Linear Algebra

October 1, 2025

Contents

Chapter 1

Vector Spaces

Linear algebra is the study of linear maps on finite-dimensional vector spaces.

1.1 Definition of Vector Spaces

Let $\mathbb{F}$ be a field, usually $\mathbb{R}$ or $\mathbb{C}$. The motivation of a vector space comes from the scalar multiplication in $\mathbb{F}^n$:

Definition 1.1.1: Vector Spaces

A Vector Field $(V, +, *)$ contains a set V and operations $+ : V \times V \to V$ and $* : \mathbb{F} \times V \to V$ such as the following proposes hold:
Addition:

- **Commutativity:** $\forall u, v \in V, u + v = v + u$
- **Associativity:** $\forall u, v, w \in V, (u + v) + w = u + (v + w)$
- **Identity:** $\exists \mathbf{0} \in V, \forall w \in V, v + \mathbf{0} = v$
- **Inverse:** $\forall v \in V, \exists w \in V, v + w = \mathbf{0}$

Multiplication:

- **Associativity:** $\forall v \in V, a, b \in \mathbb{F}, (ab)v = a(bv)$
- **Identity:** $\forall v \in V, 1v = v$, where 1 is the multiplicative identity of field $\mathbb{F}$.

Distribution: $\forall a, b \in \mathbb{F}, u, v \in V$:

- $a(u + v) = au + av$
- $(a + b)v = av + bv$

The scalar multiplication in a vector space depends on $\mathbb{F}$. We must say that V is a vector space *over* $\mathbb{F}$ to be precise.

Definition 1.1.2: Real and Complex Vector Spaces

- A vector space over $\mathbb{R}$ is called a real vector space.
- A vector space over $\mathbb{C}$ is called a complex vector space.

Vector Spaces has a very clear geometric meaning.

Definition 1.1.3: Vector and Points

Elements of a Vector Space are called vectors or points.

Notation 1.1.1: $\mathbb{F}^S$

If S is a set, then $\mathbb{F}^S$ denotes $\{f | f : S \to \mathbb{F}\}$
For $f, g \in \mathbb{F}^S$, the *sum* $f + g \in \mathbb{F}^S$ is the function defined by

$$(f+g)(x) = f(x) + g(x), \forall x \in S \tag{1.1}$$

For $\lambda \in \mathbb{F}$ and $f \in \mathbb{F}^S$, the product $\lambda f \in \mathbb{F}^S$ is the function defined by

$$(\lambda f)(x) = \lambda f(x), \forall x \in S \tag{1.2}$$

The usual notation $\mathbb{F}^n$ is actually $\mathbb{F}^{\{1,2,\cdots,n\}}$ and $\mathbb{F}^\infty$ is $\mathbb{F}^{\mathbb{N}}$

Then we have:

Example: **exp1**

$\mathbb{F}^S$ is a vector space.

Proof. Obvious. □

1.1.1 Some Properties of Vector Spaces

Theorem 1.1.1: Uniqueness of Identity

A vector space has a unique additive identity.

Proof. Let $\mathbf{0}$ and $\mathbf{0'}$ are identities, then we have

$$\mathbf{0} = \mathbf{0} + \mathbf{0'} = \mathbf{0'} + \mathbf{0} = \mathbf{0'}$$

□

Theorem 1.1.2: Uniqueness of Inverse

Every element $v \in V$ has a unique additive inverse, denotes $-v$

Proof. $\forall v \in V$, let u, w be inverses. Then we have:

$$v + u = \mathbf{0}, v + w = \mathbf{0}$$

$$w = w + \mathbf{0} = w + (v + u) = (w + v) + u = \mathbf{0} + u = u$$

$\square$

Thus, we denote $u - v = u + (-v)$

Theorem 1.1.3

$\forall v \in V, 0v = \mathbf{0}$

Proof. For $v \in V$, we have

$$0v = (0 + 0)v = 0v + 0v$$

adding the additive inverse of $0v$, we have

$$0v = \mathbf{0}$$

$\square$

Remark:
The distribution law is the only axiom connecting addition and scalar multiplication, thus must be used in the proof.
Note that the $\mathbf{0}$ is the additive identity in the vector space V, and 0 is the additive identity in the field $\mathbb{F}$

Example:

$\forall a \in \mathbb{F}, a\mathbf{0} = \mathbf{0}$

Proof. For $a \in \mathbb{F}$, we have

$$a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0} + a\mathbf{0}$$

Thus we get

$$a\mathbf{0} = \mathbf{0}$$

$\square$

Example:

1. $\forall v \in V, (-1)v = -v$
2. Suppose $a \in \mathbb{F}, v \in V, av = \mathbf{0}$, then $a = 0$ or $v = \mathbf{0}$.
3. $\emptyset$ is not a vector space

Proof. 1. Obvious

2. If $a \neq 0$, then a has inverse a^{-1}

$$a^{-1}av = a^{-1}\mathbf{0} = 0$$

$$v = \mathbf{0}$$

3. $\emptyset$ does not have additive identity.

$\square$

Remark:
In fact, in the definition of vector space, $\forall v \in V, 0v = \mathbf{0}$ can replace the additive inverse condition.

Proof. Using $\forall v = 0, 0v = \mathbf{0}$, then $(-1)v$ is the additive inverse, because

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = \mathbf{0}$$

$\square$

1.1.2 Complex Vector Spaces from Real Vector Spaces

Now we take a look at how to generate a complex vector space from a real vector space.

Suppose V is a real vector space. Denote $V_{\mathbb{C}} = V \times V$, we write $(u, v) \in V_{\mathbb{C}}$ as $u + iv$, and define

- Addition:
$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2)$$
- Complex scalar multiplication:
$$(a + bi)(u + iv) = (au - bv) + i(av + bu)$$
  for all $a, b \in \mathbb{R}$ and all $u, v \in V$

Then $V_{\mathbb{C}}$ is a vector space on $\mathbb{C}$

1.2 Subspaces

1.2.1 Definition

Definition 1.2.1: Subspaces

A subset U is the subspace of a vector space V if U is also a vector space with the same additive and scalar multiplicative property.

We can easily identify a subspace by checking:

Theorem 1.2.1: Conditions for Subspaces

A subset U of V is a subspace iff U satisfies:

- **Additive Identity:** $\mathbf{0} \in U$ (Can be replaced by $U \neq \emptyset$)
- **Closed under Addition:** $\forall u, w \in U, u + w \in U$
- **Closed under scalar multiplication:** $\forall a \in \mathbb{F}, u \in U, au \in U$

This is easily proved because other properties such as associativity and distribution dose not depends on the range.

Remark:
The Additive Identity is necessary to avoid $\emptyset$, it can be replaced by $U \neq \emptyset$ because take any $u \in U$ we shall have $0u = \mathbf{0} \in U$ by the closure under scalar multiplication.

Example: **Subspaces**

- $\{(x,0) : x \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^2$,which means the x-axis in the 2-dimensional plane.

1.2.2 Sums of subspace

We now generate bigger subspaces from smaller ones.

Definition 1.2.2: Sums of subspaces

Suppose $V_1, \ldots, V_m$ are subspaces of V. The *sum* of $V_1, \ldots, V_m$ is defined:

$$V_1 + \ldots + V_m = \{v_1 + \ldots + v_m : v_i \in V_i, i = 1, 2, \ldots, m\} \tag{1.3}$$

Theorem 1.2.2: Sums of Subspaces are Smallest

Suppose $V_1, \ldots, V_m$ are subspaces of V. Then $V_1 + \ldots + V_m$ is a subspace of V and $\forall$ subspace $U \in V$, we have $V_1 + \ldots + V_m \subseteq U$

Proof. $V_1 + \ldots + V_m$ contains $\mathbf{0} = \mathbf{0} + \ldots + \mathbf{0}$ and is closed under addition and scalar multiplication. Therefore it is a subspace.

$\forall u \in V_1 + \ldots + V_m$, let $u = v_1 + \ldots + v_m$, because $v_i \in V_i \subseteq U$, then $v_1 + \ldots + v_m \in U$ using the closure under addition. □

Remark:
Sums of subspaces are analogous to unions of sets in set theory. $\forall S_1, \ldots, S_n$, the smallest set containing all of them is $S_1 \cap S_2 \cap \ldots \cap S_n$

The following are some examples:

Example: **Sums of Subspaces**

- For the vector space $\mathbb{R}^3$, the sum of x-axis and y-axis are the x-y plane.
- For $\mathbb{R}^3$, x-axis + x-axis = x-axis

Remark:
Intuitively, the sum of subspaces are the "linear span" of subspaces. Like two axis form a

plane in $\mathbb{R}^3$

1.2.3 Direct Sums

The intuition of direct sums comes from whether the two subspaces are "independent". Every element of $V_1 + \ldots + V_m$ can be written as $v_1 + \ldots + v_m$. For "independent" subspaces, the factorization is unique.

Definition 1.2.3: Direct Sum

Suppose $V_1, \ldots, V_m$ are subspaces of V. The sum $V_1 + \ldots + V_m$ is a direct sum iff $\forall u \in V_1 + \ldots + V_m$ can be written in *only one way* as a sum $v_1 + \ldots + v_m$, where $v_i \in V_i$.
In this case, we denote $V_1 \oplus V_2 \oplus \ldots \oplus V_m$

Example: **Direct Sum**

In $\mathbb{F}^3$, let $U = \{(x, y, 0) \in \mathbb{F}^3 : x, y \in \mathbb{F}\}$ and $W = \{(0, 0, z) \in \mathbb{F}^3 : z \in \mathbb{F}\}$. Then $\mathbb{F}^3 = U \oplus W$
While $\mathbb{F} = \mathbb{F} + \mathbb{F}$ is not a direct sum.

To identify direct sums, we only need to check whether $\mathbf{0}$ can be uniquely written as an appropriate sum.

Theorem 1.2.3: Conditions for a direct sum

Suppose $V_1, \ldots, V_m$ are subspaces of V. Then $V_1 + \ldots + V_m$ is a direct sum iff:

$$\forall v_i \in V_i \ (v_1 + \ldots + v_m = \mathbf{0} \to v_1 = \ldots = v_m = \mathbf{0}) \tag{1.4}$$

Proof. If $V_1 + \ldots + V_m$ is a direct sum, then $\mathbf{0} = \mathbf{0} + \ldots + \mathbf{0}$ is unique.
The other side, if $\mathbf{0}$ can be uniquely written, $\forall v \in V$, let

$$u = v_1 + \ldots + v_m = u_1 + \ldots + u_m$$

Then we have

$$(v_1 - u_1) + \ldots + (v_m - u_m) = \mathbf{0}$$

which means $u_i = v_i$ for all $i \in \{1, 2, \ldots, m\}$ □

Remark:
Intuitively, two planes in $\mathbb{R}^3$ are not direct sums because they intersects a line, making $\mathbf{0}$-the origin, not uniquely written.
This intuition gives us another condition concerning 2 subspaces.

