#linear_algebra

The difference between a matrix equation $A\mathbf{x} = \mathbf{b}$ and the associated vector equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ is merely a matter of notation. However, a matrix equation $A\mathbf{x} = \mathbf{b}$ can arise in linear algebra (and in applications such as computer graphics and signal processing) in a way that is not directly connected with linear combinations of vectors. This happens when we think of the matrix A as an object that "acts" on a vector $\mathbf{x}$ by multiplication to produce a new vector called $A\mathbf{x}$.

For instance, the equations

$$\underset{\substack{\uparrow \\ A}}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{\substack{\uparrow \\ \mathbf{x}}}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}} = \underset{\substack{\uparrow \\ \mathbf{b}}}{\begin{bmatrix} 5 \\ 8 \end{bmatrix}} \quad \text{and} \quad \underset{\substack{\uparrow \\ A}}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}} = \underset{\substack{\uparrow \\ \mathbf{0}}}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

say that multiplication by A transforms $\mathbf{x}$ into $\mathbf{b}$ and transforms $\mathbf{u}$ into the zero vector. See Fig. 1.

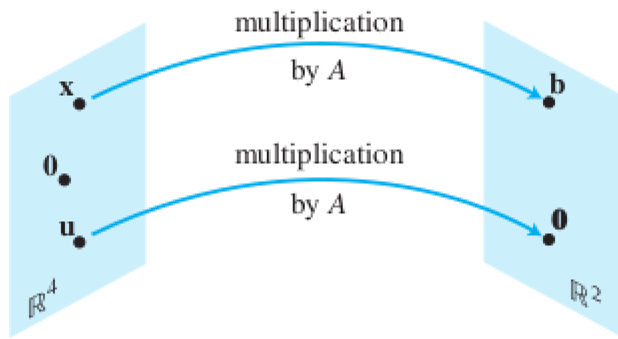

FIGURE 1 Transforming vectors via matrix multiplication.

From this new point of view, solving the equation $A\mathbf{x} = \mathbf{b}$ amounts to finding all vectors $\mathbf{x}$ in $\mathbb{R}^4$ that are transformed into the vector $\mathbf{b}$ in $\mathbb{R}^2$ under the "action" of multiplication by A.

The correspondence from $\mathbf{x}$ to $A\mathbf{x}$ is a *function* from one set of vectors to another. This concept generalizes the common notion of a function as a rule that transforms one real number into another.

A **transformation** (or **function** or **mapping**) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$. The set $\mathbb{R}^n$ is called the **domain** of T, and $\mathbb{R}^m$ is called the **codomain** of T. The notation $T : \mathbb{R}^n \to \mathbb{R}^m$ indicates that the domain of T is $\mathbb{R}^n$ and the codomain is $\mathbb{R}^m$. For $\mathbf{x}$ in $\mathbb{R}^n$, the vector $T(\mathbf{x})$ in $\mathbb{R}^m$ is called the **image** of $\mathbf{x}$ (under the action of T). The set of all images $T(\mathbf{x})$ is called the **range** of T. See Fig. 2.

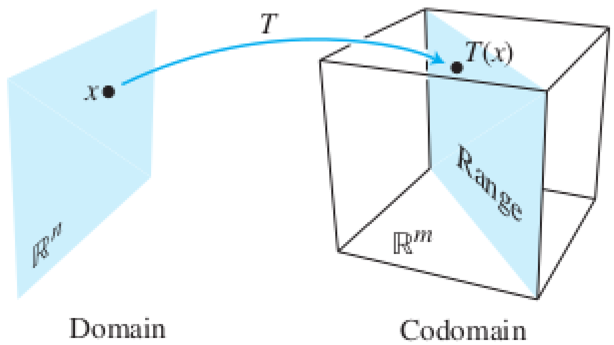

FIGURE 2 Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$.

Matrix transformation

The rest of this section focuses on mappings associated with matrix multiplication. For each $\mathbf{x}$ in $\mathbb{R}^n$, $T(\mathbf{x})$ is computed as $A\mathbf{x}$, where A is an $m \times n$ matrix. For simplicity, we sometimes denote such a *matrix transformation* by $\mathbf{x} \mapsto A\mathbf{x}$. Observe that the domain of T is $\mathbb{R}^n$ when A has n columns and the codomain of T is $\mathbb{R}^m$ when each column of A has m entries. The range of T is the set of all linear combinations of the columns of A, because each image $T(\mathbf{x})$ is of the form $A\mathbf{x}$.

EXAMPLE 2 If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ *projects* points in $\mathbb{R}^3$ onto the x_1x_2-plane because

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

See Fig. 3. ■

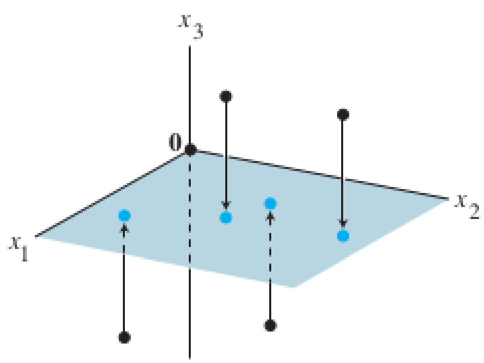

FIGURE 3
A projection transformation.

