# Linear Transformation

**Linear Transformation**
- Linear Transformation
- Change of Basis and Diagonalization

## Linear Transformation

Linear transformations can be viewed as an extension of the notion of a matrix to general vector spaces.

> *Def.* **Linear Transformation**
>
> Let $U$ and $V$ be vector spaces, and let $T$ be a function from $U$ to $V$, $T: U \to V$. We say that $T$ is a linear transformation if $\forall$ $\mathbf{u}, \mathbf{v} \in U$ and all scalars $a$, we have:
>
> 1. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.
> 2. $T(a\mathbf{u}) = aT(\mathbf{u})$
>
> *Remark.* $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v_1}) + c_2T(\mathbf{v_2}) + \cdots + c_nT(\mathbf{v_n})$.

> **Linear Transformation Using Basis in $R^n$**
>
> Let $V$ be vector space with basis $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$. Let linear transformation $T: V \to U$ . For any vector $\mathbf{v} \in V$, we have
>
> $$\begin{aligned} \mathbf{v} &= a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_p\mathbf{u}_p \\ T(\mathbf{v}) &= a_1T(\mathbf{u}_1) + a_2T(\mathbf{u}_2) + \cdots + a_pT(\mathbf{u}_p) \end{aligned} \tag{1}$$
>
> *Th.* Let $T: \mathcal{R}^n \to \mathcal{R}^m$ be a linear transformation, and let $\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_n$ be the unit vectors in $\mathcal{R}^n$. If $A$ is an $(m \times n)$ matrix defined by:
>
> $$A = [T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)] \tag{2}$$
>
> Then we have $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathcal{R}^n$.
>
> *Proof.*:
>
> $$\begin{aligned} \mathbf{x} &= \sum_{i=1}^{n} x_i\mathbf{e}_i \\ T(\mathbf{x}) &= \sum_{i=1}^{n} x_iT(\mathbf{e}_i) = A\mathbf{x} \end{aligned} \tag{3}$$

#### The Matrix of Transformation

Suppose that $U$ and $V$ both have finite dimension, say $dim(U) = n$ and $dim(V) = m$. We have that $U$ is isomorphic to $R^n$ and $V$ is isomorphic to $R^m$. Let $B$ be a basis for $U$ and $C$ be a basis for $V$. Then each vector $\mathbf{u} \in U$ and $\mathbf{v} \in V$ can be represented by the vectors $[\mathbf{u}]_B \in R^n$ and $[\mathbf{v}]_C \in R^m$. Then the linear transformation $T : U \to V$ can be represented by an $(m \times n)$ matrix $Q$ in the sense that if $T(\mathbf{u}) = \mathbf{v}$, then $Q[\mathbf{u}]_B = [\mathbf{v}]_C$.

> *Def.* **The Matrix of a Transformation**
>
> Let $T : U \to V$ be a linear transformation, where $dim(U) = n$ and $dim(V) = m$. Let $B = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a basis for $U$ and let $C = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m\}$ be a basis for $V$. The **matrix representation** for $T$ with respect to the bases $B$ and $C$ is the $(m \times n)$ matrix $Q$ defined by:
>
> $$\begin{gathered} Q = [\mathbf{Q}_1, \mathbf{Q}_2, \cdots, \mathbf{Q}_n]. \\ \mathbf{Q}_j = [T(\mathbf{u}_j)]_C \end{gathered} \tag{4}$$
>
> and we have:
>
> $$Q[\mathbf{u}]_B = [\mathbf{v}]_C \text{ if } T(\mathbf{u}) = \mathbf{v}. \tag{5}$$
>
> *Proof.*
>
> $$\begin{aligned} &[\mathbf{u}]_B = [b_1, b_2, \cdots, b_n]^T. \\ &Q[\mathbf{u}]_B = \sum_{i=1}^{n} b_i \mathbf{Q}_i = \sum_{i=1}^{n} b_i [T(\mathbf{u}_i)]_C = [T(\sum_{i=1}^{n} b_i T(\mathbf{u}_i))]_C = [T(\mathbf{u})]_C = [\mathbf{v}]_C. \end{aligned} \tag{6}$$