Theorem 1.2.4: Direct sum of 2 subspaces

Suppose $U, W \subseteq V$ are subspaces, then

$$U + W \text{ is a direct sum } \leftrightarrow U \cap W = \{\mathbf{0}\} \tag{1.5}$$

Proof. If $U+W$ is a direct sum, them let $v \in U \cap W$, then $-v \in U \cap W$ then $\mathbf{0} = v + (-v)$, where $v \in V$ and $-v \in W$, then by uniqueness, $v = \mathbf{0}$.

If $U \cap W = \{\mathbf{0}\}$, let $v = u_1 + w_1 = u_2 + w_2 \in U + W$, where $u_1, u_2 \in U, w_1, w_2 \in W$, then $(u_1 - u_2) + (w_1 - w_2) = \mathbf{0}$, where $u_1 - u_2 \in U, w_1 - w_2 \in W$, then $u_1 = u_2, w_1 = w_2$. □

Remark:
The above criterion only deals with 2 subspaces, for more subspaces, merely $V_i \cap V_j = \{\mathbf{0}\}$ is not suffice. For example, three lines in a plane that contains origin is not a direct product.

Example: **Subspaces**

1. $\mathbb{Q}$ is not a subspace of $\mathbb{R}$ under $\mathbb{R}$.
2. x-axis $\cup$ y-axis is closed under scalar multiplication but is not a subspace of $\mathbb{R}^2$.

1.2.4 Some Theorems and Properties

Theorem 1.2.5

Suppose V_1, V_2 are subspaces of V, then $V_1 \cap V_2$ is a subspace of V.
Furthermore, for any collection of subspaces $v_\alpha, \alpha \in I$, the intersection $\bigcap_{\alpha \in I} V_\alpha$ is a subspace of V.

Proof. First, as $\forall \alpha \in I, \mathbf{0} \in V_\alpha$, then $\mathbf{0} \in \bigcap V_\alpha$.
If $u, v \in \bigcap V_\alpha$, then $\forall \alpha, u, v \in V_\alpha$, then $u + v \in V_\alpha$, then $u + v \in \bigcap V_\alpha$.
scalar multiplication is similar. □

Theorem 1.2.6

Suppose U, W are subspaces of V, then

$$U \cup W \text{ is subspace of } V \;\leftrightarrow U \subseteq W \text{ or } W \subseteq U$$

Proof. If $\exists u \in U, u \notin W, w \in W, w \notin U$, let $v = u + w$.
If $v \in U$, then $w = v - u \in U$, contradicts. Similarly, $v \notin W$, therefore $v \notin U \cup W$, which means $U \cup W$ is not a subspace. □

Remark:
One may find that the sum of subspaces satisfies a commutative semi-group that has an identity. That is, for subspaces of V,

1. **Commutativity:** $U + W = W + U$
2. **Associativity:** $(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)$
3. **Identity:** $\{\mathbf{0}\}$

However, subspaces does not have additive inverse apart from $\{\mathbf{0}\}$. That is, the cancellation law does not hold.

Chapter 2

Finite-Dimensional Vector Spaces

Standing Assumptions:

- $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$
- V denotes a vector space over $\mathbb{F}$.

2.1 Span and Linear Independence

2.1.1 Linear Combination and Span

Definition 2.1.1: Linear Combination

A *linear combination* of a list $v_1, \ldots, v_m$ of vectors in V is a vector:

$$a_1 v_1 + \ldots + a_m v_m \tag{2.1}$$

where $a_1, \ldots, a_m \in \mathbb{F}$

Definition 2.1.2: Span

The set of all linear combinations of a list of vectors is called the *span* denoted

$$\operatorname{span}(v_1, \ldots, v_m) = \{a_1 v_1 + \ldots + a_m v_m : a_1, \ldots, a_m \in \mathbb{F}\} \tag{2.2}$$

The span of the empty list is defined

$$\operatorname{span}() = \{\mathbf{0}\} \tag{2.3}$$

Span is sometimes also called *linear span*

Theorem 2.1.1: Span is a subspace

For every list $v_1, \ldots, v_m$ in V, $\operatorname{span}(v_1, \ldots, v_m)$ is a subspace in V.

Proof. Let $U = \operatorname{span}(v_1, \ldots, v_m)$, we have $\mathbf{0} = 0v_1 + \ldots + 0v_m \in U$.

Also U is closed under addition and scalar multiplication. □

Theorem 2.1.2: Span is the Smallest Containing Subspace

$\forall$ subspace $U \subseteq V$ such that $v_1, \ldots, v_m \in U$, we have

$$\operatorname{span}(v_1, \ldots, v_m) \subseteq U$$

Proof. Obvious □

Definition 2.1.3: Spans

If $\operatorname{span}(v_1, \ldots, v_m) = V$, then we say the list $v_1, \ldots, v_m$ spans V.

Example: **Spans**

For $\mathbb{F}^n$, we have

$$(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1) \tag{2.4}$$

spans $\mathbb{F}^n$.

Definition 2.1.4: Finite Dimensional Vector Spaces

A vector space is called finite dimensional if some list (finite) of vectors spans the space.

The example above shows that $\mathbb{F}^n$ is finite dimensional. Note that every list has finite length.

Notation 2.1.1: Polynomials

We denote $\mathscr{P}(\mathbb{F})$ to polynomials on field $\mathbb{F}$, and $\mathscr{P}_m(\mathbb{F})$ denotes all polynomials with degree at most m on $\mathbb{F}$.

Example: **Polynomials**

Obviously the polynomials $\mathscr{P}(\mathbb{F})$ and $\mathscr{P}_m(\mathbb{F})$ are vector spaces.

- If $m \geq 0$ is an integer, then $\mathscr{P}_m(\mathbb{F}) = \operatorname{span}(1, x, \ldots, x^m)$.
- $\mathscr{P}(\mathbb{F})$ is infinite-dimensional.

  Proof. For any list of elements in $\mathscr{P}(\mathbb{F})$, Let m be the maximum degree. Then any linear combination has degree $\leq m$, which means x^{m+1} is not a linear combination. Therefore $\mathscr{P}(\mathbb{F})$ has infinite dimension. □

Remark:
We've noticed some similarities between spans and the sums of vector spaces. In fact, $\forall v \in V$, the set $\{\lambda v : \lambda \in \mathbb{F}\}$ is a subspace of V, and the span of $v_1, \ldots, v_m$ is actually $V_1 \oplus V_2 \oplus, \ldots, \oplus V_m$ where $V_i = \{\lambda v_i : \lambda \in \mathbb{F}\}$.

2.1.2 Linear Independence

Suppose $v_1, \ldots, v_m \in V$ and $v \in \operatorname{span}(v_1, \ldots, v_m)$. Then there exists $a_1, \ldots, a_m \in \mathbb{F}$ such that

$$v = a_1 v_1 + \ldots + a_m v_m$$

We want to find the condition in which the choice of the scalars is unique. This is also similar to the direct sum condition.

Definition 2.1.5: Linear Independent

A list $v_1, \ldots, v_m$ in V is linear independent if

$$\forall a_1, \ldots, a_m \in \mathbb{F}, \ (a_1 v_1 + \ldots + a_m v_m = \mathbf{0} \to a_1 = ... = a_m = 0)$$

In other words, $v_1, \ldots, v_m$ is linear dependent if $a_1, \ldots, a_m$ not all 0 such that $a_1 v_1 + \ldots + a_m v_m = \mathbf{0}$.

Remark:
We can see that $v_1, \ldots, v_m$ is linear independent is equivalent to $\{\lambda v_i : \lambda \in \mathbb{F}\}$ forms a direct sum.
Intuitively, linear independence means that the vectors do not "fall in the same plane". For example, three vectors that are coplanar are dependent in $\mathbb{R}^3$. This picture also gives us an intuition to understand the following lemma

Lemma 2.1.1: Linear Independence Lemma

Suppose $v_1, \ldots, v_m$ is a linear dependent list in V. Then there exists $k \in \{1, 2, \ldots, m\}$ such that

$$v_k \in \operatorname{span}(v_1, \ldots, v_{k-1}) \tag{2.5}$$

Furthermore, removing v_k from the list, the remaining span dose not change.

$$\operatorname{span}(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m) = \operatorname{span}(v_1, \ldots, v_m) \tag{2.6}$$

Proof. Because $v_1, \ldots, v_m$ is linear dependent, $\exists a_1, \ldots, a_m \in \mathbb{F}$ not all 0 such that

$$a_1 v_1 + \ldots + a_m v_m = \mathbf{0}$$

Let k be the largest element in $\{1, \ldots, m\}$ such that $a_k \neq 0$.

If $k = 1$ then $a_1 v_1 = \mathbf{0}$ and $a_1 \neq 0$, then $v_1 = \mathbf{0} \in \operatorname{span}()$

If $k \geq 2$, then

$$a_1 v_1 + \ldots + a_k v_k = \mathbf{0}$$

$$v_k = -\frac{a_1}{a_k} v_1 - \ldots - \frac{a_{k-1}}{a_k} v_{k-1} \tag{2.7}$$

Thus $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$

For any $v = b_1 v_1 + \ldots + b_m v_m$, we can write v_k as above. □

Remark:
Now we may guess that linear independence depicts the least number of vectors that can span the entire space. The vectors that can be written as linear combination of other vectors is indeed unnecessary.
This lemma also gives us a way to expand the independent vector list.

Corollary 2.1.1: Expanding the independent list

If $u_1, \dots, u_{n-1}$ are independent vectors, then

$$u_n \notin \operatorname{span}(u_1, \dots, u_{n-1}) \quad \leftrightarrow \quad u_1, \dots, u_n \text{ is independent} \tag{2.8}$$

Proof. If not, $\exists k \leq n-1$, such that $u_k \in \operatorname{span}(u_1, \dots, u_{k-1})$, contradicts. □

Theorem 2.1.3: Length of linearly independent list

In a finite-dimensional vector space, the length of every linearly independent list of vectors $\leq$ the length of every spanning list of vectors.
That is, $\forall u_1, \dots, u_m$ is linearly independent in V. And $\operatorname{span}(w_1, \dots, w_n) = V$, then $m \leq n$.

Proof. We do so in a process as follows.

- **Step 1**

  Because u_1 can be written as linear combination of $w_1, \dots, w_n$. The list

  $$u_1, w_1, \dots, w_n$$

  is linear dependent.

  Thus by lemma 2.1.1, one of the vectors is the span of previous vectors. For $u_1 \neq \mathbf{0}$, this vector is some w_k, delete w_k and let the remaining list be B. Then B still spans V.

- **Step k, for $k = 2, 3, \dots, m$**

  Similarly, adding u_k in front of B and delete some w_i to generate a new B (as $u_1, \dots, u_k$ are linear independent).

  During this process, there must be some w_i left before $k = m$ because if there are only u_i's is the list, they cannot span V because u_n is not a linear combination of previous u_i.

After m steps, we get $B = u_m, \dots, u_1,$ some w_i, and the list length of B does not change during the process. Thus $m \leq n$ □

Example:

No list of length $n-1$ spans $\mathbb{R}^n$.

