MA1522: Linear Algebra for Computing

Chapter 7: Linear Transformation

7.1 Introduction to Linear Transformation

Geometric Interpretation of Matrix Multiplication

Given a $m \times n$ matrix $\mathbf{A}$, we can think of it as mapping vectors $\mathbf{v}$ from $\mathbb{R}^n$ to a vector $\mathbf{Av}$ in $\mathbb{R}^m$.

Example

1. Consider the matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. It maps $\mathbb{R}^2$ to the plane in $\mathbb{R}^3$ defined by $z = 0$, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$.

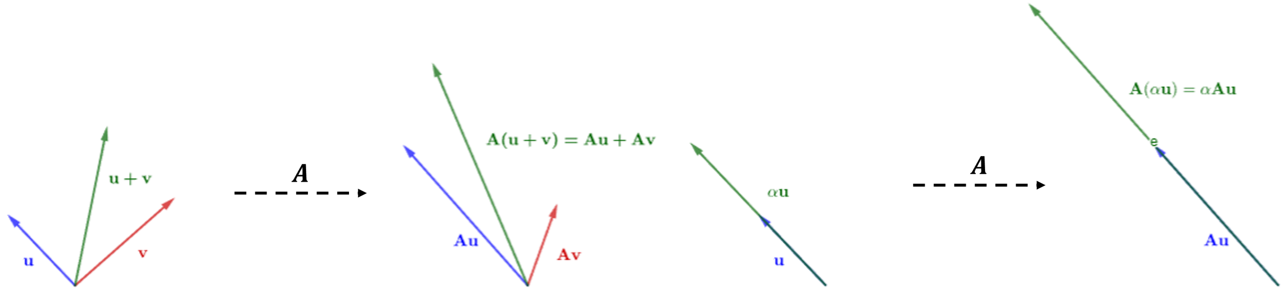
2. The matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ projects vectors in $\mathbb{R}^3$ onto the $z = 0$ plane and identifies it with $\mathbb{R}^2$, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$. That is, it can be interpreted as the map $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$.

3. The zero matrix $\mathbf{A} = \mathbf{0}_{m \times n}$ sends any vector in $\mathbb{R}^n$ to the zero vector $\mathbf{0}$ in $\mathbb{R}^m$,
$$\mathbf{Av} = \mathbf{0} \quad \text{for all } \mathbf{v} \text{ in } \mathbb{R}^n.$$

Geometric Interpretation of Matrix Multiplication

Recall that matrix multiplication commutes with scalar multiplication, and is distributive, for all vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ and scalar α,

$$\mathbf{A}(\alpha\mathbf{u}) = \alpha\mathbf{A}\mathbf{u}, \quad \text{and} \quad \mathbf{A}(\mathbf{u}+\mathbf{v}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{v}.$$



Or equivalently, matrix multiplication is linear, for all vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ and scalars α, β,

$$\mathbf{A}(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\mathbf{A}\mathbf{u} + \beta\mathbf{A}\mathbf{v}.$$

Geometrically, this means that the mapping of a linear combination is the linear combination of the mapping.

Linear Transformation

Definition

A mapping (function) $T : \mathbb{R}^n \to \mathbb{R}^m$, is a *linear transformation* if for all vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$, and scalars α, β,

$$T(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v}).$$

The Euclidean space $\mathbb{R}^n$ is called the *domain* of the mapping, and the Euclidean space $\mathbb{R}^m$ is called the *codomain* of the mapping.

Remarks

Equivalently, a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$, is a linear transformation if it satisfies the following properties.

(i) For any vector $\mathbf{u}$ in $\mathbb{R}^n$ and scalar α,

$$T(\alpha\mathbf{u}) = \alpha T(\mathbf{u}).$$

(ii) For any vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$,

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

By induction, we have that for any vectors $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ in $\mathbb{R}^n$ and scalars $c_1, c_2, ..., c_k$,

$$T(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) = c_1 T(\mathbf{u}_1) + c_2 T(\mathbf{u}_2) + \cdots + c_k T(\mathbf{u}_k).$$

The previous discussion shows that every matrix defines a linear transformation by multiplication,

$$\mathbf{A} \mapsto T_{\mathbf{A}}; \quad T_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\mathbf{u} \text{ for all } \mathbf{u} \text{ in } \mathbb{R}^n.$$

It will be shown later that this identification is one-to-one and onto, that is, every linear transformation is defined by multiplication of some matrix.

