? $(0,1)$ gets sent to $(-1,0)$. $(0,2)$ gets sent to $(-2,0)$ and so on. Points on the y-axis are sent to points on the x-axis.

What happens to the point $(1,1)$? It gets sent to $(-1,1)$. We guess that this function rotates the plane through an angle of $\pi/2$ anticlockwise. The key thing is that this map is represented by a matrix. Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

The first column is the image of the vector $(1,0)$ and the second column is the image of the vector $(0,1)$. To apply A to (x,y) simply multiply

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix},$$

as required.

What about the map

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \qquad \text{which sends} \qquad (x,y) \longrightarrow (x,-y)$$

This map fixes $(1,0)$ and in fact it fixes the whole x-axis. It sends $(0,1)$ to $(0,-1)$. It sends the y-axis to the y-axis but it flips it upside-down. $(1,1)$ gets sent to $(1,-1)$. This map represents reflection in the x-axis. The corresponding matrix is

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We check that this is correct:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix},$$

as required.

How about the matrix

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}?$$

It sends (x, y) to

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}.$$

This sends $(1, 0)$ to $(-1, 0)$ and $(0, 1)$ to $(0, -1)$. This represents rotation through an angle of π or, equivalently reflection in the origin.

The map $f(x, y) = (2x, 2y)$ represents a dilation by a factor of 2 and the map $g(x, y) = (3x, 3y)$ represents a dilation by a factor of 3. The corresponding matrices are

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}.$$

If we take the matrix

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

then the corresponding map is $f(x, y) = (2x, y)$ this magnifies distances in the x-direction by a factor of 2 and leaves heights unchanged.

How about the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}?$$

It sends (x, y) to

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ y \end{pmatrix}.$$

This sends $(1, 0)$ to $(1, 0)$ and $(0, 1)$ to $(1, 1)$. But it sends $(1, 1)$ to $(2, 1)$, $(2, 1)$ to $(3, 1)$ etc. The higher up you go, the more you move horizontally. This map is called a shear.

What about the function

$$f \colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \qquad \text{which sends} \qquad (x, y, z) \longrightarrow (x, y)?$$

This represents projection onto the xy-plane. We just forget the height. This sends $(1, 0, 0)$ to $(1, 0)$, $(0, 1, 0)$ to $(0, 1)$ and $(0, 0, 1)$ to $(0, 0)$. The corresponding matrix is

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

We check that this is correct:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix},$$

as required.

Let

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

The corresponding function is from $\mathbb{R}^3$ to $\mathbb{R}^3$, $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$. It sends $(1,0,0)$ to $(1,0,0)$, it sends $(0,1,0)$ to $(0,1,0)$ and it sends $(0,0,1)$ to $(0,0,1)$. Since A fixes $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$ it fixes everything. Let's check using matrix multiplication:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

so that $f(x, y, z) = (x, y, z)$ is the identity function.

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is called the identity matrix.

The key property about linear transformations is one just needs to know what happens to $(1,0)$ and $(0,1)$, or more generally what happens to any collection of vectors which spans.

Definition 6.1. *A function $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is called (a) linear (transformation) if*

(1) It is additive: $f(\vec{v} + \vec{w}) = f(\vec{v}) + f(\vec{w})$ for all vectors $\vec{v}$ and $\vec{w} \in \mathbb{R}^n$.
(2) $f(\lambda\vec{v}) = \lambda f(\vec{v})$, for all scalars λ and vectors $\vec{v} \in \mathbb{R}^n$.

The second condition often turns up in engineering as: "double the input, double the output".

Proposition 6.2. *If $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is given by matrix multiplication, $f(\vec{v}) = A\vec{v}$, where A an $m \times n$ matrix, then f is linear.*

Proof.

$$f(\vec{v} + \vec{w}) = A(\vec{v} + \vec{w}) = A\vec{v} + A\vec{w} = f(\vec{v}) + f(\vec{w})$$

and

$$f(\lambda\vec{v}) = A(\lambda\vec{v}) = \lambda A(\vec{v}) = \lambda f(\vec{v}). \qquad \square$$

We have seen that a linear transformation is determined by its action on $(1,0)$ and $(0,1)$, etc.

Definition 6.3. *Let*

$$\vec{e}_1 = (1, 0, \dots, 0), \qquad \vec{e}_2 = (0, 1, \dots, 0) \qquad \textit{and} \qquad \vec{e}_n = (0, 0, \dots, 1)$$

be the n unit coordinate vectors in $\mathbb{R}^n$.