Proof. with the basis $\boldsymbol{e_i}$ spans $\mathbb{R}^n$ and are linearly independent. □

Theorem 2.1.4: Finite-dimensional subspaces

Every subspace of a finite-dimensional vector space is finite-dimensional.

Proof. Suppose V is finite dimensional and U is a subspace. We construct a list of linearly independent vectors that spans U using the following process.

- **Step 1**

  If $U = \{\mathbf{0}\}$ then we are done. If not, let $u_1 \neq 0, u_1 \in U$.

- **Step** k

  If $U = \text{span}(u_1, \ldots, u_{k-1})$, then we are done. If not, let $u_k \in U, u_k \notin \text{span}(u_1, \ldots, u_{k-1})$, then by corollary 2.1.1, $u_1, \ldots, u_k$ is independent.

As the length of this list increases, it cannot surpass the length of any spanning list in V, Thus the process eventually ends with a finite step, which means that U is finite-dimensional. □

Example: **Span and Linear Independence**

1. If $v_1, \ldots, v_m$ is linear independent in V, then $\forall \lambda \in \mathbb{F}$, we have $\lambda v_1, \ldots, \lambda v_m$ in linear independent.

2. Suppose $v_1, \ldots, v_m$ is a list of vectors in V. For $k \in \{1, 2, \ldots, m\}$, let

$$w_k = v_1 + \ldots + v_k$$

   then we have

$$v_1, \ldots, v_m \text{ is linear independent } \leftrightarrow w_1, \ldots, w_m \text{ is linear independent}$$

3. A vector space V is infinite-dimensional iff $\exists$ a sequence $v_1, \ldots$ of vectors in V such that $\forall m \in \mathbb{N}, v_1, \ldots, v_m$ is linear independent.

   Proof.
   - if V is infinite-dimensional, we construct the list using the steps of theorem 2.1.4. As V cannot be the span of $v_1, \ldots, v_m$, the process would not stop.
   - if V is finite-dimensional, then let independent $u_1, \ldots, u_m$ spans V. Then $u_1, \ldots, u_m, u_1$ spans V and are linear dependent, then any independent list must have length $\leq m + 1$, contradicts.

   □

4. $\mathbb{F}^\infty$ is infinite-dimensional.

2.2 Basis

In the previous section, we give an intuition that independent depicts the "smallest" number of vectors that is needed to span the space. We now give the definition of basis.

Definition 2.2.1: Basis

A *basis* of V is a list of vectors in V that are linear independent and span V.

Example: **Basis**

- The list $(1, 0, \ldots 0), \ldots, (0, 0, \ldots, 1)$ is a basis of $\mathbb{F}^n$, called the standard basis of $\mathbb{F}^n$.

Theorem 2.2.1: Criterion for Basis

A list $v_1, \ldots, v_n$ in V is a basis of V iff $\forall v \in V$ can be uniquely written as

$$v = a_1 v_1 + \ldots + a_n v_n \tag{2.9}$$

where $a_1, \ldots, a_n \in \mathbb{F}$.

Proof. This is indeed the intuition that drives us to define linear independence in definition 2.1.5.

- First suppose $v_1, \ldots, v_n$ is a basis of V. Let $v \in V$. Because $V = \text{span}(v_1, \ldots, v_n)$, we have

  $$v = a_1 v_1 + \ldots + a_n v_n$$

  If $v = c_1 v_1 + \ldots + c_n v_n$ then

  $$(a_1 - c_1)v_1 + \ldots + (a_n - c_n)v_n = \mathbf{0}$$

  which means $a_i = c_i$ for $i = 1, 2, \ldots, n$. Meaning unique.

- the other side is exactly the same.

$\square$

We now give a method to find a basis in a linear independent list.

Theorem 2.2.2: Every Spanning List Contains a Basis

Every spanning list of a vector space V can be reduced to a basis of the vector space. (It contains a sub-list that is a basis).

Proof. Suppose $V = \text{span}(v_1, \ldots, v_n)$. We want to remove some of the vectors from the list so that the remaining still span V.

- If $v_1 = \mathbf{0}$, then delete v_1
- If $v_k \in \text{span}(v_1, \ldots, v_{k-1})$, then delete v_k. (For $k = 2, \ldots, n$)

Then the remaining list in linearly independent and still spans V. $\square$

We now come to an important corollary.

Corollary 2.2.1: Basis of finite-dimensional vector space

Every finite-dimensional vector space have a basis.

Proof. A finite dimensional vector space have a spanning list that can be reduced to a basis. $\square$

Theorem 2.2.3: Every linearly independent list extends to a basis

Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis.

Proof. let $u_1, \dots, u_m$ be linearly independent, and $w_1, \dots, w_n$ spans V. Then the list

$$u_1, \dots, u_m, w_1, \dots, w_n$$

spans V. Reduce it using the method in 2.2.2, we get a basis. Note that $u_1, \dots, u_m$ would not be deleted in the process. $\square$

Remark:
Theorem 2.2.2 and theorem 2.2.3 is somewhat dual to each other. They depict how a basis is formed in a vector space. Intuitively, a basis is like the axis of a "coordinate system" of the vector space.

Corollary 2.2.2: Every Subspace of V is a part of a direct sum equal to V

Suppose V is finite-dimensional and U is a subspace. Then $\exists$ a subspace W of V such that $V = U \oplus W$.

Proof. By 2.1.4, U is finite-dimensional. Let $u_1, \dots, u_m$ be a basis of U, then extends it to a basis of V.

$$u_1, \dots, u_m, w_1, \dots, w_n$$

Let $W = \text{span}(w_1, \dots, w_n)$ would do. $\square$

Theorem 2.2.4: Basis of Complex vector space

If $v_1, \dots, v_n$ is a basis if a real vector space V, then it is also a basis of the complex vector space $V_{\mathbb{C}}$.

Proof. $\forall v \in V, v = a_1 v_1 + \ldots + a_n v_n$, where $a_k \in \mathbb{R}$ is unique. Then $\forall v = u + iw \in V_{\mathbb{C}}, u, w \in V$ we have

$$v = z_1 v_1 + \ldots + z_n v_n$$

Let $z_k = a_k + ib_k$ where $a_k, b_k \in \mathbb{R}$. Then we have

$$u = a_1 v_1 + \ldots + a_n v_n, \quad w = b_1 v_1 + \ldots + b_n v_n$$

Then a_k, b_k exists and is unique. $\square$

2.3 Dimensions

Our intuition suggests that the dimension for $\mathbb{F}^n$ should me n, that is, the number of standard basis. To expand this intuition to an arbitrary vector space, we can show that basis length does not depend on basis.

Theorem 2.3.1: Basis Length does not depend on Basis

For a finite-dimensional vector space, any basis have same length.

Proof. Suppose V is finite-dimensional, let B_1, B_2 be two basis. Because B_1 is linear independent and B_2 spans V, then the length of $B_2 \geq$ the length of B_1. The other way is exactly the same. □

Now we define the dimension of a vector space.

Definition 2.3.1: Dimensions

The dimension of a finite-dimensional vector space V is the length of its basis, denote $\dim V$

Example: **Dimensions of vector spaces**

- $\dim \mathbb{F}^n = n$ for its standard basis has length n.
- $\dim \mathscr{P}_m(\mathbb{F}) = m + 1$ for the standard basis $1, x, \ldots, x^m$.
- The vector space $\mathbb{C}$ on $\mathbb{C}$ has dimension 1, but $\mathbb{C}$ on $\mathbb{R}$ has dimension 2.

Theorem 2.3.2: Dimension of Subspace

if V is finite-dimensional, $U \subseteq V$ is a subspace, then $\dim U \leq \dim V$.

Proof. The basis in U can be expanded to form a basis in V, as in 2.1.1. □

Using theorem 2.3.1, we can simplify our criterion for a basis.

Theorem 2.3.3: Linear independent list of length $\dim$ is a basis

Suppose V is finite-dimensional, then every independent list of length $\dim V$ is a basis of V.

Proof. Suppose $\dim V = n$, and $v_1, \ldots, v_n$ is linear independent. If $v_1, \ldots, v_n$ does not span V, it can be expanded to form a basis, which has length $\geq n$, contradicts. □

Corollary 2.3.1: Subspace of full dimension

If U is a subspace of V, and $\dim U = \dim V$, then $U = V$

Proof. Obvious. □

On the other hand, have the right length and spanning the space all does the trick.

Theorem 2.3.4: Spanning list of length $\dim$ is basis

Suppose V is finite-dimensional, then $\forall$ spanning list B that has length $\dim V$ is a basis.

Proof. Similarly, if not, reduce it. □

Remark:
Theorems 2.2.1, 2.3.3, 2.3.4 tells us that to identify a basis, any 2 of the following:

- A spanning list.
- Linear Independent.

- Has length $\dim V$.

would suffice.

Theorem 2.3.5: Dimension of sum

If V_1, V_2 are 2 subspaces of a finite-dimensional vector space V, then

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim V_1 \cap V_2. \tag{2.10}$$

Proof. For $V_1 \cap V_2$ is also a subspace of V_1 and V_2, let $v_1, \ldots, v_n$ be a basis of $V_1 \cap V_2$, expand it to a basis $v_1, \ldots, v_n, u_1, \ldots, u_m$ of V_1, and $v_1, \ldots, v_n, w_1, \ldots, w_k$ of V_2. Then we will show that

$$v_1, \ldots, v_n, u_1, \ldots, u_m, w_1, \ldots, u_k$$

is a basis of $V_1 + V_2$. Denote it with B.

First, as all of the vectors are in $V_1 \cup V_2$, then they are in $V_1 + V_2$. And the spanning contains V_1 and V_2 so $\operatorname{span} B = V_1 + V_2$, as the sum is the smallest.

Then we can show they are linear independent.

$$a_1v_1 + \ldots + a_nv_n + b_1u_1 + \ldots + b_mu_m + c_1w_1 + \ldots + c_kw_k = \mathbf{0} \tag{2.11}$$

then

$$c_1w_1 + \ldots + c_kw_k = -a_1v_1 - \ldots - a_nv_n - b_1u_1 - \ldots - b_mu_m \in V_1$$

but $c_1w_1 + \ldots + c_kw_k \in V_2$, therefore $c_1w_1 + \ldots + c_kw_k \in V_1 \cap V_2$.

$$c_1w_1 + \ldots + c_kw_k = d_1v_1 + \ldots + d_nv_n$$

then $c_1 = \ldots = c_k = 0$. Then all of the coefficient is 0. □

Corollary 2.3.2: Direst Sum Dimension

$$\dim V_1 \oplus V_2 = \dim V_1 + \dim V_2 \tag{2.12}$$

Remark:
We have noticed astounding similarities between sets and vector spaces

set	**vector space**
S is a finite set	V is a finite-dimensional vector space
$\#S$	$\dim V$
$S_1 \cup S_2$ is the smallest containing both	$V_1 + V_2$ is the smallest contain both
$\#S_1 \cup S_2 = \#S_1 + \#S_2 - \#S_1 \cap S_2$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim V_1 \cap V_2$

Example: **Dimensions**

1. **Bernstein polynomials**

Suppose $m \in \mathbb{Z}_+$, for $0 \le k \le m$, let

$$p_k(x) = x^k(1-x)^{m-k} \tag{2.13}$$

then $p_0, \ldots, p_m$ is a basis of $\mathscr{P}_m(\mathbb{F})$.