Example

1. $T : \mathbb{R}^2 \to \mathbb{R}^3,\ T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$

$$
\begin{aligned}
T\left(\alpha \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \beta \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) &= T\left(\begin{pmatrix} \alpha x_1 + \beta x_2 \\ \alpha y_1 + \beta y_2 \end{pmatrix}\right) \\
&= \begin{pmatrix} \alpha x_1 + \beta x_2 \\ \alpha y_1 + \beta y_2 \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} x_1 \\ y_1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} x_2 \\ y_2 \\ 0 \end{pmatrix} \\
&= \alpha T\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + \beta T\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right)
\end{aligned}
$$

Example

2. $T : \mathbb{R}^3 \to \mathbb{R}^2,\ T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix}.$

$$
\begin{aligned}
T\left(\alpha \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \beta \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right) &= T\left(\begin{pmatrix} \alpha x_1 + \beta x_2 \\ \alpha y_1 + \beta y_2 \\ \alpha z_1 + \beta z_2 \end{pmatrix}\right) \\
&= \begin{pmatrix} \alpha x_1 + \beta x_2 \\ \alpha y_1 + \beta y_2 \end{pmatrix} = \alpha \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \beta \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \\
&= \alpha T\left(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}\right) + \beta T\left(\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right)
\end{aligned}
$$

Example

3. $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{T}(x, y) = (2x - 3y, x, 5y)$.

$$
\begin{aligned}
\mathbf{T}(\alpha(x_1, y_1) + \beta(x_2, y_2)) &= \mathbf{T}(\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2)) \\
&= (2(\alpha x_1 + \beta x_2) - 3(\alpha y_1 + \beta y_2)), \alpha x_1 + \beta x_2, 5(\alpha y_1 + \beta y_2))) \\
&= (2\alpha x_1 - 3\alpha y_1, \alpha x_1, 5\alpha y_1) + (2\beta x_2 - 3\beta y_2, \beta x_2, 5\beta y_2) \\
&= \alpha(2x_1 - 3y_1, x_1, 5y_1) + \beta(2x_2 - 3y_2, x_2, 5y_2) \\
&= \alpha\mathbf{T}(x_1, y_1) + \beta\mathbf{T}(x_2, y_2).
\end{aligned}
$$

Question

Let $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & -1 & -1 \\ 1 & -1 & 0 & 2 \end{pmatrix}$. Then $\mathbf{A}$ defines a linear transformation defined T by matrix multiplication.

1. What are the domain and codomain of T?

2. Write down the formula of T.

Not a Linear Transformation

Observe that by linearity, a linear transformation must map the zero vector $\mathbf{0}_n$ in $\mathbb{R}^n$ to the zero vector $\mathbf{0}_m$ in $\mathbb{R}^m$, $T(\mathbf{0}_n)) = \mathbf{0}_m$. Hence, together with equivalent definition of linear transformation, we have the following.

A mapping $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ is not a linear transformation if any of the following statements hold.

(i) $\mathbf{T}$ does not map the zero vector to the zero vector, $\mathbf{T}(\mathbf{0}) \neq \mathbf{0}$.

(ii) There is a scalar α and a vector $\mathbf{u}$ in $\mathbb{R}^n$ such that $\mathbf{T}(\alpha\mathbf{u}) \neq \alpha\mathbf{T}(\mathbf{u})$.

(iii) There are vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ such that $\mathbf{T}(\mathbf{u} + \mathbf{v}) \neq \mathbf{T}(\mathbf{u}) + \mathbf{T}(\mathbf{v})$.

