# Lecture Notes on Linear Algebra

S. K. Panda

November 5, 2020

# Contents



Field:
set which has the two operations +, . defined such that they are
- associative
- commutative
- multiplicative and additive identity exist
- inverses exist (multiplicative only for non zero elements)
- distributive

Binary Operation:
Set S along with an operation (say *) such that a*b E S for all a,b E S
Written as (a,b)

Abelian Group:
Binary Operation with the following properties-
- Commutative
- Associative
- inverse and identity exist for all elements

Vector Space:
V is a set of vectors, F is a field (set of scalars)
- A vector space V(F) if there is a binary operation (V(F),+) such that it is abelian.

i.e a+b is closed, commutative, associative, inverse and identity exist

- scalar multiplication in closed

- k(a+b)=ka+kb, (k1+k2)a=k1a+k2a,(ab)k=a(bk) {distibutive}
  Ia=a if I is the multiplicative identity of F

# 1 Vector Space

## 1.1 Properties of Linearly Dependent and Independent sets

**Theorem 1.1**

Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Let $S = \{x_1, \cdots, x_n\} \subseteq \mathbb{V}$. Then $S$ is linearly dependent iff there exits $k$ s.t. $x_k$ is linear combination of $x_1, \cdots, x_{k-1}, x_{k+1}, \ldots, x_n$ (that is, there exists $x_k$ such that $x_k \in \textsc{ls}(S - \{x_k\})$.

*Proof.* Since $x_1, \cdots, x_n$ be linearly dependent, there exist scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ not all zero such that $\sum\limits_{i=1}^{n} \alpha_i x_1 = 0$. Without loss of generality we assume that $\alpha_k \neq 0$. Then $x_k = -\frac{1}{\alpha_k}(\sum\limits_{i=1, i\neq k}^{n} \alpha_i x_1)$. So $x_k$ is linear combination of $x_1, \cdots, x_{k-1}, x_{k+1}, \ldots, x_n$.

Conversely, there exits $k$ s.t. $x_k$ is a linear combination of $x_1, \cdots, x_{k-1}, x_{k+1}, \ldots, x_n$. That means $x_k = c_1x_1 + c_2x_2 + \cdots + c_{k-1}x_{k-1} + c_{k+1}x_{k+1} + \cdots + c_nx_n$ for some scalars $c_i \in \mathbb{F}$ for $i = 1, \ldots, k-1, k+1, \ldots, n$. Then $c_1x_1 + c_2x_2 + \cdots + c_{k-1}x_{k-1} - x_k + c_{k+1}x_{k+1} + \cdots + c_nx_n = 0$. This implies $S$ is linearly dependent. □

**Remark 1.1.** The above theorem gives you a technique to check whether a set is linearly dependent or not. That is, you just have to check whether a vector from that set is a linear combination of remaining vectors of $S$ or not.

**Remark 1.2.** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Then $\{\phi\}$ (empty set) is linearly independent.

*Proof.* If a finite set is linearly dependent, then there exists a vector in that set which is a linear combination of the remaining vectors. Since $\{\phi\}$ is empty, so there is no element which is a linear combination of the remaining vectors. Hence $\{\phi\}$ is linearly independent. □

The following theorem says something more.

**Theorem 1.2**

Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Let $S = \{x_1, \ldots, x_n\} \subseteq \mathbb{V}$ and $x_1 \neq 0$. Then $S$ is linearly dependent iff then $\exists k > 1$ s.t. $x_k$ is a linearly combination of $x_1, \ldots, x_{k-1}$.

*Proof.* Consider $\{x_1\}, \{x_1, x_2\}, \cdots, \{x_1, x_2, \cdots, x_n\}$ one by one. Take the smallest $k > 1$ s.t. $S_k = \{x_1, \cdots, x_k\}$ is linearly dependent. So $S_{k-1}$ is linearly independent(since $k$ is the smallest). In that case if $\sum_{i=1}^{k} \alpha_i x_i = 0$ $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ not all zero, then $\alpha_k \neq 0$. Otherwise $S_k$ will be linearly independent. Therefore $x_k$ is linear combination of $x_1, \ldots, x_{k-1}$.

There exists $k > 1$ such that $x_k$ is linear combination of $x_1, \ldots, x_{k-1}$. Therefore $x_k = c_1 x_1 + \cdots + c_{k-1} x_{k-1}$ where $c_1, c_2, \ldots, c_{k-1} \in \mathbb{F}$ not all zero. Then

$$c_1 x_1 + \cdots + c_{k-1} x_{k-1} - x_k + 0x_{k+1} + 0x_{k+2} + \cdots + 0x_n = 0$$

where $c_1, c_2, \ldots, c_{k-1}$ are not all zero. Hence $S$ is linearly dependent. $\square$

**Theorem 1.3**

Every subset of a finite linearly independent set is linearly independent.

*Proof.* Let $S$ be a linearly independent set. Let $S = \{x_1, \ldots, x_k\}$. Let $S_1 \subseteq S$. We have to show that $S_1$ is linearly independent. Assume that $S_1$ is linearly dependent. Then there exists a vector say $x_m$ in $S_1$ such that $x_m \in \text{LS}(S_1 - \{x_m\})$. Since $S_1 - \{x_m\} \subseteq S - \{x_m\}$. Then $x_m \in \text{LS}(S - \{x_m\})$. Hence $S$ is linearly independent, a contradiction. So our assumption $S_1$ is linearly dependent which is wrong. Therefore $S_1$ is linearly independent. $\square$

**Remark 1.3.** The above theorem says that if you will be able to find out a subset of a set $S$ which is linearly dependent, then $S$ is linearly dependent. But if you will find out a subset of $S$ which is linearly independent, then you can not conclude anything about $S$.

**Remark 1.4.** Every subset of a finite linearly dependent set need not be linearly dependent. In $\mathbb{R}^2$, the vectors $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, e_1, e_2$, are linearly dependent, as $x - e_1 - 2e_2 = 0$. The subset $\{e_1, e_2\}$ is linearly independent.

Till now we have seen linear independency and dependency for finite set. Next we are going to define linear independency and dependency for infinite set and we hire the concepts of linear independency and dependency of finite to define for infinite set.

**Definition 1.1.** Let $\mathbb{V}$ be a vector space and let $S \subseteq \mathbb{V}$ be infinite. We say $S$ is **linearly dependent** if it contains a finite linearly dependent set, otherwise it is **linearly independent**.

**Theorem 1.4**

Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Let $S = \{v_1, \ldots, v_n\} \subseteq \mathbb{V}$ and $T \subseteq \text{LS}(S)$ such that $m = |T| > |S|$. Then $T$ is linearly dependent.

*Proof.* Let $T = \{u_1, u_2, \ldots, u_m\}$. Since $u_j \in \text{LS}(S)$, there exists $A_{ij}$ in $\mathbb{F}$ such that $u_j = \sum_{i=1}^{n} A_{ij} v_i$ for $j = 1, \ldots, m$. For any $m$ scalars $x_1, x_2, \ldots, x_m$ we have

$$\begin{aligned} x_1u_1 + x_2u_2 + \cdots + x_mu_m &= \sum_{j=1}^{m} x_j u_j \\ &= \sum_{j=1}^{m} x_j \sum_{i=1}^{n} A_{ij} v_i \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} (A_{ij} x_j) v_i \\ &= \sum_{i=1}^{n} (\sum_{j=1}^{m} A_{ij} x_j) v_i \end{aligned}$$

There exist scalars $x_1, x_2, \ldots, x_m$ not all zero such that

$$\sum_{j=1}^{n} A_{ij} x_j = 0 \text{ for } i = 1, \ldots, n$$

(using system of homogeneous equations, here the co-efficient matrix is $A$ of size $n \times m$ and rank of $A$ is less than $m$ as $n < m$).

Therefore we have $x_1, x_2, \ldots, x_m$ in $\mathbb{F}$ not all zero such that $x_1u_1 + x_2u_2 + \cdots + x_mu_m = \sum_{j=1}^{m} x_ju_j = 0$. Hence $T = \{u_1, u_2, \ldots, u_m\}$ is linearly dependent. □

The above theorem is quite important. This theorem says that if you have a finite subset $S$ of a vector space containing $n$ elements. Then any finite subset of $\text{LS}(S)$ containing more than $n$ elements is linearly dependent. For example, consider the vector space $\mathbb{R}^3$. Take $S = \{(1,1,0),(1,0,0)\}$. Take $T = \{(1,1,0),(1,0,0),(3,1,0)\}$. You can easily check that $T \subseteq \text{LS}(S)$ and $T$ contains three elements. $T$ is linearly dependent.

**Corollary 1.1.** *Any $n+1$ vectors in $\mathbb{R}^n$ is linearly dependent.*

*Proof.* Follows as $\mathbb{R}^n = \text{LS}(e_1, \ldots, e_n)$. □

The following corollary is quite important. This theorem gives you a technique to extend a linearly independent set to a larger linearly independent set. This technique will be used frequently throughout this course. So keep it in your mind.

**Corollary 1.2.** *Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Let $S \subseteq \mathbb{V}$ be linearly independent and $x \in \mathbb{V} \setminus S$. Then $S \cup \{x\}$ is linearly independent iff $x \in \mathbb{V} - \text{LS}(S)$.*

**Remark 1.5.** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Let $S \subseteq (V)$. Then $\text{LS}(S)$ is the smallest subspace of $\mathbb{V}$ containing $S$.

*Proof.* Suppose $\text{LS}\, S$ is not the smallest subspace containing $S$. Let $W$ be a subspace containing $S$ such that $W \subsetneq \text{LS}(S)$. Then there exits $x$ in $\text{LS}(S)$ but not in $W$. Then there exist $x_1, \ldots, x_k \in S$ and $c_1, \ldots, c_k \in \mathbb{F}$ such that $x = c_1x_1 + c_2x_2 + \cdots + c_kx_k$. Since $W$ contains $S$, then $x \in W$ a contradiction. Hence $\text{LS}(S)$ is smallest subspace containing $S$. □

**Remark 1.6.** Let $\mathbb{V} = \{0\}$ be the vector space over $\mathbb{F}$. Then $\text{LS}(\phi) = \mathbb{V}$, here $\phi$ is empty set.

*Proof.* Using previous remark, $\text{LS}(\phi)$ is the smallest subspace of $\mathbb{V}$ containing $\phi$. Notice that $\mathbb{V}$ has exactly one subspace which is $\{0\}$. Hence $\text{LS}(\phi) = \{0\} = \mathbb{V}$. □

## 1.2 Properties of Basis and Dimension

**Theorem 1.5**

Every vector space has a basis.

*Proof.* Let $\mathbb{V}$ be any vector space over some field $\mathbb{F}$. Let $X$ be the set of all linearly independent subsets of $\mathbb{V}$.

The set $X$ is nonempty since the empty set is an independent subset of $\mathbb{V}$, and it is partially ordered by inclusion, which is denoted, as usual, by $\subseteq$.

Let $Y$ be a subset of $X$ that is totally ordered by $\subseteq$, and let $L_Y$ be the union of all the elements of $Y$ (which are themselves certain subsets of $\mathbb{V}$).

Since $(Y, \subseteq)$ is totally ordered, every finite subset of $L_Y$ is a subset of an element of $Y$, which is a linearly independent subset of $\mathbb{V}$, and hence $L_Y$ is linearly independent. Thus $L_Y$ is an element of $X$. Therefore, $L_Y$ is an upper bound for $Y$ in $(X, \subseteq)$ it is an element of $X$, that contains every element of $Y$.

As $X$ is nonempty, and every totally ordered subset of $(X, \subseteq)$ has an upper bound in $X$, Zorn's lemma asserts that $X$ has a maximal element. In other words, there exists some element $L_{max}$ of $X$ satisfying the condition that whenever $L_{max} \subseteq L$ for some element $L$ of $X$, then $L = L_{max}$.

It remains to prove that $L_{max}$ is a basis of $\mathbb{V}$. Since $L_{max}$ belongs to $X$, we already know that $L_{max}$ is a linearly independent subset of $\mathbb{V}$.

If there were some vector $w$ of $\mathbb{V}$ that is not in the span of $L_{max}$, then $w$ would not be an element of $L_{max}$ either. Let $L_w = L_{max} \cup \{w\}$. This set is an element of $X$, that is, it is a linearly independent subset of $\mathbb{V}$ (because $w$ is not in the span of $L_{max}$, and $L_{max}$ is independent). As $L_{max} \subseteq L_w$, and $L_{max} \neq L_w$

(because $L_w$ contains the vector $w$ that is not contained in $L_{max}$), this contradicts the maximality of $L_{max}$. Thus this shows that $L_{max}$ spans $\mathbb{V}$.

Hence $L_{max}$ is linearly independent and spans $\mathbb{V}$. It is thus a basis of $\mathbb{V}$, and this proves that every vector space has a basis. □

**Remark 1.7.** Basis of a vector space may not be unique.

**Theorem 1.6: Dimension theorem for vector spaces**

Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Any two bases of $\mathbb{V}$ have the same cardinality.

*Proof.* Let $A$ and $B$ be two bases of $\mathbb{V}$. Suppose $A$ is finite. We now show that $|A| = |B|$ ($|A|$ stands for cardinality of $A$). We consider $A$ is basis and $B$ is linearly independent, then using Theorem 1.1 $|B| \leq |A|$. Then we consider $B$ is basis and $A$ is linearly independent set, then using Theorem 1.1 $|A| \leq |B|$. Therefore $A$ and $B$ have same cardinality.
Consider $A$ is an infinite set. Then $|A| \geq \aleph_0$. We want to show that $|A| = |B|$. We just now show that $B$ is also infinite set as $A$ is infinite. Then $|B| \geq \aleph_0$. For each $a \in A$ can be expressed as a finite linear combination of elements of $B$, so let $B_a$ be the collection of these elements. Now, $B_a$ is uniquely determined by $a$, as $B$ is a basis. Also, $B_a$ is finite. Let

$$B' = \cup_{a \in A} B_a$$

Since $A$ spans $\mathbb{V}$, so does $B'$. If $B' \neq B$, pick $b \in B - B'$, so that $b$ is a linear combination of elements of $B'$. Moving $b$ to the other side of the expression and we have expressed 0 as a non-trivial linear combination of elements of $B$, contradicting the linear independence of $B$. Therefore $B' = B$. This means

$$|B| = |\cup_{a \in A} B_a| \leq \aleph_0 |A| = |A|$$

Similarly, every element in $B$ is expressible as a finite linear combination of elements in $A$, and using the same argument as above,

$$|A| \leq \aleph_0 ||B|$$

.

Therefore, $|A| = |B|$. □

**Definition 1.2.** Let $\mathbb{V}$ be a vector space. Then $\mathbb{V}$ is called **finite dimensional** if it has a basis $B$ which is finite. The dimension of $\mathbb{V}$ is the cardinality of $B$ and it is denote by $\dim(V)$. A vector space is infinite dimensional if it is not finite dimensional.

**Theorem 1.7**

Let $\mathbb{V}$ be a finite-dimensional vector space with $\dim(V) = n$. If $\mathbb{U}$ is a subspace of V , then $\dim(\mathbb{U}) \leq \dim(\mathbb{V})$, the equality holds if and only if $\dim(\mathbb{U}) = n$.

**Definition 1.3.** A subset $S \subseteq \mathbb{V}$ is called **maximal linearly independent** if
i) $S$ is linearly independent
ii) no proper super set of $S$ is linearly independent.

**Example 1.1.** The following examples must be verified by the reader.

1. In $\mathbb{R}^3$, the set $\{e_1, e_2\}$ is linearly independent but not maximal linearly independent.
2. In $\mathbb{R}^3$, the set $\{e_1, e_2, e_3\}$ is maximal linearly independent.
3. Let $S \subseteq \mathbb{R}^n$ be linearly independent and $|S| = n$. Then $S$ is maximal linearly independent.

**Theorem 1.8**

A set $S \subseteq \mathbb{V}$ is maximal linearly independent, then $\textsc{ls}(S) = \mathbb{V}$.

*Proof.* First we assume that $S$ is maximal linearly independent set. To show $\textsc{ls}(S) = \mathbb{V}$. Suppose $\textsc{ls}(S) \neq \mathbb{V}$. Then there exists $\alpha \in \mathbb{V}$ but not in $\textsc{ls}(S)$. Take $S_1 = S \cup \{\alpha\}$. Using Corollary 1.2, $S_1$ is linearly independent. A contradiction that $S$ is maximal linearly independent set. Hence $\textsc{ls}(S) = \mathbb{V}$. □

**Theorem 1.9**

A subset $S \subseteq \mathbb{V}$ is a basis of $\mathbb{V}$. Then $S$ is maximal linearly independent set.

*Proof.* Suppose that $S$ is not maximal. Then there exits a linearly independent set $S_1$ such that $S \subsetneq S_1$ and $S_1 \subseteq \textsc{ls}(S)$. By using Theorem 1.1, $S_1$ is dependent, a contradiction. Hence $S$ is maximal. □

**Remark 1.8.** It is clear that every basis is maximal linearly independent set.

**Remark 1.9.** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and let $B$ be a basis of $\mathbb{V}$. There exist unique $\alpha_1, \ldots, \alpha_k \in B$ and unique $c_1, \ldots, c_k \in \mathbb{F}$ such that $x = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_k\alpha_k$.

**Remark 1.10.** $\dim(\{0\}) = 0$. The zero vector space is spanned by the vector 0, but $\{0\}$ is a linearly dependent set and not a basis. For this reason, we shall agree that the zero vector space has dimension 0. Alternatively, we could reach the same conclusion by arguing that the empty set is a basis for the zero vector space. The empty set spans $\{0\}$, because the intersection of all subspaces containing the empty set is $\{0\}$, and the empty set is linearly independent because it contains no vectors.

**Theorem 1.10: Basis Deletion Theorem**

If $\mathbb{V} = \textsc{ls}(\{\alpha_1, \ldots, \alpha_k\})$. Then some $v_i$ can be removed to obtain a basis for $\mathbb{V}$.

*Proof.* Let $S = \{\alpha_1, \ldots, \alpha_k\}$. If $S$ is linearly independent, then $S$ is basis of $\mathbb{V}$. If $S$ is linearly dependent, then there exists $\alpha_i \in S$ such that $\alpha_i$ is linear combination of rest of the vectors of $S$. That is $\alpha_i = b_1\alpha_1 + \cdots + b_{i-1}\alpha_{i-1} + b_{i+1}\alpha_{i+1} + \cdots + b_k\alpha_k$ where $b_p \in \mathbb{F}$ for $p = 1, \ldots, i-1, i+1, \ldots, k$ and not all zero. We assume that $b_t \neq 0$. Take $S_1 = S - \{\alpha_t\}$. Notice that $\textsc{ls}(S_1) = \mathbb{V}$. If $S_1$ is linearly independent, then $S_1$ is basis of $\mathbb{V}$.

Continuing this process we get a linearly independent subset $S_p$ such that $\textsc{ls}(S_p) = \mathbb{V}$ because singleton set of non-zero element is linearly independent. ◻

consisting of $k - p = n$ vectors(since $n$ is finite existence of such $p$ is possible. Therefore $S_p$ is a basis of $\mathbb{V}$. ◻

The following is an application of Basis deletion theorem.

**Theorem 1.11**

Let $\mathbb{V}$ be a finite dimensional vector space and let $\dim \mathbb{V} = n$. Let $S = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{V}$ such that $\textsc{ls}(S) = \mathbb{V}$. Then $S$ is a basis for $\mathbb{V}$.

*Proof.* If we show that $S$ is linearly independent then we are done. If $S$ is linearly dependent, then by Deletion theorem, we can reduce $S$ to a basis for $\mathbb{V}$ which contains less than $n$ elements. A contradiction that $\dim(\mathbb{V}) = n$. Hence $S$ is basis of $\mathbb{V}$. ∎

**Theorem 1.12: Basis Extension Theorem**

Every linearly independent set of vectors in a finite-dimensional vector space $\mathbb{V}$ can be extended to a basis of $\mathbb{V}$.

*Proof.* Let $\dim(\mathbb{V}) = n$. Let $S = \{\alpha_1, \ldots, \alpha_k\}$ be a linearly independent set. If $\mathrm{LS}(S) = \mathbb{V}$, then $S$ is basis of $\mathbb{V}$. If $\mathrm{LS}(S) \neq \mathbb{V}$, then there exists a vector $\beta_1 \in \mathbb{V}$ but not in $\mathrm{LS}(S)$. Take $S_1 = S \cup \{\beta_1\}$. Using Corollary 1.2, the set $S_1$ is linearly independent. If $\mathrm{LS}(S_1) = \mathbb{V}$, then $S_1$ is basis for $\mathbb{V}$. If $\mathrm{LS}(S_1) \neq \mathbb{V}$, then there exists a vector $\beta_2 \in \mathbb{V}$ but not in $\mathrm{LS}\, S_1$. Take $S_2 = S_1 \cup \{\beta_2\}$. Using Corollary 1.2, the set $S_2$ is linearly independent. If $\mathrm{LS}(S_2) = \mathbb{V}$,then $S_2$ is basis for $\mathbb{V}$. If $\mathrm{LS}(S_2) \neq \mathbb{V}$, then there exists a vector $\beta_3 \in \mathbb{V}$ but not in $\mathrm{LS}\, S_2$. Take $S_3 = S_1 \cup \{\beta_3\}$. Continuing this process we get a linearly independent subset consisting of $k + p = n$ vectors (since $n$ is finite existence of such $p$ is possible. Therefore $S_p = \{\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_p\}$. We notice that $S_p$ is linearly independent and it contains $n$ vectors. So $S_p$ is basis of $\mathbb{V}$. ∎

The following is an application of Basis extension theorem.