Proof. let

$$c_0p_0 + \ldots + c_mp_m = \mathbf{0}$$

that is, $\forall x \in \mathbb{R}$

$$c_0p_0(x) + \ldots + c_mp_m(x) = 0$$

- let $x = 0$ we get $c_0 = 0$.
- if $c_0 = \ldots = c_{k-1} = 0$, we take k^{th} derivative and let $x = 0$, then $c_k = 0$

then $c_0 = \ldots = c_m = 0$, that is, $p_0, \ldots, p_m$ is linear independent. □

2. Suppose $V_1, \ldots, V_m$ are finite dimensional subspaces of V. Then $V_1 + \ldots + V_m$ is finite dimensional and

$$\dim(V_1 + \ldots + V_m) \le \dim V_1 + \ldots + \dim V_m \tag{2.14}$$

Proof. using $\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim V_1 \cap V_2$ and induction. □

In fact, equality hold iff all were direct sums.

2.4 The Structure of Finite-dimensional vector space

We've already have the intuition that the basis of vector space are somewhat like a coordinate system. We shall push this intuition further and prove that any vector space can be understand as $\mathbb{F}^n$.

Theorem 2.4.1: The Structure of Finite-dimensional vector space

If V is a vector space over $\mathbb{F}$ and $\dim V = n$, then $V \cong \mathbb{F}^n$.

Proof. Take $v_1, \ldots, v_n$ be a basis for V. Define a mapping $T : V \to \mathbb{F}^n$ as follows:

$$T(v_i) = \boldsymbol{e_i}, i = 1, 2, \ldots, n$$

then $\forall v \in V$, let $v = c_1v_1 + \ldots + c_nv_n$, we have

$$T(v) = \sum_{i=1}^{n} c_i\boldsymbol{e_i}$$

It is easy to check that T is a bijection. And satisfies

- $T(v_1 + v_2) = T(v_1) + T(v_2)$
- $T(\lambda v) = \lambda T(v)$

□

Chapter 3

Linear Maps

Standing Assumption

- $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.
- U, V, W denotes vector spaces.

3.1 Vector Space of Linear Maps

This is one of the key definitions of linear algebra.

Definition 3.1.1: Linear Maps

A *linear map* from V to W is a function $T : V \to W$ with the following properties.

- **Activity:** $\forall u, v \in V, T(u+v) = T(u) + T(v)$
- **Homogeneity:** $\forall \lambda \in \mathbb{F}, v \in V, T(\lambda v) = \lambda T(v)$.

We often denote Tv for $T(v)$ for linear maps.

Notation 3.1.1: $\mathscr{L}(V, W)$ and $\mathscr{L}(V)$

- The set of linear maps form V to W is denoted $\mathscr{L}(V, W)$.
- The set of linear maps from V to V is denoted $\mathscr{L}(V)$, that is, $\mathscr{L}(V) = \mathscr{L}(V, V)$.

Example: **Linear Map**

- **zero**

  The zero operator $0 \in \mathscr{L}(V, W)$ is defined by

$$\forall v \in V, 0v = 0$$

- **Identity Operator**

The identity operator $1 \in \mathscr{L}(V)$ is defined as

$$\forall v \in V, 1v = v$$

- **Differentiation**

  Define $D \in \mathscr{L}(\mathscr{P}(\mathbb{R}))$ by
  $$Dp = p'$$

- **Integration**

  Define $T \in \mathscr{L}(\mathscr{P}(\mathbb{R}), \mathbb{R})$ by
  $$Tp = \int_0^1 p(x)\mathrm{d}x$$

We now describe linear maps by the basis.

Lemma 3.1.1: Linear Map Lemma

Suppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_n \in W$. Then there exists a unique linear map $T : V \to W$ such that
$$Tv_k = w_k$$
for all $k = 1, \ldots, n$.

Proof.

- **Existence:** We define $\forall v \in V$, let $v = c_1 v_1 + \ldots + c_n v_n$
  $$Tv = c_1 w_1 + \ldots + c_n w_n \tag{3.1}$$
  It is easy to check that T is a linear map.

- **Uniqueness:** If $Tv_k = w_k$ then $\forall v = c_1 v_1 + \ldots + c_n v_n \in V$
  $$Tv = c_1 w_1 + \ldots + c_n w_n$$
  is uniquely determined.

□

3.1.1 Algebraic Operations on $\mathscr{L}(V, W)$

We begin with addition and scalar multiplication.

Definition 3.1.2: Addition and Scalar Multiplication on $\mathscr{L}(V, W)$

Suppose $S, T \in \mathscr{L}(V, W)$ and $\lambda \in \mathbb{F}$. Define

- $S + T : \forall v \in V, (S + T)v = Sv + Tv$
- $\lambda T : \forall v \in V, (\lambda T)v = \lambda(Tv)$

Note that it is easy to check that $S + T \in \mathscr{L}(V, W)$ and $\lambda T \in \mathscr{L}(V, W)$.

Theorem 3.1.1: $\mathscr{L}(V,W)$ is a vector space

$\mathscr{L}(V,W)$ is a vector space with the definition in 3.1.2.

We now define another operation on $\mathscr{L}(V,W)$: Multiplication.

Definition 3.1.3: Product of linear maps

If $T, S \in \mathscr{L}(V,W)$, then the *product* $ST \in \mathscr{L}(V,W)$ is defined as

$$\forall v \in V, (ST)v = S(Tv) \tag{3.2}$$

that is, ST is just $S \circ T$ of functions.

Theorem 3.1.2: Algebraic properties of products

- **Associativity:** $T_1(T_2T_3) = (T_1T_2)T_3$
- **Identity:** $\forall T \in \mathscr{L}(V,W), TI = IT = T$
- **Distribution:** $(S_1 + S_2)T = S_1T + S_2T$ and $T(S_1 + S_2) = TS_1 + TS_2$

Remark:
Note that the addition and multiplication on $\mathscr{L}(V,W)$ forms a commutative ring with identity.
Multiplication of linear maps are not necessary to be commutative!

Theorem 3.1.3

Suppose $T \in \mathscr{L}(V,W)$ then

$$T(0) = 0$$

Proof. T(0) = T(0+0) = T(0)+T(0) □

Theorem 3.1.4: the form of linear maps

Suppose $T \in \mathscr{L}(\mathbb{F}^n, \mathbb{F}^m)$, then $\exists A_{jk} \in \mathbb{F}$ for $j = 1, \dots, m$ and $k = 1, \dots, n$ such that

$$T(x_1, \dots, x_n) = (\sum_{i=1}^{n} A_{1i}x_i, \dots, \sum_{i=1}^{n} A_{mi}x_i) \tag{3.3}$$

Proof. Let $\boldsymbol{e_i}$ denote the i^{th} standard basis of $\mathbb{F}^n$. Let

$$T\boldsymbol{e_i} = (A_{1i}, \dots, A_{mi})$$

Because

$$(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i e_i$$

Then

$$T(x_1,\ldots,x_n)=(\sum_{i=1}^n A_{1i}x_i,\ldots,\sum_{i=1}^n A_{mi}x_i)$$

□

In the last chapter, we see that $V\cong\mathbb{F}^n$, now we state that $\mathscr{L}(V,W)\cong\mathscr{L}(\mathbb{F}^n,\mathbb{F}^m)$ similarly.

Theorem 3.1.5: The Structure of $\mathscr{L}(V,W)$

If $\dim V=n, \dim W=m$, then $\mathscr{L}(V,W)\cong\mathscr{L}(\mathbb{F}^n,\mathbb{F}^m)$ in the sense of addition, scalar multiplication and multiplication.

Proof. Let $v_1,\ldots,v_n$ be a basis of V, and $w_1,\ldots,w_m$ be a basis of W. We take a bijection $F:\mathscr{L}(V,W)\to\mathscr{L}(\mathbb{F}_n,\mathbb{F}_m)$, such that

$$\begin{aligned} T:&\quad T(c_1v_1+\ldots+c_nv_n)=a_1w_1+\ldots+a_mw_m \\ F(T):&\quad F(T)(c_1,\ldots,c_n)=(a_1,\ldots,a_m)\end{aligned} \tag{3.4}$$

would do. □

Using this isomorphism, the result of theorem 3.1.4 can be used for arbitrary $\mathscr{L}(V,W)$.

Example:

- Suppose V is finite-dimensional and $T\in\mathscr{L}(V)$. Then T is a scalar multiple of the identity iff $\forall S\in\mathscr{L}(V), ST=TS$.

  Proof. The left to right side is easy. To resolve the other side, we only need to verify the validity of $V=\mathbb{F}^n$. We let

  $$S_i(x_1,\ldots,x_n)=x_i\boldsymbol{e_i}$$

  □

3.2 Null Space and Ranges

3.2.1 Null Space and Injectivity

Definition 3.2.1: Null Space

Let $T\in\mathscr{L}(V,W)$, then the null space of T is defined as the subset of V that maps to 0.

$$T=\{v\in V:Tv=0\} \tag{3.5}$$

Theorem 3.2.1: Null Space is a Subspace

Suppose $T\in\mathscr{L}(V,W)$, then $\operatorname{null}T$ is a subspace of V.

Proof.

- Obviously, $0 \in \operatorname{null} T$.
- If $v_1, v_2 \in \operatorname{null} T$, then $T(v_1 + v_2) = Tv_1 + Tv_2 = 0$ as well.
- If $v \in \operatorname{null} T$, then $T(\lambda v) = \lambda T v = 0$.

□

Remark:
In fact, the linear map is just a homomorphism. The null space is the kernel, the ideal of the homomorphism. To understand it this way, the following statement would seem obvious.

Theorem 3.2.2: Null Space and Injectivity

Let $T \in \mathscr{L}(V, W)$, then $\operatorname{null} T = \{0\} \leftrightarrow T$ is injective .

Proof. If $\operatorname{null} T = \{0\}$, then if $Tu = Tv$, we have $T(u - v) = 0$, then $u - v \in \operatorname{null} T$, then $u = v$, thus T is injective. □

3.2.2 Range and Surjectivity

Definition 3.2.2: Range

For $T \in \mathscr{L}(V, W)$, the range of T is the subset of W defined as

$$\operatorname{range} T = \{Tu : u \in V\} \tag{3.6}$$

Theorem 3.2.3: Range is a subspace

If $T \in \mathscr{L}(V, W)$ then $\operatorname{range} T$ is a subspace of W.

Proof.

- $0 \in \operatorname{range} T$.
- addition and scalar multiplication is also closed.

□

Definition 3.2.3: Surjective

A function $T : V \to W$ is surjective if $\operatorname{range} T = W$.