Examples

1. $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{T}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ is not a linear transformation since

$$\mathbf{T}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

2. $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}$, $\mathbf{T}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = xy$ is not a linear transformation since

$$\mathbf{T}\left(2\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \mathbf{T}\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) = (2)(2) = 4 \neq 2 = 2(1)(1) = 2\mathbf{T}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$$

Question

Is the mapping $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$,

$$\mathbf{T}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \sqrt[3]{(x^3 + y^3)} \\ 0 \end{pmatrix},$$

a linear transformation?

Challenge

Find a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$T(\alpha \mathbf{u}) = \alpha T(\mathbf{u})$$

for all scalar α and vector $\mathbf{u}$ in $\mathbb{R}^n$, but is not a linear transformation.

Standard Matrix

Theorem

A mapping $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if there is a unique $m \times n$ matrix $\mathbf{A}$ such that

$$T(\mathbf{u}) = \mathbf{A}\mathbf{u} \quad \text{for all vectors } \mathbf{u} \text{ in } \mathbb{R}^n.$$

The matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{pmatrix},$$

where $E = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$. That is, the i-th column of $\mathbf{A}$ is $T(\mathbf{e}_i)$, for $i = 1, ..., n$.

Proof.

We have shown that a $m \times n$ matrix $\mathbf{A}$ defines a linear transformation by matrix multiplication.

Conversely, suppose $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. For any $\mathbf{u} = (u_i)$ in $\mathbb{R}^n$, write $\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \cdots + u_n\mathbf{e}_n$. By linearity,

$$T(\mathbf{u}) = u_1 T(\mathbf{e}_1) + u_2 T(\mathbf{e}_2) + \cdots + u_n T(\mathbf{e}_n) = \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \mathbf{A}\mathbf{u}.$$

The uniqueness of $\mathbf{A}$ is left as an exercise.

Standard Matrix

Definition

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The unique $m \times n$ matrix $\mathbf{A}$ such that

$$T(\mathbf{u}) = \mathbf{A}\mathbf{u} \quad \text{for all } \mathbf{u} \text{ in } \mathbb{R}^n$$

is called the *standard matrix*, or *matrix representation* of T.

Example

1.

$$
\begin{aligned}
\mathbf{T}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= \begin{pmatrix} 2x-3y \\ x \\ 5y \end{pmatrix} = x\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} -3 \\ 0 \\ 5 \end{pmatrix} \\
&= \begin{pmatrix} 2 & -3 \\ 1 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}
$$

So,

$$
\mathbf{A_T} = \begin{pmatrix} 2 & -3 \\ 1 & 0 \\ 0 & 5 \end{pmatrix}
$$

Examples

2.

$$\mathbf{T}\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 - 3x_2 + x_3 - 5x_4 \\ 4x_1 + x_2 - 2x_3 + x_4 \\ 5x_1 - x_2 + 4x_3 \end{pmatrix}$$

$$= x_1 \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix} + x_2 \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix} + x_4 \begin{pmatrix} -5 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

So,

$$\mathbf{A}_T = \begin{pmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{pmatrix}.$$

Question

1. Is the mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$,

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\right) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

   a linear transformation? If it is, find its standard matrix.

2. Is $T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ for some constants $a_1, a_2, ..., a_n$ a linear transformation? If it is, find its standard matrix.

3. What is the standard matrix of the following linear transformation $\mathbf{T}\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$?

Representation of Linear Transformation with Respect to a Basis

Definition

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ be a basis for $\mathbb{R}^n$. The *representation of T with respect to basis S*, denoted as $[T]_S$, is defined to be the $m \times n$ matrix

$$[T]_S = \begin{pmatrix} T(\mathbf{u}_1) & T(\mathbf{u}_2) & \cdots & T(\mathbf{u}_n) \end{pmatrix}.$$

