Chapter 2

Linear Transformations

The operations which define a vector space are vector addition and scalar multiplication. A nice function between vector spaces is one which preserves these operations, in the following sense.

Definition. For positive integers n and m, a function $L : \mathbb{R}^n \to \mathbb{R}^m$ is ***linear*** if

- $L(\boldsymbol{x}_1 + \boldsymbol{x}_2) = L(\boldsymbol{x}_1) + L(\boldsymbol{x}_2)$ for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$.
- $L(c\boldsymbol{x}) = cL(\boldsymbol{x})$ for all $c \in \mathbb{R}$, $\boldsymbol{v} \in \mathbb{R}^n$. □

Example. The function $L : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ x+y \\ x+3y \end{pmatrix}$$

is linear. ☑

We can define what it means for a function $L : V \to W$ to be linear for any vector spaces V and W, but we will restrict our attention to $\mathbb{R}^n$ and $\mathbb{R}^m$.

Note that if L is linear, then $L(\mathbf{0}) = \mathbf{0}$.

Example. The function $L : \mathbb{R}^3 \to \mathbb{R}^2$ by

$$L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ z+1 \end{pmatrix}$$

is *not* linear. Note that $L\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix} \neq \begin{pmatrix}0\\0\end{pmatrix}$. ☑

Example. For any positive integers n and m, the function $L : \mathbb{R}^n \to \mathbb{R}^m$ by $L(\boldsymbol{v}) = \mathbf{0}$ is linear. ☑

To help us analyze linear functions, we will work with the *standard basis*.

Definition. The ***standard basis*** for $\mathbb{R}^n$ is $\mathcal{S} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$, where

$$\boldsymbol{e}_1 = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}$$

$$\boldsymbol{e}_2 = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}$$

$$\boldsymbol{e}_3 = \begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix}$$

$$\vdots$$

$$\boldsymbol{e}_n = \begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}$$

□

Example. In $\mathbb{R}^3$, the standard basis is

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$\boxtimes$

Note that if $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$, then we can write

$$\begin{aligned} \boldsymbol{x} &= \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \\ &= x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3 \end{aligned}$$

The standard basis is a *basis* because any element of $\mathbb{R}^n$ can be built up from the elements; for any $\boldsymbol{x} \in \mathbb{R}^n$:

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots x_n\boldsymbol{e}_n.$$

Now, suppose that $L : \mathbb{R}^3 \to \mathbb{R}$ is linear. Let $a = L(\boldsymbol{e}_1)$, $b = L(\boldsymbol{e}_2)$ and $c = L(\boldsymbol{e}_3)$. Then, for any $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$, we have

$$\begin{aligned} L(\boldsymbol{x}) &= L(x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3) \\ &= L(x\boldsymbol{e}_1) + L(y\boldsymbol{e}_2) + L(z\boldsymbol{e}_3) \\ &= xL(\boldsymbol{e}_1) + yL(\boldsymbol{e}_2) + zL(\boldsymbol{e}_3) \\ &= ax + by + cz \end{aligned}$$

We will introduce some notation to rewrite this result in a nicer form.

Let $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ be a list of n numbers; this is known as a ***row vector***. Given a row vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ and a column vector $\boldsymbol{b} =$

$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, we define the ***product ab*** to be

$$\boldsymbol{ab} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Example.

$$(2,5,3)\begin{pmatrix} 7 \\ 9 \\ 2 \end{pmatrix} = 2 \cdot 7 + 5 \cdot 9 + 3 \cdot 2 = 65.$$

☑

What we showed before is that for a linear function $L : \mathbb{R}^3 \to \mathbb{R}$, if $\boldsymbol{a} = (L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), L(\boldsymbol{e}_3))$ then $L(\boldsymbol{x}) = \boldsymbol{ax}$ for any $\boldsymbol{x} \in \mathbb{R}^3$.

Proposition. *Let $L : \mathbb{R}^n \to \mathbb{R}$ be linear, and let*

$$\boldsymbol{a} = (L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), \ldots, L(\boldsymbol{e}_n)).$$

Then

$$L(\boldsymbol{x}) = \boldsymbol{ax}$$

for any $\boldsymbol{x} \in \mathbb{R}^n$.

Conversely, any function $L : \mathbb{R}^n \to \mathbb{R}$ of the form

$$L(\boldsymbol{x}) = \boldsymbol{ax}$$

for some row vector $\boldsymbol{a}$ is linear. □

Example.

- $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2x - 5y + 7z$ is linear. Here $\boldsymbol{a} = (2, -5, 7)$.
- $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2x + 3y + 2z + 4$ is not linear.
- $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3x + 4yz$ is not linear.

☑

There is a similar characterization of linear functions $L : \mathbb{R}^n \to \mathbb{R}^m$. We will again consider $A = (L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), \ldots, L(\boldsymbol{e}_n))$, but now each $L(\boldsymbol{e}_i)$ is a column vector. A will be an array of numbers.