**Theorem 1.13**

Let $\mathbb{V}$ be a finite dimensional vector space and $\dim \mathbb{V} = n$. Let $S = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{V}$ such that $S$ is linearly independent. Then $S$ is a basis of $\mathbb{V}$.

*Proof.* If we show that $\mathrm{LS}(S) = \mathbb{V}$, then we are done. If $\mathrm{LS}(S) \neq \mathbb{V}$, then by using Extension theorem, we can extend $S$ to be a basis for $\mathbb{V}$ and which contains at least $n + 1$ elements. A contradiction that $\dim(\mathbb{V}) = n$. Hence $S$ is a basis of $\mathbb{V}$. ∎

**Definition 1.4.** Let $\mathbb{V}$ be a finite dimensional vector spaces and let $B = \{x_1, \ldots, x_n\}$ be a basis. Let $x \in \mathbb{V}$. Then there exists unique $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$. Then $(\alpha_1, \ldots, \alpha_n)$ is called **co-ordinate** of $x$.

## 1.3 Properties of Subspaces

In this subsection we discuss some results related to subspace.

**Definition 1.5.** Let $\mathbb{V}$ be a vector space over $\mathbb{F}$ and $W \subseteq \mathbb{V}$. Then $W$ is said to be subspace of $\mathbb{V}$ if $W$ is a vector space over the filed $\mathbb{F}$ with respect to the **same** addition and scalar multiplication.

The following result supplies a way to check whether a subset is subspace or not.

**Theorem 1.14**

Let $\mathbb{V}(\mathbb{F})$ and $W \subseteq \mathbb{V}$ be non-empty. Then $W$ is subspace of $\mathbb{V}$ if and only if $\alpha x + \beta y \in W$ for all $x, y \in \mathbb{V}$ and for all $\alpha, \beta \in \mathbb{F}$.

**Remark 1.11.** It is known that arbitrary intersection of subspaces of a vector space again a subspace of that vector space. But union of two subspaces of a vector space need not be a subspace of that vector space, see the following example.

**Example 1.2.** Let $\mathbb{V} = \mathbb{R}^2$be a vector space over the filed $\mathbb{R}$. Let $\mathbb{U} = \{x \in \mathbb{R}^2 \mid x_2 = 0\}$ and let $\mathbb{V} = \{x \in \mathbb{R}^2 \mid x_1 = 0\}$. We can easily check that both are subspaces $\mathbb{R}^2$. Take $\mathbb{U} \cup \mathbb{V} = \{x \in \mathbb{R}^2 \mid x_1 x_2 = 0\}$ which is not a subspace.

**Exercise 1.1.** *Give two subspaces $\mathbb{U}, \mathbb{W}$ of $\mathbb{R}^3$ for each: a) $\mathbb{U} \cup \mathbb{W}$ is a subspace of $\mathbb{V}$. b) $\mathbb{U} \cup \mathbb{W}$ is not a subspace of $\mathbb{V}$.*

The following theorem supplies a necessary and sufficient condition to make union of two subspaces again a subspace.

**Theorem 1.15**

Let $\mathbb{U}, \mathbb{W}$ be two subspaces of $\mathbb{V}$. Then $\mathbb{U} \cup \mathbb{W}$ is a subspace of $\mathbb{V}$ if and only if either $\mathbb{U} \subseteq \mathbb{W}$ or $\mathbb{W} \subseteq \mathbb{U}$.

*Proof.* First we assume that $\mathbb{U} \cup \mathbb{W}$ is subspace of $\mathbb{V}$. To show either $\mathbb{U} \subseteq \mathbb{W}$ or $\mathbb{W} \subseteq \mathbb{U}$. Suppose that $\mathbb{U} \nsubseteq \mathbb{W}$ and $\mathbb{W} \nsubseteq \mathbb{U}$. Then there exist two vectors $x$ and $y$ in $\mathbb{V}$ such that $x \in \mathbb{U}$ but not in $\mathbb{W}$ and $y \in \mathbb{W}$ but not in $\mathbb{U}$. Therefore $x, y \in \mathbb{U} \cup \mathbb{W}$. Then $x + y \in \mathbb{U} \cup \mathbb{W}$. This implies that either $x + y \in \mathbb{U}$ or $\mathbb{W}$. If $x + y \in \mathbb{U}$, then $(x + y) - x = y \in \mathbb{U}$ which is not possible. If $x + y \in \mathbb{W}$, then $(x+y)-y = x \in \mathbb{W}$ which is not possible. Therefore either $\mathbb{U} \subseteq \mathbb{W}$ or $\mathbb{W} \subseteq \mathbb{U}$.

Conversely, $\mathbb{U} \subseteq \mathbb{W}$ or $\mathbb{W} \subseteq \mathbb{U}$. Then either $\mathbb{U} \cup \mathbb{W} = \mathbb{W}$ or $\mathbb{U} \cup \mathbb{W} = \mathbb{V}$. Hence $\mathbb{U} \cup \mathbb{W}$ is subspace of $\mathbb{V}$. $\square$

**Remark 1.12.** If $\mathbb{U}$ is a subspace of $\mathbb{W}$ and $\mathbb{W}$ is a subspace of $\mathbb{V}$, then $\mathbb{U}$ is a subspace of $\mathbb{V}$.

**Definition 1.6.** If $\mathbb{U}, \mathbb{W}$ are subspaces of $\mathbb{V}$, then $\mathbb{U} + \mathbb{W} := \{u + w \mid u \in \mathbb{U}, w \in \mathbb{W}\}$ and this is called the **sum** of $\mathbb{U}$ and $\mathbb{W}$.

**Theorem 1.16**

If $\mathbb{U}, \mathbb{W}$ are subspaces of $\mathbb{V}$. Then $\mathbb{U} + \mathbb{W}$ is the smallest subspace of $\mathbb{V}$ which contains both $\mathbb{U}$ and $\mathbb{W}$.

*Proof.* First we show that $\mathbb{U} + \mathbb{W}$ is subspace of $V$. Let $x, y \in \mathbb{U} + \mathbb{W}$. Then $x = x_1 + x_2$ for some $x_1 \in \mathbb{U}$ and $x_2 \in \mathbb{W}$ and $y = y_1 + y_2$ for some $y_1 \in \mathbb{U}$ and $y_2 \in \mathbb{W}$. Take $\alpha x + \beta y = \alpha(x_1 + x_2) + \beta(y_1 + y_2) = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2)$. Since $\mathbb{U}$ and $\mathbb{W}$ are subspaces of $\mathbb{V}$. Then $\alpha x_1 + \beta y_1 \in \mathbb{U}$ and $\alpha x_2 + \beta y_2 \in \mathbb{W}$. Therefore $\alpha x + \beta y \in \mathbb{U} + \mathbb{W}$. Hence $\mathbb{U} + \mathbb{W}$ is subspace of $\mathbb{V}$.

Now we show that $\mathbb{W}, \mathbb{U} \subseteq \mathbb{V}$. Let $x \in \mathbb{U}$. Then $x = x + 0$. It is clear that $x \in \mathbb{U} + \mathbb{W}$. Hence $\mathbb{U} \subseteq \mathbb{V}$. Similarly we can show that $\mathbb{W} \subseteq \mathbb{V}$.

Now we show that $\mathbb{U} + \mathbb{W}$ is the smallest subspace of $\mathbb{V}$ which contains both $\mathbb{U}$ and $\mathbb{W}$. Suppose that there is a subspace $\mathbb{Z}$ containing $\mathbb{U}$ and $\mathbb{W}$ such that $\mathbb{Z} \subsetneq \mathbb{U} + \mathbb{W}$. Then there exist a vector $x \in \mathbb{U} + \mathbb{W} \subseteq \mathbb{V}$ but not in $\mathbb{Z}$. Since $x \in \mathbb{U} + \mathbb{W} \subseteq \mathbb{V}$, $x = x_1 + x_2$ where $x_1 \in \mathbb{U}$ and $x_2 \in \mathbb{W}$. Notice that $x_1, x_2 \in \mathbb{Z}$. Therefore $x = x_1 + x_2 \in \mathbb{Z}$. This is a contradiction. Hence $\mathbb{U} + \mathbb{W}$ is the smallest subspace of $\mathbb{V}$ which contains both $\mathbb{U}$ and $\mathbb{W}$. ☐

**Theorem 1.17**

If $\mathbb{U}, \mathbb{W}$ are subspaces of a finite-dimensional vector space $\mathbb{V}$, then $\dim(\mathbb{U}+\mathbb{W}) = \dim(\mathbb{U}) + \dim(\mathbb{W}) - \dim(\mathbb{U} \cap \mathbb{W})$.

*Proof.* Since $\mathbb{V}$ is finite dimensional, $\mathbb{U}$ and $\mathbb{W}$ both are finite dimensional. Let $B = \{v_1, \ldots, v_k\}$ be a basis of $\mathbb{U} \cap \mathbb{W}$. By using Extension theorem we extend $B_1$ to a basis for $\mathbb{U}$ which is $\{v_1, \ldots, v_k, u_1, \ldots, u_m\}$ and basis for $\mathbb{W}$ which is $\{v_1, \ldots, v_k, w_1, \ldots, w_p, \}$.

Let $x \in \mathbb{U} + \mathbb{W}$. Then $x = x_1 + x_2$ for some $x_1 \in \mathbb{U}$ and $x_2 \in \mathbb{W}$. Therefore

$$x_1 = \sum_{i=1}^{k} a_i v_i + \sum_{j=1}^{m} b_j u_j$$

and

$$x_2 = \sum_{i=1}^{k} c_i v_i + \sum_{j=1}^{p} d_j w_j$$

.

Then $x = \sum_{i=1}^{k} a_i v_i + \sum_{j=1}^{m} b_j u_j + \sum_{i=1}^{k} c_i v_i + \sum_{j=1}^{p} d_j w_j$.

This implies $x \in \mathrm{LS}(\{v_1, \ldots, v_k, u_1, \ldots, u_m, v_1, \ldots, v_k, w_1, \ldots, w_p, \})$

We now show that $\{v_1, \ldots, v_k, u_1, \ldots, u_m, w_1, \ldots, w_p, \})$ is linearly independent.

To show that we take,

$$\sum_{i=1}^{k} f_i v_i + \sum_{j=1}^{m} g_j u_j + \sum_{l=1}^{p} h_j w_j = 0 \tag{1}$$

Therefore,

$$\sum_{i=1}^{k} f_i v_i + \sum_{j=1}^{m} g_j u_j = -\sum_{l=1}^{p} h_j w_j$$

This implies $\sum_{i=1}^{k} f_i v_i + \sum_{j=1}^{m} g_j u_j \in \mathbb{U} \cap \mathbb{W}$. Then $\sum_{i=1}^{k} f_i v_i + \sum_{j=1}^{m} g_j u_j \in \mathbb{U} \cap \mathbb{W} = \sum_{i=1}^{k} \alpha_i v_i$.

Therefore, $\sum_{i=1}^{k} (f_i - \alpha_i) v_i + \sum_{j=1}^{m} g_j u_j = 0$. Then $f_i - \alpha_i = 0$ for $i = 1, \ldots k$ and $g_j = 0$ for $j = 1, \ldots, m$ as $\{v_1, \ldots, v_k, u_1, \ldots, u_m\}$ is linearly independent. Put the values of $g_i$ in Equation (1), then

$$\sum_{i=1}^{k} f_i v_i + \sum_{l=1}^{p} h_l w_l = 0$$

.

Therefore $f_i = 0$ for $i = 1, \ldots, k$ and $h_j = 0$ for $j = 1, \ldots, p$ as $\{v_1, \ldots, v_k, w_1, \ldots, w_p, \}$. Hence $\{v_1, \ldots, v_k, u_1, \ldots, u_m, w_1, \ldots, w_p, \})$ is linearly independent.

Then $\{v_1, \ldots, v_k, u_1, \ldots, u_m, w_1, \ldots, w_p, \})$ is basis of $\mathbb{U} + \mathbb{W}$. Therefore $\dim(\mathbb{U} + \mathbb{W}) = k + m + P + k - k = \dim(\mathbb{U}) + \dim(\mathbb{W}) - \dim(\mathbb{U} \cap \mathbb{W})$. $\square$

**Definition 1.7.** When $\mathbb{U} \cap \mathbb{W} = \{0\}$, it is called the **internal direct sum** of $\mathbb{U}$ and $\mathbb{W}$. Notation: $\mathbb{U} \oplus \mathbb{W}$.

**Remark 1.13.** If $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$, then $\mathbb{W}$ is called **complement** of $\mathbb{U}$.

---

**Theorem 1.18**

Let $\mathbb{U}, \mathbb{W}$ be two subpaces of $\mathbb{V}$. Then $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ iff for each $v \in \mathbb{V}$, there exists unique $u \in \mathbb{U}$ and there exists unique $w \in \mathbb{W}$ such that $v = u + w$.

*Proof.* First we assume that $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$, that means, $\mathbb{U} \cap \mathbb{W} = \{0\}$. Let $x \in \mathbb{V}$. Then there exist $x_1 \in \mathbb{U}$ and $x_2 \in \mathbb{W}$ such that $x = x_1 + x_2$. Now we show that $x_1$ and $x_2$ are unique. Suppose there two vectors $y_1(\neq x_1) \in \mathbb{U}$ and $y_2(\neq y_1) \in \mathbb{W}$ such that $x = y_1 + y_1$. Then $x_1 + x_2 = y_1 = y_2$ this implies $x_1 - y_1 = y_2 - x_2$. Therefore $x_1 - y_1, y_2 - x_2 \in \mathbb{U} \cap \mathbb{W}$. This is only possible when $x_1 - y_1 = y_2 - x_2 = 0$. Then $x_1 = y_1$ and $x_2 = y_2$.

Conversely, $v \in \mathbb{V}$, there exists unique $u \in \mathbb{U}$ and there exists unique $w \in \mathbb{W}$ such that $v = u + w$. If we show that $\mathbb{U} \cap \mathbb{W} = \{0\}$, then we are done. Let $x \in \mathbb{U} \cap \mathbb{W}$. Then $x = x + 0$ where $x \in \mathbb{U}$ and $0 \in \mathbb{W}$, and $x = 0 + x$ where $0 \in \mathbb{U}$ and $x \in \mathbb{W}$. This is possible only when $x = 0$ otherwise the hypothesis is wrong. Then $\mathbb{U} \cap \mathbb{W} = \{0\}$. Hence $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$. □

---

The following definition says that how to create a new vector space from a two vector spaces over the same field.

**Definition 1.8.** Let $\mathbb{V}, \mathbb{W}$ be vector spaces over $\mathbb{F}$. $\mathbb{V} \times \mathbb{W}$ is Cartesian product of $\mathbb{V}$ and $\mathbb{W}$. Addition and and scalar multiplication on $\mathbb{V} \times \mathbb{W}$ defined by

$$(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2)$$

and

$$\alpha(v, w) = (\alpha v, \alpha w), \alpha \in \mathbb{F}$$

.

**Exercise 1.2.** *Then* $\mathbb{V} \times \mathbb{W}$ *is a vector space over* $\mathbb{F}$.

---

# 2 Inner Product Space

Just as the dot product on $\mathbb{R}^n$ helps us understand the geometry of Euclidean space with tools to detect angles and distances, the inner product can be used to understand the geometry of abstract vector spaces.

## 2.1 Definition and Example

Throughout this chapter we consider the filed $\mathbb{K}$ where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$.

**Definition 2.1**

Let $\mathbb{V}$ be a vector space over the field $\mathbb{K}$. An inner product is a mapping $\langle .,. \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{K}$ which satisfies the following conditions.

1. $\langle x, x \rangle \geq 0$ for all $x \in \mathbb{V}$ (**positivity**)

   and $\langle x, x \rangle = 0$ iff $x = 0$ (**definiteness**).

   <y,x> bar
2. $\langle x, y \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{V}$ (**conjugate symmetry**).
3. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ for all $x, y \in \mathbb{V}$ and for all $\alpha \in \mathbb{K}$ (**homogeneity**).
4. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ for all $x, y, z \in \mathbb{V}$ (**additivity**).

A vector space together with an inner product is called an inner product space.

**Example 2.1.** The following example must be verified by the reader.

1. Let $\mathbb{V} = \mathbb{R}^n(\mathbb{R})$. Let $\langle x, y \rangle = \sum_i^n x_i y_i$ for all $x, y \in \mathbb{V}$. Hence $(\mathbb{R}^n, \langle .,. \rangle)$ is an IPS.
2. Let $\mathbb{V} = \mathbb{C}^n(\mathbb{C})$. Let $\langle x, y \rangle = \sum_i^n x_i y_i$ for all $x, y \in \mathbb{V}$. Then this is not an IP on $\mathbb{C}^n(\mathbb{C})$.
3. Let $\mathbb{V} = \mathbb{C}^n(\mathbb{C})$. Let $\langle x, y \rangle = \sum_i^n x_i \overline{y_i}$ for all $x, y \in \mathbb{V}$. Then this an IP on $\mathbb{V}$.
4. $\mathbb{V} = \mathbb{M}_n(\mathbb{R})(\mathbb{R})$. Let $\langle A, B \rangle = \mathrm{tr}(AB^t)$ for all $A, B \in (M)_n(\mathbb{R})$.
5. $\mathbb{V} = \mathbb{C}[a, b](\mathbb{R})$. Let $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ for all $f, g \in \mathbb{C}[a, b]$. Then this an IP on $\mathbb{V}$.
6. $\mathbb{V} = \mathbb{C}(\mathbb{R})$ and $\mathbb{K} = \mathbb{R}$. Let $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ for all $f, g \in \mathbb{C}(\mathbb{R})$ where $a$ and $b$ are fixed real numbers. Then this is not an IP on $\mathbb{V}$.

7. $\mathbb{V} = \mathbb{M}_n(\mathbb{C})$ and $\mathbb{K} = \mathbb{C}$. The following is not an inner product on $\mathbb{V}$. $\langle A, B\rangle = \operatorname{tr}(AB^t)$. But $\langle A, B\rangle = \operatorname{tr}(A\bar{B}^t)$ is an inner product.

## 2.2 Existence of Inner Product

**Theorem 2.1**

There exists an inner product on every non-trivial vector space over $\mathbb{K}$.

*Proof.* Let $\mathbb{V}$ be a vector space over $\mathbb{K}$.

There are two cases.

**Case I.** $\mathbb{V}$ is finite dimensional. Let $B = \{u_1, \dots, k\}$ be a basis of $\mathbb{V}$.

Let $x, y \in \mathbb{V}$. Then $x = \sum_{i=1}^{k} a_i u_i$ and $y = \sum_{i=1}^{k} b_i u_i$ for some $a_i, b_i \in \mathbb{K}$ for $i = 1, \dots, k$.

Define

$$\langle x, y\rangle = \sum_{i=1}^{k} a_i \bar{b_i} \ .$$

We now show that $\langle , \rangle$ is an IP on $\mathbb{V}$.

i) Let $x \in \mathbb{V}$. The $x = \sum_{i=1}^{k} a_i u_i$ for some $a_i \in \mathbb{K}$ for $i = 1, \dots, k$.

$\langle x, x\rangle = \sum_{i=1}^{k} a_i \bar{a}_i = \sum_{i=1}^{k} |a_i|^2 \geq 0.$

if $\langle x, x\rangle = 0$, then $\sum_{i=1}^{k} |a_i|^2 = 0 \implies a_i = 0$ for $i = 1, \dots, k$.

Therefore $x = \sum_{i=1}^{k} a_i u_i = 0$.

ii) Let $x, y \in \mathbb{V}$. Then $x = \sum_{i=1}^{k} a_i u_i$ and $y = \sum_{i=1}^{k} b_i u_i$, $a_i, b_i \in \mathbb{K}$ for $i = 1, \dots, k$.

$$\langle x, y\rangle = \sum_{i=1}^{k} a_i\overline{b_i} = \overline{\sum_{i=1}^{k} \overline{a_i}b_i} = \overline{\langle y, x\rangle}.$$

iii) This part is same as part ii).

iv) Let $x, y, z \in \mathbb{V}$. Then $x = \sum_{i=1}^{k} a_i u_i$, $y = \sum_{i=1}^{k} b_i u_i$ and $z = \sum_{i=1}^{k} c_i u_i$ where $a_i, b_i, c_i \in \mathbb{K}$ for $i = 1, \ldots, k$.

