1 : Theory of Matrices, Vector spaces and Linear Transformations

IT5506 – Mathematics for Computing II

Level III - Semester 5

Intended Learning Outcomes

At the end of this lesson, you will be able to;

- understand the definition and the properties of a linear transformation in the context of vector spaces.
- specify a linear transformation by considering its action on a basis.
- find the matrix of a linear transformation with respect to general bases in finite dimensional vector spaces.
- use properties of linear transformations to solve problems.
- find the composite of transformations and the inverse of a transformation if it exists.

List of sub topics

1.7 Linear transformations (4 hours)

1.7.1 Linear transformations in finite dimensional spaces with examples.

1.7.2 The matrix representation of a linear transformation.

1.7.3 The rank-nullity theorem and its applications.

1.7.4 Ordered bases, the matrix of a linear transformation and similarity of matrices.

1.7.1 Linear Transformations with Examples

Definition: Let X and Y be two vector spaces over the same field F. A function $T: X \longrightarrow Y$ is called a **Linear Transformation** or **Linear function** or **linear mapping** if it satisfies the following conditions:

- $\mathrm{T}(x+y) = \mathrm{T}x + \mathrm{T}y$ for all $x, y \in X$ (T preserves addition), and
- $\mathrm{T}(\alpha x) = \alpha \mathrm{T}x$ for all $\alpha \in \mathrm{F}, x \in X$ (T preserves scalar multiplication).

Note: When $\mathrm{X} = \mathrm{Y}$, $T: \mathrm{X} \longrightarrow \mathrm{X}$ is called a **Linear Operator.**

Example 1.7.1.1: Determine whether a function $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^2 \\ x+y \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ is linear.

Solution: $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = T\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$, and $T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

Since $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \neq T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + T\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, T is not linear.

An Example of Linear Transformation

Example 1.7.1.2: Determine whether a function $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ defined by $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ x+y+z \end{bmatrix}$, for all $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ is linear.

Solution: Let $u = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$, $v = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$.

$$T(u+v) = T\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right) = T\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix} = \begin{bmatrix} (x_1+x_2)+(y_1+y_2) \\ (x_1+x_2)+(y_1+y_2)+(z_1+z_2) \end{bmatrix}.$$

$$Tu + Tv = T\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + T\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_1 + y_1 + z_1 \end{bmatrix} + \begin{bmatrix} x_2 + y_2 \\ x_2 + y_2 + z_2 \end{bmatrix}$$

$$= \begin{bmatrix} (x_1+x_2)+(y_1+y_2) \\ (x_1+x_2)+(y_1+y_2)+(z_1+z_2) \end{bmatrix}.$$

Hence, $T(u+v) = Tu + Tv$.

Similarly, we can show that $T(\alpha u) = \alpha Tu$.

Therefore, T is a linear transformation.

Basic Properties of Linear Transformations

Theorem 1.7.1.1: Let X and Y be two vector spaces over the same field F, and let $\mathbf{0}_X$, and $\mathbf{0}_Y$ be the zero vectors of X and Y respectively. let $T: X \longrightarrow Y$ be a Linear Transformation. Then

1. $\mathrm{T}(\mathbf{0}_X) = \mathbf{0}_Y$
2. $T(-v) = -Tv$ for all $v \in X$.
3. $\mathrm{T}(\alpha \mathrm{x} + \beta y) = \alpha \mathrm{T}(\mathrm{x}) + \beta T(y)$ for all $\alpha, \beta \in \mathrm{F}$, for all $x, y \in Y$.

Proof:

1. Since T is linear, $\mathrm{T}(\mathbf{0}_X) = \mathrm{T}(\mathbf{0}_X + \mathbf{0}_X) = \mathrm{T}(\mathbf{0}_X) + \mathrm{T}(\mathbf{0}_X)$.

Hence, by the additive identity law in Y, $\mathrm{T}(\mathbf{0}_X) = \mathbf{0}_Y$.

2. Since T is linear, $\mathrm{T}(\alpha v) = \alpha \mathrm{T}(\mathrm{v})$. By taking $\alpha = -1$, we get $T(-v) = -Tv$.

3. $\mathrm{T}(\alpha \mathrm{x} + \beta y) = \mathrm{T}(\alpha \mathrm{x}) + \mathrm{T}(\beta y)$ ($\because$ T is linear)

$$= \alpha T(x) + \beta T(y) \quad (\because \text{T is linear}).$$

Examples of some Trivial Linear Transformations

Example 1.7.1.3 (Zero Transformation): Let X and Y be two vector spaces over the same field F. The zero operator $\mathrm{Z}: X \longrightarrow Y$ defined by $\mathrm{Z}x = \mathbf{0}_Y$, for all $x \in X$ is a linear transformation, where $\mathbf{0}_Y$ is the zero vector of Y.

Solution: Let Let $x, y \in X$, and $\alpha \in \mathbb{R}$.

$Z(x + y) = \mathbf{0}_Y = \mathbf{0}_Y + \mathbf{0}_Y = Zx + Zy$, and

$Z(\alpha x) = \mathbf{0}_Y = \alpha\, \mathbf{0}_Y = \alpha\, Zx$.

Hence, Z is linear.

Example 1.7.1.4 (Identity Operator): Let X be a vector spaces over the field F. The identity operator $\mathrm{I}: X \longrightarrow X$ defined by $\boldsymbol{I}x = x$, for all $x \in X$, is a linear operator.

Solution: Let Let $x, y \in X$, and $\alpha \in \mathbb{R}$.

$I(x + y) = x + y = Ix + Iy$, and

$I(\alpha x) = \alpha x = \alpha\, Ix$.

Hence, I is linear.

Dot Product

Example 1.7.1.4: Define $T: \mathbb{R}^3 \longrightarrow \mathbb{R}$ by $T\boldsymbol{x} = \boldsymbol{x} \cdot \boldsymbol{a}$, for all $\mathrm{x} \in \mathbb{R}^3$, where $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ is a fixed vector in $\mathbb{R}^3$. Show that T is linear.

Solution: Dot product of two vectors x and a is defined as

$$\boldsymbol{x} \cdot \boldsymbol{a} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^{\boldsymbol{T}} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = a_1 x_1 + a_2 x_2 + a_3 x_3.$$

Let $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$.

$$T(x+y) = (x+y) \cdot \boldsymbol{a} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}^{\boldsymbol{T}} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^{\boldsymbol{T}} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}^{\boldsymbol{T}} \right) \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \boldsymbol{x} \cdot \boldsymbol{a} + y \cdot \boldsymbol{a}.$$

That is; $T(x+y) = Tx + Ty$.

Similarly, we can show that $T(\alpha x) = \alpha\, Tx$.

Therefore, T is linear.