The standard matrix or matrix representation of T is the representation of T with respect to the standard matrix,

$$\mathbf{A} = [T]_E.$$

Representation of Linear Transformation with Respect to a Basis

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ be a basis for $\mathbb{R}^n$. Then for any vector $\mathbf{v}$ in $\mathbb{R}^n$,

$$T(\mathbf{v}) = [T]_S[\mathbf{v}]_S,$$

that is, the image $T(\mathbf{v})$ is the product of the representation of T with respect to basis S with the coordinates $\mathbf{v}$ with respect to basis S. Moreover, if $\mathbf{P}$ is the transition matrix from the standard basis E of $\mathbb{R}^n$ to basis S, then the standard matrix $\mathbf{A}$ of T is given by

$$\mathbf{A} = [T]_S\mathbf{P}.$$

This means that we are able to compute the standard matrix of T if we know the image of T on a basis of $\mathbb{R}^n$, and thus from $\mathbf{A}$, we are able to reconstruct the formula for T. In fact, this is a equivalence statement; that is, we can reconstruct the formula for T if and only if we have the image of T on a basis.

Representation of Linear Transformation with Respect to a Basis

Proof.

Given any vector $\mathbf{v}$ in $\mathbb{R}^n$, write $\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$. Then $[\mathbf{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ and by linearity,

$$T(\mathbf{v}) = c_1 T(\mathbf{u}_1) + c_2 T(\mathbf{u}_2) + \cdots + c_n T(\mathbf{u}_n) = \begin{pmatrix} T(\mathbf{u}_1) & T(\mathbf{u}_2) & \cdots & T(\mathbf{u}_n) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = [T]_S[\mathbf{v}]_S.$$

Next, by definition of $\mathbf{P}$, $\mathbf{Pv} = [v]_S$. Hence, for any vector $\mathbf{v}$ in $\mathbb{R}^n$,

$$\mathbf{Av} = T(\mathbf{v}) = [T]_S[\mathbf{v}]_S = [T]_S\mathbf{Pv}.$$

Since this is true for any $\mathbf{v}$ in $\mathbb{R}^n$, we have the identity $\mathbf{A} = [T]_S\mathbf{P}$. $\square$

Example

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation given by

$$T\left(\begin{pmatrix}1\\2\\1\end{pmatrix}\right) = \begin{pmatrix}2\\6\\6\end{pmatrix}, \quad T\left(\begin{pmatrix}1\\1\\0\end{pmatrix}\right) = \begin{pmatrix}4\\8\\2\end{pmatrix}, \quad T\left(\begin{pmatrix}0\\2\\0\end{pmatrix}\right) = \begin{pmatrix}6\\6\\6\end{pmatrix}.$$

$S = \left\{\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\2\\0\end{pmatrix}\right\}$ is a basis for $\mathbb{R}^3$, so the representation of T with respect to S is

$$[T]_S = \begin{pmatrix}2 & 4 & 6\\6 & 8 & 6\\6 & 2 & 6\end{pmatrix}.$$

Example

The transition matrix $\mathbf{P}$ from the standard matrix E to S is

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 1 & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ -1/2 & 1/2 & -1/2 \end{pmatrix}.$$

Thus,

$$\mathbf{A} = [T]_S\mathbf{P} = \begin{pmatrix} 2 & 4 & 6 \\ 6 & 8 & 6 \\ 6 & 2 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ -1/2 & 1/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -5 \\ 5 & 3 & -5 \\ -1 & 3 & 1 \end{pmatrix}.$$

Hence, $T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x + 3y - 5z \\ 5x + 3y - 5z \\ -x + 3y + z \end{pmatrix}.$

Example

Let $T : \mathbb{R}^3 \to \mathbb{R}^4$ be a linear transformation such that

$$T\left(\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 3 \\ 2 \\ 2 \end{pmatrix}, \quad T\left(\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 6 \\ 2 \\ 2 \end{pmatrix}, \quad T\left(\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 8 \\ 1 \\ 0 \end{pmatrix}.$$

$S = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^3$, so the representation of T with respect to S is

$$[T]_S = \begin{pmatrix} -1 & -1 & -1 \\ 3 & 6 & 8 \\ 2 & 2 & 1 \\ 2 & 2 & 0 \end{pmatrix}.$$

