Introduction to Linear Transformations

DEFINITION: Let V and W be vector spaces, and let $f : V \to W$ be a function from V to W. (That is, for each vector $\mathbf{v} \in V$, $f(\mathbf{v})$ denotes exactly one vector of W.) Then f is a **linear transformation** if and only if both of the following are true:

(1) $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$, for all $\mathbf{v}_1, \mathbf{v}_2 \in V$

(2) $f(c\mathbf{v}) = cf(\mathbf{v})$, for all $c \in \mathbb{R}$ and all $\mathbf{v} \in V$.

EXAMPLE: The mapping $f : M_{mn} \to M_{nm}$ given by

$$f(A) = A^T$$

is a linear transformation, since

$$f(A_1 + A_2) = (A_1 + A_2)^T = A_1^T + A_2^T = f(A_1) + f(A_2)$$

and

$$f(cA) = (cA)^T = cA^T = cf(A)$$

EXAMPLE: The mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$T(\mathbf{x}) = A\mathbf{x}$$

where A is an $m \times n$ matrix, is a linear transformation, since

$$T(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y})$$

and

$$T(c\mathbf{x}) = A(c\mathbf{x}) = cA\mathbf{x} = cT(\mathbf{x})$$

EXAMPLE: The transformation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \longrightarrow \quad T(\mathbf{x}) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is linear, since

$$T(\mathbf{x}) = I_n\mathbf{x}$$

where I_n is the $n \times n$ identity matrix.

EXAMPLE: The transformation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \longrightarrow \quad T(\mathbf{x}) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 + x_2 - x_3 \\ x_1 \end{bmatrix}$$

is linear, since

$$T(\mathbf{x}) = A\mathbf{x} \quad \text{where} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{bmatrix}$$

EXAMPLE: The mapping $f : \mathbb{P}_n \to \mathbb{P}_{n-1}$ given by

$$f(p) = p'$$

is a linear transformation, since

$$f(p_1 + p_2) = (p_1 + p_2)' = p_1' + p_2' = f(p_1) + f(p_2)$$

and

$$f(cp) = (cp)' = cp' = cf(p)$$

EXAMPLE: Let V be a vector space of all differentiable functions and let

$$T(f) = f' + 2f + 1$$

Then T is not linear, since

$$\begin{aligned} T(f_1 + f_2) &= (f_1 + f_2)' + 2(f_1 + f_2) + 1 \\ &= f_1' + f_2' + 2f_1 + 2f_2 + 1 \\ &= (f_1' + 2f_1 + 1) + (f_2' + 2f_2) = T(f_1) + (f_2' + 2f_2) \neq T(f_1) + T(f_2) \end{aligned}$$

EXAMPLE: The mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{bmatrix}$$

is not a linear transformation, since

$$\begin{aligned} T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) + T\left(\begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}\right) &= \begin{bmatrix} 1^2 \\ 1^2 \\ 1^2 \end{bmatrix} + \begin{bmatrix} 2^2 \\ 2^2 \\ 2^2 \end{bmatrix} \\ &= \begin{bmatrix} 1^2 + 2^2 \\ 1^2 + 2^2 \\ 1^2 + 2^2 \end{bmatrix} \neq \begin{bmatrix} (1+2)^2 \\ (1+2)^2 \\ (1+2)^2 \end{bmatrix} = T\left(\begin{bmatrix} 1+2 \\ 1+2 \\ 1+2 \end{bmatrix}\right) = T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}\right) \end{aligned}$$

EXAMPLE: The transformation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \longrightarrow \quad T(\mathbf{x}) = \begin{bmatrix} 1 \\ x_1^2 + x_2^2 \end{bmatrix}$$

is *not* linear. Indeed, we have

$$T(2\mathbf{x}) = \begin{bmatrix} 1 \\ (2x_1)^2 + (2x_2)^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4x_1^2 + 4x_2^2 \end{bmatrix} \qquad \text{and} \qquad 2T(\mathbf{x}) = \begin{bmatrix} 2 \\ 2x_1^2 + 2x_2^2 \end{bmatrix}$$

Since

$$T(2\mathbf{x}) \neq 2T(\mathbf{x})$$

it follows that T is *not* linear.

Theorem 5.1 Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces, and let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Let $\mathbf{0}_{\mathcal{V}}$ be the zero vector in $\mathcal{V}$ and $\mathbf{0}_{\mathcal{W}}$ be the zero vector in $\mathcal{W}$. Then

(1) $L(\mathbf{0}_{\mathcal{V}}) = \mathbf{0}_{\mathcal{W}}$

(2) $L(-\mathbf{v}) = -L(\mathbf{v})$, for all $\mathbf{v} \in \mathcal{V}$

(3) $L(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) = a_1L(\mathbf{v}_1) + a_2L(\mathbf{v}_2) + \cdots + a_nL(\mathbf{v}_n)$, for all $a_1, \ldots, a_n \in \mathbb{R}$, and $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{V}$, for $n \geq 2$.

Proof.