Derivative Operator

Example 1.7.1.5: Let $P_n(\mathbb{R})$ be the vector space of real polynomials of degree at most n. The function $D: P_n(\mathbb{R}) \longrightarrow P_{n-1}(\mathbb{R})$ is defined by

$$D\big(p(x)\big) = p'(x) = \frac{d}{dx}\, p(x), \qquad \text{for all } p(x) \in P_n(\mathbb{R}),$$

where $p'(x)$ denotes the derivative of the polynomial $p(x)$. The function D is called the derivative operator. Show that D is linear.
(For example, $D(2x^3 - 3x^2 + 4x - 5 = 6x^2 - 6x + 4)$

Solution: Let $p(x), q(x) \in P_n(\mathbb{R})$ and $\alpha \in \mathbb{R}$.

$D(p(x) + q(x)) = \frac{d}{dx}\big(p(x) + q(x)\big) = \frac{d}{dx}\, p(x) + \frac{d}{dx}\, q(x) = D\big(p(x)\big) + D\big(q(x)\big)$.

$D\big(\alpha p(x)\big) = \frac{d}{dx}\big(\alpha p(x)\big) = \alpha\, \frac{d}{dx}\, p(x) = \alpha\, D\big(p(x)\big)$.

Hence, D is linear.

Matrix Transformations are Linear Transformationsns

Theorem 1.7.1.2: Let A be an $m \times n$- real matrix. Consider the vector function $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined by $T(v) = Av$, for all $v \in \mathbb{R}^n$. Then T is a linear transformation from the vector space $\mathbb{R}^n$ to the vector space $\mathbb{R}^m$.

Proof: Let $x, y \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$.

$T(x+y) = A(x+y)$ ($\because$ Definition of T)

$= Ax + Ay$ ($\because$ matrix multiplication is distributive over matrix addition)

$= Tx + Ty$.

$T(\alpha x) = A(\alpha x)$

$= \alpha Ax$

$= \alpha Tx$.

Hence, T is a linear transformation.

Thus, every $m \times n$- real matrix induces a linear transformation from the vector space $\mathbb{R}^n$ to the vector space $\mathbb{R}^m$.

Matrix Transformations are Linear Transformationsns

Example 1.7.1.6: Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}$. Then A induces a linear transformation

$T_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ defined by $T_A(x) = Ax$, for all $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$.

That is; $T_A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 \\ x_2 - 2x_3 \end{bmatrix}$.

Activity

Define a linear transformation from $\mathbb{R}$ to $\mathbb{R}$ and define a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}$.

1.7.2 The matrix representation of a linear transformation

Theorem 1.7.2.1: Let $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation from the vector space $\mathbb{R}^n$ to the vector space $\mathbb{R}^m$. Then there exists an $m \times n$- real matrix ($A \in \mathbb{R}^{m\times n}$) such that $T(x) = A$x, for all $x \in \mathbb{R}^n$. In other words, T is a matrix transformation.

Proof: . Suppose $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation and consider the standard ordered basis $\{e_1, e_2, \cdots, e_n\}$ of $\mathbb{R}^n$.

For all i $(1 \le i \le n)$, define $u_i = T(e_i)$, and let A be the matrix that has $u_1, u_2, \cdots, u_n$ as its columns. We claim that A is the desired matrix such that $T(x) = A$x, for all $x \in \mathbb{R}^n$.

Let $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ be arbitrary. Then $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$ and we have

$$T(x) = T(x_1e_1 + x_2e_2 + \cdots + x_ne_n)$$

$$= x_1\, T(e_1) + x_2\, T(e_2) + \cdots + x_n\, T(e_n) \text{ (}\because\text{ T is linear)}$$

$$= x_1u_1 + x_2u_2 + \cdots + x_nu_\mathrm{n} \quad \text{(}\because\text{ definition of } u_\mathrm{i}\text{)}$$

$$= Ax \text{ (}\because\text{ by the column method of matrix multiplication)}$$

Note that, the matrix corresponding to the linear transformation T has as its columns the vectors $T(e_1), T(e_2), \cdots, T(e_n)$.

Linear transformations are matrix transformations

Example 1.7.2.1: Let $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be the linear transformation defined by

$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 + 2x_2 - x_3 \end{bmatrix}$, for all $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$. Find the matrix A of this linear transformation with respect to the standard ordered basis.

Solution: First, we compute the images of the standard basis vectors:

$T(e_1) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $T(e_2) = T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and $T(e_3) = T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$.

The matrix A has $T(e_1), T(e_2)$, and $T(e_3)$ as its columns. Therefore, the matrix A representing the transformation T is $A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix}$.

Rotation by 90^0 in $\mathbb{R}^2$

Example 1.7.2.2: Consider the linear transformation $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ that is rotating every vector in $\mathbb{R}^2$ through 90^0 degrees in the counterclockwise direction. Find the matrix A corresponding to this linear transformation. Find a formula for T and find the image of the vector $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

Solution: First, we compute the images of the standard ordered basis vectors:

Under this rotation, the image of $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. That is; $T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Under this rotation, the image of $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$. That is; $T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

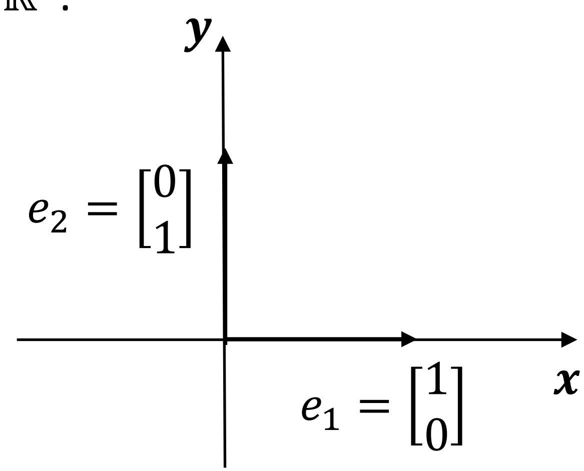
Therefore, the matrix of T is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Hence, the formula for T is $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$.

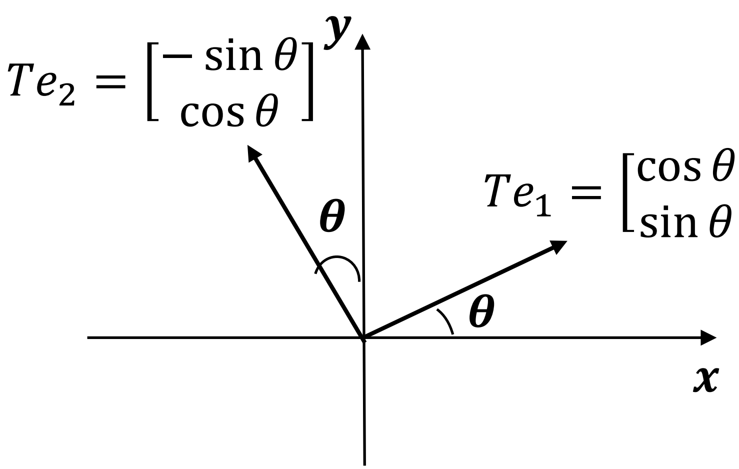
The image of $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is $T\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$.

