CHAPTER 2

# Linear Transformations and Matrices

## 2.1 Linear Transformations

**Definition 2.1.** A **linear transformation** $T$ from a vector space $V$ to a vector space $W$ is a map

$$T : V \longrightarrow W$$

such that for all vectors $\mathbf{u}, \mathbf{v} \in V$ and scalar $c \in K$:

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

(2) $T(c \cdot \mathbf{u}) = c \cdot T(\mathbf{u})$

$V$ is called **domain** and $W$ is called **codomain** of $T$.

The set of all such linear transformations $T : V \longrightarrow W$ is denoted by $L(V, W)$.

**Fact.** Any $T \in L(V, W)$ is uniquely determined by the image on any basis $\mathcal{B}$ of $V$.

**Example 2.1.** The **identity** map $\mathbf{Id} : V \longrightarrow V$ given by $\mathbf{Id}(\mathbf{v}) = \mathbf{v}$.

**Example 2.2.** Differential operators on the space of real-valued differentiable functions.

**Example 2.3.** $Tr : M_{3\times 3}(\mathbb{R}) \longrightarrow \mathbb{R}$ on the space of $3 \times 3$ matrices with real entries.

**Example 2.4.** Matrix multiplication: $\mathbb{R}^n \longrightarrow \mathbb{R}^m$.

**Theorem 2.2.** Let $V, W$ be **finite dimensional** vector spaces with $\dim V = n, \dim W = m$. Then every linear transformation $T \in L(V, W)$ can be represented by an $m \times n$ matrix.

Let $\mathcal{B} \subset V, \mathcal{B}' \subset W$ be their bases. If $T(\mathbf{u}) = \mathbf{w}$, then the matrix is given by

$$\begin{aligned} [T]_{\mathcal{B}'}^{\mathcal{B}} : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ [\mathbf{u}]_{\mathcal{B}} &\mapsto [\mathbf{w}]_{\mathcal{B}'} \end{aligned}$$

It is called the **matrix of $T$ with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$**.

**Definition 2.3.**

- The **kernel** or **null space** of $T$ is:

$$\mathrm{Ker}(T) := \{\mathbf{u} \in V : T(\mathbf{u}) = \mathbf{0}\}$$

- The **image** or **range** of $T$ is:

$$\mathrm{Im}(T) := \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{u}) \text{ for some } \mathbf{u} \in V\}$$

**Theorem 2.4.** Let $T \in L(V, W)$. Then

- The kernel of $T$ is a subspace of $V$.
- The image of $T$ is a subspace of $W$.

**Example 2.5.** If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is represented by a matrix $\mathbf{A}$, then

- The kernel of $\mathbf{A}$ is the **null space** Nul$\mathbf{A}$. It is the set of all solutions to $\mathbf{Ax} = \mathbf{0}$ of $m$ homogeneous linear equations in $n$ unknown. It is a subspace of $\mathbb{R}^n$.
- The image of $\mathbf{A}$ is the **column space** Col$\mathbf{A}$. It is the set of all linear combinations of the columns of $\mathbf{A}$. It is a subspace of $\mathbb{R}^m$.

**Rule.** For a matrix $\mathbf{A}$, the row operations do not affect linear dependence of the columns.
Use the *reduced echelon form* to find the basis of Nul$\mathbf{A}$ and Col$\mathbf{A}$.

**Example 2.6.** The kernel of $\frac{d}{dx}$ on the space of differentiable functions is the set of all constant functions.

## 2.2 Injection, Surjection and Isomorphism

**Definition 2.5.** A linear transformation $T : V \longrightarrow W$ is called

- **one-to-one** or **injective** if $T(\mathbf{u}) = T(\mathbf{v})$ implies $\mathbf{u} = \mathbf{v}$
- **onto** or **surjective** if for every $\mathbf{w} \in W$, there exists $\mathbf{u} \in V$ such that $T(\mathbf{u}) = \mathbf{w}$
- **isomorphism** if $T$ is one-to-one and onto.