Example. Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ by $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+3y-z \\ x+5z \end{pmatrix}$. Then $L(\boldsymbol{e}_1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $L(\boldsymbol{e}_2) = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ and $L(\boldsymbol{e}_3) = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$. We get

$$A = (L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), \ldots, L(\boldsymbol{e}_n)) = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & 5 \end{pmatrix}$$

☑

Definition. An $m \times n$ ***matrix*** is an array of numbers with m rows and n columns. □

Example.

$$\begin{pmatrix} 2 & 4 & -3 & 4 \\ 0 & 1 & 3 & 9 \\ 2 & 5 & 1 & 1 \end{pmatrix}$$

is a 3×4 matrix. ☑

If A is a matrix, we write a_{ij} to represent the number in the ith row and the jth column.

Example. If

$$A = \begin{pmatrix} 2 & 4 & 9 \\ 1 & -3 & 8 \end{pmatrix},$$

then $a_{2,1} = 1$. ☑

If $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear and $A = (L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), \ldots, L(\boldsymbol{e}_n))$, then A is an $m \times n$ matrix. For $\boldsymbol{x} \in \mathbb{R}^n$ we want to define the product $A\boldsymbol{x}$ so that $L(\boldsymbol{x}) = A\boldsymbol{x}$.

Definition. Let A be an $m \times n$ matrix, we can write $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ for vectors $\boldsymbol{a}_i \in \mathbb{R}^m$. For $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, we define

$$A\boldsymbol{x} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n.$$

□

Example.

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & 5 & 9 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} = 3 \begin{pmatrix} 3 \\ 2 \end{pmatrix} + 5 \begin{pmatrix} 2 \\ 5 \end{pmatrix} + 7 \begin{pmatrix} 4 \\ 9 \end{pmatrix} \quad = \begin{pmatrix} 9 \\ 6 \end{pmatrix} + \begin{pmatrix} 10 \\ 25 \end{pmatrix} + \begin{pmatrix} 28 \\ 63 \end{pmatrix} = \begin{pmatrix} 47 \\ 94 \end{pmatrix} \quad \square$$

Notice that to multiply $A\boldsymbol{x}$, the number of columns of A must equal the number of elements in $\boldsymbol{x}$.

Proposition. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be linear, and let*

$$A = (L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), \ldots, L(\boldsymbol{e}_n)).$$

Then A is an $m \times n$ matrix, and

$$L(\boldsymbol{x}) = A\boldsymbol{x}$$

for any $\boldsymbol{x} \in \mathbb{R}^n$.

Conversely, any function $L : \mathbb{R}^n \to \mathbb{R}^m$ of the form

$$L(\boldsymbol{x}) = A\boldsymbol{x}$$

for some $m \times n$ matrix A is linear. $\square$

Example. The function

$$L : \mathbb{R}^3 \to \mathbb{R}^2$$

by

$$L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y \\ x + 2y + 3z \end{pmatrix}$$

is linear. We have

$$L \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{2.0}$$

$$L \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{2.0}$$

$$L \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \tag{2.0}$$

so letting

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

we get

$$L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

☑

Example. The function $L : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3x - y \\ 4y \end{pmatrix}$$

is linear. We have

$$A = \left(L \begin{pmatrix} 1 \\ 0 \end{pmatrix} L \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 4 \end{pmatrix}$$

and $L(\boldsymbol{x}) = A\boldsymbol{x}$. ☑

Example. The function

$$L : \mathbb{R}^3 \to \mathbb{R}^2$$

by

$$L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 4x + 5y + 6z \end{pmatrix}.$$

is linear. (Note the connection between the matrix and the resulting value of L.) ☑

Example. Suppose $L : \mathbb{R}^2 \to \mathbb{R}^3$ is linear, $L\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\\1\end{pmatrix}$ and $L\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}2\\1\\3\end{pmatrix}$. Then

$$L\begin{pmatrix}3\\1\end{pmatrix} = 3L\begin{pmatrix}1\\0\end{pmatrix} + 1L\begin{pmatrix}0\\1\end{pmatrix} \tag{2.0}$$

$$= 3\begin{pmatrix}1\\1\\1\end{pmatrix} + 1\begin{pmatrix}2\\1\\3\end{pmatrix} \tag{2.0}$$

$$= \begin{pmatrix}5\\4\\6\end{pmatrix}. \tag{2.0}$$

More generally, we have $L(\boldsymbol{x}) = A\boldsymbol{x}$, where

$$A = \left(L\begin{pmatrix}1\\0\end{pmatrix} \; L\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}1 & 2\\1 & 1\\1 & 3\end{pmatrix}$$

☑

We now see that $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear exactly when each component of $L(\boldsymbol{x})$ is of the form $a_1x_1 + \cdots + a_nx_n$ for constants $a_1, \ldots, a_n$.