3.2.3 Fundamental Theorem of Linear Maps

Theorem 3.2.4: Fundamental Theorem of Linear Maps

Suppose V is finite-dimensional and $T \in \mathscr{L}(V, W)$, then $\operatorname{range} T$ is finite dimensional and

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T \tag{3.7}$$

Proof. Let $u_1, \ldots, u_m$ be a basis of $\operatorname{null} T$, then $\dim \operatorname{null} T = m$. Expand it to a basis of V

$$u_1, \ldots, u_m, v_1, \ldots, v_n$$

We shall prove that

$$Tv_1, \dots, Tv_n$$

is a basis of range T.

$\forall v \in V$, let $v = a_1u_1 + \dots + a_mu_m + b_1v_1 + \dots + b_nv_n$. Then we have

$$Tv = b_1Tv_1 + \dots + b_nTv_n$$

That is, $Tv_1, \dots, Tv_n$ spans range T.

To show that $Tv_1, \dots, Tv_n$ is linearly independent, if

$$c_1Tv_1 + \dots + c_nTv_n = 0$$

then

$$T(c_1v_1 + \dots + c_nv_n) = 0$$

that is,

$$c_1v_1 + \dots + c_nv_n \in \operatorname{null} T$$

let

$$c_1v_1 + \dots + c_nv_n = d_1u_1 + \dots + d_mu_m$$

then $c_1 = \dots = c_n = 0$ as desired. □

Remark:
This proof is quite elegant, and it also gives us a picture of linear maps. We can see linear maps from a higher dimensional vector space to a lower one (note that the range dimension is always $\leq$ the domain) as losing some dimension of freedom. The dimensions that are lost forms the null space.

Theorem 3.2.5: Linear maps to a lower dimensional space

Suppose V and W are finite-dimensional. $\dim V > \dim W$. Then there are no injective linear map from V to W.

Proof. $\dim \operatorname{null} T > 0$ would suffice. □

Theorem 3.2.6: Linear maps to a higher dimensional space

Suppose V and W are finite-dimensional. $\dim V < \dim W$. Then there are no surjective linear map from V to W.

Proof.

$$\dim \operatorname{range} T \leq \dim V < \dim W$$

□

Linear maps have important consequences for linear equations. We can express the theory of linear equations in terms of linear maps. Consider the homogeneous system of linear equations

$$\begin{cases} \sum_{k=1}^{n} A_{1k} x_k = 0 \\ \vdots \\ \sum_{k=1}^{n} A_{mk} x_k = 0 \end{cases} \tag{3.8}$$

Clearly $x_1 = \ldots = x_n = 0$ is a trivial solution to the equation. The question is whether other solution exists.

Define a linear map $T : \mathbb{F}^n \to \mathbb{F}^m$ by

$$T(x_1, \ldots, x_n) = \left(\sum_{k=1}^{n} A_{1k} x_k, \ldots, \sum_{k=1}^{n} A_{mk} x_k \right) \tag{3.9}$$

Then the equation becomes $T(x_1, \ldots, x_n) = 0$. That is, we want to find $\operatorname{null} T$.

Using the result of theorem 3.2.5, we have:

Theorem 3.2.7: Homogeneous system of linear equations

A homogeneous system of linear equations with more variables the equations has nonzero solutions.

For an inhomogeneous system of linear equations,

$$\begin{cases} \sum_{k=1}^{n} A_{1k} x_k = c_1 \\ \vdots \\ \sum_{k=1}^{n} A_{mk} x_k = c_m \end{cases} \tag{3.10}$$

This is analogous to $T(x_1, \ldots, x_n) = (c_1, \ldots, c_m)$. Using theorem 3.2.6, we have

Theorem 3.2.8: Inhomogeneous system of linear equations

An inhomogeneous system of linear equations with more equations than variables has no solutions for some choice of constant terms.

It is obvious that the equation has solution iff $(c_1, \ldots, c_m) \in \operatorname{range} T$.

3.2.4 Some Result

- Suppose V is finite dimensional and $T \in \mathscr{L}(V, W)$, then $\exists$ a subspace U such that

$$U \cap \operatorname{null} T = \{0\} \text{ and } \operatorname{range} T = \{Tu : u \in U\}$$

- Suppose there is a linear map on V whose null space and range in finite dimensional, then V is finite dimensional.

Proof. Suppose $u_1, \ldots, u_n$ spans $\operatorname{null} T$ and $w_1, \ldots, w_m$ spans $\operatorname{range} T$. Let $w_k = Tv_k$ for $k = 1, \ldots, m$. Then $u_1, \ldots, u_n, v_1, \ldots, v_m$ spans V. □

- Suppose V is finite dimensional and $T \in \mathscr{L}(V, W)$, U is a subspace of W. Then $\{v \in V : Tv \in U\}$ is a subspace of V, and

$$\dim \{v \in V : Tv \in U\} = \dim \operatorname{null} T + \dim (U \cap \operatorname{range} T) \tag{3.11}$$

Proof. In fact, we shall interpret T as $T : \{v \in V : Tv \in U\} \to W$ would suffice. □

- Suppose U and V are finite dimensional and $S \in \mathscr{L}(V, W)$ and $T \in \mathscr{L}(U, V)$, then

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T \tag{3.12}$$

Proof.

$$\begin{aligned} \dim \operatorname{null} ST &= \dim \{u \in U : Tu \in \operatorname{null} S\} \\ &= \dim \operatorname{null} T + \dim(\operatorname{null} S \cap \operatorname{range} T) \\ &\leq \dim \operatorname{null} S + \dim \operatorname{null} T \end{aligned}$$

□

- Suppose U and V are finite-dimensional, and $S \in \mathscr{L}(V, W), T \in \mathscr{L}(U, V)$, then

$$\dim \operatorname{range} ST \leq \min \{\dim \operatorname{range} S, \dim \operatorname{range} T\} \tag{3.13}$$

Proof. Obvious. □

- Suppose V, W are real vector spaces and $T \in \mathscr{L}(V, W)$, define $T_{\mathbb{C}} : V_{\mathbb{C}} \to W_{\mathbb{C}}$ such that $\forall u, v \in V$

$$T_{\mathbb{C}}(u + iv) = Tu + iTv \tag{3.14}$$

then

1. $T_{\mathbb{C}}$ is a complex linear map from $V_{\mathbb{C}}$ to $W_{\mathbb{C}}$.
2. $T_{\mathbb{C}}$ is injective iff T is injective.
3. $\operatorname{range} T_{\mathbb{C}} = W_{\mathbb{C}}$ iff $\operatorname{range} T = W$.

3.3 Matrices

3.3.1 Representing a Linear Map by a Matrix

Matrices is a good way to visualize the structure of linear maps as is shown in theorem 3.1.4.

Definition 3.3.1: Matrices A_{jk}

Suppose $m, n \in \mathbb{Z}_{\geq 0}$. An m-by-n matrix is a rectangular array of elements is $\mathbb{F}$ with m rows and n columns.

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}$$

Now we define matrix of linear map.

Definition 3.3.2: Matrix of a linear map

Suppose $T \in \mathscr{L}(V, W)$, and $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ is a basis of W. Then the matrix of T respect of these basis is $\mathscr{M}(T)$

$$\mathscr{M}(T) = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix} \tag{3.15}$$

where $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}w_1 + \ldots + A_{m,k}w_m \tag{3.16}$$

Remark:
Note that the matrix $\mathscr{M}(T)$ depends on the basis which we choose.
If T is a linear map $T : \mathbb{F}^n \to \mathbb{F}^m$, then usually we take the standard basis.

3.3.2 Addition and Scalar Multiplication of Matrices

We define the sum of two matrix of the same size as follows

Definition 3.3.3: Matrix Addition

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdots & C_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{m,1} & C_{m,2} & \cdots & C_{m,n} \end{pmatrix} = \begin{pmatrix} A_{1,1}+C_{1,1} & A_{1,2}+C_{1,2} & \cdots & A_{1,n}+C_{1,n} \\ A_{2,1}+C_{2,1} & A_{2,2}+C_{2,2} & \cdots & A_{2,n}+C_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1}+C_{m,1} & A_{m,2}+C_{m,2} & \cdots & A_{m,n}+C_{m,n} \end{pmatrix} \tag{3.17}$$

This is because

Theorem 3.3.1: Matrix of the sum of linear maps

Suppose $S, T \in \mathscr{L}(V, W)$, then $\mathscr{M}(S + T) = \mathscr{M}(S) + \mathscr{M}(T)$.

Definition 3.3.4: Scalar Multiplication of Linear Maps

We define

$$\lambda \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \lambda A_{1,2} & \cdots & \lambda A_{1,n} \\ \lambda A_{2,1} & \lambda A_{2,2} & \cdots & \lambda A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda A_{m,1} & \lambda A_{m,2} & \cdots & \lambda A_{m,n} \end{pmatrix} \tag{3.18}$$

Theorem 3.3.2: Matrix of the scalar product of linear maps

Suppose $T \in \mathscr{L}(V, W)$, then $\mathscr{M}(\lambda T) = \lambda \mathscr{M}(T)$.

As addition and scalar multiplication is defined, then not surprisingly, matrices is a vector space.

Notation 3.3.1: $\mathbb{F}^{m,n}$

For positive integers m, n, the set of all m-by-n matrices with entries in $\mathbb{F}$ in denoted by $\mathbb{F}^{m,n}$. It is a vector space.

Theorem 3.3.3: Dimension of $\mathbb{F}^{m,n}$

$$\dim \mathbb{F}^{m,n} = mn \tag{3.19}$$

Remark:
as any linear map can be represented as a matrix, we have $\mathscr{L}(V, W) \cong \mathbb{F}^{m,n}$ if $\dim V = n$ and $\dim W = m$.

3.3.3 Matrix Multiplication

Consider finite-dimensional vector spaces U, V, W. Let basis

- $u_1, \ldots, u_p$ for U.
- $v_1, \ldots, v_n$ for V.
- $w_1, \ldots, w_m$ for W.

Consider linear maps $T : U \to V$ and $S : V \to W$. We want to define a matrix multiplication such that

$$\mathscr{M}(ST) = \mathscr{M}(S)\mathscr{M}(T)$$

Suppose $\mathscr{M}(S) = A$ and $\mathscr{M}(T) = B$. For $1 \le k \le p$, we have

$$\begin{aligned}
(ST)u_k &= S\left(\sum_{r=1}^{k} B_{r,k} v_r\right) \\
&= \sum_{r=1}^{n} B_{r.k} S v_r \\
&= \sum_{r=1}^{n} B_{r,k} \sum_{j=1}^{m} A_{j,r} w_j \\
&= \sum_{j=1}^{m} \left(\sum_{r=1}^{n} A_{j,r} B_{r,k}\right) w_j
\end{aligned}$$

Therefore, we have

Definition 3.3.5: Matrix Multiplication

Suppose $A \in \mathbb{F}^{m,n}$ and $B \in \mathbb{F}^{n,p}$. Then define $AB \in \mathbb{F}^{m,p}$ such that

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k}. \tag{3.20}$$

Remark:
Matrix multiplication is not commutative.