Example

Now the transition matrix from the standard basis E to S is

$$\mathbf{P} = \begin{pmatrix} 2 & 3 & 3 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -3 & -3 \\ -2 & 4 & 3 \\ 1 & -2 & -1 \end{pmatrix}.$$

Thus,

$$\mathbf{A} = [T]_S \mathbf{P} = \begin{pmatrix} -1 & -1 & -1 \\ 3 & 6 & 8 \\ 2 & 2 & 1 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 2 & -3 & -3 \\ -2 & 4 & 3 \\ 1 & -2 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 2 & 0 \end{pmatrix}.$$

Hence, $T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} -x + y + z \\ 2x - y + z \\ x - z \\ 2y \end{pmatrix}.$

Example

Suppose it is given that $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation such that

$$T\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right) = \begin{pmatrix}1\\2\\1\end{pmatrix}, \quad T\left(\begin{pmatrix}0\\1\\0\end{pmatrix}\right) = \begin{pmatrix}0\\1\\1\end{pmatrix}, \quad T\left(\begin{pmatrix}1\\1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\3\\2\end{pmatrix}.$$

Is it possible to find the formula for T?

We do not have enough information to reconstruct the formula for T as $S = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right\}$ is linearly dependent; it is not a basis for $\mathbb{R}^3$. For one can check that for any real numbers a, b, c

$$T\left(\begin{pmatrix}x\\y\\z\end{pmatrix}\right) = \begin{pmatrix}1 & 0 & a\\2 & 1 & b\\1 & 1 & c\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x + az\\2x + y + bz\\x + y + cz\end{pmatrix}$$

satisfies the given conditions above.

7.2 Range and Kernel of Linear Transformation

Range of Linear Transformation

Definition

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The *range* of T is

$$R(T) = T(\mathbb{R}^n) = \{ \ \mathbf{v} \in \mathbb{R}^m \mid \mathbf{v} = T(\mathbf{u}) \text{ for some } \mathbf{u} \in \mathbb{R}^n \ \}.$$

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The range of T is a subspace.

Let $\mathbf{A}$ be the standard matrix of T. Recall that $T(\mathbf{u}) = \mathbf{A}\mathbf{u}$ for all $\mathbf{u}$ in $\mathbb{R}^n$. Hence,

$$R(T) = \{ \ \mathbf{v} = T(\mathbf{u}) \mid \mathbf{u} \in \mathbb{R}^n \ \} = \{ \ \mathbf{v} = \mathbf{A}\mathbf{u} \mid \mathbf{u} \in \mathbb{R}^n \ \} = \mathrm{Col}(\mathbf{A}).$$

That is, the range of T is the column space of its standard matrix, and therefore is a subspace of the codomain $\mathbb{R}^m$. The abstract proof can be found in the appendix.

Range of Linear Transformation

Definition

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The *rank* of T is the dimension of the range of T

$$\operatorname{rank}(T) = \dim(\mathrm{R}(T)).$$

Let $\mathbf{A}$ be the standard matrix of T. Since the range of T is the column space of $\mathbf{A}$, $\mathrm{R}(T) = \mathrm{Col}(\mathbf{A})$, therefore

$$\operatorname{rank}(T) = \dim(\mathrm{R}(T)) = \dim(\mathrm{Col}(\mathbf{A})) = \operatorname{rank}(\mathbf{A}).$$

Kernel of Linear Transformation

Definition

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The set of all vectors in $\mathbb{R}^n$ that maps to the zero vector $\mathbf{0}$ by T is called the *kernel* of T, and is denoted as

$$\ker(T) = \{ \ \mathbf{u} \in \mathbb{R}^n \ \big| \ T(\mathbf{u}) = \mathbf{0} \ \}.$$

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The kernel of T is a subspace.

Let $\mathbf{A}$ be the standard matrix of T. Then

$$\ker(T) = \{ \ \mathbf{u} \in \mathbb{R}^n \ \big| \ \mathbf{A}\mathbf{u} = T(\mathbf{u}) = \mathbf{0} \ \} = \mathrm{Null}(\mathbf{A}).$$

That is, the kernel of T is the nullspace of its standard matrix, and is thus a subspace.