Then $x + y = \sum_{i=1}^{k}(a_i + b_i)u_i$. Therefore,

$$\begin{aligned}
\langle x + y, z\rangle &= \sum_{i=1}^{k}(a_i + b_i)\overline{c_i} \\
&= \sum_{i=1}^{k}(a_i\overline{c_i} + b_i\overline{c_i}) \\
&= \langle x, z\rangle + \langle y, z\rangle.
\end{aligned}$$

**Case II.** $\mathbb{V}$ is infinite dimensional.

Let $B = \{u_\alpha : \alpha \in I\}$ be a basis of $\mathbb{V}$ where $I$ is an index set. Take $f : B \times B \to \mathbb{K}$ define by

$$f(u_\alpha, u_\beta) = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \neq \beta \end{cases}$$

Let $x, y \in \mathbb{V}$. Then there exist two finite subsets $\{u_{\alpha_1}, \ldots, u_{\alpha_k}\}$ and $\{u_{\beta_1}, \ldots, u_{\beta_m}\}$ of $B$ such that

$x = a_1u_{\alpha_1} + \cdots + a_ku_{\alpha_k}$ where $a_i \in \mathbb{K}$ for $i = 1, \ldots, k$.

$y = b_1u_{\beta_1} + \cdots + b_ku_{\beta_m}$ where $b_j \in \mathbb{K}$ for $j = 1, \ldots, m$.

Define $\langle x, y\rangle = \sum_{j=1}^{m}\sum_{i=1}^{k} a_i\overline{b_j}f(u_{\alpha_i}, u_{\beta_j})$.

Check $\langle x, y\rangle = \sum_{j=1}^{m}\sum_{i=1}^{k} a_i\overline{b_j}f(u_{\alpha_i}, u_{\beta_j})$ is an inner product on $\mathbb{V}$.

1. $\langle x, x\rangle = \sum_{i=1}^{k}\sum_{i=1}^{k} a_i\overline{a_j}f(u_{\alpha_i}, u_{\alpha_j})$.

$= \sum_{i=1}^{k} |a_i|^2$ (using the definition of $f$)

It is easy to show that $\langle x, x\rangle = 0$ if and only if $x = 0$.

2. It is easy.

3. It is easy

4. Let $x, y, z \in \mathbb{V}$. Then there exits two finite subsets $\{u_{\alpha_1}, \dots, u_{\alpha_k}\}$ $\{u_{\beta_1}, \dots, u_{\beta_m}\}$ and $\{u_{\gamma_1}, \dots, u_{\gamma_n}\}$ of $B$ such that

$x = a_1 u_{\alpha_1} + \cdots + a_k u_{\alpha_k}$ where $a_i \in \mathbb{K}$ for $i = 1, \dots, k$.

$y = b_1 u_{\beta_1} + \cdots + b_m u_{\beta_m}$ where $b_j \in \mathbb{K}$ for $j = 1, \dots, m$.

$z = c_1 u_{\gamma_1} + \cdots + c_n u_{\gamma_n}$ where $c_l \in \mathbb{K}$ for $l = 1, \dots, n$.

$x + y = \sum_{i=1}^{k} a_i u_{\alpha_i} + \sum_{j=1}^{m} b_j u_{\beta_j}$

$\langle x + y, z\rangle = \sum_{i=1}^{k} \sum_{l=1}^{n} a_i \overline{c_l} f(u_{\alpha_i}, u_{\gamma_l}) + \sum_{j=1}^{m} \sum_{l=1}^{n} b_i \overline{c_l} f(u_{\beta_i}, u_{\gamma_l})$.

$\langle x + y, z\rangle = \langle y, z\rangle + \langle y, z\rangle$.

**Remark 2.1.** 1. There exists an inner product on a non-trivial vector space over the filed $\mathcal{K}$.

2. There is more than one inner product can be defined on a non-trivial vector space over the filed $\mathcal{K}$.

**Definition 2.2**

A vector space together with an inner product is called an inner product space.

**Remark 2.2.** 1. A finite dimensional inner product space over the field $\mathcal{R}$ is called an Euclidean space.

2. An inner product space over the field $\mathbb{C}$ is called a unitary space.

## 2.3 Basic properties of an inner product

**Theorem 2.2**

1. $\langle 0, u\rangle = 0$ for every $u \in \mathbb{V}$.
2. $\langle u, 0\rangle = 0$ for every $u \in \mathbb{V}$.
3. $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for every $u, v, w \in \mathbb{V}$.
4. $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$ for all $u, v \in \mathbb{V}$ and $\lambda \in \mathbb{K}$.

1. $$\begin{aligned}\langle 0, u\rangle &= \langle u-u, u\rangle \\ &= \langle u, u\rangle + \langle -u, u\rangle \\ &= \langle u, u\rangle - \langle u, u\rangle \\ &= 0.\end{aligned}$$
2. $$\begin{aligned}\langle 0, u\rangle &= \overline{\langle 0, u\rangle} \\ &= \overline{\langle u-u, u\rangle} \\ &= \overline{\langle u, u\rangle + \langle -u, u\rangle} \\ &= \overline{\langle u, u\rangle - \langle u, u\rangle} \\ &= \overline{0} \\ &= 0. \; .\end{aligned}$$
3. Suppose $u, v, w \in \mathbb{V}$. Then
$$\begin{aligned}\langle u, v+w\rangle &= \overline{\langle v+w, u\rangle} \\ &= \overline{\langle v, u\rangle + \langle w, u\rangle} \\ &= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} \\ &= \langle u, v\rangle + \langle u, w\rangle.\end{aligned}$$

4. Suppose $\lambda \in \mathbb{K}$ and $u, v \in \mathbb{V}$. Then

$$\begin{aligned} \langle u, \lambda v\rangle &= \overline{\langle \lambda v, u\rangle} \\ &= \overline{\lambda\langle \lambda v, u\rangle} \\ &= \overline{\lambda}\overline{\langle v, u\rangle} \\ &= \overline{\lambda}\langle u, v\rangle. \end{aligned}$$

as desired.

## 2.4 Orthogonal and Orthonormal

**Definition 2.3**

Let $(\mathbb{V}, \langle ., .\rangle)$ be an inner product space. Let $u, v \in \mathbb{V}$. Then $u$ and $v$ is called **orthogonal** if $\langle u, v\rangle = 0$.

**Remark 2.3.** Orthogonality of two vectors depends on the inner product. That means, inner product plays an important role to decide whether two vectors are orthogonal or not. The following is an example.

Consider $\mathbb{V} = \mathbb{R}^2$. Take $\langle x, y\rangle = x_1y_1 + x_2y_2$ where $x = (x_1, x_2)$, $y = (y_1, y_2)$. This an inner product on $\mathbb{R}^2$. Take $e_1 = (1, 0)$ and $e_2 = (0, 1)$. Notice that these two vectors are orthogonal with respect to the above inner product. Consider $\langle x, y\rangle_1 = x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2$. This is an inner product on $\mathbb{R}^2$ (check!). Notice that $\langle e_1, e_2\rangle_1 = -1$. These two vectors are not orthogonal with respect to $\langle ., .\rangle_1$

**Definition 2.4**

Let $(\mathbb{V}, \langle ., .\rangle)$ be an inner product space. A subset $S$ of $\mathbb{V}$ is said to be **orthogonal** if $\langle u, v\rangle = 0$ for all $u, v \in S$.

**Theorem 2.3**

Let $(\mathbb{V}, \langle ., . \rangle)$ be an inner product space. Let $S = \{\alpha_1, \dots, \alpha_k\}$ be an orthogonal subset of $\mathbb{V}$ and let $\alpha_i \neq 0$ for $i = 1, \dots, k$. Then $S$ is linearly independent.

*Proof.* To show $S = \{\alpha_1, \dots, \alpha_k\}$ is LI.

Take $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_k\alpha_k = 0$.

Take inner product with $\alpha_i$ both side.

$\langle c_1\alpha_1 + c_2\alpha_2 + \cdots + c_k\alpha_k, \alpha_i \rangle = \langle 0, \alpha_i \rangle$

$\sum\limits_{j=1}^{k} c_i \langle \alpha_j, \alpha_i \rangle = 0.$

$c_i \langle \alpha_i, \alpha_i \rangle = 0$

$c_i = 0$. This is true for $i = 1, \dots, k$.

Hence $S = \{\alpha_1, \dots, \alpha_k\}$ is linearly independent. □

**Remark 2.4.**

1. The above result is true if $S$ is infinite set. That is, if $S$ is an infinite orthogonal set of non-zero element, then $S$ is linearly independent.

2. The converse of the above theorem is not true. That means, linearly independent set may not be an orthogonal set. See the following example. Let $(\mathbb{R}^2, \langle ., . \rangle)$ be an inner product space where $\langle x, y \rangle = \sum\limits_{i=1}^{2} x_i y_i$. Take $u = (1, 0)$ and $v = (1, 1)$. These two vectors are linearly independent but not orthogonal.

3. Let $(\mathbb{V}, \langle , \rangle)$ be a finite dimensional inner product space. Let $S$ be an infinite subset of $\mathbb{V}$. Then $S$ never be an orthogonal set.

4. Let $(\mathbb{V}, \langle , \rangle)$ be a finite dimensional inner product space. Let $S$ be a finite subset of $\mathbb{V}$. Then $S$ may or may not be orthogonal.

**Definition 2.5**

Let $(\mathbb{V}, \langle ., . \rangle)$ be an inner product space. Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\} \subseteq \mathbb{V}$. Then $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ is called **orthonormal** if $\langle \alpha_i, \alpha_j \rangle = 0$ for $i \neq j$ and $\langle \alpha_i, \alpha_i \rangle = 1$.

The next immediate question is that can we construct an orthogonal set from a finite linearly independent set. The answer is yes. Gram Schimdt supplied a process to construct an orthogonal set from a linearly independent finite set.

**Discussions:** Let $(\mathbb{V}, \langle ., . \rangle)$ be an inner product space.

Let $u_1, u_2, \dots, u_k \in \mathbb{V} - \{0\}$ be LI. How to construct orthogonal vectors $v_1, \dots, v_k$ using $u_1, u_2, \dots, u_k$?

**Ans:** Step 1. Take $v_1 = u_1$.

Step 2. We now construct $v_2$ using $v_1$ and $u_2$. Take $v_2 = u_2 + cv_1$ where $c \in \mathbb{K}$. We have to calculate the value of $c$ such that $v_2$ is orthogonal to $v_1$. That means

$\langle v_2, v_1 \rangle = \langle u_2, v_1 \rangle + c\langle v_1, v_1 \rangle$
$0 = \langle u_2, v_1 \rangle + c\langle v_1, v_1 \rangle.$
$c = -\frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle}.$
Therefore $v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} u$ is a vector which is perpendicular to $v_1$. You can easily check that $v_2 \neq 0$, otherwise $u_1$ and $u_2$ are linearly dependent which is not possible.

Step 3. We now construct $v_3$ using $v_1, v_2, u_3$.
Take $v_3 = u_3 + c_1 v_2 + c_2 v_1$ is an element in $\mathbb{V}$ where $c_1, c_2 \in \mathbb{K}$. We have to calculate the values of $c_1, c_2$ such that $v_3$ is orthogonal to $v_1$ and $v_2$. That means

$\langle v_3, v_1 \rangle = \langle u_3, v_1 \rangle + c_1\langle v_2, v_1 \rangle + c_2\langle v_1, v_1 \rangle$
$0 = \langle u_3, v_1 \rangle + c_1 \times 0 + c_2\langle v_1, v_1 \rangle.$
$c_2 = -\frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle}.$
Similarly, we have $c_1 = -\frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle}$. Therefore $v_3 = u_3 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1$ is a vector which is perpendicular to $v_1$ and $v_2$. You can easily check that $v_3 \neq 0$, otherwise $u_3$ is linear combination of $u_1$ and $u_2$ which is not possible.

Step 4. Similar way you can calculate $v_4$ using $u_4, v_1, v_2$ and $v_3$.

Step k. $v_k = u_k - \frac{\langle u_k, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \cdots - \frac{\langle u_k, v_{k-1} \rangle}{\langle v_{k-1}, v_{k-1} \rangle} v_{k-1}$. You can easily check that $v_k \neq 0$.

**Theorem 2.4: Gram Schmidt Orthogonalization Theorem**

Let $(\mathbb{V}, \langle ., . \rangle)$ be an inner product space. Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a linearly independent set. Then there exists an orthogonal set $\{\beta_1, \beta_2, \dots, \beta_n\}$ such that $\text{LS}(\{\beta_1, \beta_2, \dots, \beta_n\}) = \text{LS}(\{\alpha_1, \alpha_2, \dots, \alpha_n\})$.

*Proof.* Since $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a linearly independent set, $\alpha_i \neq 0$ for $i = 1, \dots, n$.

Step 1.

Put
$\beta_1 = \alpha_1$.

$\beta_2 = \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1$

$\vdots$

$\beta_n = \alpha_n - \frac{\langle \alpha_n, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1 - \cdots - \frac{\langle \alpha_n, \beta_{n-1} \rangle}{\langle \beta_{n-1}, \beta_{n-1} \rangle} \beta_{n-1}$.

Step 2.

It is easy to check $\langle \beta_i, \beta_j \rangle = 0$ for $i \neq j$ for $i, j \in \{1, \dots, n\}$.

Step 3.

To show $\text{LS}(\{\beta_1, \beta_2, \dots, \beta_n\}) = \text{LS}(\{\alpha_1, \alpha_2, \dots, \alpha_n\})$.

It is clear that $\alpha_i$ is a linear combination of $\beta_1, \dots, \beta_i$ for $i = 1, \dots, n$.
You can write each $\beta_i$ as a linear combination of $\alpha_1, \dots, \alpha_i$ for $i = 1, \dots, n$.
I just prove it for $\beta_3$ is a linear combination of $\alpha_1, \alpha_2, \alpha_3$.
We know that:

$\beta_1 = \alpha_1$.

$\beta_2 = \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1$

$\beta_2 = \alpha_2 + a_1 \alpha_1$ where $a_1 = -\frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle}$. Therefore $\beta_2$ is a linear combination of $\alpha_1$ and $\alpha_2$.

$\beta_3 = \alpha_3 - \frac{\langle\alpha_3,\beta_2\rangle}{\langle\beta_2,\beta_2\rangle}\beta_2 - \frac{\langle\alpha_3,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle}\beta_1.$

$\beta_3 = \alpha_3 + b_1\beta_2 + b_2\beta_1$ where $b_2 = -\frac{\langle\alpha_3,\beta_2\rangle}{\langle\beta_2,\beta_2\rangle}$ and $b_1 = \frac{\langle\alpha_3,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle}$.

$\beta_3 = \alpha_3 + b_1(\alpha_2 + a_1\alpha_1) + b_2\alpha_1$

Similar way you can show that $\beta_i$ is a linear combination of $\alpha_1, \dots, \alpha_i$ for $i = 1, \dots,$.

This follows that $\text{LS}(\{\beta_1, \beta_2, \dots, \beta_n\}) = \text{LS}(\{\alpha_1, \alpha_2, \dots, \alpha_n\})$. □

**Remark 2.5.** Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a linearly independent set. Then there exists an **orthonormal** set $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$ such that $\text{LS}(\{\gamma_1, \gamma_2, \dots, \gamma_n\}) = \text{LS}(\{\alpha_1, \alpha_2, \dots, \alpha_n\})$.

**Proof:** Using Gram Schimdt process we have an orthogonal set $\{\beta_1, \beta_2, \dots, \beta_n\}$ such that $\text{LS}(\{\beta_1, \beta_2, \dots, \beta_n\}) = \text{LS}(\{\alpha_1, \alpha_2, \dots, \alpha_n\})$.
Then put $\gamma_i = \frac{1}{\sqrt{\langle\beta_i,\beta_i\rangle}}\beta_i$ for $i = 1, \dots, n$.
Then $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$ is an orthonormal set. You can easily check that $\text{LS}(\{\gamma_1, \gamma_2, \dots, \gamma_n\}) = \text{LS}(\{\alpha_1, \alpha_2, \dots, \alpha_n\})$.

**Theorem 2.5**

Every non-trivial finite dimensional inner product space has an orthonormal basis.

*Proof.* We know that every finite dimensional vector space has a basis $B$. Then $B$ is linearly independent set. Using Remark 2.5, we can transform $B$ to an orthonormal set $B'$. Then $B'$ is linear independent and $\text{LS}(B') = \text{LS}(B)$. Hence $B'$ is an orthonormal basis of $\mathbb{V}$. □

**Theorem 2.6**

Let $(\mathbb{V}, \langle .,. \rangle)$ be an inner product space. Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be an orthonormal basis. Then for each $x \in \mathbb{V}$ we have $x = \sum_{i=1}^{n} \langle x, \alpha_i\rangle\alpha_i$

*Proof.* Since $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is an orthonormal basis. Let $x \in \mathbb{V}$. Then there exist $c_1, \ldots, c_n$ such that $x = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$.

Take $\langle x, \alpha_i \rangle = \langle c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n, \alpha_i \rangle$
$= c_1\langle \alpha_1, \alpha_i \rangle + \cdots + c_i\langle \alpha_i, \alpha_i \rangle + \cdots c_n\langle \alpha_n, \alpha_i \rangle$
$= c_i$ for $i = 1, \ldots, n$. □

## 2.5 Orthogonal Complement

**Definition 2.6**

Let $(\mathbb{V}, \langle ., . \rangle)$ be an inner product space. Let $S \subseteq \mathbb{V}$. Then the set $\{u \in \mathbb{V} \mid \langle u, v \rangle = 0 \text{ for all } v \in S\}$ is called **S perpendicular** of $S$ and it is denoted by $S^{\perp}$.

**Remark 2.6.** The following are true.

1. One can easily check that $S^{\perp}$ is a subspace of $\mathbb{V}$ for any $S \subseteq \mathbb{V}$.

2. $\{\phi\}^{\perp} = \mathbb{V}$.

**Theorem 2.7**

Let $(\mathbb{V}, \langle ., . \rangle)$ be a **finite dimensional** inner product space. Let $\mathbb{W}$ be a subspace of $\mathbb{V}$. Then $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^{\perp}$.

*Proof.* Let $\dim(\mathbb{V}) = n$ and let $\dim(\mathbb{W}) = k$. Let $B = \{u_1, \ldots, u_k\}$ be an orthonormal basis of $\mathbb{W}$. Using extension theorem and Gram Schmidt process we extend $B$ to an othonormal basis of $\mathbb{V}$ which is $\{u_1, \ldots, u_k, u_{k+1}, \ldots, u_n\}$. It is clear that $u_{k+1}, \ldots, u_n \in \mathbb{W}^{\perp}$.

We now show that $\mathbb{W}^{\perp} = \text{LS}\{u_{k+1}, \ldots, u_n\}$. It is clear that $\text{LS}\{u_{k+1}, \ldots, u_n\} \subseteq \mathbb{W}^{\perp}$ as $u_{k+1}, \ldots, u_n \in \mathbb{W}^{\perp}$ and $\mathbb{W}^{\perp}$ is a subspace. Let $x \in \mathbb{W}^{\perp}$. Then

$$x = \sum_{i=1}^{n} \langle x, x_i \rangle x_i$$
$$= \sum_{i=k+1}^{n} \langle x, x_i \rangle x_i; \text{ as } u_1, \ldots, u_k \in \mathbb{W} \text{ and } u_{k+1}, \ldots, u_n \in \mathbb{W}^{\perp}.$$

Hence $x \in \text{LS}\{u_{k+1}, \ldots, u_n\}$. Therefore $\mathbb{W}^{\perp} = \text{LS}\{u_{k+1}, \ldots, u_n\}$. It is clear that $\mathbb{W} \cap \mathbb{W}^{\perp} = \{0\}$. Hence $\mathbb{V} = \mathbb{W} \oplus \mathbb{W}^{\perp}$. □

**Theorem 2.8**

Let $(\mathbb{V}, \langle ., . \rangle)$ be a **finite dimensional** inner product space. Let $S$ is subspace of $\mathbb{V}$. Then $(S^\perp)^\perp = S$.

*Proof.* Let $\dim(\mathbb{V}) = n$. Using the proof of above theorem we have an orthonormal basis $\{u_1, \ldots, u_k, u_{k+}, \ldots, u_n\}$ of $\mathbb{V}$ such that $\{u_1, \ldots, u_k\}$ is a basis of $S$ and $\{u_{k+1}, \ldots, u_n\}$ is a basis of $S^\perp$.

We now show that $(S^\perp)^\perp = S$. Let $x \in S$. Then $\langle x, y \rangle = 0$ for all $y \in S^\perp$ $\Bigg(\text{using definition of } S^\perp\Bigg)$. Since $x$ is perpendicular to each vector of $S^\perp$. Hence $x \in (S^\perp)^\perp$. Therefore $S \subseteq (S^\perp)^\perp$.