Theorem 3.3.4: Matrix of product of linear map

If $T \in \mathscr{L}(U, V)$ and $S \in \mathscr{L}(V, W)$, then $\mathscr{M}(S, T) = \mathscr{M}(S)\mathscr{M}(T)$

Notation 3.3.2: Row and Column vectors

Suppose $A \in \mathbb{F}^{m,n}$.

- If $1 \leq j \leq m$ then $A_{j,.}$ denotes the j^{th} row of A.

$$A_{j,.} = \begin{pmatrix} A_{j,1} & A_{j,2} & \cdots & A_{j,n} \end{pmatrix}$$

- If $1 \leq k \leq n$ then $A_{.,k}$ denotes the k^{th} column of A.

$$A_{.,k} = \begin{pmatrix} A_{1,k} \\ A_{2,k} \\ \vdots \\ A_{m,k} \end{pmatrix}$$

Theorem 3.3.5: Entry equals row times column

Suppose $A \in \mathbb{F}^{m,n}$, $B \in \mathbb{F}^{n,p}$. Then

$$(AB)_{j,k} = A_{j,.} \cdot B_{.,k}$$

Example:

We have:

- $(AB)_{.,k} = AB_{.,k}$, that is

$$A \begin{pmatrix} B_{.,1} & B_{.,2} & \cdots & B_{.,p} \end{pmatrix} = \begin{pmatrix} AB_{.,1} & AB_{.,2} & \cdots & AB_{.,p} \end{pmatrix}$$

- $(AB)_{j,.} = A_{j,.}B$, that is

$$\begin{pmatrix} A_{1,.} \\ A_{2,.} \\ \vdots \\ A_{m,.} \end{pmatrix} B = \begin{pmatrix} A_{1,.}B \\ A_{2,.}B \\ \vdots \\ A_{m,.}B \end{pmatrix}$$

Theorem 3.3.6: Linear Combination

Suppose $A \in \mathbb{F}^{m,n}$ and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, then

$$Ab = b_1 A_{.,1} + \ldots + b_n A_{.,n}$$

Similarly, if $c = \begin{pmatrix} c_1 & \cdots & c_m \end{pmatrix}$, then

$$cA = c_1 A_{1,.} + \ldots + c_m A_{m,.}$$

We expand the full content of the theorem as follows:

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = b_1 \begin{pmatrix} A_{1,1} \\ A_{2,1} \\ \vdots \\ A_{m,1} \end{pmatrix} + \ldots + b_n \begin{pmatrix} A_{1,n} \\ A_{2,n} \\ \vdots \\ A_{m,n} \end{pmatrix} \tag{3.21}$$

And that

$$\begin{pmatrix} c_1 & c_2 & \cdots & c_m \end{pmatrix} \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix} = c_1 \begin{pmatrix} A_{1,1} & \cdots & A_{a,n} \end{pmatrix} + \ldots + c_m \begin{pmatrix} A_{m,1} & \cdots & A_{m,n} \end{pmatrix}$$

3.3.4 Column-Row Factorization and Rank of a Matrix

Definition 3.3.6: Column and Row Rank

Suppose $A \in \mathbb{F}^{m,n}$

- The column rank of A is the dimension of the span of row of A in $\mathbb{F}^{1,n}$.
- The row rank of A is the dimension of the span of columns of A in $\mathbb{F}^{m,1}$.

We now define the transpose of a matrix.

Definition 3.3.7: Transpose

Suppose $A \in \mathbb{F}^{m,n}$, define the transpose of A be $A^t \in \mathbb{F}^{n,m}$.

$$(A^t)_{k.j} = A_{j,k}$$

That is if

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix}$$

then

$$A^t = \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{m,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1,n} & A_{2,n} & \cdots & A_{n,m} \end{pmatrix}$$

Remark:
In fact, transposition is a linear map.

Theorem 3.3.7: Algebraic Properties of Transposition

$(A+B)^t = A^t + B^t$.

$(\lambda A)^t = \lambda A^t$.

$(AB)^t = B^t A^t$.

Theorem 3.3.8: Column-Row Factorization

Suppose $A \in \mathbb{F}^{m,n}$, and column rank $c > 1$, then there exists $C \in \mathbb{F}^{m,c}$ and $R \in \mathbb{F}^{c,n}$, such that $A = CR$.

Proof. As the definition of column rank, the list $A_{.,1}, \ldots, A_{.,n}$ can be reduced to a basis of its span, which has length c. Putting together, we denote it $C \in \mathbb{F}^{m,c}$.

Let $A_{.,k} = CR_{.,k}$, we get a $R \in \mathbb{F}^{c,n}$. □

Theorem 3.3.9: Column Rank equals Row Rank

Suppose $A \in \mathbb{F}^{m,n}$, then the column rank of A equals the row rank.

Proof. Let c be the column rank. Let $A = CR$ be the column-row factorization. Then we have

$$A_{j,.} = C_{j,.}R$$

therefore, row rank $\leq c$.

However, taking A^t we have

$$\text{column rank of } A = \text{ row rank of } A^t \leq \text{ column rank of } A^t = \text{ row rank of } A$$

that is, column rank = row rank. □

Definition 3.3.8: Rank

The rank of $A \in F^{m,n}$ is the column/row rank of A, denoted $\operatorname{rank} A$

3.4 Invertibility and Isomorphisms

3.4.1 Invertible Linear Maps

Definition 3.4.1: Invertible and Inverse

- A linear map $T \in \mathscr{L}(V, W)$ is invertible if there exists a linear map $S \in \mathscr{L}(W, V)$ such that $ST = I_V$ and $TS = I_W$.
- This S is called the inverse of T, denoted T^{-1}.

Theorem 3.4.1: Inverse is Unique

An invertible linear map has a unique inverse.

Proof. Suppose S_1, S_2 are inverse of T.

$$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2$$

□

Theorem 3.4.2: Invertibility Conditions

A linear map in invertible iff it is injective and surjective.
Futhermore, if V, W are finite dimensional and $\dim V = \dim W$, then injectivity $\leftrightarrow$ surjectivity.

Proof. Because from the fundamental theorem of linear map

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$$

If T is injective, then $\dim \operatorname{null} T = 0$, then $\dim \operatorname{range} T = \dim V = \dim W$, then T is surjective. □

Theorem 3.4.3: $ST = I \leftrightarrow TS = I$

Suppose $\dim V = \dim W < \infty$, $S \in \mathscr{L}(V, W), T \in \mathscr{L}(W, V)$, then

$$ST = I \leftrightarrow TS = I$$

Proof. If $ST = I$, then if $v \in V$ and $Tv = 0$, then

$$v = Iv = STv = S0 = 0$$

Thus T is injective. Then by 3.4.2, T is invertible, we have

$$S = STT^{-1} = T^{-1}$$

thus $TS = TT^{-1} = I$ as desired. □

3.4.2 Isomorphic Vector Spaces

Definition 3.4.2: Isomorphism

An **isomorphism** is an invertible linear map. Two vector spaces are called isomorphic if there there is an isomorphism from one to the other.

Remark:
This definition of isomorphism is a simplified version. The intuition of isomorphism is that two algebraic structures are indistinguishable. That is, for vector spaces V and W. They are isomorphic iff there exists a bijection $T : V \to W$ such that:

- $T(v_1 + v_2) = Tv_1 + Tv_2$.
- $T(\lambda v) = \lambda Tv$.

This is exactly the definition of linear maps. Furthermore, every linear map is a homomorphism.

Theorem 3.4.4: Dimension and Isomorphism

Two finite dimensional vector spaces over $\mathbb{F}$ are isomorphic iff they have the same dimension.

Proof. This is extremely obvious. If $v_1, \ldots, v_n$ is a basis of V, then $Tv_1, \ldots, Tv_n$ is a basis of W. $\square$

Therefore we have the earlier statement

- $\dim V = n$ then $V \cong \mathbb{F}^n$.
- $\dim V = n$ and $\dim W = m$ then $\mathscr{L}(V, W) \cong \mathbb{F}^{m,n}$

3.4.3 Linear Map Thought of as Matrix Multiplication

Definition 3.4.3: Matrix of Vectors

If V is a finite dimensional vector space, and $v_1, \ldots, v_n$ is a basis. Let $v = b_1v_1 + \ldots + b_nv_n$, then the matrix of v respect to these basis is

$$\mathscr{M}(v) = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{3.22}$$

Note that the matrix of v depends on the basis we choose.

Remark:
We can understand the matrix $\mathscr{M}$ as an isomorphism from V to $\mathbb{F}^{n,1}$.

If we define this way, linear maps are like matrix multiplication. For a linear map $T : V \to W$, we first choose the basis $v_1, \ldots, v_n$ for V and $w_1, \ldots, w_m$ for W. If we have

$$\mathscr{M}(T) = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix}, \quad \mathscr{M}(v) = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Then we have $\mathscr{M}(Tv) = \mathscr{M}(T)\mathscr{M}(v)$ as expected.

Remark:
Each linear map indicates a linear map from $\mathbb{F}^{n,1}$ to $\mathbb{F}^{m,1}$. As matrix representations depends on the choice of basis, one objection of the later chapters is how to find a basis that makes the matrix as simply as possible.

No bases are in sight in the statement of the next result. Although $\mathscr{M}(T)$ in the next result depends on a choice of bases of V and W, the next result shows that the column rank of $\mathscr{M}(T)$ is the same for all such choices (because range T does not depend on a choice of basis).

Theorem 3.4.5: Dimension and Column Rank

For $T \in \mathscr{L}(V, W)$, dimension of range $T = \operatorname{rank} \mathscr{M}(T)$

3.4.4 Change of Basis

In the part we denote

$$\mathscr{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m)) \tag{3.23}$$

To the matrix representation of T under basis $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$. We also simplify notation if $T \in \mathscr{L}(V)$ and we use the same basis for domain and range.

$$\mathscr{M}(T, (v_1, \ldots, v_n)) = \mathscr{M}(T, (v_1, \ldots, v_n), (v_1, \ldots, v_n))$$

Definition 3.4.4: Identity Matrix

We denote identity matrix I for the identity map which use the same basis. That is, $\mathscr{M}(I) = I$

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{3.24}$$

Definition 3.4.5: Inverse of Matrix

If square matrix A, B satisfies $AB = BA = I$ then B is the inverse of A, denote

$$B = A^{-1}$$

Following are some algebraic properties of inverse.

- $(A^{-1})^{-1} = A$.
- $(AB)^{-1} = B^{-1}A^{-1}$.

Theorem 3.4.6: Matrix of Product of linear maps

Suppose $T \in \mathscr{L}(U, V)$ and $S \in \mathscr{L}(V, W)$, $u_1, \ldots, u_m$ is a basis of U, $v_1, \ldots, v_m$ is a basis of V, $w_1, \ldots, w_p$ is a basis of W.

$$\mathscr{M}(ST, u_i, w_i) = \mathscr{M}(S, v_i, w_i)\mathscr{M}(T, u_i, v_i)$$

Theorem 3.4.7: matrix of identity operator

Suppose $v_1, \ldots, v_n$ and $u_1, \ldots, u_n$ are basis of V, then

$$\mathscr{M}(I, u_i, v_i) \quad \text{and} \quad \mathscr{M}(I, v_i, u_i) \tag{3.25}$$

are inverses to each other.