Rotation by $\boldsymbol{\theta}^{\mathbf{0}}$ in $\mathbb{R}^2$

Example 1.7.2.3: Find the matrix A for a counterclockwise rotation by angle $\boldsymbol{\theta}$ in $\mathbb{R}^2$.



Before the rotation



After the rotation

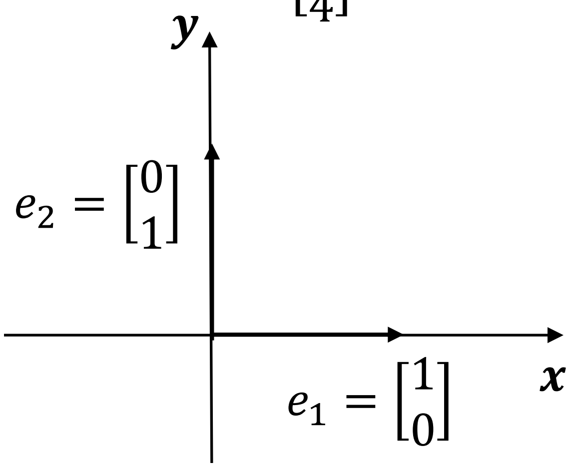
This transformation is rotating every vector in $\mathbb{R}^2$ through $\boldsymbol{\theta}^{\mathbf{0}}$ in the counterclockwise direction:

The matrix A for the rotation operator T is $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

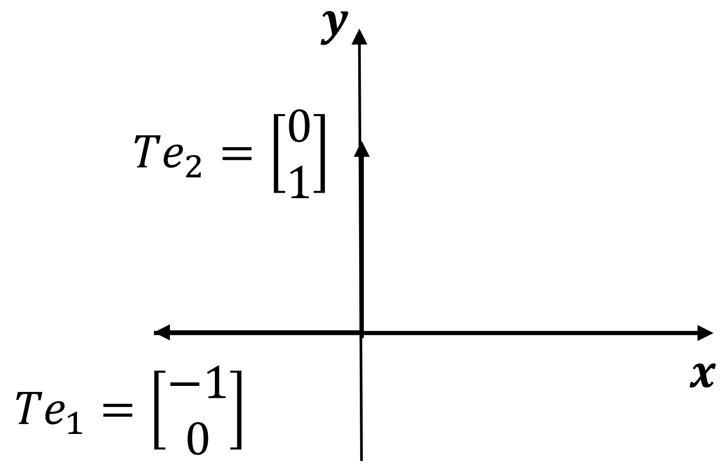
Hence, the formula for T is $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}$.

Reflection about the y-axis in in $\mathbb{R}^2$

Example 1.7.2.4: Let $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a reflection operator about the $y-$axis (transforming every vector in $\mathbb{R}^2$ to its image about the $y-$axis). Find the matrix A corresponding to this linear operator T, and a formula for T. Also find the image of the vector $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$.



Before the reflection



After the reflection

The matrix A for the reflection operator about the $y-$axis T is $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

Hence, the formula for T is $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}$.

$T\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$.

1.7.3 The Rank-nullity Theorem

Subspaces associated with a linear Transformation:

Definition (Null space): Let $\boldsymbol{T}: X \longrightarrow Y$ be a linear Transformation, where X and Y are two vector spaces over the same field F. Let $\mathbf{0}_X$, and $\mathbf{0}_Y$ be the zero vectors of X and Y respectively. The **null space** (or **kernel**) of T is denoted by $Null(T)$ or $N(T)$ and defined as

$$N(T) = \left\{x \in X \,\middle|\, T(x) = \mathbf{0}_Y\right\}.$$

Definition (Range): Let $\boldsymbol{T}: X \longrightarrow Y$ be a linear Transformation, where X and Y are two vector spaces over the same field F. The **range** of T is denoted by $\mathcal{R}(T)$ and defined as

$$\mathcal{R}(T) = \{T(x) \mid x \in X\}.$$

Theorem 1.7.3.1: Let $\mathrm{T}: X \longrightarrow Y$ be a linear Transformation, where X and Y are two vector spaces over the same field F. Then $\mathrm{N(T)}$, and $\mathcal{R}(\mathrm{T})$ are subspaces of X and Y respectively.

Proof of Theorem 1.7.3.1

Proof: Since $T(\mathbf{0}_X) = \mathbf{0}_Y$, we have $\mathbf{0}_X \in N(T)$. Hence $N(T)$ is non-empty.
Suppose that $x_1, x_2 \in N(T)$, and $\alpha \in \mathrm{F}$. Then
$T(x_1 + x_2) = T(x_1) + T(x_2)$ ($\because$ T is linear)

$$= \mathbf{0}_Y + \mathbf{0}_Y = \mathbf{0}_Y \quad (\because x_1, x_2 \in N(T)).$$

Hence, $x_1 + x_2 \in N(T)$.

$T(\alpha x_1) = \alpha\, T(x_1)$ ($\because$ T is linear)

$$= \alpha \mathbf{0}_Y = \mathbf{0}_Y \ (\because x_1 \in N(T)).$$

Hence, $\alpha x_1 \in N(T)$. Therefore $N(T)$ is a subspace of X.

Since $T(\mathbf{0}_X) = \mathbf{0}_Y$, we have $\mathbf{0}_Y \in \mathcal{R}(T)$. Hence $\mathcal{R}(T)$ is non-empty.

Suppose that $y_1, y_2 \in \mathcal{R}(T)$, and $\alpha \in \mathrm{F}$. Then, there exist $x_1, x_2 \in X$ such that $T(x_1) = y_1$ and $T(x_2) = y_2$.
Now, $y_1 + y_2 = T(x_1) + T(x_2) = T(x_1 + x_2)$ ($\because$ T is linear).

Hence, $y_1 + y_2 \in \mathcal{R}(T)$.

Similarly, one can show that $\alpha y_1 \in \mathcal{R}(T)$. Therefore, $\mathcal{R}(T)$ is a subspace of Y.

One-to-one and Onto Functions

Definition: Let $f: X \longrightarrow Y$ be a function. f is said to be **one–to-one**, or an **injunction**, if and only if $f(a) = f(b)$ implies that $a = b$ for all a and b in X.

Definition: Let $f: X \longrightarrow Y$ be a function. f is said to be **onto**, or a **surjection**, if and only if for every element $b \in B$ there is an element $a \in A$ such that $f(a) = b$.

Example 1.7.3.1: Determine whether the function $f(x) = x + 1$ from the set of real numbers to itself is one-to-one.

Solution: Suppose that $f(x_1) = f(x_2)$ for some $x_1, x_2 \in \mathbb{R}$. Then from the definition of f, we have $x_1 + 1 = x_2 + 1$. Hence, $x_1 = x_2$. Therefore, f is one-to-one.

Example 1.7.3.2: Is the function $f(x) = 2x + 1$ from the set of real numbers to the set of real numbers onto?