Part (1):

$$\begin{aligned} L(\mathbf{0}_{\mathcal{V}}) &= L(0\mathbf{0}_{\mathcal{V}}) && \text{part (2) of Theorem 4.1, in } \mathcal{V} \\ &= 0L(\mathbf{0}_{\mathcal{V}}) && \text{property (2) of linear transformation} \\ &= \mathbf{0}_{\mathcal{W}} && \text{part (2) of Theorem 4.1, in } \mathcal{W} \end{aligned}$$

Part (2):

$$\begin{aligned} L(-\mathbf{v}) &= L(-1\mathbf{v}) && \text{part (3) of Theorem 4.1, in } \mathcal{V} \\ &= -1(L(\mathbf{v})) && \text{property (2) of linear transformation} \\ &= -L(\mathbf{v}) && \text{part (3) of Theorem 4.1, in } \mathcal{W} \end{aligned}$$

Part (3): (Abridged) This part is proved by induction. We prove the Base Step $(n = 2)$ here and leave the Inductive Step as Exercise 29. For the Base Step, we must show that $L(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) = a_1L(\mathbf{v}_1) + a_2L(\mathbf{v}_2)$. But,

$$\begin{aligned} L(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) &= L(a_1\mathbf{v}_1) + L(a_2\mathbf{v}_2) && \text{property (1) of linear transformation} \\ &= a_1L(\mathbf{v}_1) + a_2L(\mathbf{v}_2) && \text{property (2) of linear transformation.} \end{aligned}$$ □

Theorem 5.2 Let $\mathcal{V}_1, \mathcal{V}_2$, and $\mathcal{V}_3$ be vector spaces. Let $L_1: \mathcal{V}_1 \to \mathcal{V}_2$ and $L_2: \mathcal{V}_2 \to \mathcal{V}_3$ be linear transformations. Then $L_2 \circ L_1: \mathcal{V}_1 \to \mathcal{V}_3$ given by $(L_2 \circ L_1)(\mathbf{v}) = L_2(L_1(\mathbf{v}))$, for all $\mathbf{v} \in \mathcal{V}_1$, is a linear transformation.

Theorem 5.3 Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation.

(1) If $\mathcal{V}'$ is a subspace of $\mathcal{V}$, then $L(\mathcal{V}') = \{L(\mathbf{v}) \mid \mathbf{v} \in \mathcal{V}'\}$, the image of $\mathcal{V}'$ in $\mathcal{W}$, is a subspace of $\mathcal{W}$. In particular, the range of L is a subspace of $\mathcal{W}$.

(2) If $\mathcal{W}'$ is a subspace of $\mathcal{W}$, then $L^{-1}(\mathcal{W}') = \{\mathbf{v} \mid L(\mathbf{v}) \in \mathcal{W}'\}$, the pre-image of $\mathcal{W}'$ in $\mathcal{V}$, is a subspace of $\mathcal{V}$.

Proof. Part (1): Suppose that $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and that $\mathcal{V}'$ is a subspace of $\mathcal{V}$. Now, $L(\mathcal{V}')$, the image of $\mathcal{V}'$ in $\mathcal{W}$ (see Figure 5.5), is certainly nonempty (why?). Hence, to show that $L(\mathcal{V}')$ is a subspace of $\mathcal{W}$, we must prove that $L(\mathcal{V}')$ is closed under addition and scalar multiplication.

First, suppose that $\mathbf{w}_1, \mathbf{w}_2 \in L(\mathcal{V}')$. Then, by definition of $L(\mathcal{V}')$, we have $\mathbf{w}_1 = L(\mathbf{v}_1)$ and $\mathbf{w}_2 = L(\mathbf{v}_2)$, for some $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}'$. Then, $\mathbf{w}_1 + \mathbf{w}_2 = L(\mathbf{v}_1) + L(\mathbf{v}_2) = L(\mathbf{v}_1 + \mathbf{v}_2)$ because L is a linear transformation. However, since $\mathcal{V}'$ is a subspace of $\mathcal{V}$, $(\mathbf{v}_1 + \mathbf{v}_2) \in \mathcal{V}'$. Thus, $(\mathbf{w}_1 + \mathbf{w}_2)$ is the image of $(\mathbf{v}_1 + \mathbf{v}_2) \in \mathcal{V}'$, and so $(\mathbf{w}_1 + \mathbf{w}_2) \in L(\mathcal{V}')$. Hence, $L(\mathcal{V}')$ is closed under addition.

Next, suppose that $c \in \mathbb{R}$ and $\mathbf{w} \in L(\mathcal{V}')$. By definition of $L(\mathcal{V}')$, $\mathbf{w} = L(\mathbf{v})$, for some $\mathbf{v} \in \mathcal{V}'$. Then, $c\mathbf{w} = cL(\mathbf{v}) = L(c\mathbf{v})$ since L is a linear transformation. Now, $c\mathbf{v} \in \mathcal{V}'$, because $\mathcal{V}'$ is a subspace of $\mathcal{V}$. Thus, $c\mathbf{w}$ is the image of $c\mathbf{v} \in \mathcal{V}'$, and so $c\mathbf{w} \in L(\mathcal{V}')$. Hence, $L(\mathcal{V}')$ is closed under scalar multiplication. □

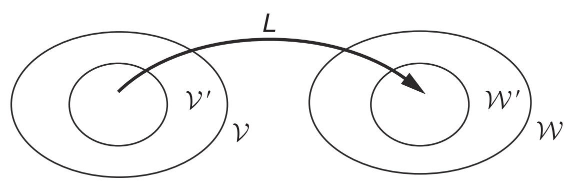

FIGURE 5.5

Subspaces of $\mathcal{V}$ correspond to subspaces of $\mathcal{W}$ under a linear transformation $L: \mathcal{V} \to \mathcal{W}$