Let $u \in (S^\perp)^\perp$. Then $\langle u, v \rangle = 0$ for all $v \in S^\perp$ $\Bigg(\text{using definition of } (S^\perp)^\perp\Bigg)$. Since $u \in \mathbb{V}$, then $u = c_1 u_1 + \cdots + c_k u_k + c_{k+1} u_{k+1} + \cdots + c_n u_n$. If we are able to show that $c_{k+i} = 0$ for all $i = 1, \ldots, n-k$, then $u \in S$.

$$\langle u, u_{k+i} \rangle = \langle c_1 u_1 + \cdots + c_k u_k + c_{k+1} u_{k+1} + \cdots + c_n u_n, u_{k+i} \rangle$$

$$0 = c_{k+i} \Bigg( u_{k+i} \text{ is an element in } S^\perp, \text{ hence } \langle u_j, u_{k+i} \rangle = 0 \text{ for } j = 1, \ldots, k \Bigg)$$

Hence $c_{k+i} = 0$ for $i = 1, \ldots, n-k$. Therefore $u \in S$. Then $(S^\perp)^\perp = S$. □

**Remark 2.7.** The above two theorems are not true for infinite dimensional inner product space.

**Theorem 2.9**

Let $(\mathbb{V}, \langle ., . \rangle)$ be a finite dimensional inner product space. Let $S$ and $T$ be two subspaces of $\mathbb{V}$. Then the following are true.

1. If $S \subseteq T \implies T^\perp \subseteq S^\perp$.
2. $(S+T)^\perp = S^\perp \cap T^\perp$ and $(S \cap T)^\perp = S^\perp + T^\perp$

**Remark 2.8.** Let $(\mathbb{V}, \langle ., . \rangle)$ be a finite dimensional inner product space. Let $\mathbb{W}$ be a subspace of $\mathbb{V}$. For each $x \in \mathbb{V}$ there exists unique $x_1 \in \mathbb{W}$ and $x_2 \in \mathbb{W}^{\perp}$ such that $x = x_1 + x_2$. The vector $x_1$ is called the **orthogonal projection** of $x$ into $W$.

**Theorem 2.10**

Let $(\mathbb{V}, \langle ., . \rangle)$ be a finite dimensional inner product space. Let $\mathbb{W}$ be a subspace of $\mathbb{V}$. Let $\{\alpha_1, \ldots, \alpha_k\}$ be an orthogonal basis of $\mathbb{W}$. Then the orthogonal projection of any vector $x \in \mathbb{V}$ on $\mathbb{W}$ is $\sum\limits_{i=1}^{k} \frac{\langle x, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \alpha_i$.

**Definition 2.7**

Let $\mathbb{V}$ be a vector space over the field $\mathbb{K}$. A map $||, || : \mathbb{V} \to \mathbb{R}$ is said to be a norm on $\mathbb{V}$ if it satisfies the following condition.

1. $||x|| \geq 0$ for all $x \in \mathbb{V}$ and $||x|| = 0$ if and only if $x = 0$.
2. $||\alpha x|| = |\alpha| ||x||$ for all $\alpha \in \mathbb{K}$ and for all $x \in \mathbb{V}$.
3. $||x + y|| \leq ||x|| + ||y||$ for all $x, y \in \mathbb{V}$.

A vector space to gather with a norm $||, ||$ is called a **normed linear space.**

**Example 2.2.** The following are examples of norm.

1. Let $x \in \mathbb{R}^n$. Then $||x|| = \sum\limits_{i=1}^{n} |x_i|$ is a norm on $\mathbb{R}^n$.
2. Let $x \in \mathbb{R}^n$. Then $||x|| = \max\limits_{1 \leq i \leq n} |x_i|$ is a norm on $\mathbb{R}^n$.

**Theorem 2.11: Cauchy Schwarz Inequality**

Let $(\mathbb{V}, \langle ., . \rangle)$ be an inner product space. Let $x, y \in \mathcal{V}$. Then $|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$. Equality holds if and only if $x$ and $y$ are linearly dependent.

*Proof.* If $y = 0$, then it is trivial.
If $y \neq 0$, take $x - ty$ where $t \in \mathbb{C}$. Then

$0 \leq \langle x - ty, x - ty \rangle$

$= \langle x, x\rangle + \langle x, -ty\rangle + \langle -ty, x\rangle + \langle -ty, -ty\rangle$

$= \langle x, x\rangle - \bar{t}\langle x, y\rangle - t\langle y, x\rangle + |t|^2\langle y, y\rangle$

Put $t = \frac{\langle x,y\rangle}{\langle y,y\rangle}$.

$\langle x, x\rangle - \frac{\bar{\langle x,y\rangle}}{\langle y,y\rangle}\langle x, y\rangle - \frac{\langle x,y\rangle}{\langle y,y\rangle}\langle y, x\rangle + |t|^2\langle y, y\rangle$

$= \langle x, x\rangle - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} + \frac{|\langle x,y\rangle|^2}{(\langle y,y\rangle)^2}\langle y, y\rangle$

$= \langle x, x\rangle - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle}$

$\langle x, x\rangle - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} \geq 0$

Then $|\langle x, y\rangle| \leq (\langle x, x\rangle)^{1/2}(\langle y, y\rangle)^{1/2}$.

The first inequality which we have used in this proof is $0 \geq \langle x - ty, x - ty\rangle$. If $|\langle x, y\rangle|^2 = \langle x, x\rangle\langle y, y\rangle$ hold, then $\langle x - ty, x - ty\rangle = 0$. This says that $x = ty$. Then $x$ and $y$ are linearly dependent. $\square$

**Remark 2.9.** Let $(\mathbb{V}, \langle ., .\rangle)$ be an inner product space. Let $x \in \mathbb{V}$. Then we can easily check that $||x|| = (\langle x, x\rangle)^{1/2}$ is a norm on $\mathbb{V}$.

# 3 Linear Transformation

**Definition 3.1**

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over the same field $\mathbb{F}$. A mapping $T : \mathbb{V} \to \mathbb{W}$ is said to be a linear transformation if $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for all $x, y \in \mathbb{V}$ and for all $\alpha, \beta \in \mathbb{F}$.

**Convention 3.1.** *If $T : \mathbb{V} \to \mathbb{W}$ is a linear transformation and $x \in \mathbb{V}$, then the $T(x)$ is usually denoted by $Tx$, i.e., $Tx := T(x)$ for all $x \in \mathbb{V}$.*

**Example 3.1.** The following are examples of linear transformation.

1. Let $T : \mathbb{R}^2 \to \mathbb{R}$ defined by $T(x, y) = x$. Then $T$ is a linear transformation.
2. Let $A \in \mathbb{M}_{m\times n}(R)$, let $T : \mathbb{R}^n \to \mathbb{R}^m$ be defined by $T(x) = Ax$, $x \in \mathbb{R}^n$. Then $T$ is a linear transformation.
3. Let $T : C[a, b] \to \mathbb{R}$ be defined by $T(f) = \int\limits_a^b f(x)dx$. Then $T$ is a linear transformation.
4. Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over the same field $\mathbb{F}$. Let $T : \mathbb{V} \to \mathbb{W}$ be defined by $T(x) = 0$, $x \in \mathbb{V}$. This transformation is called **Zero transformation**.
5. Let $\mathbb{V}$ be vector space over the field $\mathbb{F}$. Let $T : \mathbb{V} \to \mathbb{V}$ be defined by $T(x) = x$, $x \in \mathbb{V}$. This transformation is called **identity transformation**.
6. Let $\mathbb{V}$ and $\mathbb{W}$ be vector space over the field $\mathbb{F}$ and let $\lambda \in \mathbb{F}$. Let $T : \mathbb{V} \to \mathbb{V}$ be defined by $T(x) = \lambda x$, $x \in \mathbb{V}$. This transformation is called **scalar transformation**.

**Theorem 3.1**

Let $T : \mathbb{R}^n \to \mathbb{R}$ be a linear transformation. Then $T$ has the following form. $T(x_1, \ldots, x_n) = \sum\limits_{i=1}^{n} \alpha_i x_i$ for some $\alpha_i \in \mathbb{R}$ for $i = 1, \ldots, n$ and for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

*Proof.* Let $\{e_1, \ldots, e_n\}$ be a basis of $\mathbb{R}^n$. Let $T(e_i) = \alpha_i$ for $i = 1, \ldots, n$. Let $x = [x_1, x_2, \ldots, x_n] \in \mathbb{V}$. Then $x = \sum\limits_{i=1}^{n} x_i e_i$.

$$\begin{aligned} T(x) &= T(\sum_{i=1}^{n} x_i e_i) \\ &= \sum_{i=1}^{n} x_i T(e_i) \\ &= \sum_{i=1}^{n} \alpha_i x_i. \ \square \end{aligned}$$

**Theorem 3.2**

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exist linear transformations $T_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ such that $T(x) = (T_1(x), \ldots, T_m(x))$ for all $x \in \mathbb{R}^n$.

*Proof.* Let $\{e_1, \ldots, e_n\}$ be a basis of $\mathbb{R}^n$. Let $T(e_i) = (\alpha_1^i, \ldots, \alpha_m^i)$ for $i = 1, \ldots, n$. Let $x = [x_1, x_2, \ldots, x_n] \in \mathbb{V}$. Then $x = \sum\limits_{i=1}^{n} x_i e_i$.

$$\begin{aligned} T(x) &= T(\sum_{i=1}^{n} x_i e_i) \\ &= \sum_{i=1}^{n} x_i T(e_i) \\ &= \sum_{i=1}^{n} (\alpha_1^i, \ldots, \alpha_m^i) x_i \\ &= (\sum_{i=1}^{n} x_i \alpha_i^1, \ldots, \sum_{i=1}^{n} x_i \alpha_i^m) \end{aligned}$$

$T(x) = (T_1(x), \ldots, T_n(x))$, where $T_i(x) = \sum\limits_{i=1}^{n} x_i \alpha_i^i$. $\square$

**Theorem 3.3**

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over the field $\mathbb{F}$. Let $\{u_1, \ldots, u_n\}$ be an **ordered basis** of $\mathbb{V}$. Let $w_1, \ldots, w_n$ be any set (not necessarily distinct) vectors in $\mathbb{W}$. Then there is a unique linear transformation $T : \mathbb{V} \to \mathbb{W}$ such that $T(u_j) = w_j$, for $j = 1, \ldots, n$. Furthermore, if $\{w_1, \ldots, w_n\}$ is a basis of $\mathbb{W}$, then $T$ is bijective.

*Proof.* Let $x \in \mathbb{V}$. Then $x = \sum_{i=1}^{n} c_i u_i$.

Define $T(x) = \sum_{i=1}^{n} c_i w_i$. It is clear that $T$ is well defined because $x = \sum_{i=1}^{n} c_i u_i$, this expression unique.

We first show that $T$ is a linear transformation. Take $x, y \in \mathbb{V}$. Then $x = \sum_{i=1}^{n} c_i u_i$ and $y = \sum_{i=1}^{n} d_i u_i$.

Let $\alpha, \beta \in \mathbb{F}$. $T(\alpha x + \beta y) = T(\sum_{i=1}^{n} (\alpha c_i + \beta d_i) u_i)$.

$T(\alpha x + \beta y) = \sum_{i=1}^{n} (\alpha c_i + \beta d_i) w_i$.

$= \alpha \sum_{i=1}^{n} c_i w_i + \beta \sum_{i=1}^{n} d_i w_i$.

$= \alpha T(x) + \beta T(y)$.

Hence $T$ is linear.

Suppose that there is another linear transformation $U$ such that $U(u_i) = w_i$. To show that $U = T$. Let $x \in \mathbb{V}$. Then $x = \sum_{i=1}^{n} a_i u_i$. Using definition of $T$ we have $T(x) = T(\sum_{i=1}^{n} a_i u_i) = \sum_{i=1}^{n} a_i w_i$.

$U(x) = U(\sum_{i=1}^{n} a_i u_i)$

$= \sum_{i=1}^{n} a_i U(u_i)$ (applying the definition of linear transformation)

$= \sum_{i=1}^{n} a_i w_i$.

Then $U(x) = T(x)$ for all $x \in \mathbb{V}$. Hence $U = T$.

The last part is trivial. □

**Example 3.2.** 1. Take the basis $\{e_1, e_2, e_3\}$ in $\mathbb{R}^3$. Take $1, 2, 3 \in \mathbb{R}$. Then using previous theorem we have a unique linear transformation $T$ from $\mathbb{R}^3$ to $\mathbb{R}$ such that $T(e_1) = 1, T(e_2) = 2, T(e_3) = 3$ and $T(x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3$.

2. Take the basis $\{e_1, e_2, e_3\}$ in $\mathbb{R}^3$. Take $1, 2, 3 \in \mathbb{R}$. Then using previous theorem we have a unique linear transformation $T$ from $\mathbb{R}^3$ to $\mathbb{R}$ such that $T(e_1) = 2, T(e_2) = 1, T(e_3) = 3$ and $T(x_1, x_2, x_3) = 2x_1 + x_2 + 3x_3$. Since we change the images, so this

transformation is different from the previous transformation.

3. The previous theorem gives a technique to construct a linear transformation from a finite dimensional vector space to another vector space over the same filed $\mathbb{F}$.

The following is generalization of the above theorem.

**Theorem 3.4**

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over the field $\mathbb{F}$. Let $\{u_\alpha : \alpha \in I\}$ be an **ordered basis** of $\mathbb{V}$. Let $\{w_\alpha : \alpha \in I\} \subseteq \mathbb{W}$ be any set (not necessarily distinct) vectors in $\mathbb{W}$. Then there is a unique linear transformation $T : \mathbb{V} \to \mathbb{W}$ such that $T(u_\alpha) = w_\alpha$, for $\alpha \in I$. Furthermore, if $\{w_\alpha : \alpha \in I\}$ is a basis of $\mathbb{W}$, then $T$ is bijective.

*Proof.* The proof is similar to the proof of Theorem 3.3

**Definition 3.2**

Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation.

1. $Ker(T) := \{x \in \mathbb{V} : T(x) = 0\}$. This is called **Kernel** of $T$.
2. $R(T) := \{T(x) : x \in \mathbb{V}\}$.

**Remark 3.1.**

1. $Ker(T)$ is a subspace of $\mathbb{V}$. This subspaces is called **null space** of $T$ and sometimes it is denoted by $N(T)$.
2. $R(T)$ is a subspace of $\mathbb{W}$. The subspace $R(T)$ is called the **range space** of $T$.

**Definition 3.3**

The $\dim(R(T))$ is called the **rank** of $T$ and $\dim(Ker(T))$ is called the **nullity** of $T$.

**Example 3.3.** Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a map defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2, x_1 - x_3)$$

$Ker(T) := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : T(x_1, x_2, x_3) = 0\}$.

$T(x_1, x_2, x_3) = 0$

$\implies (x_1 - x_2, x_1 - x_3) = (0, 0)$

$\implies x_1 - x_2 = 0, x_1 - x_3 = 0$

$\implies x_1 = x_2, x_1 = x_3$

$\implies x_1 = x_2 = x_3 = k,\ k \in \mathbb{R}^3$

$\implies (x_1, x_2, x_3) = k(1, 1, 1)$

$\implies Ker(T) \subseteq \text{LS}(\{(1, 1, 1)\})$. Since $(1, 1, 1) \in Ker(T)$, then $Ker(T) = \text{LS}(\{(1, 1, 1)\})$.

Hence $Ker(T) = \{k(1, 1, 1) : k \in \mathbb{R}\}$ and nullity of $T$ is 1.

$R(T) := \{T(x) : x \in \mathbb{R}^3\}$. Let $y = (y_1, y_2) \in R(T)$. Then there exists $(x_1, x_2, x_3) \in \mathbb{R}^3$ such that $(y_1, y_2) = T(x_1, x_2, x_3)$.

$(y_1, y_2) = (x_1 - x_2, x_1 - x_3)$.

$(y_1, y_2) = (x_1 - x_2)(1, 0) + (x_1 - x_3)(0, 1)$

$(y_1, y_2) = k_1(1, 0) + k_2(0, 1)$

$\implies R(T) \subseteq \text{LS}(\{(1, 0), (0, 1)\})$.

Since $(1, 0), (0, 1) \in R(T)$. Hence $R(T) = \text{LS}(\{(1, 0), (0, 1)\})$

Notice that $\dim(R(T)) = 2$. Hence rank of $T$ is 2.

**Definition 3.4**

Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation.

1. $T$ is called one-one (injective) if $T(x_1) = T(x_2) \implies x_1 = x_2$ for all $x_1, x_2 \in \mathbb{V}$.

2. $T$ is called onto (surjective) if $T(\mathbb{V}) = \mathbb{W}$, that is $R(T) = \mathbb{W}$.

**Theorem 3.5**

Let $T : \mathbb{V} \to \mathbb{V}$ be a linear transformation. Then $T$ is one-one if and only if $Ker(T) = \{0\}$.

*Proof.* We first assume that $T$ is one-one. To show $Ker(T) = \{0\}$. Let $x \in ker(T)$. Then $T(x) = 0$ and we know that $T(0) = 0$. Since $T$ is one-one, then $x = 0$. Hence $Ker(T) = \{0\}$.

We now assume that $ker(T) = \{0\}$. To show $T$ is one-one. Let $T(x_1) = T(x_2)$. This implies $T(x_1 - x_2) = 0$. Hence $x_1 - x_2 \in Ker(T)$. Therefore $x_1 - x_2 = 0$. This implies $x_1 = x_2$. Hence $T$ is one-one. □

**Theorem 3.6**

Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $u_1, \ldots, u_n$ are in $\mathbb{V}$ such that $T(u_1), \ldots, T(u_n)$ are linearly independent in $\mathbb{W}$, then $u_1, \ldots, u_n$ are linearly independent in $\mathbb{V}$.

*Proof.* Given that $T(u_1), \ldots, T(u_n)$ are LI. To show $u_1, \ldots, u_n$ are linearly independent.

$c_1u_1 + \cdots + c_nu_n = 0_{\mathbb{V}}$.

$T(c_1u_1 + \cdots + c_nu_n) = 0_{\mathbb{W}}$

$c_1T(u_1) + \cdots + c_nT(u_n) = 0_{\mathbb{W}}$.

$c_1 = c_2 = \ldots = c_n = 0$ as $T(u_1), \ldots, T(u_n)$ are linearly independent. □

**Corollary 3.1.** *Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $\{u_\alpha : \alpha \in I\} \subseteq \mathbb{V}$ such that $\{T(u_\alpha) : \alpha \in I\}$ is an infinite linearly independent set in $\mathbb{W}$, then $\{u_\alpha : \alpha \in I\} \subseteq \mathbb{V}$ is also an infinite linearly independent set in $\mathbb{W}$.*

**Remark 3.2.** Converse of Theorem 3.5 is not true in general. That is, if $u_1, \ldots, u_n$ are linearly independent, then $T(u_1), \ldots, T(u_n)$ may not be linearly independent. The problem is that the image of some $u_i$ may be zero. Consider the following example.

$T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation defined by $T(x_1, x_2) = (x_1 - x_2, x_2 - x_1)$.

Take $u_1 = (1, 0)$ and $u_2 = (1, 1)$. Notice that $u_1, u_2$ are linearly independent.

But $T(u_1) = (1, -1)$ and $T(u_2) = (0, 0)$ are linearly dependent.

We have to put some condition on $T$ such that $T(u_i) \neq 0$ for $i = 1, \ldots, n$. We have seen that if $T$ is one-one, then the pre-image of $0_{\mathbb{W}}$ is exactly $0_{\mathbb{V}}$. So we have to put one-one condition on $T$.

**Theorem 3.7**

Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $T$ is one-one and $u_1, \ldots, u_n$ are linearly independent in $\mathbb{V}$, then $T(u_1), \ldots, T(u_n)$ are linearly independent in $\mathbb{W}$.

*Proof.* Given $u_1, \ldots, u_n$ are linearly independent and $T$ is one-one. To show that $T(u_1), \ldots, T(u_n)$ are LI.

$c_1T(u_1) + \cdots + c_nT(u_n) = 0_{\mathbb{W}}$.

$T(c_1u_1 + \cdots + c_nu_n) = 0_{\mathbb{W}}$.

$c_iu_1 + \cdots + c_nu_n = 0_{\mathbb{V}}$ as $T$ is one-one.

$c_1 = c_2 = \ldots = c_n = 0$ as $u_1, \ldots, u_n$ are linearly independent. □

**Corollary 3.2.** *Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. If $T$ is one-one and $\{u_\alpha : \alpha \in I\}$ is an infinite linearly independent set in $\mathbb{V}$, then $\{T(u_\alpha) : \alpha \in I\}$ is also an infinite linearly independent set in $\mathbb{W}$.*

**Theorem 3.8**

Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then the following are true.

1. If $B$ is a basis of $\mathbb{V}$, then $R(T) \overset{\text{= hoga}}{\subseteq} \text{LS}(T(B))$.

2. $\dim(R(T)) \leq \dim(\mathbb{V})$.