Solution: Let Let y be any real number. Then $x = \frac{y-1}{2}$ is also a real number and $f(x) = f\left(\frac{y-1}{2}\right) = 2\left(\frac{y-1}{2}\right) + 1 = y$. Hence, for all $y \in \mathbb{R}$, there exists a $x \in \mathbb{R}$ such that $f(x) = y$. Therefore, f is onto.

Results Associated with Null Space and Range

Theorem 1.7.3.2: Let V, W be finite dimensional vector spaces over the same field F. Let $\boldsymbol{T}: V \longrightarrow W$ be a linear Transformation. Suppose that $\{v_1, v_2, \cdots, v_n\}$ is an ordered basis for V. Then

1. $\mathcal{R}(T) = Span\ (\{Tv_1, \mathrm{T}v_2, \cdots, Tv_n\})$.
2. $\dim(\mathcal{R}(T)) \leq \dim(W)$.
3. $\dim(N(T)) \leq \dim(V)$.
4. T is one-to-one $\Leftrightarrow$ $N(T) = \{0_V\} \Leftrightarrow \{Tv_1, \mathrm{T}v_2, \cdots, Tv_n\}$ is a basis of $\mathcal{R}(T)$.
5. $\dim(\mathcal{R}(T)) = \dim(V) \Leftrightarrow N(T) = \{0_V\}$.

Remarks:

1. We write $\rho(T) = \dim(\mathcal{R}(T))$ and $\nu(T) = \dim(N(T))$.
2. $\rho(T)$ is called the rank of the linear transformation T and $\nu(T)$ is called the nullity of the linear transformation T.

Proof of Theorem 1.7.3.2

Proof of (1), (2) and (3): Let $y \in \mathcal{R}(T)$. Then there exists $x \in V$ such that $y = Tx$. Since $\{v_1, v_2, \cdots, v_n\}$ is a basis for V, x can be written as $x = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ for some scalars $a_1, a_2, \cdots, a_n$.

Now $T(x) = T(a_1v_1 + a_2v_2 + \cdots + a_nv_n)$

$$= a_1T(v_1) + a_2T(v_2) + \cdots + a_nT(v_n) \ (\because \text{ T is linear}).$$

Hence, $y = Tx \in Span\ (\{Tv_1, \mathrm{T}v_2, \cdots, Tv_n\})$ and $\mathcal{R}(T) \subseteq Span\ (\{Tv_1, \mathrm{T}v_2, \cdots, Tv_n\})$.
We have $Span\ (\{Tv_1, \mathrm{T}v_2, \cdots, Tv_n\}) \subseteq \mathcal{R}(T)$ because $Tv_1, \mathrm{T}v_2, \cdots, Tv_n \in \mathcal{R}(T)$ and $\mathcal{R}(T)$ is a subspace of W.
Therefore, $\mathcal{R}(T) = Span\ (\{Tv_1, \mathrm{T}v_2, \cdots, Tv_n\})$.

Since $\mathcal{R}(T)$ is a subspace of W and W is finite dimensional, we have $\dim(\mathcal{R}(T)) \leq \dim(W)$.

Since $N(T)$ is a subspace of V and V is finite dimensional, we have $\dim(N(T)) \leq \dim(V)$.

Proof of Theorem 1.7.3.2 Continue

Proof of (4) and (5):

Assume that T is one-to-one. We need to show that $N(T) = \{0_V\}$.

Let $u \in N(T)$. Then by definition, $T(u) = 0_W$. Also for any linear transformation $T(0_V) = 0_W$. Thus $T(u) = T(0_V)$. So, T is one-to-one-one implies $u = 0_V$. That is; $N(T) = \{0_V\}$.

Conversely, suppose that $N(T) = \{0_V\}$. We need to show that T is one-to-one-one. So, let us assume that for some $u, v \in V$, $T(u) = T(v)$. Then, by linearity of T, $T(u - v) = 0_W$. This implies, $u - v \in N(T) = \{0_V\}$. This in turn implies $u = v$. Hence, T is one-to-one.

Remaining part of the proof is left as an exercise.

Finding the range and null space

Example 1.7.3.3: Determine the range and null space of the linear transformation $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ defined by $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_2 - x_3 \\ x_1 \\ 2x_1 - 5x_2 + 5x_3 \end{bmatrix}$, for all $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$. Also find $\dim(\mathcal{R}(T))$ and $\dim(N(T))$.

Solution:

$$\mathcal{R}(T) = \left\langle \left\{ \begin{bmatrix} x_1 - x_2 + x_3 \\ x_2 - x_3 \\ x_1 \\ 2x_1 - 5x_2 - 5x_3 \end{bmatrix} \right\} \right\rangle = \left\langle \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 5 \end{bmatrix} \right\} \right\rangle = \left\langle \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ -5 \end{bmatrix} \right\} \right\rangle$$

$$= \left\{ \alpha \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 1 \\ 0 \\ -5 \end{bmatrix} : \ \alpha, \beta \in \mathbb{R} \right\}.$$

Hence, $\dim(\mathcal{R}(T)) = 2$.

Finding the range and null space

Solution of example 1.7.3.3 continue:

$$\boldsymbol{N(T)} = \left\{\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \ : \boldsymbol{Tx} = \boldsymbol{0}\right\} = \left\{\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : \begin{bmatrix} x_1 - x_2 + x_3 \\ x_2 - x_3 \\ x_1 \\ 2x_1 - 5x_2 + 5x_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ 0 \\ 0 \end{bmatrix} \right\}$$

$$= \left\{\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : x_1 = \boldsymbol{0} \ \& \ x_2 = x_3\right\}$$

$$= \left\{\begin{bmatrix} 0 \\ y \\ y \end{bmatrix} : \boldsymbol{y} \in \boldsymbol{R}\right\} = \left\langle \left\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\} \right\rangle.$$

Hence, $\dim(N(T)) = 1$.

Also $\left\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$ is a basis of $N(T)$ and $\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ -5 \end{bmatrix}\right\}$ is a basis of $\mathcal{R}(T)$.

Note that $\dim(N(T)) + \dim(\mathcal{R}(T)) = \dim(\mathbb{R}^3)$.

Activity

1. A linear transformation $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is defined by $T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}0\\0\\1\end{bmatrix}$, $T\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\1\\1\end{bmatrix}$, and $T\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) = \begin{bmatrix}1\\1\\0\end{bmatrix}$.
   a. Find $\mathcal{R}(T)$ and calculate $\dim(\mathcal{R}(T))$.
   b. Find $N(T)$ and calculate $\dim(N(T))$.

2. A linear transformation $T: \mathbb{R}^5 \longrightarrow \mathbb{R}^3$ is defined by

$$T\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix} = \begin{bmatrix}2x_1 - 3x_3 + 4x_4\\ x_2 - x_3 + 2x_5\\ x_1 - 2x_3 + x_4\end{bmatrix}.$$

   a. Find $\mathcal{R}(T)$ and a basis for $\mathcal{R}(T)$.
   b. Find $N(T)$ and a basis for $N(T)$.