> *Th.* **The Representation Theorem**
>
> Let $T : U \to V$ be a linear transformation, where $dim(U) = n$ and $dim(V) = m$. Let $B$ and $C$ be bases for $U$ and $V$, and let $Q$ be the matrix of $T$ relative to $B$ and $C$. If $\mathbf{u}$ is a vector in $U$ and $T(\mathbf{u}) = \mathbf{v}$, then:
>
> $$Q[\mathbf{u}]_B = [\mathbf{v}]_C \tag{7}$$
>
> Moreover, $Q$ is the unique matrix that satisfies the equation above.
>
> *Corollary.* Let $T_1$, $T_2$ and $T$ be transformation from U to V and let $Q_1$, $Q_2$ and $Q$ be the matrix representations with respect to $B$ and $C$ for $T_1$, $T_2$ and $T$. Then:
>
> 1. $Q_1 + Q_2$ is the matrix representation for $T_1 + T_2$ with respect to $B$ and $C$.
> 2. For a scalar $a$, $aQ$ is the matrix representation for $aT$ with respect to $B$ and $C$.

> **Composition**
>
> *Th.* Let $T : U \to V$ and $S : V \to W$ be linear transformations, and suppose $dim(U) = n$, $dim(V) = m$ and $dim(W) = k$. Let $B$, $C$, $D$ be bases for $U$, $V$ and $W$. If the matrix for $T$ relative to $B$ and $C$ is $Q_{m\times n}$ and the matrix for $S$ relative to $C$ and $D$ is $P_{k\times m}$, then the matrix representation for $S \circ T$ is $PQ$.

> *Def.* **Null Space and Range**

Let $V$ and $W$ be subspaces, and let $T : V \to W$ be a linear transformation. The **null space** of $T$, denoted by $\mathcal{N}(T)$, is the subset of $V$ given by:

$$\mathcal{N}(T) = \{\mathbf{v}|\mathbf{v} \in V \text{ and } T(\mathbf{v}) = \theta_{\mathbf{W}}\} \tag{8}$$

The **range** of $T$, denoted by $\mathcal{R}(T)$, is the subset of $W$ defined by:

$$\mathcal{R}(T) = \{\mathbf{w}|\mathbf{w} \in W \text{ and for some } \mathbf{v} \in V, \text{ we have } \mathbf{w} = T(\mathbf{v})\}. \tag{9}$$

We know that there exists a matrix $A$ s.t. $T(\mathbf{x}) = A\mathbf{x}$. In this case we have:

1. $\mathcal{N}(T) = \mathcal{N}(A)$.
2. $\mathcal{R}(T) = \mathcal{R}(A)$.

*Def.* **One to One Linear Transformation**

We say a linear transformation $T : U \to V$ in **one to one** if $T(\mathbf{u}) = T(\mathbf{v})$ implies $\mathbf{u} = \mathbf{v}$ for all $\mathbf{u}$ and $\mathbf{v}$ in $U$.

*Remark.* One to one linear transformation has strong relation with unique solution $A\mathbf{x} = \mathbf{b}$.

*Th.* Let $T : U \to V$ be a linear transformation. Then:

1. $T(\theta_U) = \theta_V$.
2. $\mathcal{N}(T)$ is a subspace of $U$.
3. $\mathcal{R}(T)$ is a subspace of $V$.
4. $T$ is one to one if and only if $\mathcal{N}(T) = \{\theta_U\}$; that is, $T$ is one to one if and only if $\text{nullity}(T) = 0$.

*Th.* [**for finite-dimensional vector space only**]

Let $T : U \to V$ be a linear transformation and let $U$ be $p$-dimensional, where $B = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ is a basis for $U$.