*Proof.* 1. Let $y \in R(T)$. Then there exists $x \in \mathbb{V}$ such that $T(x) = y$. Since $x \in \mathbb{V}$, then there exist unique $c_1, \ldots, c_k \in \mathbb{F}$ and $u_{\alpha_1}, \ldots, u_{\alpha_k} \in B$ such that $x = c_1u_{\alpha_1} + \cdots + c_ku_{\alpha_k}$.

$$\begin{aligned} y &= T(x) \\ &= T(c_1u_{\alpha_1} + \cdots + c_ku_{\alpha_k}) \\ &= c_1T(u_{\alpha_1}) + \cdots + c_kT(u_{\alpha_k}). \end{aligned}$$

This says that $y$ is a linear combination of $T(u_{\alpha_1}), \dots, T(u_{\alpha_k})$. Therefore $y \in \text{LS}\, T(B)$. Hence $R(T) \subseteq \text{LS}(T(B))$.

2. Using Theorem 3.5, we have $\dim(R(T)) \leq \dim(\mathbb{V})$. □

**Theorem 3.9: Rank Nullity Theorem**

Let $\mathbb{V}$ be a finite dimensional vector space over the field $\mathbb{F}$. Let $\mathbb{W}$ be a vector space over the field $\mathbb{F}$. Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then $Nullity(T) + Rank(T) = \dim(\mathbb{V})$.

*Proof.* Since $\mathbb{V}$ is finite dimensional, then $Ker(T)$ is finite dimensional. Let $\dim(\mathbb{V}) = n$ and let $Ker(T) = k$.

Let $\{u_1, \dots, u_k\}$ be a basis of $\dim(Ker(T))$.

Using extension theorem we extend $\{u_1, \dots, u_k\}$ to a basis of $\mathbb{V}$ which is $\{u_1, \dots, u_k, u_{k+1}, \dots, u_n\}$.

Let $y \in R(T)$. Then there exists $x \in \mathbb{V}$ such that $T(x) = y$.

Since $x \in \mathbb{V}$, then there exist unique $c_1, \dots, c_n \in \mathbb{F}$ such that $x = c_1u_1 + \cdots + c_nu_n$.

$$\begin{aligned} T(x) &= T(c_1u_1 + \cdots + c_nu_n) \\ &= c_1T(u_1) + \cdots + c_nT(u_n) = c_{k+1}T(u_{k+1}) + \cdots + c_nT(u_n). \end{aligned}$$

Each vector of $T(x)$ is a linear combination of $T(u_{k+1}), \dots, T(u_n)$ and $T(u_{k+1}), \dots, T(u_n) \in R(T)$. Hence $\text{LS}(\{T(u_{k+1}), \dots, T(u_n)\}) = R(T)$.

To show that $T(u_{k+1}), \dots, T(u_n)$ are LI. Take $a_1T(u_{k+1}) + \cdots + a_{n-k}T(u_n) = 0$.

$T(a_1u_{k+1} + \cdots + a_{n-k}u_n) = 0$.

Then $a_1u_{k+1} + \cdots + a_{n-k}u_n \in Ker(T)$.

$a_1u_{k+1} + \cdots + a_{n-k}u_n = b_1u_1 + \cdots + b_ku_k$

$a_1u_{k+1} + \cdots + a_{n-k}u_n - b_1u_1 - \cdots - b_ku_k = 0$

Since $\{u_1, \dots, u_k, u_{k+1}, \dots, u_n\}$ is basis of $\mathbb{V}$. Then $a_1 = \cdots = a_{n-k} = 0$.

Therefore $T(u_{k+1}), \dots, T(u_n)$ are LI. Hence $\{T(u_{k+1}), \dots, T(u_n)\}$ is a basis of $R(T)$. Then $\dim(R(T)) = n - k$.

Hence $\dim(\mathbb{V}) = nullity(T) + rank(T)$. $\square$

**Theorem 3.10**

Then the following are true.

1. Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Assume that $\dim(\mathbb{V}) = \dim(\mathbb{W})$. Then the following are equivalent.
   (a) $T$ is one-one.
   (b) $T$ is onto.
2. If $\dim(\mathbb{V}) < \dim(\mathbb{W})$, then there is no onto linear transformation from $\mathbb{V}$ to $\mathbb{W}$.
3. If $\dim(\mathbb{V}) > \dim(\mathbb{W})$, then there is no one-one linear transformation from $\mathbb{V}$ to $\mathbb{W}$.

*Proof.* Using rank-nullity theorem, we can prove. $\square$

**Remark 3.3.** Rank-nullity theorem is true only when $\mathbb{V}$ is finite dimensional. Otherwise there is no meaning of rank nullity theorem.

**Question:** Is there a linear transformation $T$ from an infinite dimensional vector space $\mathbb{V}$ to another vectors space $\mathbb{W}$ such that $rank(T)$ and $nullity(T)$ are finite?

**Answer:** Not possible to have such type of LT. Let $nullity(T) = k$ and $rank(T) = m$.

Let $\{u_1, \ldots, u_k\}$ be a basis of $Ker(T)$. We extend $\{u_1, \ldots, u_k\}$ to a LI set of $m + k$ vectors (this is possible as $\mathbb{V}$ is infinite dimensional) which is $\{u_1, \ldots, u_k, \ldots, u_{k+1}, \ldots, u_{m+k+1}\}$.

Let $\mathbb{S} = \text{LS}(\{u_1, \ldots, u_k, \ldots, u_{k+1}, \ldots, u_{m+k+1}\})$. So $\mathbb{S}$ is a subspace of $\mathbb{V}$ and $\dim(\mathbb{S}) = m + k + 1$.

$T_{\mathbb{S}}$ is a LT from $\mathbb{S}$ to $\mathbb{W}$. Then $Ker(T_{\mathbb{S}}) \subseteq Ker(T)$ and $R(T_{\mathbb{S}}) \subseteq R(T)$. $nullity(T_{\mathbb{S}}) \leq k$ and $rank(T_{\mathbb{S}}) \leq m$. Then $nullity(T_{\mathbb{S}}) + rank(T_{\mathbb{S}}) \leq k + m$ and $\dim(\mathbb{S}) = m + k + 1$. Then rank nullity theorem is not true on $\mathbb{S}$.

**Definition 3.5**

Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then $T$ is said to be isomorphism if $T$ is bijective (one-one+onto).

**Example 3.4.** 1. Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. Let $T : \mathbb{V} \to \mathbb{V}$ be defined by $T(x) = \alpha x$, $\alpha \neq 0$. Then $T$ is a linear isomorphism.

2. Let $\mathbb{V} = M_{n\times m}(\mathbb{R})$ be the set of $n \times m$ matrices with real entries and let $\mathbb{W} = \mathbb{R}^{mn}$. Define $T : \mathbb{V} \to \mathbb{W}$ by

   $T(A) = (a_{11}, \ldots, a_{1m}, a_{21}, \ldots, a_{2m}, \ldots, a_{n1}, \ldots, a_{nm})$. Here $A = (a_{ij}) \in \mathbb{V}$. Then $T$ is an isomorphism.

3. Let $\mathbb{V} = M_{n\times n}(\mathbb{R})$ be the set of $n \times n$ matrices with real entries. Define $T : \mathbb{V} \to \mathbb{R}$ by $T(A) = trace(A)$. Then $T$ is not a linear transformation as $T$ is not one-one.

4. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation define by $T(x_1, x_2) = (x_1, x_1 - x_2)$. Then $T$ is isomorphism.

**Definition 3.6**

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over the same field $\mathbb{F}$. Then $\mathbb{V}$ and $\mathbb{W}$ are said to be isomorphic if there is an isomorphism from $\mathbb{V}$ to $\mathbb{W}$.

**Example 3.5.** 1. $\mathbb{R}^n$ and $\mathbb{R}^m$ are isomorphic if and only if $m = n$.

2. $\mathbb{R}^{mn}$ are isomorphic to $\mathbb{V} = M_{n\times m}(\mathbb{R})$ , the set of $n \times m$ matrices with real entries.

3. $\mathbb{R}^n$ is isomorphic to $\mathbb{P}_n(x, \mathbb{R})$, set of all real polynomials of degree at most $n$.

**Theorem 3.11**

Let $\mathbb{V}$ and $\mathbb{W}$ be two finite dimensional vector spaces over the same field $\mathbb{F}$. Then $\mathbb{V}$ and $\mathbb{W}$ are isomorphic if and only if $\dim(\mathbb{V}) = \dim(\mathbb{W})$.

*Proof.* We first assume that $\mathbb{V}$ and $\mathbb{W}$ are isomorphic. Let $T$ be an isomorphism from $\mathbb{V}$ to $\mathbb{W}$. Let $\{u_1, \ldots, u_n\}$ be a basis of $\mathbb{V}$. Since $T$ is one-one, then using Theorem 3.7 $\{T(u_1), \ldots, T(u_n)\}$ is linearly independent. Using Theorem 3.8, $\mathbb{W} = R(T) = \text{LS}(\{T(u_1), \ldots, T(u_n)\})$. Hence $\{T(u_1), \ldots, T(u_n)\}$ is a basis of

$\mathbb{W}$. Then $\dim(\mathbb{V}) = \dim(\mathbb{W})$.

We now assume that $\dim(\mathbb{V}) = \dim(\mathbb{W})$. Let $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_n\}$ be two bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. Using Theorem 3.3, we have a linear transformation $T$ such that $T(u_i) = v_i$ for $i = 1, \ldots, n$. We can easily check that $T$ is bijective. □

The following is generalization of the above theorem for infinite dimensional case.

**Theorem 3.12**

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over $\mathbb{F}$. Then $\mathbb{V}$ is isomorphic to $\mathbb{W}$ if and only if cardinality of basis of $\mathbb{V}$ is equal to the cardinality of basis of $\mathbb{W}$.

*Proof.* We first assume that $\mathbb{V}$ is isomorphic to $\mathbb{W}$. Let $T$ be an isomorphism from $\mathbb{V}$ to $\mathbb{W}$. Let $B = \{u_\alpha : \alpha \in I\}$ be a basis of $\mathbb{V}$. Since $T$ is bijective, then $\{T(u_\alpha) : \alpha \in I\}$ is a basis of $\mathbb{W}$. Since $T$ is bijective, we have $card(B) = card(T(B))$.

We now assume that cardinality of basis of $\mathbb{V}$ is equal to the cardinality of basis of $\mathbb{W}$. Let $B$ and $B'$ be bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. Since $B$ and $B'$ have same cardinality, then there is a bijective map from $B$ to $B'$. Let $f$ be a bijective map from $B$ to $B'$. Let $B = \{u_\alpha : \alpha \in I\}$. Then $B = \{f(u_\alpha) : \alpha \in I\}$ Using Theorem 3.4, we have a bijective map $T$ such that $T(u_\alpha) = f(u_\alpha)$ for all $\alpha \in I$. Hence $\mathbb{V}$ and $\mathbb{W}$ are isomorphic. □

**Definition 3.7**

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces over the same field $\mathbb{F}$. Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then $T$ is invertible if $T$ is bijective (one-one+onto).

**Theorem 3.13**

Let $T : \mathbb{V} \to \mathbb{W}$ be an invertible linear transformation. Then $T^{-1}$ is also a bijective linear transformation from $\mathbb{W}$ to $\mathbb{V}$.

*Proof.* The proof is trivial. □

**Notation:** We use $\mathcal{L}(\mathbb{V}, \mathbb{W})$ to denote set of all linear transformation from $\mathbb{V}$ to $\mathbb{W}$.

**Theorem 3.14**

Let $S, T \in \mathcal{L}(\mathbb{V}, \mathbb{W})$ and $\alpha \in \mathbb{R}$. Then $S+T$ and $\alpha S$ defined by $(S+T)(x) = S(x) + T(x)$ (vector addition) and $\alpha S)x = \alpha S(x)$ (scalar multiplication) are again a linear maps. With these operations $\mathcal{L}(\mathbb{V}, \mathbb{W})$ is a vector space over the field $\mathbb{F}$.

*Proof.* The proof is straightforward. □

## 3.1 Dual Space

**Definition 3.8**

1. A linear transformation $T$ from $\mathbb{V}$ to $\mathbb{V}$ is called a **linear operator**.
2. A linear transformation $T$ from $\mathbb{V}$ to $\mathbb{F}$ is called a **linear functional**.

**Example 3.6.** The following are examples of linear operator and linear functional.

1. Let $T : \mathbb{M}_{n\times n} \to \mathbb{R}$ defined as $T(A) = \text{tr}(A)$, $A \in \mathbb{M}_{n\times n}$. Then $T$ is linear functional.
2. Let $T : C[0,1] \to \mathbb{R}$ defined as $T(f) = \int\limits_0^1 f(x)dx$. Then $T$ is linear functional.
3. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined as $T(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$. Then $T$ is linear operator.

**Definition 3.9**

The space $\mathcal{L}(\mathbb{V}, \mathbb{F})$ is called the **dual space** of $\mathbb{V}$ and it is denoted by $\mathbb{V}^*$. Elements of $\mathbb{V}^*$ are usually denoted by lower case letters $f$, $g$, etc.

**Theorem 3.15**

Let $\mathbb{V}$ be a finite dimensional space and $B = \{v_1, \ldots, v_n\}$ be an ordered basis of $\mathbb{V}$. For each $j \in \{1, \ldots, n\}$, let $f_j : \mathbb{V} \to \mathbb{F}$ be defined by $f_j(x) = \alpha_j$ for $x = \sum\limits_{j=1}^{n} \alpha_j v_j$. Then the following are true.

1. $f_1, \ldots, f_n$ are in $\mathbb{V}^*$ and they satisfy $f_i(v_j) = \delta_{ij}$ for $i, j \in \{1, \ldots, n\}$.
2. $\{f_1, \ldots, f_n\}$ is a basis of $\mathbb{V}^*$.

*Proof.*

1. We first show that $f_i(v_j) = \delta_{ij}$ for $i, j \in \{1, \ldots, n\}$. Here $\delta_{ij}$ is Kronecker delta. Since $\{v_1, \ldots, v_n\}$ is a basis of $\mathbb{V}$, then $v_j = 0v_1 + \cdots + 0v_{j-1} + v_j + 0v_{j+1} \cdots + 0v_n$ (this expression is unique). Using the definition of $f_i$, we have
$$f_i(v_j) = \begin{cases} 1 & \text{if } i = j) \\ 0 & \text{if} i \neq j \end{cases}$$

   We now show that $f_i \in \mathbb{V}^*$, that is, $f_i$ is a linear functional for $i = 1, \ldots, n$. Let $x, y \in \mathbb{V}$. Then

   $x = a_1 v_1 + \cdots + c_n v_n$ (unique expression) and

   $y = b_1 v_1 + \cdots + c_n v_n$ (unique expression).

   Using definition of $f_i$ we have, $f_i(x) = a_i$ and $f_i(y) = b_i$. Let $\alpha, \beta \in \mathbb{F}$. Then $\alpha x + \beta y = (\alpha a_1 + \beta b_1)v_1 + \cdots + (\alpha a_n + \beta b_n)v_n$.

   Using definition of $f_i$ we have

   $f_i(\alpha x + \beta y)$

   $= \alpha a_i + \beta b_i$

   $= \alpha f_i(x) + \beta f(y)$.

   Hence $f_i$ is a linear transformation from $\mathbb{V}$ to $\mathbb{F}$ for $i = 1, \ldots, n$. We have proved that $f_i \in \mathbb{V}^*$ for $i = 1, \ldots, n$.

2. We mow show that $\{f_1, \ldots, f_n\}$ is a basis of $\mathbb{V}^*$. We first show that $\{f_1, \ldots, f_n\}$ is linearly independent.

   $c_1 f_1 + c_2 f_2 + \cdots + c_n f_n = 0$.

   $(c_1 f_1 + c_2 f_2 + \cdots + c_n f_n)v_1 = 0(v_1)$.

$c_1 f_1(v_1) + \cdots + c_n f_n(v_1) = 0.$

$c_1 = 0.$

Similarly we can show that $c_2 = \cdots = c_n = 0$. Hence $\{f_1, \ldots, f_n\}$ is linearly independent.

We now show that $\text{LS}(\{f_1, \ldots, f_n\}) = \mathbb{V}^*$. Let $f \in \mathbb{V}^*$. Let $f(v_i) = c_i$ for $i = 1, \ldots, n$ where $c_1, \ldots, c_n \in \mathbb{F}$.

Let $x \in \mathbb{V}$. Then $x = b_1 v_1 + \cdots + b_n v_n$ (unique expression).

$f(x) = f(b_1 v_1 + \cdots + b_n v_n)$

$= b_1 f(v_1) + \cdots + b_n f(v_n)$

$= c_1 b_1 + \cdots + c_n b_n$

$= c_1 f_1(x) + \cdots + c_n f(x)$

$f(x) = (c_1 f_1 + \cdots + c_n f)(x)$ for all $x \in \mathbb{V}$. Therefore $f = c_1 f_1 + \cdots + c_n f$. Then $\text{LS}(\{f_1, \ldots, f_n\}) = \mathbb{V}^*$. Hence $\{f_1, \ldots, f_n\}$ is basis of $\mathbb{V}^*$. $\square$

**Definition 3.10**

Let $\mathbb{V}$ be a finite dimensional space and $B = \{v_1, \ldots, v_n\}$ be an ordered basis of $\mathbb{V}$. A basis $\{f_1, \ldots, f_n\}$ of $\mathbb{V}^*$ such that $f_i(v_j) = \delta_{ij}$ for $i, j \in \{1, \ldots, n\}$. Then $\{f_1, \ldots, f_j\}$ is called **dual basis** of $\mathbb{V}^*$.

**Example 3.7.** Let $\mathbb{V} = \mathbb{R}^2$. Let $B = \{(1,0), (0,1)\}$ be an ordered basis of $\mathbb{V}$. Then find the dual basis of $\mathbb{V}^*$ corresponding $B$.

**Sol:** Let $\{f_1, f_2\}$ be the dual basis of $\mathbb{V}^*$ corresponding to $B$. Therefore $f_1(x, y) = \alpha_1 x + \alpha_2 y$ and $f_2 = \beta_1 x + \beta_2 y$ where $\alpha_1, \alpha_2 . \beta_1, \beta_2 \in \mathbb{R}$. Applying the definition of dual basis we have,

$f_1(1,0) = 1 \implies \alpha_1 = 1$ and $f_1(0,1) = 0 \implies \alpha_2 = 0$

Similarly $f_2(1,0) = 0 \implies \beta_1 = 0$ and $f_2(0,1) = 0 \implies \beta_2 = 1$.

Hence $f_1(x,y) = x$ and $f_2(x,y) = y$

**Theorem 3.16**

If $\mathbb{V}$ is finite dimensional, then $\mathbb{V}$ and $\mathbb{V}^*$ are isomorphic.

*Proof.* Using above theorem we have $\dim(\mathbb{V}) = \dim(\mathbb{V}^*)$. Then using Theorem 3.7, $\mathbb{V}$ is isomorphic to $\mathbb{V}^*$. □

## 3.2 Matrix Representation

Let $\mathbb{V}$ and $\mathbb{W}$ be two finite dimensional vector spaces over the same field $\mathbb{F}$ and let $B_1 = \{u_1, \ldots, u_n\}$ and $B_2 = \{v_1, \ldots, v_m\}$ be two ordered bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. $T(u_j) \in \mathbb{W}$. Then there exist unique $a_{ij} \in \mathbb{F}$ for $i = 1, \ldots, m$ such that $T(u_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{mj}v_m$ for $j = 1, \ldots, n$.

Let $x \in \mathbb{V}$. There exist unique $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $x = \sum\limits_{j=1}^{n} \alpha_j u_j$. Then

$$\begin{aligned}
T(x) &= T(\sum_{j=1}^{n} \alpha_j u_j) \\
&= \sum_{j=1}^{n} \alpha_j T(u_j) \\
&= \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{m} a_{ij} v_i \\
&= \sum_{i=1}^{m} \Big( \sum_{j=1}^{n} \alpha_j a_{ij} \Big) v_i.
\end{aligned}$$

$[T(x)]_{B_2} = A[x]_{B_1}$ where $A = [a_{ij}]_{m \times n}$. That is co-ordinate of $T(x)$ with respect to the basis $B_2$ is $[T(x)]_{B_2}$ which can be calculated using the co-ordinate of $x$ with respect to basis $B_1$.

Here $A$ is called the matrix representation of $T$ with respect to the bases $B_1$ and $B_2$ and it is denoted by $[T]_{B_1 B_2}$.

**Example 3.8.** Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-z \end{bmatrix}$. Let $B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ and $B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ be bases of $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively.

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\-1\end{bmatrix} = 0\cdot\begin{bmatrix}1\\1\end{bmatrix} - 1\cdot\begin{bmatrix}0\\1\end{bmatrix}$$

$A = \begin{bmatrix}1 & 1 & 0\\0 & -2 & -1\end{bmatrix}$ is the matrix representation of $T$ with respect to the given bases.

Let $\begin{bmatrix}1\\1\\1\end{bmatrix} \in \mathbb{R}^3$. Then $\begin{bmatrix}1\\1\\1\end{bmatrix} = 1\cdot\begin{bmatrix}1\\0\\0\end{bmatrix} + 1\cdot\begin{bmatrix}0\\1\\1\end{bmatrix} + 0\cdot\begin{bmatrix}0\\0\\1\end{bmatrix}$.