Kernel of Linear Transformation

Definition

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The *nullity* of T is the dimension of the kernel of T,

$$\text{nullity}(T) = \dim(\ker(T)).$$

Let $\mathbf{A}$ be the standard matrix of T. Then

$$\text{nullity}(T) = \dim(\ker(T)) = \dim(\text{Null}(\mathbf{A})) = \text{nullity}(\mathbf{A}).$$

Examples

1. $T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix}$. The standard matrix is $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. So rank$(T) = 2$, nullity$(T) = 1$.

2. $T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 - 3x_2 + x_3 - 5x_4 \\ 4x_1 + x_2 - 2x_3 + x_4 \\ 5x_1 - x_2 + 4x_3 \end{pmatrix}$. The standard matrix is

$$\mathbf{A} = \begin{pmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & 1/73 \\ 0 & 1 & 0 & 133/73 \\ 0 & 0 & 1 & 32/73 \end{pmatrix}.$$

So rank$(T) = 3$, nullity$(T) = 1$.

Question

What are the rank and nullity of the following linear transformation?

1. $T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

2. $T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

Injectivity of Linear Transformation

Definition

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is *injective*, or *one-to-one* if for every vector $\mathbf{v}$ in the range of T, $\mathbf{v} \in \mathrm{R}(T)$, there is a unique $\mathbf{u}$ in $\mathbb{R}^n$ such that $T(\mathbf{u}) = \mathbf{v}$.

Alternatively, T is injective if whenever $T(\mathbf{u}_1) = T(\mathbf{u}_2)$, then $\mathbf{u}_1 = \mathbf{u}_2$.

Theorem

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is injective if and only if the kernel is trivial, $\ker(T) = \{\mathbf{0}\}$.

Let $\mathbf{A}$ be the standard matrix of T. Then recall that since the general solution to the consistent system $\mathbf{Ax} = \mathbf{v}$ is a particular solution plus the general solution to the homogeneous system $\mathbf{Ax} = \mathbf{0}$, $\mathbf{Ax} = \mathbf{v}$ has a unique solution if and only if $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution. Hence, T is injective if and only if $\mathbf{Ax} = \mathbf{v}$ has a unique solution for every $\mathbf{v}$ in $\mathrm{R}(T) = \mathrm{Col}(\mathbf{A})$, if and only if $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution, or $\ker(T) = \mathrm{Null}(\mathbf{A}) = \{\mathbf{0}\}$. The abstract proof is given in the appendix. We will add this to the equivalent statements for full rank matrices.

Full Rank Equals Number of Columns

Theorem

Suppose $\mathbf{A}$ *is a* $m \times n$ *matrix. The following statements are equivalent.*

(i) $\mathbf{A}$ *is full rank, where the rank is equal to the number of columns,* $\text{rank}(\mathbf{A}) = n$.

(ii) *The rows of* $\mathbf{A}$ *spans* $\mathbb{R}^n$, $Row(\mathbf{A}) = \mathbb{R}^n$.

(iii) *The columns of* $\mathbf{A}$ *are linearly independent.*

(iv) *The homogeneous system* $\mathbf{Ax} = \mathbf{0}$ *has only the trivial solution, that is,* $\text{Null}(\mathbf{A}) = \{\mathbf{0}\}$.

(v) $\mathbf{A}^T\mathbf{A}$ *is an invertible matrix of order* n.

(vi) $\mathbf{A}$ *has a left inverse.*

(vii) *The linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$ *defined by* $\mathbf{A}$ *is injective.*

Example

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 3x + y \\ 5x + 7y \\ x + 3y \end{pmatrix}$.