Proof of the Rank Nullity Theorem

Theorem 1.7.3.3: Let $\boldsymbol{T}: V \longrightarrow W$ be a linear transformation and V be a finite dimensional vector space. Then $\dim\big(R(T)\big) + \dim\big(N(T)\big) = \dim(V)$.
That is; $\rho_T \; + \; \nu_T = \dim(V)$.

Proof: Let $\dim(V) = n$ and $\dim\big(N(T)\big) = r$. Suppose that $\mathcal{B}_N = \{u_1, u_2, \cdots, u_r\}$ is a basis of $N(T)$. Since $\{u_1, u_2, \cdots, u_r\}$ is a linearly independent set in V, we can extend $\mathcal{B}_N$ to basis $\mathcal{B}_V = \; \{u_1, u_2, \cdots, u_r, v_1, v_2, \cdots, v_{n-r}\}$ for V by appending $n - r$ vectors form $V - N(T)$.
Since $\mathcal{B}_V$ is a basis for V, any $x \in V$ can be uniquely expressed as a linear combination of vectors in $\mathcal{B}_V$. Let $x = \sum_{i=1}^{r} \alpha_i u_i + \; \sum_{j=1}^{n-r} \beta_j v_j$. Hence,

$$T(x) = T\left(\sum_{i=1}^{r} \alpha_i u_i + \sum_{j=1}^{n-r} \beta_j v_j\right) = \sum_{i=1}^{r} \alpha_i \, T(u_i) + \sum_{j=1}^{n-r} \beta_j \, T(v_j)$$
$$= \sum_{j=1}^{n-r} \beta_j \, T(v_j) \, .$$

$$\Rightarrow T(x) \in Span\{T(v_1), T(v_2), \cdots, T(v_{n-r})\}, \; \forall x \in V.$$
$$\Rightarrow \;\; R(T) = Span\{w_1, w_2, \cdots, w_{n-r}\}, \text{where } \; w_j = T\big(v_j\big), \; for \;\; 1 \le j \; \le n - r.$$

Hence, If we show that $\{w_1, w_2, \cdots, w_{n-r}\}$ is a linearly independent set, then it is a basis for $R(T)$.

Proof Continued

Suppose that $\gamma_1 w_1 + \gamma_2 w_2 + \cdots + \gamma_{n-r} w_{n-r} = \mathbf{0}_w$, for some $\gamma_1, \gamma_2, \cdots, \gamma_{n-r} \in \mathrm{F}$, where $\mathbf{0}_w$ is the zero vector of W.

$\Rightarrow \gamma_1 T(v_1) + \gamma_2 T(v_2) + \cdots + \gamma_{n-r} T(v_{n-r}) = \mathbf{0}_w$.

$\Rightarrow T\left(\gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_{n-r} v_{n-r}\right) = \mathbf{0}_w$ ($\because$ T is linear).

$\Rightarrow \sum_{i=1}^{n-r} \gamma_i v_i \in N(T)$.

$\Rightarrow \sum_{i=1}^{n-r} \gamma_i v_i = \sum_{j=1}^{r} \eta_j u_j$ for some $\eta_1, \eta_2, \cdots, \eta_r \in \mathrm{F}$ ($\because \mathcal{B}_N = \{u_1, u_2, \cdots, u_r\}$ is a basis of $N(T)$.

$\Rightarrow \sum_{j=1}^{r} \eta_j u_j + \sum_{i=1}^{n-r} -\gamma_i v_i = \mathbf{0}_V$, where $\mathbf{0}_V$ is the zero vector of V.

$\Rightarrow \eta_j = 0, (1 \leq j \leq r)$, and $-\gamma_i = 0, (1 \leq i \leq n-r)$. ($\because \{u_1, \cdots, u_r, v_1, \cdots, u_{n-r}\}$ is a basis of V.

$\Rightarrow \gamma_i = 0, \forall i \ (1 \leq i \leq n-r)$.

$\Rightarrow \{w_1, w_2, \cdots, w_{n-r}\}$ is linearly independent.

Hence, $\dim(R(T)) = n - r$. Therefore, $\dim\big(R(T)\big) + \dim\big(N(T)\big) = \dim(V)$.

Applications of Rank Nullity Theorem

Corollary 1.7.3.4: Let $T: V \longrightarrow V$ be a linear transformation on a finite dimensional vector space V. Then the following statements are equivalent:

1. T is one-to-one.

2. T is onto.

3. T is invertible.

Proof: By Theorem 7.3.2 (4), T is one-to-one if and only if $N(T\ ) = \{0_V\}$. By the rank-nullity Theorem 7.3.3, $N(T\ ) = \{0_V\}$ is equivalent to the condition $\dim(R(T)) = \dim(V)$. Or equivalently T is onto.

By definition, T is invertible if T is one-one and onto. But we have shown that T is one-to-one if and only if T is onto. Thus, T is invertible.

1.7.4 Ordered bases, Matrix of Linear Transformation, Similarity of Matrices

Ordered Bases and coordinate systems:

Definition: Let X be an n dimensional vector space over the field F.
Let $\mathcal{B} = \{b_1, b_2, \cdots, b_n\}$ be an ordered base for X. Then any $x \in X$ can be uniquely written as $x = x_1 b_1 + x_2 b_2 + \cdots + x_n b_n$, where $x_i \in F$ for each i. We say that $x_1, x_2, \cdots, x_n$ are the coordinates of x with respect to the ordered base $\mathcal{B}$,

and we write

$$x_{[\mathcal{B}]} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathrm{F}^n$$ ($x_{[\mathcal{B}]}$ is the coordinates of x w.r.t the base $\mathcal{B}$).

Example: Vector $v \in \mathbb{R}^3$ has the coordinate $v_{[\mathcal{B}]} = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$ with respect to the standard ordered basis $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, since $v = \mathbf{2}e_1 + \mathbf{1}e_2 + (\mathbf{-3})e_3$.

Finding coordinates of a vector w.r.t given basis

Example 1.7.4.1: Find the coordinates of the vector $v = \begin{bmatrix}1\\2\\3\end{bmatrix}$ with respect to an ordered basis $\mathcal{B}' = \{b_1, b_2, b_3\}$ of $\mathbb{R}^3$, where $b_1 = \begin{bmatrix}1\\2\\1\end{bmatrix}, b_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, and $b_3 = \begin{bmatrix}-1\\0\\1\end{bmatrix}$.