EXAMPLE 3 Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is called a **shear transformation**. It can be shown that if T acts on each point in the 2×2 square shown in Fig. 4, then the set of images forms the shaded parallelogram. The key idea is to show that T maps line segments onto line segments (as shown in Exercise 27) and then to check that the corners of the square map onto the vertices of the parallelogram. For instance, the image of the point $\mathbf{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ is $T(\mathbf{u}) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$, and the image of $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$. T deforms the square as if the top of the square were pushed to the right while the base is held fixed. Shear transformations appear in physics, geology, and crystallography. ■

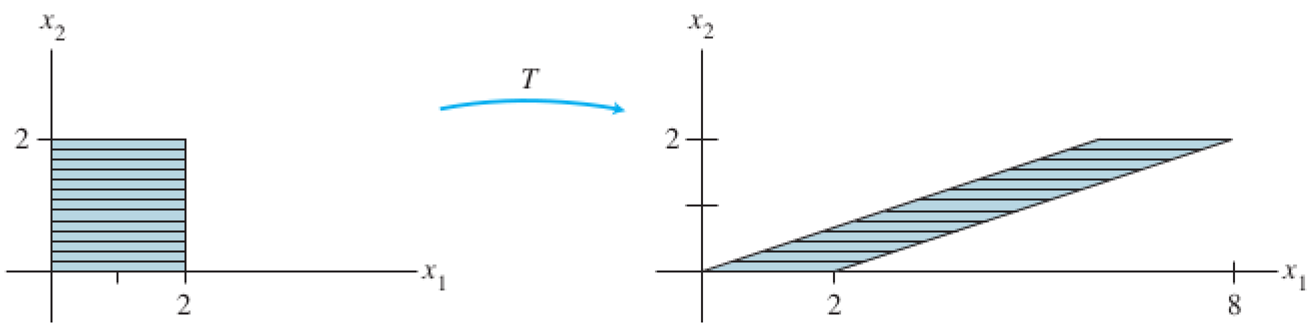

FIGURE 4 A shear transformation.

Linear transformations

A transformation (or mapping) T is **linear** if:

(i) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T;

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all $\mathbf{u}$ in the domain of T.

If T is a linear transformation, then

$$T(\mathbf{0}) = \mathbf{0} \tag{3}$$

and

$$T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v}) \tag{4}$$

for all vectors $\mathbf{u}, \mathbf{v}$ in the domain of T and all scalars c, d.

Property (3) follows from condition (ii) in the definition, because $T(\mathbf{0}) = T(0\mathbf{u}) = 0T(\mathbf{u}) = \mathbf{0}$. Property (4) requires both (i) and (ii):

$$T(c\mathbf{u} + d\mathbf{v}) = T(c\mathbf{u}) + T(d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$$

Observe that *if a transformation satisfies* (4) *for all* $\mathbf{u}, \mathbf{v}$ *and* c, d, *it must be linear*. (Set $c = d = 1$ for preservation of addition, and set $d = 0$ for preservation of scalar multiplication.) Repeated application of (4) produces a useful generalization:

$$T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p) \tag{5}$$

In engineering and physics, (5) is referred to as a *superposition principle*. Think of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ as signals that go into a system and $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)$ as the responses of that system to the signals. The system satisfies the superposition principle if whenever an input is expressed as a linear combination of such signals, the system's response is *the same* linear combination of the responses to the individual signals. We will return to this idea in Chapter 4.

EXAMPLE 4 Given a scalar r, define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{x}) = r\mathbf{x}$. T is called a **contraction** when $0 \le r \le 1$ and a **dilation** when $r > 1$. Let $r = 3$, and show that T is a linear transformation.

SOLUTION Let $\mathbf{u}, \mathbf{v}$ be in $\mathbb{R}^2$ and let c, d be scalars. Then

$$\begin{aligned} T(c\mathbf{u} + d\mathbf{v}) &= 3(c\mathbf{u} + d\mathbf{v}) && \text{Definition of } T \\ &= 3c\mathbf{u} + 3d\mathbf{v} && \\ &= c(3\mathbf{u}) + d(3\mathbf{v}) && \text{Vector arithmetic} \\ &= cT(\mathbf{u}) + dT(\mathbf{v}) \end{aligned}$$

Thus T is a linear transformation because it satisfies (4). See Fig. 5. ■

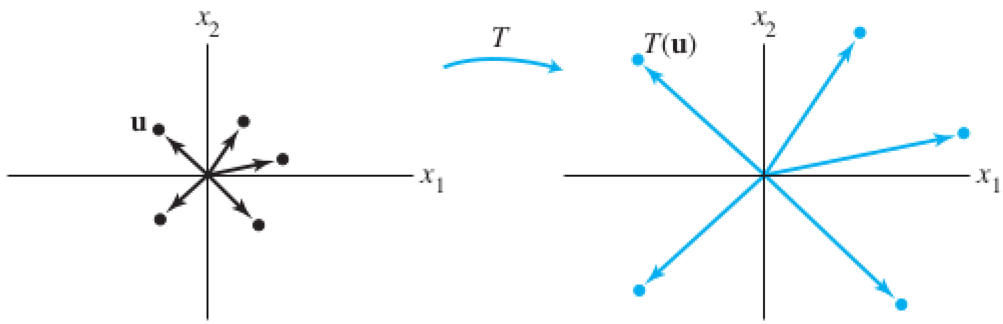

FIGURE 5 A dilation transformation.