The next result shows how to change basis.

Theorem 3.4.8: Change Basis Formula

Suppose $T \in \mathscr{L}(V)$, and $v_1, \ldots, v_n$ and $u_1, \ldots, u_n$ are basis of V. Let

- $A = \mathscr{M}(T, u_i)$.
- $B = \mathscr{M}(T, v_i)$.
- $C = \mathscr{M}(T, u_i, v_i)$

Then we have

$$A = C^{-1}BC \tag{3.26}$$

3.5 Products and Quotients of Vector Spaces

3.5.1 Products of Vector Spaces

Definition 3.5.1: Products of vector space

Let $V_1, \ldots, V_m$ be vector spaces over $\mathbb{F}$. Then define the product $V_1 \times \ldots \times V_m$ be:

$$V_1 \times \ldots \times V_m = \{(v_1, \ldots, v_m) : v_i \in V_i\} \tag{3.27}$$

with

- **Addition:** $(v_1, \ldots, v_m) + (u_1, \ldots, u_m) = (v_1 + u_1, \ldots, v_m + u_m)$
- **Scalar Multiplication:** $\lambda(v_1, \ldots, v_m) = (\lambda v_1, \ldots, \lambda v_m)$

It is obvious that the product of vector spaces is a vector space

Theorem 3.5.1: Dimension of Products of Vector Spaces

$$\dim(V_1 \times \ldots \times V_m) = \dim V_1 + \ldots + \dim V_m \tag{3.28}$$

Proof. For each V_i there is a basis B_i. Then the i^{Th} slot is filled with some vector in B_i and others are 0 forms a basis of $V_1 \times \ldots \times V_k$. □

Theorem 3.5.2: Products and sums

Suppose that $V_1, \ldots, V_m$ are subspaces of V. Define a linear map $\Gamma : V_1 \times \ldots \times V_m \to V_1 + \ldots + V_m$ as:

$$\Gamma(v_1, \ldots, v_m) = v_1 + \ldots + v_m$$

Then $V_1 + \ldots + V_m$ is a direct sum iff Γ is injective.

Proof. Γ is injective iff $\operatorname{null}\Gamma = \{(0, \ldots, 0)\}$, that is equivalent to

$$v_1 + \ldots + v_m = 0 \to v_1 = \ldots = v_m = 0$$

□

Corollary 3.5.1: a sum is a direct sum if and only if dimensions add up

Suppose $V_1, \ldots, V_m$ is subspaces of V. Then $V_1 + \ldots + V_m$ is direct sum iff

$$\dim(V_1 + \ldots + V_m) = \dim V_1 + \ldots + \dim V_m$$

Proof. Note that Γ is already surjective.

If $V_1 + \ldots + V_m$ is a direct sum, then Γ is injective, that is, $\dim(V_1 + \ldots + V_m) = \dim(V_1 \times \ldots \times V_m)$.

The other way, if the dimension adds, then by the fundamental theorem of linear maps, $\operatorname{null}\Gamma = 0$, that is, Γ is injective. □

3.5.2 Quotient Spaces

Subspaces are kernels of sum homomorphisms. To defined Quotient Spaces, we begin with translates.

Definition 3.5.2: Translate

For $v \in V$ and U a subset of V. Let

$$v + U = \{v + u : u \in U\} \tag{3.29}$$

be a translate of U.

Theorem 3.5.3: two translates of a subspace are equal or disjoint

Suppose U is a subspace of V and $v, w \in V$. Then

$$v + U = w + U \vee v + U \cap w + U = \emptyset$$

In fact, $v + U = w + U$ iff $v - w \in U$.

This is just the same as groups. As the addition in vector spaces is an Abelian group, we are not surprised to see that every subspace is a normal subgroup.

Definition 3.5.3: Quotient Spaces

If U is a subspace of V. Defined quotient space

$$V/U = \{v + U : v \in V\} \tag{3.30}$$

We define Addition and Scalar multiplication as follows.

- $(v_1 + U) + (v_2 + U) = (v_1 + v_2) + U$
- $\lambda(v + U) = (\lambda v) + U$

Then quotient space becomes a vector space. And we have

$$V \cong U \times V/U$$

The homomorphism is called the quotient map

Definition 3.5.4: Quotient map

The quotient map $\pi : V \to V/U$ is defined

$$\pi(v) = v + U$$

Theorem 3.5.4: Dimension of quotient space

If U is a subspace of V, then

$$\dim V/U = \dim V - \dim U$$

We clearly have the isomorphism theorems.

Theorem 3.5.5: The First Isomorphism Theorem

Suppose $T \in \mathscr{L}(V, W)$, then

$$V/\operatorname{null} T \cong \operatorname{range} T \tag{3.31}$$

3.6 Duality

3.6.1 Dual Space and Dual Map

Definition 3.6.1: Linear Functional

A linear functional is an element of $\mathscr{L}(V, \mathbb{F})$

Definition 3.6.2: Dual Space

For a vector space V, its dual space V' is defined as

$$V' = \mathscr{L}(V, \mathbb{F})$$

Therefore, we have

Theorem 3.6.1: Dimension of Dual Space

$$\dim V' = \dim V \tag{3.32}$$

Next we consider the basis of dual space. Given a basis of V. The matrix representation of a dual space is

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

Like the row vector while elements in V is column vector. In this way, its standard basis is very clear.

Definition 3.6.3: Dual Basis

If $v_1, \ldots, v_n$ is a basis of V. The standard basis of dual space is

$$\varphi_j(v_k) = \delta_{jk} = \begin{cases} 1, & \text{if } k = j \\ 0, & \text{if } k \neq j \end{cases} \tag{3.33}$$

Remark:
The standard basis of dual space acts as extraction of coefficients of linear combination. For $v \in V$, we have

$$v = \sum_{i=1}^{n} \varphi_i(v) v_i$$

We do the same for linear maps.

Definition 3.6.4: Dual map

Suppose $T \in \mathscr{L}(V, W)$, The dual map of T is the linear map $T' \in \mathscr{L}(W', V')$ defined for each $\varphi \in W'$ by

$$T'(\varphi) = \varphi \circ T \tag{3.34}$$

Example: **Dual Map of Differentiation linear map**

Define $D : \mathscr{P}(\mathbb{R}) \to \mathscr{P}(\mathbb{R})$ by $Dp = p'$. Suppose $\varphi(p) = \int_0^1 p(x)\mathrm{d}x$, then

$$D'(\varphi)(p) = (\varphi \circ D)(p) = \varphi(p') = p(1) - p(0)$$

Theorem 3.6.2: Algebraic Properties of Dual Maps

Suppose $S, T \in \mathscr{L}(V, W)$ then

- $(S+T)' = S' + T'$
- $(\lambda T)' = \lambda T'$
- $(ST)' = T'S'$

Proof. the last one

$$(ST)'(\varphi) = \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = (T'S')(\varphi)$$

□

3.6.2 Null Space and Range of Dual Linear Maps

Our goal is to describe $\operatorname{null} T'$ and $\operatorname{range} T'$ in terms of $\operatorname{null} T$ and $\operatorname{range} T$.

Definition 3.6.5: Annihilator

For $U \subseteq V$, the annihilator of U, denoted U^0, is defined by

$$U^0 = \{\varphi \in V' : \forall u \in U, \varphi(u) = 0\} \tag{3.35}$$

Remark:
We can see annihilators as some generalization of orthogonality without specifying an inner product on the vector space. $\varphi(u)$ corresponds to $\varphi \cdot u$.

Theorem 3.6.3: Annihilator is a subspace

Suppose $U \subseteq V$, then U^0 is a subspace of V.

Proof.

- First $0 \in U^0$.
- The addition and multiplication is easy to verify.

□

Theorem 3.6.4: The dimension of the annihilator

Suppose V is a finite dimensional vector space and U is a subspace of V, then

$$\dim U^0 = \dim V - \dim U$$

Proof. Let $i \in \mathscr{L}(U, V)$ be the inclusion map defined $\forall u \in U, i(u) = u$. Thus $i' \in \mathscr{L}(V', U')$ is a linear map and $\operatorname{null} i' = U^0$. Because $\operatorname{null} i' = \{\varphi \in V' : \varphi \circ i = 0\}$, that is, $\varphi(u) = 0$

We have

$$\dim \operatorname{range} i' + \dim \operatorname{null} i' = \dim V'$$

then

$$\dim U + \dim U^0 = \dim V$$

□

Corollary 3.6.1

Suppose V is a finite dimensional vector space, U is a subspace of V, then

- $U^0 = \{0\} \leftrightarrow U = V$
- $U^0 = V' \leftrightarrow U = \{0\}$

Theorem 3.6.5: The Null Space of T'

Suppose V, W are finite dimensional and $T \in \mathscr{L}(V, W)$, then

- $\operatorname{null} T' = (\operatorname{range} T)^0$ (this is also right when V, W are not finite dimensional)
- $\dim \operatorname{range} T' = \dim \operatorname{range} T$
- $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$

Proof.

1. First suppose $\varphi \in \operatorname{null} T'$, then $\varphi \circ T = 0$. Hence

$$0 = (\varphi \circ T)(v) = \varphi(Tv) \forall v \in V$$

thus $\varphi \in (\operatorname{range} T)^0$. The other way is the same.

(An intuitive understanding is that $A^t v = 0 \Leftrightarrow v^t$ is orthogonal to range A.)

2. We have

$$\dim \operatorname{range} T' = \dim \operatorname{range} T$$

From above

□

Remark:

We can see T and T' as the same matrix

$$T = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{pmatrix}$$

But T takes column vectors from the right, while T' takes row vectors from the left. The null space of T' and the annihilator of range T both is the row vectors that makes T zero. The second and third results are merely row rank = column rank.

Theorem 3.6.6: T is surjective iff T' is injective

Suppose V, W are finite dimensional vector spaces, $T \in \mathscr{L}(V, W)$, then

$$T \text{ is surjective } \leftrightarrow T' \text{ is injective}$$

Proof. T is surjective iff $\operatorname{range} T = W$ iff $(\operatorname{range} T)^0 = \operatorname{null} T' = \{0\}$ iff T' is injective. □

Theorem 3.6.7: Range of T'

Suppose V, W are finite dimensional vector spaces. $T \in \mathscr{L}(V, W)$, then

- $\dim \operatorname{range} T' = \dim \operatorname{range} T$.
- $\operatorname{range} T' = (\operatorname{null} T)^0$.

Proof. The second one, if $\varphi \in \operatorname{range} T$, then $\exists \psi \in W'$ such that $\varphi = T'(\psi)$, If $v \in \operatorname{null} T$, then

$$\varphi(v) = (T'(\psi))(v) = (\psi \circ T)(v) = \psi(Tv) = \psi(0) = 0$$

Hence $\varphi \in (\operatorname{null} T)^0$. And we have

$$\dim \operatorname{range} T' = \dim(\operatorname{null} T)^0$$

□

Remark:
This is easily understand because dual spaces are relative.