Solution: To find the coordinates, we must solve the system of equations $v = v_1 b_1 + v_2 b_2 + v_3 b_3$. Using Gauss-Jordan method (Section 1.4.2), we solve this system:

$$\begin{bmatrix} 1 & 0 & -1 & \vdots & 1 \\ 2 & 1 & 0 & \vdots & 2 \\ 1 & 0 & 1 & \vdots & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 & \vdots & 1 \\ 0 & 1 & 2 & \vdots & 0 \\ 0 & 0 & 2 & \vdots & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & \vdots & 2 \\ 0 & 1 & 0 & \vdots & -2 \\ 0 & 0 & 1 & \vdots & 1 \end{bmatrix}.$$

Hence, the unique solution is $v_1 = 2, v_2 = -2$ and $v_3 = 1$. Therefore, the coordinates of v with respect to the ordered basis $\mathcal{B}'$ are 2, -2 and 1.

Also $v_{[\mathcal{B}']} = \begin{bmatrix}2\\-2\\1\end{bmatrix}$.

Matrix of a Composition of transformations

Definition (Composition of linear transformations): Let $S: \mathbb{R}^k \longrightarrow \mathbb{R}^n$ and $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be linear transformations. Then the composition of S and T (also called the composite transformation of S and T) is the linear transformation $T \circ S: \mathbb{R}^k \longrightarrow \mathbb{R}^m$ that is defined by

$$(T \circ S)\ (v) = T\big(S(v)\big), \text{for all } v \in \mathbb{R}^k.$$

Theorem 1.7.4.1 (Matrix of a composite transformation): Let $S: \mathbb{R}^k \longrightarrow \mathbb{R}^n$ and $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be linear transformations. Let A be the matrix corresponding to S, and let B be the matrix corresponding to T. Then the matrix corresponding to the composite linear transformation $T \circ S$ is BA.

Proof: We have

$$(T \circ S)(v) = T\big(S(v)\big) = T(Av) = B(Av) = (BA)v, \text{ for all } v \in \mathbb{R}^k.$$

Therefore, BA is the matrix corresponding to $T \circ S$.

Composition of linear transformations

Example 1.7.4.2: Let $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ and $S: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be two linear transformations defined respectively by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x + y \\ -x + 2y \end{bmatrix}$, and $S\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2y \\ 4x + 2y \end{bmatrix}$, $\forall \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$. Find the matrix of $S \circ T$ and compute $(S \circ T)\,(v)$ for $v = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

Solution: The matrix of T is $\mathrm{A} = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$ and the matrix of S is $\mathrm{B} = \begin{bmatrix} 0 & -2 \\ 4 & 2 \end{bmatrix}$.

Hence, the matrix of $S \circ T$ is $BA = \begin{bmatrix} 0 & -2 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ 10 & 8 \end{bmatrix}$.

$(S \circ T)\,(v) = (BA)v = \begin{bmatrix} 2 & -4 \\ 10 & 8 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix}$.

Inverse of a Linear Transformation

Definition: Let T: $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ and S: $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ be linear transformations. Suppose that for each $v \in \mathbb{R}^n$

$$(S \circ T)\,(v) = v \quad \text{and} \quad (T \circ S)\,(v) = v.$$

Then S is called the inverse of T, and we write $S = T^{-1}$.

Theorem 1.7.4.2: Let $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation and let A be the $n \times n$ matrix corresponding to T. Then T has an inverse if and only if the matrix A is invertible. In this case, the matrix of T^{-1} is A^{-1}.

Inverse of a transformation

Example 1.7.4.3: Find the inverse of the linear transformation $T\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined

by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ 7x + 4y \end{bmatrix}$, for all $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$.

Solution: The matrix of T is $\mathrm{A} = \begin{bmatrix} 2 & 1 \\ 7 & 4 \end{bmatrix}$.

The inverse of A is $A^{-1} = \begin{bmatrix} 2 & 1 \\ 7 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix}$.

Therefore, T^{-1} is the linear transformation defined by

$$T^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4x - y \\ -7x + 2y \end{bmatrix}, \text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2.$$

Similarity of Matrices

Definition: Let X be a real $n \times n$ matrix ($X \in M_{n,n}(\mathbb{R})$). We say $Y \in M_{n,n}(\mathbb{R})$ is similar to X if, there exists an invertible matrix $P \in M_{n,n}(\mathbb{R})$ such that

$$P^{-1}XP = Y.$$

Definition: Two vector spaces V and W over the same field F are said to be **isomorphic** if there is a linear transformation $T: V \longrightarrow W$ such that T is both one-to-one and onto. The linear transformation T is called an isomorphism of vector spaces.

n- Dimensional real Vector space is Isomorphic to $\mathbb{R}^n$

Theorem 1.7.4.3: Let V be an n dimensional real vector space. Then V is isomorphic to the vector space $\mathbb{R}^n$.

Proof: Let $\mathcal{B}_V = \{v_1, v_2, \cdots, v_n\}$ be an ordered basis of V. Then any vector $x \in V$ has a unique representation $x = x_1v_1 + x_2v_2 + \cdots + x_nv_n$, where $x_1, x_2, \cdots, x_n \in \mathbb{R}$.

Define a function $T: V \longrightarrow \mathbb{R}^n$ by $T(x) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

One can easily show that T is linear, one-to-one and onto.

Therefore, V is isomorphic to $\mathbb{R}^n$.

Note: The column vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is known as the coordinate of x w.r.t the standard basis of $\mathbb{R}^n$.

The Matrix representation of a Linear Transformation

In sections 1.7.1 and 1.7.2, we have shown that there is a one-to-one correspondence between the set of all linear transformations $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and the set of all real $m \times n$ matrices (Theorem 1.7.1.2 & Theorem 1.7.2.1).

In this section, we will generalize this correspondence to arbitrary finite-dimensional vector spaces. While $\mathbb{R}^n$ comes with a natural coordinate system (with respect to the standard ordered basis $e_1, e_2, \cdots, e_n$), there is no distinguished coordinate system on an arbitrary n dimensional vector space.

Let V, W be two finite dimensional vector spaces. To define the matrix of a linear transformation $T: V \longrightarrow W$, we must first choose a basis, or equivalently a coordinate system, for the vector spaces V and W. Different choices of basis will give rise to different matrices.

The Matrix of a Linear Transformation

Theorem 1.7.4.4: Let V, W be two finite dimensional vector spaces over the field F with $\dim(V) = n$ and $\dim(W) = m$. Let $T: X \longrightarrow W$ be a linear transformation. Then there exists a unique $\mathrm{m} \times n$ matrix A such that for all $x \in V$, $A\,[x]_{[B_V]} = [Tx]_{[B_W]}$, where $B_V = \{b_1, b_2, \cdots, b_n\}$ and $B_W = \{c_1, c_2, \cdots, c_m\}$ are ordered basis for V and W respectively. Moreover, the j^{th} column of A holds the coordinates of the vector $T(b_i)$ in W with respect to basis B_W.

Proof: Let $x \in V$ be arbitrary. Since B_V is an ordered basis for V, x can be uniquely written as $x = x_1 b_1 + x_2 b_2 + \cdots + x_n b_n$. That is; $[x_1 \quad x_2 \quad \cdots \quad x_n]^T = [x]_{[B_V]}$.

$$T(x) = T(x_1 b_1 + x_2 b_2 + \cdots + x_n b_n)$$

$$= x_1\, T(b_1) + x_2\, T(b_2) + \cdots + x_n\, T(b_n) \; (\because \text{ T is linear}).$$

Hence, the linear transformation T is completely determined by the images of the basis vectors, $T(b_1), T(b_2), \cdots, T(b_n) \in W$.