In $\mathbb{R}^2$ we have $\vec{e}_1 = (1, 0)$ and $\vec{e}_2 = (0, 1)$ and in $\mathbb{R}^3$ we have $\vec{e}_1 = (1, 0, 0)$, $\vec{e}_2 = (0, 1, 0)$ and $\vec{e}_3 = (0, 0, 1)$. Note that $\vec{e}_i$ has a 1 in the ith position and zeroes everywhere else.

The key point is that:

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n.$$

For example

$$(x, y) = (x, 0) + (0, y) = x(1, 0) + y(0, 1) = x\vec{e}_1 + y\vec{e}_2.$$

Similarly

$$(x, y, z) = (x, 0, 0)+(0, y, 0)+(0, 0, z) = x(1, 0, 0)+y(0, 1, 0)+z(0, 0, 1) = x\vec{e}_1+y\vec{e}_2+z\vec{e}_3.$$

Theorem 6.4. *If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear then there is an $m \times n$ matrix A such that $f(\vec{v}) = A\vec{v}$.*

Proof. Let $\vec{v}_i = f(\vec{e}_i) \in \mathbb{R}^m$. Let A be the matrix whose columns are the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$. Then

$$\begin{aligned} f(\vec{x}) &= f(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n) \\ &= f(x_1\vec{e}_1) + f(x_2\vec{e}_2) + \cdots + f(x_n\vec{e}_n) \\ &= x_1 f(\vec{e}_1) + x_2 f(\vec{e}_2) + \cdots + x_n f(\vec{e}_n) \\ &= x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n \\ &= A\vec{x}. \end{aligned}$$

□

Consider the matrix equation

$$A\vec{x} = \vec{b}.$$

Suppose that no matter the choice of $\vec{b}$ we can always solve this equation, that is, this equation is always consistent. In terms of functions, given any $\vec{b} \in \mathbb{R}^m$ we may find $\vec{x} \in \mathbb{R}^n$ such that $f(\vec{x}) = \vec{b}$.

Definition 6.5. *We say that $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is onto if given any $\vec{b} \in \mathbb{R}^m$ we can find $\vec{x} \in \mathbb{R}^n$ such that $f(\vec{x}) = \vec{b}$.*

All of the functions $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ above are onto. Projection from $\mathbb{R}^3 \longrightarrow \mathbb{R}^2$ is onto. The identity is onto. The map

$$\mathbb{R}^2 \longrightarrow \mathbb{R}^2 \qquad \text{given by} \qquad (x, y) \longrightarrow (x, 0)$$

is not onto. The vector $(0, 1)$ is not in the image, that is, we cannot find a vector mapping to $(0, 1)$.

There is another natural question we can ask about matrix equations:

$$A\vec{x} = \vec{b}.$$

Suppose that no matter the choice of $\vec{b}$ there are never infinitely many solutions. This is the same as saying there are never two or more solutions. In terms of functions there are never two vectors $\vec{x}_1$ and $\vec{x}_2$ such that $f(\vec{x}_1) = f(\vec{x}_2) = \vec{b}$.

Definition 6.6. *We say that $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is one to one if given any $\vec{b} \in \mathbb{R}^m$ there is at most one vector $\vec{x}$ such that $f(\vec{x}) = \vec{b}$.*

All of the functions $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ above are one to one. The function

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \qquad \text{given by} \qquad (x, y) \longrightarrow (x, y, 0)$$

is one to one. Projection from $\mathbb{R}^3 \longrightarrow \mathbb{R}^2$ is not one to one. The vector $(0, 0, 1)$ and the vector $(0, 0, 2)$ are both sent to $(0, 0)$. In fact

Theorem 6.7. *Let A be an $n \times m$ matrix. The function*

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m \qquad \text{given by} \qquad \vec{x} \longrightarrow A\vec{x}$$

is one to one if and only if the homogeneous equation

$$A\vec{x} = \vec{0}$$

has only the trivial solution $\vec{x} = \vec{0}$.

Proof. f is one to one if and only if

$$A\vec{x} = \vec{b}$$

has at most one solution, for any $\vec{b}$.

If f is one to one then every equation has at most one solution and so the homogeneous has only the trivial solution.

Now suppose that the homogeneous has only the trivial solution. Then the solution to any consistent equation

$$A\vec{x} = \vec{b}$$

is of the form a particular solution plus any solution to the homogeneous. As the homogeneous has only one solution there is only one solution to $A\vec{x} = \vec{b}$. □