1. $\mathcal{R}(T) = Sp\{T(\mathbf{u}_1), T(\mathbf{u}_2), \cdots, T(\mathbf{u}_p)\}$.
2. $T$ is one to one if and only if $\{T(\mathbf{u}_1), T(\mathbf{u}_2), \cdots, T(\mathbf{u}_p)\}$ is linearly independent in $V$.
3. $rank(T) + nullity(T) = p$.

*Def.* **Operations with Linear Transformations**

Let $U$ and $V$ be vector spaces and let $T_1$ and $T_2$ be linear transformations, where $T_1 : U \to V$ and $T_2 : U \to V$.

1. **Sum**: $T_3 : U \to V$, $T_3 = T_1 + T_2$ means $T_3(\mathbf{u}) = T_1(\mathbf{u}) + T_2(\mathbf{u})$.
2. **Scalar Product**: $aT$ denotes $(aT)(\mathbf{u}) = a(T(\mathbf{u}))$.
3. **Composition**: Let $U, V, W$ be vector space, let $S$ and $T$ be linear transformations, where $S : U \to V$ and $T : V \to W$. The composition, $L = T \circ S$, of $S$ and $T$ is defined to be function $L : U \to W$ given by $L(\mathbf{u}) = T(S(\mathbf{u}))$.

*Def.* **Onto**(满射)

A function $f : X \to Y$ is **onto** provided that $\mathcal{R}(f) = Y$. A linear transformation $T : U \to V$ is **onto** provided that $\mathcal{R}(T) = V$.

*Def.* **Inverse**

Let $f : X \to Y$ be a function. If $f$ is both one to one and onto, then the **inverse** of $f$, denoted by $f^{-1} : Y \to X$, is the function defined by:

$$f^{-1}(y) = x \text{ if and only if } f(x) = y. \tag{10}$$

Therefore, if $T : U \to V$ is a linear transformation that is both one to one and onto, then the inverse function $T^{-1} : V \to U$ is defined.

*Def.* **Invertible Linear Transformations**

A linear transformation $T : U \to V$ that is both one to one and onto is called an **invertible** linear transformation.

*Th.* Let $U$ and $V$ be vector spaces, and let $T : U \to V$ be an invertible linear transformation. Then:

1. $T^{-1} : V \to U$ is a linear transformation.
2. $T^{-1}$ is invertible and $(T^{-1})^{-1} = T$.
3. $T^{-1} \circ T = I_U$ and $T \circ T^{-1} = I_V$, where $I_U$ and $I_V$ are the identity transformations on $U$ and $V$.
4. For each vector $\mathbf{b} \in V$, $\mathbf{x} = T^{-1}(\mathbf{b})$ is the unique solution of $T(\mathbf{x}) = \mathbf{b}$.
5. If $S$ and $T$ are invertible and $S \circ T$ is defined, then $S \circ T$ is invertible and $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.

*Note.* Almost all properties of invertible matrix can be used in invertible linear transformation.

**What is isomorphic**

Suppose invertible transformation $T : U \to V$ is both one to one and onto, $T$ established an exact pairing between elements of $U$ and $V$. Moreover, because $T$ is a linear transformation, this pairing preserves algebraic properties. Therefore, although $U$ and $V$ may be different sets, they may be regraded as indistinguishable(or equivalent) algebraically. Stated another way, $U$ and $V$ both represent just one underlying vector space but perhaps with different "labels" for the elements. The invertible linear transformation $T$ acts as a translation from one set of labels to another.

*Def.* **Isomorphic Vector Space**

If $U$ and $V$ are vector spaces and if $T : U \to V$ is an invertible linear transformation, then $U$ and $V$ are said to be **isomorphic vector space**. Also, an invertible transformation $T$ is called an **isomorphism**.

*Th.* If $U$ is a real $n$-dimensional vector space, then $U$ and $R^n$ are isomorphic.

> *Corollary.* If $U$ and $V$ are real $n$-dimensional vector spaces, then $U$ and $V$ are isomorphic.

# Change of Basis and Diagonalization

A linear transformation from $U$ to $V$ can be represented as an $(m \times n)$ matrix when $dim(U) = n$ and $dim(V) = m$. A consequence of this representation is that properties of transformations can be studied by examining their corresponding matrix representations.