Since $B_W = \{c_1, c_2, \cdots, c_m\}$ is a basis of W, each $T(b_i)$ can be uniquely written as

as $T(b_i) = a_{1i}c_1 + a_{2i}c_2 + \cdots + a_{mi}c_m$ for some scalars $a_{1i}, a_{2i}, \cdots, a_{mi}$.

That is; $[a_{1i} \quad a_{2i} \quad \cdots \quad a_{mi}]^T = [T(b_i)]_{[B_W]}$.

Proof Cont.

That is; $T(b_1) = a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m$

$T(b_2) = a_{12}c_1 + a_{22}c_2 + \cdots + a_{m2}c_m$

$\vdots$

$T(b_n) = a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m.$

Now,

$$T(x) = T\left(\sum_{j=1}^{n} x_j b_j\right) = \sum_{j=1}^{n} x_j\, T(b_j) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\, x_j\right) c_i.$$

Define a $m \times n$ matrix A by $A = [a_{ij}]$. Then the coordinates of the vector $T(x)$ with respect to the ordered basis $\mathcal{B}_W$ of W is

$$[T(x)]_{[\mathcal{B}_W]} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}\, x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}\, x_j \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A[x]_{[\mathcal{B}_V]}.$$

The matrix A is called the matrix of the linear transformation T with respect to the ordered bases $\mathcal{B}_V$ and $\mathcal{B}_W$ of V and W respectively, and is denoted by $T[\mathcal{B}_W, \mathcal{B}_V]$.

Some Observations

Let $\boldsymbol{T}: V \longrightarrow W$ be a linear Transformation, where V and W are two finite dimensional vector spaces over the same field F with respective dimensions m and n.
Suppose that $\mathcal{B}_1 = \{v_1, v_2, \cdots, v_n\}$ is an ordered basis of V, and $\mathcal{B}_2 = \{w_1, w_2, \cdots, w_m\}$ is an ordered basis of W.

1. $T[\mathcal{B}_2, \mathcal{B}_1] = [[T(v_1)]_{[\mathcal{B}_2]} \quad [T(v_2)]_{[\mathcal{B}_2]} \quad \cdots \quad [T(v_n)]_{[\mathcal{B}_2]}]$.

2. It is important to note that $[T(x)]_{\mathcal{B}_2} = T[\mathcal{B}_2, \mathcal{B}_1][x]_{\mathcal{B}_1}$.
   That is, we multiply the matrix of the linear transformation with the coordinates $[x]_{\mathcal{B}_1}$, of the vector $x \in V$ to obtain the coordinates of the vector $T(x) \in W$ w.r.t the ordered basis $\mathcal{B}_2$.

3. If A is an $m \times n$ real matrix, then A induces a linear transformation $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, defined by $T_A(x) = Ax$.
   Suppose that the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ are the ordered bases $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively. Then observe that $T[\mathcal{B}_2, \mathcal{B}_1] = A$.

Finding the matrix of a linear transformation

Example 1.7.4.4: Let $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix}$, for all $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$. Let $B_1 = \left\{ b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ and $B_2 = \left\{ c_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ be two ordered basis for $\mathbb{R}^2$. Find the matrix $\mathrm{A} = T[\mathcal{B}_2, \mathcal{B}_1]$ of T with respect to the bases B_1 and B_2.

Solution:

$$T\left(b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 1+0 \\ 1-0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1c_1 + 0\, c_2, \text{ and}$$

$$T\left(b_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 1+1 \\ 1-1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1c_1 + 1\, c_2.$$

Therefore, the matrix A that is representing T is $\mathrm{A} = T[\mathcal{B}_2, \mathcal{B}_1] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Finding the matrix of a linear transformation

Example 1.7.4.5: Let $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix}$, for all $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$. Let $B_1 = \left\{ b_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, b_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$ and $B_2 = \left\{ c_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ be two ordered basis for $\mathbb{R}^2$. Find the matrix $\mathrm{A} = T[\mathcal{B}_2, \mathcal{B}_1]$ of T with respect to the bases B_1 and B_2.

Solution:

$T\left(b_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1+1 \\ 1-1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2c_1 + 2\, c_2,$ and

$T\left(b_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 0-1 \\ 0+1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0\begin{bmatrix} 0 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0c_1 - 1\, c_2.$

Therefore, the matrix A that is representing T is $\mathrm{A} = T[\mathcal{B}_2, \mathcal{B}_1] = \begin{bmatrix} 2 & 0 \\ 2 & -1 \end{bmatrix}$.

Remark: Even though the linear transformations given in the last two examples are the same, matrices representing them are different. This is because the matrix depends not only on the linear transformation, but also on the given bases. The art of linear algebra often lies in choosing “convenient” bases for a given application.

Finding the matrix of a linear transformation

Example 1.7.4.6: Find the matrix of the derivative operator $D: P_3(\mathbb{R}) \longrightarrow P_2(\mathbb{R})$ with respect to the basis $B_1 = \{1, x+1, x^2+x+1, x^3+x^2+x+1\}$ of $P_3(\mathbb{R})$ and the basis $B_2 = \{1, x-1, x^2-1\}$ of $P_2(\mathbb{R})$.

Solution: Let us denote the basis vectors of B_1 as $b_1 = 1, b_2 = x+1, b_3 = x^2 + x + 1$, and $b_4 = x^3 + x^2 + x + 1$, and the basis vectors of B_2 as $c_1 = 1, c_2 = x - 1$, and $c_3 = x^2 - 1$.

$$
\begin{aligned}
D(b_1) &= D(1) && = \mathbf{0} && = 0c_1 + 0c_2 + 0c_3 \\
D(b_2) &= D(x+1) && = 1 && = 1c_1 + 0c_2 + 0c_3 \\
D(b_3) &= D(x^2+x+1) && = 2x+1 && = 3c_1 + 2c_2 + 0c_3 \\
D(b_4) &= D(x^3+x^2+x+1) && = 3x^2+2x+1 && = 6c_1 + 2c_2 + 6c_3.
\end{aligned}
$$

Therefore, the matrix is $T[\mathcal{B}_2, \mathcal{B}_1] = \begin{bmatrix} 0 & 1 & 3 & 6 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

Change of Coordinates

Let $\mathcal{B}_1 = \{v_1, v_2, \cdots, v_n\}$ be an ordered basis of a real vector space V and $T_1 \colon V \longrightarrow \mathbb{R}^n$ be the coordinate mapping corresponding to this basis.