For any $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$, $T(\mathbf{x}) = \mathbf{v}$ if and only if

$$\begin{pmatrix} 3x + y \\ 5x + 7y \\ x + 3y \end{pmatrix} = \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad \text{or} \quad \left\{ \begin{array}{ccccc} 3x & + & y & = & v_1 \\ 5x & + & 7y & = & v_2 \\ x & + & 3y & = & v_3 \end{array} \right.$$

$$\left(\begin{array}{cc|c} 3 & 1 & v_1 \\ 5 & 7 & v_2 \\ 1 & 3 & v_3 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_3} \xrightarrow[R_3 - 3R_1]{R_2 - 5R_1} \xrightarrow{R_3 - R_2} \xrightarrow{-\frac{1}{8}R_2} \xrightarrow{R_1 - 3R_2} \left(\begin{array}{cc|c} 1 & 0 & (3v_2 - 7v_3)/8 \\ 0 & 1 & (5v_3 - v_2)/8 \\ 0 & 0 & v_1 - v_2 + 2v_3 \end{array}\right)$$

tells us that $\mathbf{v}$ is in the range of T if and only if $v_1 - v_2 + 2v_3 = 0$. In this case, $T(\mathbf{x}) = \mathbf{v}$ has only a unique solution

$$x = \frac{3v_2 - 7v_3}{8}, \quad y = \frac{5v_3 - v_2}{8};$$

that is, T is injective. Let $v_1 = v_2 = v_3 = 0$, we conclude that the kernel of T is trivial $\ker(T) = \{\mathbf{0}\}$.

Exercise

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Show that if T is injective, then necessary $n \leq m$.

Surjectivity of Linear Transformation

Definition

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is *surjective* or *onto* if for every $\mathbf{v}$ in the codomain $\mathbb{R}^m$, there exists a $\mathbf{u}$ in the domain $\mathbb{R}^n$ such that $T(\mathbf{u}) = \mathbf{v}$.

Alternatively, T is surjective if the range is the codmain, $\mathrm{R}(T) = \mathbb{R}^m$, which is equivalent to $\mathrm{rank}(T) = m$. This means that if $\mathbf{A}$ is the standard matrix of T, then $\mathbf{A}$ is full rank, where the rank is equal to its number of rows. We will add this to the equivalent statements for full rank matrices.

Full Rank Equals Number of Rows

Theorem

Suppose $\mathbf{A}$ *is a* $m \times n$ *matrix. The following statements are equivalent.*

(i) $\mathbf{A}$ *is full rank, where the rank is equal to the number of rows,* $\text{rank}(\mathbf{A}) = m$.

(ii) *The columns of* $\mathbf{A}$ *spans* $\mathbb{R}^m$, $Col(\mathbf{A}) = \mathbb{R}^m$.

(iii) *The rows of* $\mathbf{A}$ *are linearly independent.*

(iv) *The linear system* $\mathbf{Ax} = \mathbf{b}$ *is consistent for every* $\mathbf{b} \in \mathbb{R}^m$.

(v) $\mathbf{AA}^T$ *is an invertible matrix of order* m.

(vi) $\mathbf{A}$ *has a right inverse.*

(vii) *The linear transformation* T *defined by* $\mathbf{A}$ *is surjective.*

Exercise

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Show that if T is surjective, then necessary $n \geq m$.

Example

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation $T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+y+z \\ x+3y \\ y+z \end{pmatrix}$.

$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ if and only if

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ -1/3 & 1/3 & 1/3 \\ 1/3 & -1/3 & 2/3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 - v_3 \\ (-v_1 + v_2 + v_3)/3 \\ (v_1 - v_2 + 2v_3)/3 \end{pmatrix}.$$

In this case, T is both injective and surjective, which follows from the fact that the standard matrix $\mathbf{A}$ is invertible.

Equivalent Statements of Invertibility

Theorem (Equivalent Statements for Invertibility)

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

(i) $\mathbf{A}$ *is invertible.*

(ii) $\mathbf{A}^T$ *is invertible.*

(iii) *(left inverse) There is a matrix* $\mathbf{B}$ *such that* $\mathbf{BA} = \mathbf{I}$.

(iv) *(right inverse) There is a matrix* $\mathbf{B}$ *such that* $\mathbf{AB} = \mathbf{I}$.

