Linear transformations

Andrew Taylor

April 16 2022

Definition 1. *A function T from $\mathbb{R}^m$ to $\mathbb{R}^n$ is called a linear transformation if there exists an $n \times m$ matrix A such that*

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

for all $\vec{x}$ in the vector space $\mathbb{R}^m$.

Theorem 1. *A transformation T from $\mathbb{R}^m$ to $\mathbb{R}^n$ is linear if and only if*

$$\begin{aligned} T(\vec{v} + \vec{w}) &= T(\vec{v}) + T(\vec{w}) \quad \textit{for all vectors } \vec{v}, \vec{w} \in \mathbb{R}^m \\ T(k\vec{v}) &= kT(\vec{v}) \quad \textit{for all vectors } \vec{v} \in \mathbb{R}^m \textit{ and for all scalars } k \in \mathbb{R} \end{aligned}$$

Proof. Let $T(\vec{x}) = A\vec{x}$ be a linear transformation. Then we know that

$$T(\vec{v} + \vec{w}) = A(\vec{v} + \vec{w}) = A\vec{v} + A\vec{w} = T(\vec{v}) + T(\vec{w})$$

from the distributive property for matrices. We also know that

$$T(k\vec{v}) = A(k\vec{v}) = k(A\vec{v}) = kT(\vec{v})$$

from the rules of matrix multiplication.

Let T be a transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ that satisfies the properties

$$\begin{aligned} T(\vec{v} + \vec{w}) &= T(\vec{v}) + T(\vec{w}) & \forall \vec{v}, \vec{w} \in \mathbb{R}^m \\ T(k\vec{v}) &= kT(\vec{v}) & \forall \vec{v} \in \mathbb{R}^m \text{ and } \forall k \in \mathbb{R} \end{aligned}$$

Let $e_1, e_2, ..., e_m$ be the unit vectors of $\mathbb{R}^m$. We can write

$$
\begin{aligned}
T(\vec{x}) &= T(x_1e_1 + x_2e_2 + ... + x_me_m) \\
&= T(x_1e_1) + T(x_2e_2) + ... + T(x_me_m) \\
&= x_1T(e_1) + x_2T(e_2) + ... + x_mT(e_m) \\
&= \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_m) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \\
&= \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_m) \end{bmatrix} \vec{x}
\end{aligned}
$$

Thus $T(\vec{x})$ is a linear transformation that has a matrix of transformation

$$
A = \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_m) \end{bmatrix}
$$

and can be written as $T(\vec{x}) = A\vec{x}$.

$\square$