The co-ordinate of $\begin{bmatrix}1\\1\\1\end{bmatrix}$ with respect to $B_1$ is $\begin{bmatrix}1\\1\\0\end{bmatrix}$.

$$\left[T\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right)\right]_{B_2} = A\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right)_{B_1} = \begin{bmatrix}1 & 1 & 0\\0 & -2 & -1\end{bmatrix}\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}2\\-2\end{bmatrix}$$

$$T\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) = 2\cdot\begin{bmatrix}1\\1\end{bmatrix} + -2\cdot\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$$

**Definition 3.11**

The matrix $A = (a_{ij})$ in the above discussion is called the **matrix representation** of $T$ with respect to the ordered bases $B_1$ and $B_2$ of $\mathbb{V}$ and $\mathbb{W}$, respectively. This matrix is usually denoted by $[T]_{B_1B_2}$, that is, $[T]_{B_1B_2} = (a_{ij})$.

**Remark 3.4.** The matrix representation of $T$ is unique up to $B_1$ and $B_2$. If you change the bases $B_1$ and $B_2$ you will have another matrix representation.

**Remark 3.5.** Let $\mathbb{V}$ and $\mathbb{W}$ be finite dimensional vector spaces over the same field $\mathbb{F}$. Let $\dim(\mathbb{V}) = n$ and $\dim(\mathbb{W}) = m$. Assume that $B = \{u_1, \ldots, u_n\}$ and $B' = \{v_1, \ldots, v_m\}$ are ordered basis of $\mathbb{V}$ and $\mathbb{W}$ respectively.

1. We have seen that for each linear transformation $T : \mathbb{V} \to \mathbb{W}$, we have a matrix $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ such that $[T]_{BB'} = A$.

2. Let $A \in \mathbb{M}_{m\times n}(\mathbb{F})$. Then there exists a linear transformation $T : \mathbb{V} \to \mathbb{W}$ such that $A = [T]_{BB'}$ and such linear transformation is $T(u_j) = \sum_{i=1}^{m} a_{ij} v_i$ for $j = 1 \ldots, n$.

3. Let $T, S : \mathbb{V} \to \mathbb{W}$ be two linear transformation. Let $B_1$ and $B_2$ be two bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. Then $[T + S]_{B_1B_2} = [T]_{B_1B_2} + [S]_{B_1B_2}$ and $[\alpha T]_{B_1B_2} = \alpha[T]_{B_1B_2}$.

**Theorem 3.17**

Let $\mathbb{V}$ and $\mathbb{W}$ be finite dimensional vector spaces over the same field $\mathbb{F}$. Let $\dim(\mathbb{V}) = n$ and $\dim(\mathbb{W}) = m$. Then $\mathcal{L}(\mathbb{V}, \mathbb{W})$ is isomorphic to $\mathbb{M}_{m\times n}(\mathbb{F})$.

*Proof.* Define $\zeta : \mathcal{L}(\mathbb{V}, \mathbb{W}) \to \mathbb{M}_{n\times m}(\mathbb{F})$ such that $\zeta(T) = [T]_{BB'}$ where $B = \{u_1, \ldots, u_n\}$ and $B' = \{v_1, \ldots, v_m\}$ are ordered bases of $\mathbb{V}$ and $\mathbb{W}$ respectively. Using previous remark it is cleared that $\zeta$ is a linear transformation from $\mathcal{L}(\mathbb{V}, \mathbb{W})$ to $\mathbb{M}_{m\times n}(\mathbb{F})$. We now show that $\zeta$ is bijective.

Let $T \in Ker(\zeta)$. Then $\zeta(T) = 0_{m\times n}$. This implies that $[T]_{BB'} = 0_{m\times n}$. This implies $[T(x)]_{B'} = 0_{m\times 1}$ with respect to $B'$ is zero, co-ordinate of $T(x)$ for each $x \in \mathbb{V}$. Hence $T(x) = 0_{\mathbb{W}}$ for each $x \in \mathbb{V}$. Then $T = 0$. Therefore $Ker(\zeta) = \{0\}$.

We now show that $\zeta$ is onto. Let $A \in \mathbb{M}_{m\times n}$. Define $T(u_j) = \sum_{i=1}^{m} a_{ij} v_i$ for $j = 1 \ldots, n$. It is clear that $[T]_{BB'} = A$. Hence $\zeta$ is onto.

Therefore $\mathcal{L}(\mathbb{V}, \mathbb{W})$ is isomorphic to $\mathbb{M}_{m\times n}(\mathbb{F})$. □

**Corollary 3.3.** *Let $\mathbb{V}$ and $\mathbb{W}$ be finite dimensional vector spaces over the same field $\mathbb{F}$. Let $\dim(\mathbb{V}) = n$ and $\dim(\mathbb{W}) = m$. Then dimension of $\mathcal{L}(\mathbb{V}, \mathbb{W}) = mn$.*

**Theorem 3.18**

Let $\mathbb{V}$ be a finite dimensional vector space over the same field $\mathbb{F}$. Let $S$ and $T$ be two linear transformations from $\mathbb{V}$ and to $\mathbb{V}$. Let $B$ be an ordered basis of $\mathbb{V}$. Then $[S \circ T]_{BB} = [S]_{BB}[T]_{BB}$.

*Proof.* Let $B = \{v_1, \dots, v_n\}$. Let $[T]_{BB} = A$ and $[S]_{BB} = C$.

Then $T(v_i) = a_{1i}v_1 + a_{2i}v_2 + \cdots + a_{ni}v_n$ for $i = 1, \dots, n$.

$S(v_i) = b_{1i}v_1 + b_{2i}v_2 + \cdots + b_{ni}v_n$ for $i = 1, \dots, n$.

$(S \circ T)(v_1) = S(T(v_1))$

$= S(a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n)$

$= a_{11}S(v_1) + a_{21}S(v_2) + \cdots + a_{n1}S(v_n)$

$= a_{11}(b_{11}v_1 + b_{21}v_2 + \cdots + b_{n1}v_n) + a_{21}(b_{12}v_1 + b_{22}v_2 + \cdots + b_{n2}v_n) + \cdots + a_{n1}(b_{1n}v_1 + b_{2n}v_2 + \cdots + b_{nn}v_n)$

$= (a_{11}b_{11} + a_{21}b_{12} + \cdots a_{n1}b_{1n})v_1 + \cdots + (a_{11}b_{n1} + a_{21}b_{n2} + \cdots a_{n1}b_{nn})v_n$

$$= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}$$

$$\big[(S \circ T)(u_i)\big] = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix}$$

$$\big[S \circ T\big]_{BB} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \quad \square$$

**Remark 3.6.** Let $\dim V = n$ and $\dim W = m$. Let $T : \mathbb{V} \to \mathbb{W}$ and $S : \mathbb{W} \to \mathbb{V}$ be two linear transformation. Let $B$ and $B'$ be bases of $\mathbb{V}$ and $\mathbb{W}$, respectively. Then $\big[S \circ T\big]_{BB'} = [S]_{BB'}[T]_{BB'}$.

**Theorem 3.19**

Let $\mathbb{V}$ be a finite dimensional vector space over the same field $\mathbb{F}$. Let $T$ be an invertible linear transformation from $\mathbb{V}$ and to $\mathbb{V}$. Let $B$ be an ordered basis of $\mathbb{V}$. Then $[T^{-1}]_{B,B} = [T]^{-1}_{B,B}$.

## 3.3 Matrix Representation Under Change of Basis

**Theorem 3.20**

Let $\mathbb{V}$ and $\mathbb{W}$ be two finite dimensional vector spaces and let $\dim(\mathbb{V}) = n$ and $\dim(\mathbb{W}) = m$. Let $B$ and $C$ be two ordered bases of $\mathbb{V}$ and let $B'$ and $C'$ be two bases of $\mathbb{W}$. Let $T : \mathbb{V} \to \mathbb{W}$ be a linear transformation. Then there exist two non-singular matrix $P \in \mathbb{M}_m(\mathbb{F})$ and $Q \in \mathbb{M}_n(\mathbb{F})$ such that $[T]_{BB'} = P^{-1}[T]_{CC'}Q$.

The following is a particular case of above theorem.

**Theorem 3.21**

Let $\mathbb{V}$ be an $n$-dimensional vector space over $\mathbb{F}$. Let $T$ be a linear transformation from $\mathbb{V}$ and to $\mathbb{V}$. Let $B$ and $B'$ be two ordered basis of $\mathbb{V}$. Then there exists a non-singular matrix $P$ such that $[T]_{BB} = P^{-1}[T]_{B'B'}P$.

*Proof.* Let $B = \{u_1, \ldots, u_n\}$ and let $B' = \{v_1, \ldots, v_n\}$. Let $S : \mathbb{V} \to \mathbb{V}$ be a linear transformation such that $S(u_i) = v_i$ for $i = 1, \ldots, n$. It is clear that $S$ is bijective. Let $[T]_{BB} = A = (a_{ij})$.

Then $T(u_j) = \sum\limits_{i=1}^{n} a_{ij}u_i$.

Therefore $S \circ T \circ S^{-1}(v_j) = S \circ T(u_j) = S(T(u_j)) = S(\sum\limits_{i=1}^{n} a_{ij}u_i) = \sum\limits_{i=1}^{n} a_{ij}v_i$.

$[STS^{-1}]_{B'B'} = (a_{ij}) = [T]_{BB}$.

$\implies [S]_{B'B'}[T]_{B',B'}[S^{-1}]_{B'B'} = [T]_{BB}$.

$\implies [S]_{BB}[T]_{BB}[S]^{-1}_{B,B} = [T]_{BB}$.

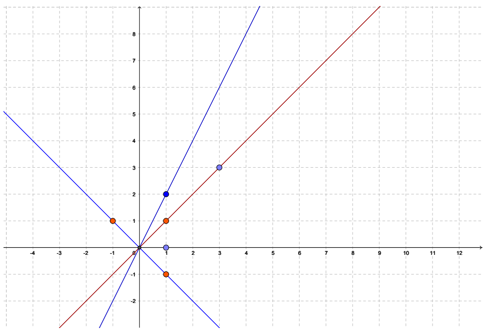
$\implies [T]_{BB} = [S]^{-1}_{B'B'}[T]_{B'B'}[S]_{B'B'} = P^{-1}[T]_{B'B'}P$ where $P = [S]_{B'B'}$. ▯

Eigen values can be "0"
Eigenvectors CANNOT be "0"

# 4 Eigenvalues and Eigenvectors

Let $A \in \mathbb{M}_n(\mathbb{F})$ and let $x \in \mathbb{F}^n$. Then $Ax$ is also a vector in $\mathbb{F}^n$. For example consider the following matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. $Ax = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.



An eigenvector of a matrix $A$ is a nonzero vector that changes by a scalar factor when the matrix $A$ is multiplied with that vector.

**Definition 4.1**

Let $A \in \mathbb{M}_n(\mathbb{F})$. A scalar $\lambda$ is said to be an **eigenvalue** of $A$ if there exists a non-null vector $x \in \mathbb{F}^n$ such that $Ax = \lambda x$. Any such (non-zero) $x$ is called an **eigenvector** of $A$ corresponding to the eigenvalue $\lambda$.

**Remark 4.1.** Let $A \in \mathbb{M}_n(\mathbb{F})$. Then $\det(xI - A)$ is a polynomial of $x$ whose co-efficients are coming from $\mathbb{F}$. For example consider the following matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$. Then $\det(xI - A) = x^3 - 6x^2 + 11x - 6$.

**Theorem 4.1**

Let $A \in \mathbb{M}_n(\mathbb{F})$ and $\lambda \in \mathbb{F}$. Then $\lambda$ is an eigenvalue of $A$ if and only if $\lambda$ is a root of the polynomial $\det(xI - A)$.

*Proof.* We first assume that $\lambda$ is an eigenvalue of $A$. To show $\lambda$ is a root of the polynomial $\det(xI - A)$. Then there exists a non-zero vector $x \in \mathbb{F}^n$ such that

$Ax = \lambda x.$

$\implies Ax - \lambda x = 0$

$\implies (A - \lambda I)x = 0$

This says that the system of homogeneous equations $(A - \lambda I)y = 0$ has non-trivial solution. Hence $rank(A - \lambda I) < n$. Then $det(A - \lambda I) = 0 = det(\lambda I - A)$. This implies $\lambda$ is a root of $\det(xI - A)$.

**Converse: We now assume that $\lambda$ is a root of the polynomial $\det(xI - A)$. To show $\lambda$ is an eigenvalue of $A$.**

$\det(\lambda I - A) = 0$

$rank(A - \lambda I) < n$

$(A - \lambda I)x = 0$ **has non-trivial solution. There is a non-zero $y \in \mathbb{F}^n$ such that $Ay = \lambda y$. Hence $\lambda$ is an eigenvalue of $A$.** □

**Definition 4.2**

Let $A \in \mathbb{M}_n(\mathbb{F})$. Then the polynomial $\det(xI - A)$ is called characteristic polynomial and the equation $\det(xI - A) = 0$ is called characteristic equation.

**Example 4.1.** Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. Calculate the eigenvalues of $A$. We have to calculate the roots of $\det(xI - A)$. Then $\det(xI - A) = \begin{vmatrix} x-1 & -2 & -1 \\ -2 & x-1 & -1 \\ -1 & -1 & x-2 \end{vmatrix} = x^3 - 4x^2 - x + 4$.

The roots of $x^3 - 4x^2 - x + 4$ are $1, -1, 4$. Since $A \in \mathbb{M}_n \underline{\mathbb{R}}$ and $1, -1, 4 \in \mathbb{R}$, then $1, -1, 4$ are the eigenvalues.

**Remark 4.2. Question:** Does every matrix have eigenvalue?

This question is similar to the following question. Let $P(x) \in \mathbb{P}(x, \mathbb{F})$, where $\mathbb{P}(x, \mathbb{F})$ set of all polynomials with coefficients are coming from $\mathbb{F}$. Does $P(x)$ have root in $\mathbb{F}$?

**Answer:** The answer is no. For example, consider the polynomial $x^2 + 1$ in

$\mathbb{P}(x, \mathbb{R})$, this polynomial does not have root in $\mathbb{R}$.

Now we are able to answer our question. If $x^2+1$ is the characteristic polynomial of a matrix $A \in \mathbb{M}_2(\mathbb{R})$, then $A$ does not have eigenvalues. Here is that $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

**Theorem 4.2**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A$ has at least one eigenvalue.

*Proof.* The characteristic polynomial of $A$ is $P_A(x) = \det(A - xI)$, here $P_A(x) \in \mathbb{P}_n(x, \mathbb{C})$. The Fundamental Theorem of Algebra says that $P_A(x)$ has at least one root in $\mathbb{C}$. Hence $A$ has at least one eigenvalue. □

**Remark 4.3.** Let $A \in \mathbb{M}_n(\mathbb{C})$. The $A$ has exactly $n$ number of eigenvalues.

**Remark 4.4.** We know how to calculate eigenvalues of a given matrix $A \in \mathbb{F}$. Just we have to find out all those roots of the characteristic polynomial of $A$ which are in $\mathbb{F}$.

**Question:** If you know an eigenvalue of a given matrix $A \in \mathbb{F}$ then how do you calculate eigenvectors corresponding to that eigenvalue?

Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. We have already calculated eigenvalues of $A$ which are $1, -1, 4$.

Eigenvectors corresponding to $\lambda = 1$.

The eigenvectors of $A$ corresponding to $\lambda = 1$ is the set of all non-zero solutions of $(A - I)x = 0$.

$A - I = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Then REF of $A - I$ is $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. The set $\left\{ k \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} : k \in \mathbb{R} - \{0\} \right\}$ is the set of all eigenvectors.

**Definition 4.3**

Let $x$ be an eigenvector of $A$ corresponding to the eigenvalue $\lambda$. Then the pair $(\lambda, x)$ is called eigenpair of $A$.

**Theorem 4.3**

Similar matrices have the same characteristic polynomial. But the converse is not true.

*Proof.* Let $A$ and $B$ be two similar matrices. Then there exists a nonsingular matrix $P$ such that $P^{-1}AP = B$.

Then $\det(B - xI)$

$= \det(P^{-1}AP - xI)$

$= \det(P^{-1}AP - xP^{-1}P)$

$= \det(P^{-1})\det(A - xI)\det(P) = \det(A - xI)$.

Converse is not true. Consider the following two matrices.

$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. These two matrices have the same characteristic polynomial but they are not similar.

**Theorem 4.4**

Let $C = \begin{bmatrix} A_{n\times n} & D_{n\times m} \\ 0_{m\times n} & B_{m\times m} \end{bmatrix}$. Then characteristic polynomial of $C$, $P_C(x) = P_A(x)P_B(x)$.

*Proof.* We have $C - xI_{n+m} = \begin{bmatrix} A_{n\times n} - xI_n & D_{n\times m} \\ 0_{m\times n} & B_{m\times m} - xI_m \end{bmatrix}$. Then $\det(C) = \begin{vmatrix} A_{n\times n} - xI_n & D_{n\times m} \\ 0_{m\times n} & B_{m\times m} - xI_m \end{vmatrix} = \det(A - xI_n)\det(B - xI_m)$.

**Theorem 4.5**

Let $f(x)$ be a polynomial and let $(\lambda, x)$ be an eigenpair of $A$. Then $(f(\lambda), x)$ is an eigenpair of $f(A)$. But the converse is not true.

*Proof.* If $(\lambda, x)$ is an eigenpair of $A$, then $(\lambda^k, x)$ is an eigenpair of $A^k$ where $k$ is natural number. Using this we can easily prove that $(f(\lambda), x)$ is an eigenpair of $f(A)$.

Converse is not true. Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{M}_n(\mathbb{R})$ and $f(x) = x^2$. The characteristic polynomial of $A$ is $x^2+1$. Hence $A$ has no eigenvalue. Then $f(A) = A^2 = \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix}$. It is clear that $-1$ is an eigenvalue of $A^2$ and $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of $A^2$ corresponding to $\lambda = -1$. But $(-1, x)$ is not an eigenpair of $A$.

**Theorem 4.6**

$$\begin{bmatrix} A_{n\times n} & B_{n\times m} \\ C_{m\times n} & D_{m\times m} \end{bmatrix} \begin{bmatrix} E_{n\times n} & F_{n\times m} \\ G_{m\times n} & H_{m\times m} \end{bmatrix} = \begin{bmatrix} AE+BG & AF+BH \\ CE+DG & CF+DH \end{bmatrix}$$

**Theorem 4.7**

Let $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ and $B \in \mathbb{M}_{n\times m}$, where $m \geq n$.. Then $P_{AB}(x) = x^{m-n} P_{BA}(x)$.

*Proof.* We can write $\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix} \begin{bmatrix} AB & 0 \\ B & 0_n \end{bmatrix} \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} 0_m & 0 \\ B & BA \end{bmatrix}$

The matrix $\begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix}$ is invertible and $\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix}$ is the inverse.

$P^{-1} \begin{bmatrix} AB & 0 \\ B & 0_n \end{bmatrix} P = \begin{bmatrix} 0_m & 0 \\ B & BA \end{bmatrix}$ where $P = \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix}$.

**Definition 4.4**

Let $A \in \mathbb{M}_n(\mathbb{F})$ and let $\lambda$ be an eigenvalue of $A$. The number of times $\lambda$ appears as a root of the characteristic polynomial of $A$ is called **algebraic multiplicity.**

**Example 4.2.** Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}$. The eigenvalues of $A$ are $1, 1, 2$. The algebraic multiplicity of 1 is 2 and 2 is 1.

**Theorem 4.8**

Let $\lambda_1, \ldots, \lambda_p$ be distinct eigenvalues of $A$. Let $x_1, \ldots, x_p$ be eigenvectors of $A$ corresponding to $\lambda_1, \ldots, \lambda_p$. Then $x_1, x_2 \ldots, x_p$ are linearly independent.

*Proof.* Consider $\{x_1\}, \{x_1, x_2\}, \cdots, \{x_1, x_2, \cdots, x_p\}$ one by one. Take the largest $k$ s.t. $S_k = \{x_1, \cdots, x_k\}$ is linearly independent. Then we show that $k = p$. Assume that $k < p$. Then $S_{k+1} = \{x_1, \cdots, x_k, x_{k+1}\}$ is linearly dependent.

In that case if $\sum\limits_{i=1}^{k+1} \alpha_i x_i = 0$ $\alpha_1, \ldots, \alpha_{k+1} \in \mathbb{F}$ not all zero, then $\alpha_{k+1} \neq 0$. Otherwise $S_{k+1}$ will be linearly independent.

$x_{k+1} = c_1 x_1 + c_2 x_2 + \cdots + c_k x_k.$

$Ax_{k+1} = c_1 Ax_1 + c_2 x_2 + \cdots + c_k Ax_k.$

$\lambda_{k+1} = c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \cdots + c_k \lambda_k x_k.$

$\lambda_{k+1}(c_1 x_1 + c_2 x_2 + \cdots + c_k x_k) = c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \cdots + c_k \lambda_k x_k.$

$(\lambda_{k+1} - \lambda_1) c_1 x_1 + \cdots + (\lambda_{k+1} - \lambda_k) c_k x_k = 0$

$(\lambda_{k+1} - \lambda_1) c_1 = \cdots = (\lambda_{k+1} - \lambda_k) c_k =$

$c_1 = \cdots = c_k = 0$ as $\lambda_{k+1} - \lambda_i \neq 0$ for $i = 1, \ldots k$.