Let $\mathcal{B}_2 = \{u_1, u_2, \cdots, u_n\}$ be an another ordered basis of V and $T_2 \colon V \longrightarrow \mathbb{R}^n$ be the coordinate mapping corresponding to this basis.

$$\begin{tikzcd} & V \arrow[dl, "T_1"'] \arrow[dr, "T_2"] & \\ \mathbb{R}^n \arrow[rr] & & \mathbb{R}^n \end{tikzcd}$$

Consider the composition of transformations $\left(T_2 \circ {T_1}^{-1}\right) \colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Since both T_1 and T_2 isomorphisms, $T_2 \circ {T_1}^{-1}$ is also an isomorphism.

It is represented as $\left(T_2 \circ {T_1}^{-1}\right)(v) = Uv$, where U is an $n \times n$ matrix.

U is called the transition matrix from $v_1, v_2, \cdots, v_n$ to $u_1, u_2, \cdots, u_n$.

Columns of U are coordinates of the vectors $v_1, v_2, \cdots, v_n$ with respect to the basis $u_1, u_2, \cdots, u_n$.

Matrix of a Linear Transformation

Let $\boldsymbol{T}: V \longrightarrow W$ be a linear transformation, where V, W are finite dimensional vector spaces over the field $\mathbb{R}$.

Let $\mathcal{B}_V = \{v_1, v_2, \cdots, v_n\}$ be an ordered basis of V and $T_1: V \longrightarrow \mathbb{R}^n$ be the coordinate mapping corresponding to this basis.

Let $\mathcal{B}_W = \{u_1, u_2, \cdots, u_m\}$ be an ordered basis of W and $T_2: W \longrightarrow \mathbb{R}^m$ be the coordinate mapping corresponding to this basis.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "T_1"] & W \arrow[d, "T_2"'] \\ \mathbb{R}^n \arrow[r] & \mathbb{R}^m \end{tikzcd}$$

The composition of transformations $\left(T_2 \circ T \circ {T_1}^{-1}\right)$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

It is represented as $\left(T_2 \circ T \circ {T_1}^{-1}\right)(v) = Av$, where A is an $m \times n$ matrix.

A is called the matrix of the $\boldsymbol{T}$ with respect to a ordered bases $v_1, v_2, \cdots, v_n$ of V and $u_1, u_2, \cdots, u_m$ of W.

Columns of A are coordinates of vectors $Tv_1, Tv_2, \cdots, Tv_n$ respect to the basis $u_1, u_2, \cdots, u_m$ of W.

Change of Basis for a Linear Operator

Let $\boldsymbol{T}: V \longrightarrow V$ be a linear operator, where V is a finite dimensional vector spaces over the field $\mathbb{R}$.

Let A be the matrix of $\boldsymbol{T}$ relative to an ordered basis $\mathcal{B}_1 = \{v_1, v_2, \cdots, v_n\}$ of V.

Let B be the matrix of $\boldsymbol{T}$ relative to another ordered basis $\mathcal{B}_2 = \{u_1, u_2, \cdots, u_n\}$ of V.

Let U be the transition matrix from the basis $v_1, v_2, \cdots, v_n$ to $u_1, u_2, \cdots, u_n$.

Let $x \in V$ be arbitrary.

$$
\begin{tikzcd}
\boxed{\text{Coordinates of } v \text{ w.r.t } \mathcal{B}_1} \arrow[r, "A"] \arrow[d, "U"] & \boxed{\text{Coordinates of } \mathrm{T}v \text{ w.r.t } \mathcal{B}_1} \arrow[d, "U"] \\
\boxed{\text{Coordinates of } v \text{ w.r.t } \mathcal{B}_2} \arrow[r, "B"] & \boxed{\text{Coordinates of } \mathrm{T}v \text{ w.r.t } \mathcal{B}_2}
\end{tikzcd}
$$

Since U is invertible, it is obvious from the diagram that $B = U^{-1}A\,U$ and $\mathrm{A} = UAU^{-1}$.

Note that A and B are similar matrices.

Applications

Example 1.7.4.7: Let $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. Find the matrix of the linear operator T: $\mathbb{R}^3 \longrightarrow \mathbb{R}^3$

defined by $Tx = Ax$ w.r.t the ordered basis $v_1 = [-1 \quad 1 \quad 0]^T$, $v_2 = [1 \quad 1 \quad 0]^T$, $v_3 = [-1 \quad 0 \quad 1]^T$.

Solution: Let B be the desired matrix. The columns of B are coordinates of the vectors $T(v_1), T(v_2)$, and $T(v_3)$.

$$T(v_1) = Av_1 = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \; T(v_2) = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = 2v_2,$$

and $T(v_3) = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix} = 2v_3$.

Hence, $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

Applications

Example 1.7.4.8: Let $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. Find A^8.

Solution: From the last example, $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

Transition matrix $U = \begin{bmatrix} -1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

It follows from the solution of the previous problem that $A = U^{-1}B\,U$.

Hence, $A^8 = (U^{-1}B\,U)^8 = (U^{-1}B\,U)(U^{-1}B\,U)\cdots(U^{-1}B\,U)$

$$= U^{-1}B^8U = U^{-1}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2^8 & 0 \\ 0 & 2^8 & 2 \end{bmatrix} U = 2^{14}(U^{-1}B\,U) = 2^{14}A.$$

Therefore, $A^8 = 2^{14}\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2^{14} & 2^{14} & -2^{14} \\ 2^{14} & 2^{14} & 2^{14} \\ 0 & 0 & 2^{15} \end{bmatrix}$.

Some Results about Similarity Matrices

Theorem 1.7.4.5: Let $\mathrm{A}, \mathrm{B}, \mathrm{C}$ be three real square matrices of order n. Then

1. If A is similar to B, then B is similar to A.
2. If A is similar to B and B is similar to C, then A is similar to C.
3. If A and B are similar matrices then $|A| = |B|$.
4. If A and B are similar matrices then trace$\mathrm{of\,A} =$ trace $\mathrm{of\,B}$.
5. If A and B are similar matrices then rank of A = rank of B.
6. If A and B are similar matrices then $\mathrm{Null(A)} = \mathrm{Null(B)}$.

Proof:

1. If $B = U^{-1}A\,U$ then $A = UBU^{-1} = \ (U^{-1})^{-1}BU^{-1}$.
2. If $A = U^{-1}B\,U$ and $B = V^{-1}C\,V$, then $A = U^{-1}V^{-1}C\,VU = (VU)^{-1}C(VU)$.
3. If $A = U^{-1}B\,U$, then $|A| = |U^{-1}B\,U| = [U^{-1}]\,|B|\,|U| = |B|$.

Other parts of the theorem can be proved similarly.

Activity

Which of the following is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$?

a. $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x^2 + z \\ 0 \end{bmatrix}$

b. $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2 \\ x + y - z \end{bmatrix}$

c. $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2z \\ x + y + z \end{bmatrix}$

d. $Tx = Ax$, for all $x \in \mathbb{R}^3$, where $A = \begin{bmatrix} 1 & 3 \\ 0 & 4 \\ 1 & 2 \end{bmatrix}$.