(v) *The reduced row-echelon form of* $\mathbf{A}$ *is the identity matrix.*

(vi) $\mathbf{A}$ *can be expressed as a product of elementary matrices.*

(vii) *The homogeneous system* $\mathbf{Ax} = \mathbf{0}$ *has only the trivial solution.*

(viii) *For any* $\mathbf{b}$, *the system* $\mathbf{Ax} = \mathbf{b}$ *has a unique solution.*

(ix) *The determinant of* $\mathbf{A}$ *is nonzero,* $\det(\mathbf{A}) \neq 0$.

(x) *The columns/rows of* $\mathbf{A}$ *are linearly independent.*

(xi) *The columns/rows of* $\mathbf{A}$ *spans* $\mathbb{R}^n$.

(xii) $\text{rank}(\mathbf{A}) = n$ *(*$\mathbf{A}$ *has full rank).*

(xiii) $\text{nullity}(\mathbf{A}) = 0$.

(xiv) 0 *is not an eigenvalue of* $\mathbf{A}$.

(xv) *The linear transformation* T *defined by* $\mathbf{A}$ *is injective.*

(xvi) *The linear transformation* T *defined by* $\mathbf{A}$ *is surjective.*

Exercise

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is bijective if it is both injective and surjective.

Show that $T : \mathbb{R}^n \to \mathbb{R}^n$ is bijective if and only if there is a linear transformation $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{and} \quad S(T(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

Appendix

Range of Linear Transformation is a Subspace

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The range of T is a subspace.

Proof.

(i) Since $T(\mathbf{0}) = \mathbf{0}$, the range of T contains the zero vector, $\mathbf{0} \in \mathrm{R}(T)$.

(ii) Suppose now $\mathbf{v}_1, \mathbf{v}_2$ are in the range of T. This means that there are some $\mathbf{u}_1, \mathbf{u}_2$ in $\mathbb{R}^n$ such that

$$T(\mathbf{u}_1) = \mathbf{v}_1 \quad \text{and} \quad T(\mathbf{u}_2) = \mathbf{v}_2.$$

Therefore, for any scalars α, β,

$$T(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) = \alpha T(\mathbf{u}_1) + \beta T(\mathbf{u}_2) = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2,$$

which shows that $\alpha\mathbf{v}_1 + \beta\mathbf{v}_2$ is also in the range of T.

$\square$

Kernel of Linear Transformation is a Subspace

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The kernel of T is a subspace.

Proof.

(i) Since $T(\mathbf{0}) = \mathbf{0}$, it is clear that the zero vector is in the kernel of T, $\mathbf{0} \in \ker(T)$.

(ii) Suppose now $\mathbf{u}_1, \mathbf{u}_2$ are in the kernel of T, $T(\mathbf{u}_i) = \mathbf{0}$ for $i = 1, 2$. Then for any scalars α, β,

$$T(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) = \alpha T(\mathbf{u}_1) + \beta T(\mathbf{u}_2) = \alpha\mathbf{0} + \beta\mathbf{0} = \mathbf{0},$$

which shows that $\alpha\mathbf{u}_1 + \beta\mathbf{u}_2$ is in the kernel of T too.

$\square$

Injectivity of Linear Transformation

Theorem

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is injective if and only if the kernel is trivial, $\ker(T) = \{\mathbf{0}\}$.

Proof.

Suppose T is injective. Then for any $\mathbf{u}$ in the kernel of T, $T(\mathbf{u}) = \mathbf{0} = T(\mathbf{0})$, which shows that $\mathbf{u} = \mathbf{0}$ by the injectivity of T.

Conversely, suppose $\ker(T) = \{\mathbf{0}\}$. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ are such that $T(\mathbf{u}_1) = T(\mathbf{u}_2)$. Then by linearity,

$$\mathbf{0} = T(\mathbf{u}_1) - T(\mathbf{u}_2) = T(\mathbf{u}_1 - \mathbf{u}_2),$$

which shows that $\mathbf{u}_1 - \mathbf{u}_2$ is in the kernel of T. Since the kernel is the zero space, necessarily $\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{0}$, or $\mathbf{u}_1 = \mathbf{u}_2$. $\square$