**Definition 2.6.** If there exists an isomorphism $T \in L(V, W)$, we say $V$ is **isomorphic** to $W$, written as $V \simeq W$.

**Properties 2.7.** Let $T \in L(V, W)$.

- If $T$ is injective, then $\mathcal{N}(T) = \{\mathbf{0}\}$, i.e. $T(\mathbf{u}) = \mathbf{0} \Longrightarrow \mathbf{u} = \mathbf{0}$.
- If $T$ is injective, it maps linearly independent set to linearly independent set.
- If $T$ is injective, and $H \subset V$ is a subspace, then $\dim T(H) = \dim H$.
- If $T$ is surjective, then $\mathcal{R}(T) = W$.
- If $T$ is isomorphism, then $\dim V = \dim W$.

**Theorem 2.8.** If $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_n\}$ is a basis for a vector space $V$, then the **coordinate mapping**

$$\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$$

is an isomorphism $V \simeq K^n$.

**Example 2.7.** $\{\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R}\} \neq \mathbb{R}^2$ but $\{\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R}\} \simeq \mathbb{R}^2$

**Example 2.8.** $\mathbb{P}_n(K) \simeq K^{n+1}$

**Example 2.9.** $M_{m\times n}(K) \simeq K^{mn}$

**Example 2.10.** $\mathbb{C} \simeq \mathbb{R}^2$ as vector spaces over $\mathbb{R}$.

## 2.3 Rank

**Definition 2.9.**

- The **rank** of $T$ is the dimension of the image of $T$.
- The **nullity** of $T$ is the dimension of the kernel of $T$.

Below we have the **Fundamental Theorem of Linear Algebra**, which consists of the Rank-Nullity Theorem and the Theorem of Column Rank = Row Rank :

**Theorem 2.10** (Rank–Nullity Theorem)**.** Let $T \in L(V, W)$. Then

$$\dim \operatorname{Im}(T) + \dim \operatorname{Ker}(T) = \dim V$$

Let $T$ be represented by a $m \times n$ matrix $\mathbf{A}$ (i.e. $V = \mathbb{R}^n, W = \mathbb{R}^m$). Then **rank** of $\mathbf{A}$ is the dimension of Col$\mathbf{A}$. The **row space** is the space spanned by the rows of $\mathbf{A}$. It is a subspace of $\mathbb{R}^n$. The **row rank** is the dimension of the row space.

The row space of $\mathbf{A}$ is the column space of $\mathbf{A}^T$, hence the row rank of $\mathbf{A}$ equals the rank of $\mathbf{A}^T$.

**Theorem 2.11** (Column Rank = Row Rank)**.** Rank of $\mathbf{A}$ = Rank of $\mathbf{A}^T$.

**Theorem 2.12** (Invertible Matrix Theorem)**.** Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\mathbf{A}$ is invertible iff any one of the statements hold:

(1) Columns of $\mathbf{A}$ form a basis of $\mathbb{R}^n$

(2) Col$\mathbf{A} = \mathbb{R}^n$

(3) Rank of $\mathbf{A} = n$

(4) Nul$\mathbf{A} = \{\mathbf{0}\}$

(5) Nullity of $\mathbf{A} = 0$.

## 2.4 Change of Basis

Recall that for $\mathbf{x} \in V$ and $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_n\}$ a basis of $V$,

$$[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$$

is the $\mathcal{B}$-coordinate vector of $\mathbf{x}$ if

$$\mathbf{x} = c_1\mathbf{b}_1 + ... + c_n\mathbf{b}_n.$$

If $\mathcal{B}' = \{\mathbf{b}'_1, ..., \mathbf{b}'_n\}$ is another basis of $V$, then

$$[\mathbf{x}]_{\mathcal{B}'} = \begin{pmatrix} c'_1 \\ \vdots \\ c'_n \end{pmatrix} \in \mathbb{R}^n$$

is the $\mathcal{B}'$-coordinate vector of $\mathbf{x}$ if

$$\mathbf{x} = c'_1\mathbf{b}'_1 + ... + c'_n\mathbf{b}'_n.$$

The relationship between the vectors $[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{x}]_{\mathcal{B}'}$ is given by