Hence $x_{k+1}$ is zero a contradiction. Then $k = p$. □

**Theorem 4.9**

Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A$. Then $\prod\limits_{i=1}^{n} \lambda_i = \det(A)$ and $\sum\limits_{i=1}^{n} \lambda_i$.

*Proof.* The characteristic polynomial of $A$ is $P_A(x) = \det(xI - A)$ and $\lambda_1, \ldots, \lambda_n$ are the roots of $P_A(x)$. Then $P_A(x) = (x - \lambda_1) \cdots (x - \lambda_n)$. Then equating the co-efficient of $x^{n-1}$ and constant term from both side.

Then we have $\sum\limits_{i=1}^{n} \lambda_i = trace(A)$ and $\prod\limits_{i=1}^{n} \lambda_i = \det(A)$.

**Remark 4.5.** For example consider the following matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$. Then $\det(xI - A) = x^3 - 6x^2 + 11x - 6$. The $\operatorname{tr}(A) = 6$ and $\det(A) = 6$. It is clear that coefficient of $x^2$ is $-\operatorname{tr}(A)$ and constant term is $(-1)^3 \det(A)$.

**Definition 4.5**

Let $A \in \mathbb{M}_n(\mathbb{F})$ and let $\lambda$ be an eigenvalue of $A$. Let $E_\lambda(A) = \{x \in \mathbb{F}^n : (A - \lambda I)x = 0\}$. Then $E_\lambda(A)$ is called eigenspace of $A$ corresponding to the eigenvalue of $\lambda$. It is clear that $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$.

The dimension of $E_\lambda$ is called the **geometric multiplicity** of $\lambda$ with respect to $A$.

**Example 4.3.** Consider the matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. We have already calculated the eigenvalues of $A$ which are $1, -1, 4$. We have calculated the set of all eigenvectors of $A$ corresponding to $\lambda = 1$.

Then $E_{(\lambda=1)} = \left\{ k \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} : k \in \mathbb{R} \right\}$.

The $\dim(E_{(\lambda=1)}) = 1$. Hence geometric multiplicity of $\lambda = 1$ is 1. Similarly you can calculated geometric multiplicities of $-1$ and 4.

**Theorem 4.10**

Let $A \in \mathbb{M}_n(\mathbb{F})$ and $\lambda$ be an eigenvalue of $A$. Then the geometric multiplicity of $\lambda$ is $n - rank(A - \lambda I)$.

**Theorem 4.11**

Let $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ and let $\lambda_1$ and $\lambda_2$ be two distinct eigenvalues. Then $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{0\}$.

*Proof.* Let $x \in E_{\lambda_1}(A) \cap E_{\lambda_2}(A)$. Then $x \in E_{\lambda_1}(A)$ implies $Ax = \lambda_1 x$ and $x \in E_{\lambda_2}(A)$ implies $Ax = \lambda_2 x$. This implies $\lambda_1 x = \lambda_2 x$. Since $\lambda_1$ and $\lambda_2$ are distinct, then $x = 0$. Hence $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{0\}$. □

**Discussions:** Let $B \in \mathbb{M}_n(\mathbb{F})$. Then we can write $B$ in the following way $B = [b_1 : b_2 : \cdots : b_n]$ where $b_i$ is the ith column of $B$ for $i = 1, \ldots, n$. Let $A \in \mathbb{M}_n(\mathbb{F})$. Then $AB = A[b_1 : b_2 : \cdots : b_n]$

$= [Ab_1 : Ab_2 : \cdots : Ab_n]$ where $Ab_i$ is the ith column of $AB$ for $i = 1, \ldots, n$.

I will use this notation frequently.

**Theorem 4.12**

Let $A \in \mathbb{M}_n(\mathbb{F})$ and let $\lambda$ be an eigenvalue of $A$. Then the algebraic multiplicity of $\lambda$ is greater or equal to the geometric multiplicity of $\lambda$.

*Proof.* Let the geometric multiplicity of $\lambda$ be $m$. That is $\dim(E_\lambda(A)) = m$. Let $\mathcal{B}_1 = \{x_1, \ldots, x_m\}$ be a basis of $E_\lambda(A)$. Since $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$. We extend $\mathcal{B}_1$ to a basis for $\mathbb{F}^n$, say, $\mathcal{B} = \{x_1, \ldots, x_m, x_{m+1}, \ldots, x_n\}$.

Thus $P = [x_1 : x_2 : \cdots : x_n]$ is a non-singular matrix. Therefore

$P^{-1}AP = P^{-1}[Ax_1 : Ax_2 : \cdots : Ax_n]$.

$= P^{-1}[\lambda x_1 : \cdots : \lambda x_m : Ax_{m+1} : \cdots : Ax_n]$.

We can show that $P^{-1}(\lambda x_j) = \lambda P^{-1} x_j = \lambda e_j$ (Here $x_j$ is the jth column of $P$ and $P^{-1}$ is the inverse) $j = 1, \ldots, m$. Then

$P^{-1}AP = \begin{bmatrix} \lambda I_m & B \\ 0 & C \end{bmatrix}$ for some matrices $B$ and $C$. Hence $P_A(x) = P_{P^{-1}AP}(x) = (x-\lambda)^m P_C(x)$. So the algebraic multiplicity of $\lambda$ with respect to $A$ is at least $m$ and the theorem follows. □

**Definition 4.6**

A matrix is said to be diagonalizable if there exists a nonsingular matrix $P$ such that $P^{-1}AP$ is a diagonal matrix. That is, $A$ is similar to a diagonal matrix.

**Example 4.4.** Every diagonal matrix is diagonalizable. Use the argument every matrix is similar to itself. Consider the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Take $P = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Then you can check that $P^{-1}AP = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$. Hence $A$ is diagonalizable matrix.

**Theorem 4.13**

Let $A \in \mathbb{M}_n(\mathbb{F})$ and let $P^{-1}AP = D$ where $D = (d_i)$ is a diagonal matrix. Then $d_i$ is an eigenvalue of $A$ and $i$th column of $P$ is an eigenvector corresponding to $d_i$.

*Proof.* We have $P^{-1}AP = D$. That is $AP = PD$. Let $P = [P_1 : \cdots : P_n]$ where $P_i$ is the $i$th column of $P$. Then $PD = [d_1P_1 : \cdots : d_nP_n]$ and $AP = [AP_1 : \cdots : AP_n]$. Therefore $AP_i = d_iP_i$ for $i = 1, \ldots, n$. Since $P$ is non-singular, we have each $P_i$ is non-zero for each $i = 1, \ldots, n$. This implies that $d_i$ is an eigenvalue of $A$ and corresponding eigenvector $P_i$ for $i = 1, \ldots, n$. □

**Remark 4.6.** There is a square matrix which is non diagonalizable. For example, consider the matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. It is not diagonalizable. If $A$ is diagonalizable then there exists a non-singular matrix $P$ such that $P^{-1}AP = D$ where $D$ is a diagonal matrix and each diagonal entry is the eigenvalue of $A$. Hence $D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Therefore $P^{-1}AP = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. This implies that $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ which is not possible.

**We have seen that all the matrices are not diagonalizable. The following questions are natural.**

**Q1. How do we check whether a matrix is diagonalizable or not?**

**Q2. If $A$ is diagonalizable, then how do I calculate such $P$ matrix?**

Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of $A$ with algebraic multiplicities $m_1, \ldots, m_k$ respectively. Then the characteristic polynomial of $A$ is $\prod_{i=1}^{k}(x - \lambda_i)^{m_i}$. This implies that $m_1 + \cdots + m_k = n$.

If we consider some other field $\mathbb{F}$ instead of $\mathbb{C}$, $\prod_{i=1}^{k}(x - \lambda_i)^{m_i}$ may not be the characteristic polynomial of $A$. Consider the following matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$ in $\mathbb{M}_n(\mathbb{R})$. Then $A$ has exactly one eigenvalue which is 1.

The following are necessary and sufficient conditions for a matrix $A$ to be diagonalizable. The following theorems give the answer of Q1.

**Theorem 4.14**

Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of $A$ with algebraic multiplicities $m_1, \ldots, m_k$ respectively. Then $A$ is diagonalizable if and only if $A$ has $n$ linearly independent eigenvectors.

*Proof.* We first assume that $A$ is diagonalizable. Then there exists a non-singular matrix $P$ such that $P^{-1}AP = D$ where $D$ is a diagonal matrix. Using previous theorem, each column of the matrix $P$ is an eigenvector of $A$. Since $P$ is non-singular, the columns vectors of $P$ are linearly independent. Hence $A$ has $n$ linearly independent eigenvectors.

We now assume that $A$ has $n$ linearly independent eigenvectors those are $x_1, \ldots, x_n$. Take the matrix $P = [x_1 : x_2 : \cdots : x_n]$. Then $P$ is non-singular. Therefore

$$AP = A[x_1 : x_2 : \cdots : x_n]$$

$$AP = [Ax_1 : Ax_2 : \cdots : Ax_n]$$

$$AP = [\lambda_1 x_1 : \cdots : \lambda_n x_n]$$

$$AP = [\lambda_1 Pe_1 : \cdots : \lambda_n Pe_n]$$

$$AP = P[\lambda_1 e_1 : \cdots : \lambda_n e_n]$$

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$ Hence $A$ is diagonalizable. □

**Corollary 4.1.** *Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $A$ have $n$ distinct eigenvalues. Then $A$ is diagonalizable.*

The following theorem is quite important to check whether a matrix is diagonalizable or not.

**Theorem 4.15**

Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of $A$ with algebraic multiplicities $m_1, \ldots, m_k$, respectively. Then $A$ is diagonalizable if and only if the geometric multiplicity of $\lambda_i$ is $m_i$ for $i = 1, \ldots, k$.

*Proof.* We first assume that $A$ is diagonalizable. To show the geometric multiplicity of $\lambda_i$ is $m_i$ for $i = 1, \ldots, k$. Suppose that there exists $1 \le t \le k$ such that the geometric multiplicity of $\lambda_t$ is strictly less that $m_t$. The eigenspace $E_{\lambda_i}(A)$ gives us at most $m_i$ linearly independent eigenvectors for $i = 1, \ldots, k$ and $i \neq t$. The eigenspace $E_{\lambda_t}(A)$ gives us at most $m_t - 1$ linearly independent eigenvectors. So we have at most $n-1$ eigenvector. A contradiction that a diagonalizable matrix must have $n$ linearly eigenvectors. Contradiction because we assume that the geometric multiplicity of $\lambda_t$ is strictly less that $m_t$. Hence the geometric multiplicity of $\lambda_i$ is $m_i$ for $i = 1, \ldots, k$.

We now assume that the geometric multiplicity of $\lambda_i$ is $m_i$ for $i = 1, \ldots, k$. The eigenspace $E_{\lambda_i}(A)$ gives us at most $m_i$ linearly independent eigenvectors for $i = 1, \ldots, k$. So we have $n$ linearly independent eigenvectors. Hence $A$ is diagonalizable. □

**Problem;** Using above theorem check whether the following matrices are diagonalizable or not. $\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 2 & 6 & 0 \\ 3 & 2 & 1 \end{bmatrix}$.

**Theorem 4.16**

Let $A \in \mathbb{M}_n(\mathbb{C})$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of $A$ with algebraic multiplicities $m_1, \ldots, m_k$ respectively. Then $A$ is diagonalizable if and only if $E_{\lambda_1}(A) \oplus E_{\lambda_2}(A) \oplus E_{\lambda_3}(A) \oplus \cdots \oplus E_{\lambda_k}(A) = \mathbb{C}^n$.

*Proof.* We first assume that $A$ is diagonalizable. Then the geometric multiplicity of $\lambda_i$ is $m_i$ for $i = 1, \ldots, k$. We have already proved that $E_{\lambda_i}(A) \cap E_{\lambda_j}(A) = \{0\}$ for $1 \le i, j \le k$ and $i \neq j$. Then $\dim(E_{\lambda_1}(A) + E_{\lambda_2}(A) + E_{\lambda_3}(A) + \cdots + E_{\lambda_k}(A) = \sum_{i=1}^{k} \dim(E_{\lambda_i}(A)) = \sum_{i=1}^{k} m_i = n$. Hence $E_{\lambda_1}(A) \oplus E_{\lambda_2}(A) \oplus E_{\lambda_3}(A) \oplus \cdots \oplus E_{\lambda_k}(A) = \mathbb{C}^n$.

Assume that $E_{\lambda_1}(A) \oplus E_{\lambda_2}(A) \oplus E_{\lambda_3}(A) \oplus \cdots \oplus E_{\lambda_k}(A) = \mathbb{C}^n$. Let $B_i$ be a basis of $E_{\lambda_i}(A)$ for $i = 1, \ldots, k$. It can be proved that $\cup_{i=1}^{k} B_i$ is a basis of

$E_{\lambda_1}(A) \oplus E_{\lambda_2}(A) \oplus E_{\lambda_3}(A) \oplus \cdots \oplus E_{\lambda_k}(A) = \mathbb{C}^n$. The cardinality of $\cup_{i=1}^k B_i$ is $n$. Hence we have $n$ linearly independent vectors. Thus $A$ is diagonalizable. $\square$

**Remark 4.7.** How to check whether a matrix is diagonalizable or not? If it is diagonalizable, then how to find such matrix $P$?

**Step 1.** First you will have to find out all the eigenvalues if it has. Let $\lambda_i$ $i = 1, \dots, k$ be the distinct eigenvalues of $A$ with algebraic multiplicities $m_i$ for $i = 1, \dots, k$.

**Step 2.** You will have to find basis of each eigenspace $E_{\lambda_i}(A)$. Let $B_i$ be a basis of $E_{\lambda_i}(A)$ for $i = 1, \dots, k$. If the cardinality $|B_i| = m_i$ for $i = 1, \dots, k$ (that is algebraic multiplicity= geometric multiplicity), then $A$ is diagonalizable.

**Step 3.** Let $B_i = \{x_{i1}, \dots, x_{im_i}\}$ for $i = 1, \dots, k$. Then

$$P = (x_{11}, x_{12}, \dots, x_{1m_1}, x_{21}, \dots, x_{2m_2}, \dots, x_{k1}, \dots, x_{km_k})$$

.

**Problem:** Find out a matrix $P$ such that $P^{-1}AP = D$ where $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$

**Definition 4.7**

A matrix $U \in \mathbb{M}_n(\mathbb{C})$ is said to be a unitary matrix such that $UU^* = U^*U = I$ where $U^* = \overline{U}^t$.

**Theorem 4.17**

A matrix $U$ is unitary if and only if the set of all column vectors (resp. row vectors) form an orthonormal set.

*Proof.* Since $A$ is unitary, we have $A^*A = AA^* = I$. $A^*A = I$ implies that the set of all column vectors is orthonormal and $AA^* = I$ implies that the set of all row vectors is orthonormal.

**Theorem 4.18: Schur Theorem**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then there exists a unitary matrix $U$ such that $U^*AU$ is an upper triangular matrix.

*Proof.* We prove the theorem by induction on $n$. If $n = 1$ the result holds trivially. So assume it for matrices of order $n-1$ and let $A$ be of order $n$. Let $\lambda$ be an eigenvalue of $A$ and let $x$ be an eigenvector of $A$ corresponding to $\lambda$. We extend $\{x\}$ to a basis of $\mathbb{C}^n$ which is $\{x, x_1, \ldots, x_{n-1}\}$. Applying Gram Schimdt process on $\{x, x_1, \ldots, x_{n-1}\}$, we have an orthonormal set $\{y_1, \ldots, y_n\}$. Let $P = [y_1 : \cdots : y_n]$. It is clear that $P$ is a unitary matrix. Then

$$AP = A[y_1 : \cdots : y_n]$$

$$AP = [Ay_1 : \cdots : Ay_n]$$

$$AP = [\lambda y_1 : \cdots : Ay_n].$$

$$P^*AP = P^*[\lambda y_1 : \cdots : Ay_n].$$

$$P^*AP = [\lambda P^* y_1 : \cdots : P^*Ay_n].$$

$$P^*AP = [\lambda e_1 : \cdots : z_n]. \text{ Where } z_i = \begin{bmatrix} z_{1i} \\ z_{2i} \\ z_{3i} \\ \vdots \\ z_{ni} \end{bmatrix} = P^*Ay_i \text{ for } i = 2, \ldots, n$$

$$P^*AP = \begin{bmatrix} \lambda & z_{12} & z_{13} & z_{14} & \cdots & z_{1n} \\ 0 & z_{22} & z_{23} & z_{24} & \cdots & z_{2n} \\ 0 & z_{32} & z_{33} & z_{34} & \cdots & z_{3n} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & z_{n2} & z_{n3} & z_{n4} & \cdots & z_{nn} \end{bmatrix}.$$

$$P^*AP = \begin{bmatrix} \lambda & z^t \\ 0 & C \end{bmatrix}, \text{ where } z = \begin{bmatrix} z_{12} \\ z_{13} \\ z_{14} \\ \vdots \\ z_{1n} \end{bmatrix} \text{ and } C = \begin{bmatrix} z_{22} & z_{23} & z_{24} & \cdots & z_{2n} \\ z_{32} & z_{33} & z_{34} & \cdots & z_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ z_{n2} & z_{n3} & z_{n4} & \cdots & z_{nn} \end{bmatrix}$$

Since $C$ is a matrix of size $n-1$, by induction hypothesis there exists a unitary matrix $Q$ of size $n-1$ such that $Q^*CQ = T$ is upper triangular. Consider $W = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}$. You can easily check that $W$ is unitary.

Then $(PW)^*A(PW) = W^*P^*APW = W^* \begin{bmatrix} \lambda & y^t \\ 0 & C \end{bmatrix} W = \begin{bmatrix} \lambda & y^tQ \\ 0 & Q^*CQ \end{bmatrix} = \begin{bmatrix} \lambda & y^tQ \\ 0 & T \end{bmatrix}$.

Consider $U = PW$, it is easy to check that $U$ is unitary and $U^*AU$ is upper triangular. □

**Definition 4.8**

Let $A, B \in \mathbb{M}_n(\mathbb{C})$. Then $A$ and $B$ are called unitarily similar if there exists a unitary matrix $U$ such that $U^*AU = B$.

**Remark 4.8.** The Schur theorem says that every complex square matrix is unitarily similar to an upper triangular matrix. This result is quite important and very significant result.

## 4.1 Annihilating Polynomials

**Definition 4.9**

Let $A \in \mathbb{M}_n(\mathbb{F})$. A polynomial $f(x)$ is said to annihilate $A$ if $f(A) = 0_n$.

**Theorem 4.19: Cayley Hamilton Theorem**

Let $A \in \mathbb{M}_n(\mathbb{F})$. Then $A$ satisfies its own characteristic polynomial.

**Definition 4.10**

A polynomial $f(x)$ is said to monic polynomial if the co-efficient of the highest degree term is 1.

**Definition 4.11**

The monic polynomial of the least degree which annihilates $A$ is called the **minimal polynomial** of $A$. We use $m_A(x)$ to denote the minimal polynomial of $A$.

**Example 4.5.** Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The minimal polynomial of $A$ is $x - 1$. Consider the matrix $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. The minimal polynomial of $B$ is $(x-1)(x+1)$.

**Theorem 4.20**

The minimal polynomial of $A$ divides every polynomial which annihilates $A$.

Let $m_A(x)$ be the minimal polynomial of $A$ and let $P(x)$ be a annihilate polynomial of $A$. Using division algorithm there exist polynomials $q(x)$ and $r(x)$ such that $P(x) = m_A(x)q(x) + r(x)$, where degree of $r(x)$ is less than the degree of $m_A(x)$. Then $P(A) = m_A(A)q(A) + r(A)$. This implies that $r(A) = 0_n$ as $P(A) = m_A(A) = 0_n$. Thus $r(x)$ annihilates $A$. By the minimality of $m_A(x)$, it follows $r(x) = 0$. Hence $m_A(x)$ divides $P(x)$. □

**Corollary 4.2.** *The minimal polynomial of $A$ divides characteristic polynomial of $A$.*

**Theorem 4.21**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $\lambda$ is a root of the minimal polynomial $m_A(x)$ if and only if $\lambda$ is a root of the characteristic polynomial of $A$.

*Proof.* We first assume that $\lambda$ is a root of $m_A(x)$. Using previous corollary, $m_A(x)$ divides the characteristic polynomial $P_A(x)$ of $A$. Then there exists a polynomial $q(x)$ such that $P_A(x) = m_A(x)q(x)$. This implies $\lambda$ is a root of $P_A(x)$.