For change of basis, we consider only transformations from $V$ to $V$. If we are interested in the properties of $T$, then it is reasonable to search for a basis for $V$ that makes the matrix representation of $T$ as simple as possible. Finding such a basis is the subject of this section.

> *Def.* **Diagonalizable Transformations**
>
> If $T$ is a linear transformation with a matrix representation that is diagonal, then $T$ is called diagonalizable.

> *Def.* **Eigenvalue and Eigenvector in General Vector Space Setting**
>
> A scalar $\lambda$ is called an **eigenvalue** for a linear transformation $T : V \to V$ provided that there is a **nonzero** vector $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \lambda \mathbf{v}$. $\mathbf{v}$ is called an **eigenvector** for $T$ corresponding to $\lambda$.

> *Th.* Let $V$ be an $n$-dimensional vector space. A linear transformation $T : V \to V$ is diagonalizable if and only if there exists a basis for $V$ consisting of eigenvectors for $T$.

> *Th.* **Change of Basis**
>
> Let $B$ and $C$ be bases for the vector space $V$, with $B = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$, and let $P_{n \times n}$ be the matrix given by $P = [\mathbf{P}_1, \mathbf{P}_2, \cdots, \mathbf{P}_n]$, where the $i$th column of $P$ is:
>
> $$\mathbf{P}_i = [\mathbf{u}_i]_C \tag{11}$$
>
> Then $P$ is nonsingular matrix and
>
> $$[\mathbf{v}]_C = P[\mathbf{v}]_B \tag{12}$$
>
> The matrix $P$ is called the **transition matrix**. Since $P$ is nonsingular, we have
>
> $$[\mathbf{v}]_B = P^{-1}[\mathbf{v}]_C \tag{13}$$
>
> *Proof.*: The linear transformation here is $T : V \to V$, or $T(\mathbf{v}) = \mathbf{v}$. Then the **The Representation Theorem** applies. For the fact that $P$ is nonsingular, we have:
>
> $$\sum_{i=1}^{n} a_i \mathbf{P}_i = \sum_{i=1}^{n} a_i [\mathbf{u}_i]_C = \sum_{i=1}^{n} [a_i \mathbf{u}_i]_C = [\theta]_C \tag{14}$$
>
> $$\text{And we have: } \sum_{i=1}^{n} a_i \mathbf{u}_i = \theta$$

### Matrix Representation and Change of Basis

Given a basis $B$, the relationship between the matrix representations of a linear transformation with respect to two different bases suggests how to determine a basis $C$ such that the matrix relative to $C$ is simpler matrix.

> *Th.* Let $B$ and $C$ be bases for the $n$-dimensional vector space $V$, and let $T : V \to V$ be a linear transformation. If $Q_1$ is the matrix of $T$ with respect to $B$ and if $Q_2$ is the matrix of $T$ with respect to $C$, then:
>
> $$Q_2 = P^{-1} Q_1 P \tag{15}$$
>
> where P is the transition matrix from $C$ to $B$.
>
> *Proof.* $\forall\ \mathbf{v} \in V$, we have
>
> $$Q_2[\mathbf{v}]_C = [\mathbf{v}]_C = P^{-1}[\mathbf{v}]_B = P^{-1}Q_1[\mathbf{v}]_B = P^{-1}Q_1P[\mathbf{v}]_C \tag{16}$$
>
> and $Q_2$ is unique, so
>
> $$Q_2 = P^{-1} Q_1 P \tag{17}$$

### Application in Eigenvalue

If $Q$ is similar to a diagonal matrix $R$. If we choose $\{\mathbf{S}_1, \mathbf{S}_2, \cdots, \mathbf{S}_n\}$ to be a basis of $R^n$ consisting of eigenvectors for $Q$, then

$$R = S^{-1}QS = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \tag{19}$$

where $d_i$ are eigenvalues for $Q$ and where $Q\mathbf{S}_i = d_i\mathbf{S}_i$.