1.8 Introduction to Linear Transformations

A transformation (or function or mapping) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector 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.

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 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 1. *Let* $A = \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & -4 \\ 3 & -5 & -9 \end{bmatrix}$ *and* $\mathbf{b} = \begin{bmatrix} 6 \\ -7 \\ -9 \end{bmatrix}$. *With T defined by $T(\mathbf{x}) = A\mathbf{x}$, find a vector $\mathbf{x}$ whose image under T is $\mathbf{b}$. Determine whether $\mathbf{x}$ is unique.*

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

$$T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \tag{1}$$

$$T(c\mathbf{v}) = cT(\mathbf{v}) \tag{2}$$

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

Linear transformations preserve the operations of vector addition and scalar multiplication. Property (1) says that the result $T(\mathbf{u}+\mathbf{v})$ of first adding $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and then applying T is the same as first applying T to $\mathbf{u}$ and $\mathbf{v}$ and then adding $T(\mathbf{u})$ and $T(\mathbf{v})$ in $\mathbb{R}^m$. These two properties lead to the following useful facts. If T is a linear transformation, then

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

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

for all vectors $\mathbf{u}, \mathbf{v}$ in the domain of T and all scalars c, d. Property (3) follows from (2) because

$$T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}.$$

Property (4) requires both (1) and (2):

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

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.

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 $r > 1$.

Example 2. *Use a rectangular coordinate system to plot* $\mathbf{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, *and their images underthe transformation* $T(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.