We now assume that $\lambda$ is a root of $P_A(x)$. That means $\lambda$ is eigenvalue of $A$. Then $m_A(\lambda)$ is an eigenvalue of $m_A(A)$. Since $m_A(x)$ is a minimal polynomial, we have $m_A(A) = 0_n$. Therefore $m_A(\lambda) = 0$. Hence $\lambda$ is a root of $m_A(x)$. □

**Discussions:** Above theorem help us to calculate the minimal polynomial of a given matrix $A \in \mathbb{M}_n(\mathbb{C})$. Consider the following matrix $A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$.

We first calculate all the roots of the characteristic polynomial of $A$.

The characteristic polynomial of $A$ is $x^2(x^2-4)$. So the roots of the characteristic polynomial of $A$ are $0, -2, 2$ with algebraic multiplicities of $2, 1, 1$, respectively.

Using above theorem we know that $0, -2, 2$ are the roots of the minimal polynomial of $A$. The following are the possibilities to be minimal polynomial.

1. $x^2(x-2)(x+2)$.

2. $x(x-2)(x+2)$.

The first one is the characteristic polynomial of $A$, hence it annihilates the matrix $A$. We can check that second polynomial also annihilates $A$. Among these two polynomials the second one is minimal. Hence the minimal polynomial is $x(x-2)(x+2)$.

Note: If $x(x-2)(x+2)$ does not annihilate $A$ then the minimal polynomial is $x^2(x-2)(x+2)$.

The following is quite important result which helps to check whether a matrix $A$ is diagonalizable or not.

**Theorem 4.22**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A$ is diaqgonalizable if and only if the minimal polynomial of $A$ is the product of distinct linear factors.

## 4.2 Spectral Theorem

In this section, we discuss about the diagonalizability of four special matrices which are Hermitian Matrix, Skew-Hermitian Matrix, Unitary Matrix and Normal Matrix.

**Notation:** Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A^* = \bar{A}^T$. For real case it should be $A^* = A^T$.

**Definition 4.12**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A$ is called **Hermitian** if $A^* = A$. For real case it is called **real-symmetric** and $A^T = A$.

**Theorem 4.23: Hermitian Matrix**

Let $A$ be Hermitian (or real-symmetrix). Then the following are true.

1. Each eigenvalue of Hermitian ( or real symmetric) matrix is real.

2. The eigenvectors corresponding to distinct eigenvalues of Hermitian (or real-symmetric) matrix are orthogonal.

*Proof.*

1. Let $(\lambda, x)$ be an eigenpair of $A$. To show $\lambda$ is real. $Ax = \lambda x$.

   $(Ax)^* = (\lambda x)^*$

   $x^* A^* = \overline{\lambda} x^*$

   $x^* A = \overline{\lambda} x^*$ (as $A^* = A$)

   $x^* A x = \overline{\lambda} x^* x$ (multiplying both side by $x$)

   $\lambda x^* x = \overline{\lambda} x^* x$

   $(\lambda - \overline{\lambda}) x^* x = 0.$

   $\lambda = \overline{\lambda}$. Hence $\lambda$ is real.

2. Let $x$ and $y$ be two eigenvectors of $A$ corresponding to the eigenvalues $\lambda$ and $\mu$ where $\lambda \neq \mu$.

   $Ax = \lambda x$

$(Ax)^* = (\lambda x)^*$

$x^* A^* = \overline{\lambda} x^*$

$x^* A = \lambda x^*$ (as $\lambda$ is real which we have proved just now)

$x^* Ay = \lambda x^* y$ (multiplying both side by $y$)

$\mu x^* y = \lambda x^* y$

$(\mu - \lambda) x^* y = 0$

$x^* y = 0$. (as $\lambda \neq \mu$). Hence they are orthogonal.

$\square$

**Definition 4.13: Skew-Hermitian Matrix**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A$ is called **Skew-Hermitian** if $A^* = -A$. For real case it is called **real-Skew-symmetric** and $A^T = -A$.

**Theorem 4.24**

Each eigenvalue of Skew-Hermitian ( or real symmetric) matrix either 0 or purely imaginary.

*Proof.* Let $(\lambda, x)$ be an eigenpair of $A$. To show $\lambda$ is either zero or purely imaginary.

$Ax = \lambda x$.

$(Ax)^* = (\lambda x)^*$

$x^* A^* = \overline{\lambda} x^*$

$-x^* A = \overline{\lambda} x^*$ (as $A^* = -A$)

$-x^* Ax = \overline{\lambda} x^* x$ (multiplying both side by $x$)

$-\lambda x^* x = \overline{\lambda} x^* x$

$(-\lambda - \overline{\lambda}) x^* x = 0.$

Real part of $\lambda$ zero. Hence $\lambda$ is purely imaginary or zero. □

**Definition 4.14: Unitary Matrix**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A$ is called **Unitary** if $A^* A = AA^* = I$. For real case it is called **orthogonal** and $A^T A = AA^T = I$.

**Theorem 4.25**

Let $A$ be unitary (real orthogonal). Then each eigenvalue of $A$ us unit modulus.

*Proof.* Let $(\lambda, x)$ be an eigenpair of $A$. To show $\lambda$ is either zero or purely imaginary.

$Ax = \lambda x.$

$(Ax)^* = (\lambda x)^*$

$x^* A^* = \overline{\lambda} x^*.$

$x^* A^* A = \overline{\lambda} x^* A.$

$x^* = \overline{\lambda} x^* A.$

$x^* x = \overline{\lambda} x^* A x.$

$x^* x = \overline{\lambda} x^* \lambda x.$

$x^* x = |\lambda|^2 x^* x.$

$|\lambda| = 1.$

**Definition 4.15: Normal Matrix**

Let $A \in \mathbb{M}_n(\mathbb{C})$. Then $A$ is called **Normal** matrix if $A^* A = AA^*$. For real case it should be $A^T A = AA^T$.

**Remark 4.9.** We can notice that Hermitian (real-symmetric), Skew-Hermitian (Skew-Symmetric), Unitary (orthogonal) matrices are normal matrices.

**Definition 4.16**

A matrix $A$ is said to be unitarily (resp. orthogonally in real case) diagonalizable if there exists a unitary matrix (resp. orthogonal in real case) $S$ such that $S^*AS = D$.

**Theorem 4.26: Spectral Theorem**

Let $A \in \mathbb{M}_n(\mathbb{C})$ be a normal matrix. Then there exists a unitary matrix $U$ such that $U^*AU = D$, where $D$ is a diagonal matrix. That is each normal matrix is unitarily (real case orthogonally) diagonalizable.

*Proof.* By using Schur theorem, we have $U^*AU = T$, where $T$ is upper triangular. Since $A$ is normal you can easily prove that $T$ is normal. The matrix $T$ is normal and upper triangular, then $T$ is diagonal matrix. □

**Corollary 4.3.** *Hermitian, Skew-Hermitina, Unitary matrices are normal matrices. Hence they are unitarily diagonalizable. ( real case. Similarly real-symmetric. skew-symmetric and orthogonal matrices are orthogonally diagonalizable.)*

## 4.3 Eigenvalues and eigenvectors of operators

**Definition 4.17**

Let $\mathbb{V}$ be a vector space (no matter $\mathbb{V}$ is finite dimensional or infinite dimensional )over the field $\mathbb{F}$. Let $T$ be a linear operator from $\mathbb{V}$ to $\mathbb{V}$. Let $\lambda \in \mathbb{F}$ is called eigenvalue of $T$ if there exists a non-zero vector $v \in \mathbb{V}$ such that $T(v) = \lambda v$.

**Remark 4.10.** Let $\mathbb{V}$ be a vector space (no matter $\mathbb{V}$ is finite dimensional or infinite dimensional ) over the field $\mathbb{F}$. Let $T$ be a linear operator from $\mathbb{V}$ to $\mathbb{V}$. Then $\lambda \in \mathbb{F}$ is an eigenvalue of $T$ if and only if $T - \lambda I$ is not one

**Example 4.6.** $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x_1, x_2) = (-x_2, x_1)$. This linear operator does not have eigenvalues.

**Example 4.7.** $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x_1, x_2) = (x_2, 0)$. This linear operator has two eigenvalues which are 0 and 1.

The following example says that it is possible to have linear operators $T$ from $\mathbb{V}(\mathbb{F})$ to $\mathbb{V}(\mathbb{F})$ such that $\mathbb{F}$ is the set of all eigenvalues.

**Example 4.8.** Let $\mathbb{V}$ be the vector space of all real sequences. Let $T : \mathbb{V} \to \mathbb{V}$ be a linear operator defined by $T(x_1, x_2, \ldots) = (x_2, \ldots)$ (left shift operator). Then find all the eigenvalues of $T$.

**Sol:** We have to find all the values of $\lambda \in \mathbb{R}$ such that $T - \lambda I$ is not injective.

Let $(x_1, x_2, \ldots) \in Ker(T - \lambda I)$. Then $T - \lambda I(x_1, x_2, \ldots) = (0, 0, \ldots)$.

$T(x_1, x_2 \ldots) - \lambda(x_1, x_2, \ldots) = (0, 0, \ldots)$

$(x_2, x_3, \ldots) - \lambda(x_1, x_2, \ldots) = (0, 0, \ldots)$

$x_2 - \lambda x_1 = 0, x_3 - \lambda x_2 = 0 \ldots$

$x_2 = \lambda x_1, x_3 = \lambda^2 x_1, x_4 = \lambda^3 x_1, \ldots.$

Hence $(x_1, x_2, \ldots) = (x_1, \lambda x_2, \lambda^2 x_1, \lambda^3 x_1, \ldots)$.

Take $x_1 = 1$. Then $(1, \lambda, \lambda^2, \ldots) \in Ker(T - \lambda I)$ for each $\lambda \in \mathbb{R}$.

Hence $T - \lambda I$ is not injective for each $\lambda \in \mathbb{R}$.
$\mathbb{R}$ is the set of all eigenvalues of $T$.

We have seen that each matrix $A \in \mathbb{M}_n(\mathbb{C})$ has exactly $n$ eigenvalues. But there is a linear operator $T$ from a vectors space $\mathbb{V}(\mathbb{C})$ to $\mathbb{V}(\mathbb{C})$ which has no eigenvalues. The following is an example.

**Example 4.9.** Let $\mathbb{V}$ be the vector space of all the complex sequences. Let $T : \mathbb{V} \to \mathbb{V}$ be a linear operator defined by $T(x_1, x_2, \ldots) = (x_2, \ldots)$ (left shift operator). Then $T$ has no eigenvalues.

**Sol:** 0 can not be an eigenvalue of $T$. If 0 is an eigenvalue of $T$, then there exists a non-zero sequence $(x_1, x_2, \ldots)$ such that $T(x_1, x_2, \ldots) = 0(x_1, x_2, \ldots)$. Then $(0, x_1, x_2, \ldots) = (0, 0, \ldots)$. This implies $x_1 = x_2 = \cdots = 0$. A contradiction.

Any non-zero scalar $\lambda$ is not eigenvalue. If $\lambda$ is an eigenvalue of $T$, then there exists a non-zero sequence $(x_1, x_2, \ldots)$ such that $T(x_1, x_2, \ldots) = \lambda(x_1, x_2, \ldots)$. Then $(0, x_1, x_2, \ldots) = (\lambda x_1, \lambda x_2, \ldots)$. This implies $\lambda x_1 = 0$ and $\lambda x_i = x_{i-1}$ for $i = 2, 3, \ldots.$ This implies $x_1 = x_2 = \cdots = 0$. A contradiction. Hence $T$ has no eigenvalue.

**Theorem 4.27**

Let $T$ be a linear operator on the finite-dimensional space $\mathbb{V}$ over the field $\mathbb{F}$. Let $A$ be a matrix representation of $T$ with respect to some basis of $\mathbb{V}$. Then $\lambda$ is an eigenvalue of $T$ if and only if $\lambda$ is an eigenvalue of $A$.

*Proof.* First we assume that $\lambda$ is an eigenvalue of $T$. Let $x$ be an eigenvector of $T$ corresponding to the eigenvalue of $\lambda$. Then $T(x) = \lambda x$. Let $B$ be the basis of $\mathbb{V}$ such that $[T]_B = A$. Then co-ordinate of $[T(x)]_B = A[x]_B \implies \lambda[x]_B = A[x]_B$. Since $x$ is non-zero vector hence co-ordinate of $x$ is also non-zero and $[x]_B$ is a vector in $\mathbb{F}^n$. Hence $\lambda$ is an eigenbvalue of $A$ and corresponding eigenvector is co-ordinate of $x$ with respect to the basis $B$.

Similarly we can show that if $\lambda$ is eigenvalue of $A$, then $\lambda$ is an eigenvalue of $T$. $\square$

**Corollary 4.4.** *Let $T$ be a linear operator on the finite-dimensional space $\mathbb{V}$ over the field $\mathbb{F}$. The eigenvalues of $T$ are the zeros of its characteristic polynomial.*

**Definition 4.18**

Let $\mathbb{V}(\mathbb{F})$ be a finite dimensional vector space and let $T \in \mathbb{L}(\mathbb{V}, \mathbb{V})$. Let $A$ be a matrix representation of $T$ with respect to a basis of $\mathbb{V}$. Then the characteristic and minimal polynomials of $T$ are the characteristic and minimal polynomial of $A$.

**Remark 4.11.** The above definition is make sense only because similar matrices have same characteristic polynomial. Otherwise for different different bases there are different different matrix representation of $T$, and hence $P_T(x)$ is not unique.

**Definition 4.19**

Let $\mathbb{V}(\mathbb{F})$ be a vector space and $T \in \mathbb{L}(\mathbb{V}, \mathbb{V})$. Then $T$ is called diagonalizable operator if there is a basis of $B$ of $\mathbb{V}$ such that each vector of $B$ is an eigenvector of $T$.

For finite dimensional case:

1. An operator is diagonalizable if there is a basis of $B$ of $\mathbb{V}$ such that $[T]_B$ is a diagonal matrix.

2. Let $A$ be a matrix representation of $T$. Then $T$ is diagonalizable iff $A$ is diagonalizable.

# 5 Types of Operators

**Theorem 5.1**

Let $(\mathbb{V}, \langle , \rangle)$ be a finite dimensional inner product space over $\mathbb{K}$ and let $f$ be a linear functional. Then there exists a unique vector $y \in \mathbb{V}$ such that $f(x) = \langle x, y \rangle$ for all $x \in \mathbb{V}$.

*Proof.* Let $\{u_1, \dots, u_n\}$ be an orthonormal basis of $\mathbb{V}$. Consider the vector $y = \sum_{i=1}^{n} \overline{f(u_i)} u_i$. Let $f_y(x) = \langle x, y \rangle$ be a linear functional.

$$f_y(u_k) = \langle u_k, y \rangle$$

$$= \langle u_k, \sum_{i=1}^{n} \overline{f(u_i)} u_i \rangle$$

$$= f(u_k)$$

This is true for each $u_k$, it follows that $f_y = f$.

We now prove the uniqueness. Suppose there exist $y_1$ and $y_2$ such that $f(x) = \langle x, y_1 \rangle$ and $\langle x, y_2 \rangle$ for all $x \in \mathbb{V}$. This implies $\langle x, y_1 \rangle = \langle x, y_1 \rangle$ for all $x \in \mathbb{V}$. Hence $y_1 = y_2$. □

**Definition 5.1**

Let $T$ be a linear operator on an inner product space $\mathbb{V}$. Then we say that $T$ has an **adjoint** on $\mathbb{V}$ if there exists a linear operator $T^*$ on $\mathbb{V}$ such that $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$ for all $x$ and $y$ in $\mathbb{V}$.

**Theorem 5.2**

For any linear operator $T$ on a finite-dimensional inner product space $(\mathbb{V}, \langle , \rangle)$, there exists a unique linear operator $T^*$ on $\mathbb{V}$ such that $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$ for all $x, y$ in $\mathbb{V}$.

*Proof.* Let $y$ be any vector in $\mathbb{V}$. Then $f(x) = \langle T(x), y \rangle$ for all $x \in \mathbb{V}$ is a linear functional on $\mathbb{V}$. Using previous theorem there is a unique vector $y'$ in $\mathbb{V}$ such that $f(x) = \langle x, y' \rangle$ for all $x \in \mathbb{V}$. We have seen that for each $y \in \mathbb{V}$ there is $y'$ in $\mathbb{V}$. We define a map $T^* : \mathbb{V} \to \mathbb{V}$ such that $T^*(y) = y'$.

We now show that $T^*$ is linear. Let $y_1, y_2 \in \mathbb{V}$ and $\alpha, \beta i$
$in \mathbb{K}$. Then

$$\begin{aligned}
\langle x, T^*(\alpha y_1 + \beta y_2)\rangle &= \langle T(x), (\alpha y_1 + \beta y_2)\rangle \\
&= \langle T(x), \alpha y_1\rangle + \langle T(x), \beta y_2\rangle \\
&= \overline{\alpha}\langle T(x), y_1\rangle + \overline{\beta}\langle T(x), y_2\rangle \\
&= \overline{\alpha}\langle x, T^*(y_1)\rangle + \overline{\beta}\langle x, T^*(y_2)\rangle \\
&= \langle x, \alpha T^*(y_1)\rangle + \langle x, \beta T^*(y_2)\rangle
\end{aligned}$$

This is true for all $x \in \mathbb{V}$. Hence $T^*(\alpha y_1 + \beta y_2) = \alpha T^*(y_1) + \beta T^*(y_2)$.

The uniqueness of $T^*$ is clear. For any $y$ in $\mathbb{V}$, the vector $T^*(y)$ is uniquely determined as the vector $y'$ such that $\langle T(x), y\rangle = \langle x, y'\rangle$ for all $x$ in $\mathbb{V}$. □

**Theorem 5.3**

Let $(\mathbb{V}, \langle,\rangle)$ be a finite dimensional inner product space over $\mathbb{K}$ and let $B = \{u_1, \ldots, u_n\}$ be an orthonormal basis for $\mathbb{V}$. Let $T$ be a linear operator on $\mathbb{V}$ and let $A$ be the matrix of $T$ in the ordered basis $B$. Then $A_{kj} = \langle T(\alpha_j), \alpha_k\rangle$.

**Theorem 5.4**

Let $(\mathbb{V}, \langle,\rangle)$ be a finite dimensional inner product space over $\mathbb{K}$ and let $T$ be a linear operator on $\mathbb{V}$. In any orthonormal basis for $\mathbb{V}$, the matrix of $T^*$ is the conjugate transpose of the matrix of $T$.

*Proof.* Let $\mathcal{B} = \{u_1, \ldots, u_n\}$ be an orthonormal basis for $\mathbb{V}$, let $A = [T]_{\mathcal{B}}$ and $B = [T^*]_{\mathcal{B}}$. Then

$A_{kj} = \langle T(u_j), u_k\rangle$.

$B_{kj} = \langle T^*(u_j), u_k\rangle$.

By the definition of $T^*$ we have

$B_{kj} = \langle T^*(u_j), u_k\rangle$

$$= \overline{\langle u_k, T^*(u_j)\rangle}$$

$$= \overline{\langle T(u_k), u_j\rangle}$$

$$= \overline{A_{kj}}. \square$$

**Definition 5.2**

Let Let $(\mathbb{V}, \langle , \rangle)$ be an inner product space over $\mathbb{K}$. A linear operator $T$ is called **self-adjoint operator** (or **Hermitian operator**) if $T = T^*$.

**Definition 5.3**

Let Let $(\mathbb{V}, \langle , \rangle)$ be an inner product space over $\mathbb{K}$. A linear operator $T$ is called **Skew-Hermitian operator** if $T = -T^*$.

**Definition 5.4**

Let Let $(\mathbb{V}, \langle , \rangle)$ be an inner product space over $\mathbb{K}$. A linear operator $T$ is called **Unitary operator** if $TT^* = T^*T = I$.

**Definition 5.5**

Let Let $(\mathbb{V}, \langle , \rangle)$ be an inner product space over $\mathbb{K}$. A linear operator $T$ is called **Normal operator** if $TT^* = T^*T$.

**Remark 5.1.** Let $(\mathbb{V}, \langle , \rangle)$ be a finite dimensional inner product space over $\mathbb{K}$ and let $T$ be a linear operator on $\mathbb{V}$.

1. $T$ is self-adjoint if $[T]_B$ is Hermitian for any orthonormal basis $B$.
2. $T$ is Skew-Hermitian if $[T]_B$ is Skew-Hermitian for any orthonormal basis $B$.
3. $T$ is Unitary if $[T]_B$ is Unitary for any orthonormal basis $B$.
4. $T$ is normal if $[T]_B$ is Normal for any orthonormal basis $B$.

**Important:** For each of the above cases, we must have an orthonormal basis. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ define by $T(x_1, x_2) = (x_2, x_1)$. We can easily check that $T$ is a Hermitian matrix. Take a basis $B = \{(1, 0), (1, 1)\}$. The basis $B$ is not orthonormal. Then

$[T]_B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$ which is not Hermitian. Only for othonormal basis $\mathcal{B}$ we have $[T]_{\mathcal{B}}$ is Hermitian (resp. Skew-Hermitian, Normal, Unitary) if $T$ is Hermitian (resp. Skew-Hermitian, Normal, Unitary).