**Theorem 2.13** (Change of Basis formula)**.** There exists an $n \times n$ matrix $P^{\mathcal{B}}_{\mathcal{B}'}$ such that

$$[\mathbf{x}]_{\mathcal{B}'} = P^{\mathcal{B}}_{\mathcal{B}'} \cdot [\mathbf{x}]_{\mathcal{B}}$$

where *column-wise* it is given by

$$P^{\mathcal{B}}_{\mathcal{B}'} = \begin{pmatrix} [\mathbf{b}_1]_{\mathcal{B}'} & [\mathbf{b}_2]_{\mathcal{B}'} & \cdots & [\mathbf{b}_n]_{\mathcal{B}'} \end{pmatrix}$$

$P^{\mathcal{B}}_{\mathcal{B}'}$ is called the **change-of-coordinate matrix** from $\mathcal{B}$ to $\mathcal{B}'$.

In other words, it is the matrix of the identity map **Id** with respect to the basis $\mathcal{B}$ and $\mathcal{B}'$ (see Theorem 2.2).

**Properties 2.14.** The $n \times n$ matrix $\mathcal{P}^{\mathcal{B}}_{\mathcal{B}'}$ is invertible. We have

$$[\mathbf{x}]_{\mathcal{B}} = \left(P^{\mathcal{B}}_{\mathcal{B}'}\right)^{-1} \cdot [\mathbf{x}]_{\mathcal{B}'}$$

Hence

$$P^{\mathcal{B}'}_{\mathcal{B}} = \left(P^{\mathcal{B}}_{\mathcal{B}'}\right)^{-1}$$

**Example 2.11.** If $\mathcal{B}$ is a basis of $\mathbb{R}^n$ and $\mathcal{E}$ is the standard basis of $\mathbb{R}^n$, then

$$[\mathbf{b}_i]_{\mathcal{E}} = \mathbf{b}_i.$$

Hence we simply have

$$P^{\mathcal{B}}_{\mathcal{E}} = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{pmatrix}$$

$P^{\mathcal{B}} := P^{\mathcal{B}}_{\mathcal{E}}$ is called the **change-of-coordinate matrix** from $\mathcal{B}$ to the standard basis of $\mathbb{R}^n$.

**Properties 2.15.** We have

- $P^{\mathcal{B}'}_{\mathcal{B}''} \cdot P^{\mathcal{B}}_{\mathcal{B}'} = P^{\mathcal{B}}_{\mathcal{B}''}$
- $P^{\mathcal{B}}_{\mathcal{B}'} = (P^{\mathcal{B}'})^{-1} \cdot P^{\mathcal{B}}$

**Example 2.12.** Let $\mathcal{E} = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ be the standard basis of $\mathbb{R}^2$. Let

$$\mathcal{B} = \{\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\},$$
$$\mathcal{B}' = \{\mathbf{b}'_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{b}'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$$

be two other bases of $\mathbb{R}^2$. Then

$$\begin{aligned} P^{\mathcal{B}} &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ P^{\mathcal{B}'} &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ P^{\mathcal{B}}_{\mathcal{B}'} &= (P^{\mathcal{B}'})^{-1} \cdot P^{\mathcal{B}} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix} \end{aligned}$$

One can check that this obeys the formula from Theorem 2.13.

$$\begin{aligned} \mathbf{b}_1 &= 0 \cdot \mathbf{b}'_1 + 1 \cdot \mathbf{b}'_2 \\ \mathbf{b}_2 &= 2 \cdot \mathbf{b}'_1 + (-1) \cdot \mathbf{b}'_2. \end{aligned}$$