5.1: Linear Transformations

Outcomes

A. Understand the definition of a linear transformation, and that all linear transformations are determined by matrix multiplication.

Recall that when we multiply an $m \times n$ matrix by an $n \times 1$ column vector, the result is an $m \times 1$ column vector. In this section we will discuss how, through matrix multiplication, an $m \times n$ matrix **transforms** an $n \times 1$ column vector into an $m \times 1$ column vector.

Recall that the $n \times 1$ vector given by

$$\vec{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \right]$$

is said to belong to $\mathbb{R}^n$, which is the set of all $n \times 1$ vectors. In this section, we will discuss transformations of vectors in $\mathbb{R}^n$.

Consider the following example.

Example 5.1.1:A Function Which Transforms Vectors

Consider the matrix $A = \left[\begin{array}{ccc} 1 & 2 & 0 \\ 2 & 1 & 0 \end{array} \right]$. Show that by matrix multiplication A transforms vectors in $\mathbb{R}^3$ into vectors in $\mathbb{R}^2$.

Solution

First, recall that vectors in $\mathbb{R}^3$ are vectors of size 3×1, while vectors in $\mathbb{R}^2$ are of size 2×1. If we multiply A, which is a 2×3 matrix, by a 3×1 vector, the result will be a 2×1 vector. This what we mean when we say that A *transforms* vectors.

Now, for $\left[\begin{array}{c} x \\ y \\ z \end{array} \right]$ in $\mathbb{R}^3$, multiply on the left by the given matrix to obtain the new vector. This product looks like

$$\left[\begin{array}{ccc} 1 & 2 & 0 \\ 2 & 1 & 0 \end{array} \right] \left[\begin{array}{c} x \\ y \\ z \end{array} \right] = \left[\begin{array}{c} x + 2y \\ 2x + y \end{array} \right]$$

The resulting product is a 2×1 vector which is determined by the choice of x and y. Here are some numerical examples.

$$\left[\begin{array}{ccc} 1 & 2 & 0 \\ 2 & 1 & 0 \end{array} \right] \left[\begin{array}{c} 1 \\ 2 \\ 3 \end{array} \right] = \left[\begin{array}{c} 5 \\ 4 \end{array} \right]$$

Here, the vector $\left[\begin{array}{c} 1 \\ 2 \\ 3 \end{array} \right]$ in $\mathbb{R}^3$ was transformed by the matrix into the vector $\left[\begin{array}{c} 5 \\ 4 \end{array} \right]$ in $\mathbb{R}^2$.

Here is another example:

$$\left[\begin{array}{ccc} 1 & 2 & 0 \\ 2 & 1 & 0 \end{array} \right] \left[\begin{array}{r} 10 \\ 5 \\ -3 \end{array} \right] = \left[\begin{array}{c} 20 \\ 25 \end{array} \right]$$

The idea is to define a function which takes vectors in $\mathbb{R}^3$ and delivers new vectors in $\mathbb{R}^2$. In this case, that function is multiplication by the matrix A.

Let T denote such a function. The notation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ means that the function T transforms vectors in $\mathbb{R}^n$ into vectors in $\mathbb{R}^m$. The notation $T(\vec{x})$ means the transformation T applied to the vector $\vec{x}$. The above example demonstrated a transformation achieved by matrix multiplication. In this case, we often write

$$T_A(\vec{x}) = A\vec{x}$$

Therefore, T_A is the transformation determined by the matrix A. In this case we say that T is a matrix transformation.

Recall the property of matrix multiplication that states that for k and p scalars,

$$A(kB + pC) = kAB + pAC$$

In particular, for A an $m \times n$ matrix and B and C, $n \times 1$ vectors in $\mathbb{R}^n$, this formula holds.

In other words, this means that matrix multiplication gives an example of a linear transformation, which we will now define.

Definition 5.1.1: Linear Transformation

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a function, where for each $\vec{x} \in \mathbb{R}^n, T(\vec{x}) \in \mathbb{R}^m$. Then T is a **linear transformation** if whenever k, p are scalars and $\vec{x}_1$ and $\vec{x}_2$ are vectors in $\mathbb{R}^n$ ($n \times 1$ vectors),

$$T(k\vec{x}_1 + p\vec{x}_2) = kT(\vec{x}_1) + pT(\vec{x}_2)$$

Consider the following example.

Example 5.1.2: Linear Transformation

Let T be a transformation defined by $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y \\ x - z \end{bmatrix} \text{ for all } \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$$

Show that T is a linear transformation.

Solution

By Definition 5.1.1 we need to show that $T(k\vec{x}_1 + p\vec{x}_2) = kT(\vec{x}_1) + pT(\vec{x}_2)$ for all scalars k, p and vectors $\vec{x}_1, \vec{x}_2$. Let

$$\vec{x}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$$

Then

$$\begin{aligned}
T(k\vec{x}_1 + p\vec{x}_2) &= T\left(k\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + p\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right) \\
&= T\left(\begin{bmatrix} kx_1 \\ ky_1 \\ kz_1 \end{bmatrix} + \begin{bmatrix} px_2 \\ py_2 \\ pz_2 \end{bmatrix}\right) \\
&= T\left(\begin{bmatrix} kx_1 + px_2 \\ ky_1 + py_2 \\ kz_1 + pz_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} (kx_1 + px_2) + (ky_1 + py_2) \\ (kx_1 + px_2) - (kz_1 + pz_2) \end{bmatrix} \\
&= \begin{bmatrix} (kx_1 + ky_1) + (px_2 + py_2) \\ (kx_1 - kz_1) + (px_2 - pz_2) \end{bmatrix} \\
&= \begin{bmatrix} kx_1 + ky_1 \\ kx_1 - kz_1 \end{bmatrix} + \begin{bmatrix} px_2 + py_2 \\ px_2 - pz_2 \end{bmatrix} \\
&= k\begin{bmatrix} x_1 + y_1 \\ x_1 - z_1 \end{bmatrix} + p\begin{bmatrix} x_2 + y_2 \\ x_2 - z_2 \end{bmatrix} \\
&= kT(\vec{x}_1) + pT(\vec{x}_2)
\end{aligned}$$

Therefore T is a linear transformation.

Two important examples of linear transformations are the zero transformation and identity transformation. The zero transformation defined by $T\left(\vec{x}\right) = \vec(0)$ for all $\vec{x}$ is an example of a linear transformation. Similarly the identity transformation defined by $T\left(\vec{x}\right) = \vec(x)$ is also linear. Take the time to prove these using the method demonstrated in Example 5.1.2.

We began this section by discussing matrix transformations, where multiplication by a matrix transforms vectors. These matrix transformations are in fact linear transformations.

Theorem 5.1.1:Matrix Transformations are Linear Transformations

Let $T:\mathbb{R}^n \mapsto \mathbb{R}^m$ be a transformation defined by $T(\vec{x}) = A\vec{x}$. Then T is a linear transformation.

It turns out that every linear transformation can be expressed as a matrix transformation, and thus linear transformations are exactly the same as matrix transformations.