Theorem 3.6.8: T is injective $=$ T' is surjective

Suppose V, W are finite dimensional vector spaces, $T \in \mathscr{L}(V, W)$, then

$$T \text{ is injective } \leftrightarrow T' \text{ is surjective}$$

3.6.3 Matrix of Dual of Linear Maps

We've used many intuition from matrices so far. Here is a summary.

Theorem 3.6.9: Matrix of T'

Suppose V, W are finite dimensional vector spaces, $T \in \mathscr{L}(V, W)$, then

$$\mathscr{M}(T') = (\mathscr{M}(T))^t$$

Chapter 4

Polynomials

Standing Assumptions:

$\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.

4.1 Zeros of Polynomials

Theorem 4.1.1: zeros and degree-one factor

If $m \in \mathbb{Z}_{\geq 0}$ and $p \in \mathscr{P}(\mathbb{F})$ with $\deg p = m$, we have

Chapter 5

Eigenvalues and Eigenvectors

Standing Assumptions:

- $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.
- V denotes a vector space of $\mathbb{F}$.

We begin our investigation from operators, that is, linear maps from V to itself.

5.1 Invariant Subspaces

5.1.1 Eigenvalues

Definition 5.1.1: operator

A linear map from a vector space to itself is called an operator.

Notation 5.1.1: Restriction of maps

If $f : V \to W$ and $U \subseteq V$. Then $f|_U$ is the function $U \to W$ with

$$\forall x \in U, f|_U(x) = f(x).$$

Suppose $T \in \mathscr{L}(V)$, if

$$V = V_1 \oplus \ldots \oplus V_m$$

To analyze the properties of T, we only need to analyze $T|_{V_k}$ for each k.

However, the restriction of T to V_k may not be an operator anymore. That is, $T(V_k)$ may map outside V_k.

Definition 5.1.2: Invariant Subspace

Suppose $T \in \mathscr{L}(V)$, then a subspace U of V is called invariant under T if $\forall u \in U, Tu \in U$.

Thus, if U is an invariant subspace then $T|_U$ is an operator on U.

Example: **Invariant Subspaces**

If $T \in \mathscr{L}(V)$, then the following are invariant subspace of V.

- $\{0\}$
- V
- $\operatorname{null} T$, because if $u \in \operatorname{null} T$, $Tu = 0 \in \operatorname{null} T$.
- $\operatorname{range} T$, because $Tu \in \operatorname{range} T$.

Proposition: **Intersection of Invariant Subspaces**

If U, W are invariant subspaces of V, then $U \cap W$ is an invariant subspace of V.

We begin with invariant subspaces of dimension one. Take any $v \in V$ and let

$$U = \{\lambda v : \lambda \in \mathbb{F}\} = \operatorname{span}(v)$$

Every one dimension subspace has this form. If U is invariant under $T \in \mathscr{L}(V)$, then $Tv \in U$, that is,

$$\exists \lambda \in \mathbb{F}, Tv = \lambda v.$$

Conversely, if $\exists \lambda \in \mathbb{F}, Tv = \lambda v$, then $\forall u \in U, Tu = \lambda u$. (Set $u = \mu v$ would do).
Therefore, we have

$$\operatorname{span}(v) \text{ is a one dimension subspace } \leftrightarrow \exists \lambda \in \mathbb{F}, Tv = \lambda v$$

Definition 5.1.3: Eigenvalue

Suppose $T \in \mathscr{L}(V)$, a number $\lambda \in \mathbb{F}$ is called an *eigenvalue* of T if there exists $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$.

Theorem 5.1.1: Conditions for an eigenvalue

Suppose V is finite dimensional. $T \in \mathscr{L}(V)$ and $\lambda \in \mathbb{F}$. Then the following are equivalent.

1. λ is an eigenvalue of T.
2. $T - \lambda I$ is not invertible.
3. $T - \lambda I$ is not injective (surjective).

Proof. We've already seen that the last two conditions are equivalent for an operator (has the same dimension).

If λ is an eigenvalue of T, then $Tv = \lambda v$ for some $v \in V, v \neq 0$, that is, $(T - \lambda I)v = 0$, which is equivalent to T being not invertible. $\square$

Definition 5.1.4: Eigenvectors

Suppose $T \in \mathscr{L}(V)$ and λ is an eigenvalue of T, A vector is called an eigenvector of T corresponding to λ if $v \neq 0$ and $Tv = \lambda v$.

We frequently use the correspondence

$$Tv = \lambda v \leftrightarrow (T - \lambda I)v = 0$$

Theorem 5.1.2: Linearly Independent Eigenvectors

Suppose $T \in \mathscr{L}(V)$, then every list of eigenvectors corresponding to distinct eigenvalues are linear independent.

Proof. We can do this by induction, but we can also do by the least number property.

Suppose the result is false, then $\exists$ the smallest positive integer m such that a linearly dependent list $v_1, \ldots, v_m$ corresponds to different eigenvalues $\lambda_1, \ldots, \lambda_m$. We have $a_1, \ldots, a_m \in \mathbb{F}$, none of which are 0 because of minimality of m.

$$a_1 v_1 + \ldots + a_m v_m = 0$$

Apply $T - \lambda_m I$ to both sides we get

$$a_1(\lambda_1 - \lambda_m)v_1 + \ldots + a_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0$$

making $a_1, \ldots, a_{m-1}$ a linear dependent list with length less than m. That contradicts. $\square$

Corollary 5.1.1: An upper bound of number of eigenvalues

Suppose V is finite dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

5.1.2 Polynomials Applied to Operators

The main reason an operator is far more interesting than arbitrary linear maps is that it can be raised to powers. That is, if T is an operator, then $T^2 = TT$ makes sense.

Notation 5.1.2: T^m

Suppose $T \in \mathscr{L}(V)$, $m \in \mathbb{Z}_+$. Define

- $T^m \in \mathscr{L}(V)$ is defined by $T^m = \underbrace{T \ldots T}_{m \text{ times}}$.
- T^0 is defined to be the identity operator I on V.
- If T is invertible, $T^{-m} = (T^{-1})^m$.

Then we can define polynomials of operators.

Notation 5.1.3: $p(T)$

Suppose $T \in \mathscr{L}(V)$, $p \in \mathscr{P}(\mathbb{F})$ is a polynomial

$$p(x) = a_0 + a_1 x + \ldots + a_m z^m$$

Then define $p(T) \in \mathscr{L}(V)$ to be

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \ldots + a_m T^m.$$

We earlier observe if $T \in \mathscr{L}(V)$, then $\operatorname{null} T$ and $\operatorname{range} T$ are invariant subspaces under T. Of course this still hold for polynomials.

Theorem 5.1.3: Polynomials of operator are invariant

Suppose $T \in \mathscr{L}(V)$ and $p \in \mathscr{P}(\mathbb{F})$. Then $\operatorname{null} p(T)$ and $\operatorname{range} p(T)$ are invariant subspaces under T.

We know from intuition that change of basis does not influence eigenvalues.

Theorem 5.1.4: Change of basis and eigenvalues

Suppose $T \in \mathscr{L}(V)$, and $S \in \mathscr{L}(V)$ invertible. Then

- T and $S^{-1}TS$ has the same eigenvalues.
- v is an eigenvector corresponding to λ in T, then $S^{-1}v$ is an eigenvector corresponding to λ in $S^{-1}TS$.

Proof. If $\lambda$is an eigenvalue, then $Tv = \lambda v$ for some $v \in V$. Then

$$S^{-1}TS(S^{-1}v) = \lambda(S^{-1}v).$$

□

The same goes for duals.

Theorem 5.1.5: Dual Maps and Eigenvalues

Suppose V is finite dimensional. $T \in \mathscr{L}(V)$. Then λ is an eigenvalue of T iff λ is an eigenvalue of $T' \in \mathscr{L}(V')$.

For complexification.

Theorem 5.1.6: Complexification and Eigenvalues

$\mathbb{F} = \mathbb{R}$, $T \in \mathscr{L}(V)$, and $\lambda \in \mathbb{C}$. Then λ is an eigenvalue of $T_{\mathbb{C}}$ iff $\overline{\lambda}$ is an eigenvalue of $T_{\mathbb{C}}$.

Proof. If $\lambda = a + ib$ is an eigenvalue of $T_{\mathbb{C}}$. Then $T(v + iu) = (a + ib)(v + iu)$. Then $T(v) = av - bu$ and $T(u) = au + bv$. Then $T(u - iv) = av - bu - i(au + bv) = (a - ib)(v - iu)$. □

Example: **Infinite-dimensional vector spaces**

- The forward shift operator $T \in \mathscr{L}(\mathbb{F}^\infty)$ defined by

$$T(z_1, z_2, \ldots) = (0, z_1, z_2, \ldots)$$

has no eigenvalues.

5.2 The Minimal Polynomial

5.2.1 Existence of Eigenvalues on Complex Vector Spaces

Theorem 5.2.1: Existence of Eigenvalues

Every operator on a finite dimensional nonzero complex vector space has an eigenvalue.

Proof. Suppose V is a finite dimensional complex vector space of dimension $n > 0$ and $T \in \mathscr{L}(V)$. Let $v \in V$ with $v \neq 0$, then

$$v, Tv, T^2v, \ldots, T^nv$$

is not linear independent. Thus, there exists a non-constant polynomial p of the smallest degree such that $p(T)v = 0$. But $\exists \lambda \in \mathbb{C}$ such that $p(\lambda) = 0$. That is $\exists q \in \mathscr{P}(\mathbb{C})$ such that

$$p(z) = (z - \lambda)q(z), \forall z \in \mathbb{C}$$

Then

$$0 = p(T)v = (T - \lambda I)q(T)v$$

Because q has smaller degree then p, $q(T)v \neq 0$. Thus λ is an eigenvalue of T. $\square$

Remark:
Both finite dimensional and complex are essential. For example, the forward shift operator has no eigenvalues.
This theorem can also be proved by determinants, with the eigenvalues being the roots of the characteristic polynomial $\det(T - \lambda I) = 0$.

5.2.2 Eigenvalues and Minimal Polynomials

Definition 5.2.1: Minimal Polynomials

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$. Then there is a unique monic polynomial $p \in \mathscr{P}(\mathbb{F})$ of smallest degree such that $p(T) = 0$. This polynomial is called the minimal polynomial of T. Furthermore, $\deg p \leq \dim V$.

Proof. If $\dim V = 0$, then I is the zero operator and we take $p = 1$.

Now use induction on $\dim V$. Let $\dim V = n$, assume that the result is true for all vector spaces of smaller dimension. Let $v \in V$, $v \neq 0$. Then $v, Tv, \ldots, T^nv$ is linear dependent. Thus there exists a smallest positive integer m such that $\exists c_0, c_1, \ldots, c_{m-1} \in \mathbb{F}$ with

$$c_0 v + c_1 T v + \ldots + c_{m-1} T^{m-1} v + T^m v = 0$$

Define a polynomial $q \in \mathscr{P}_m(\mathbb{F})$ such that

$$q(z) = c_0 + c_1 z + \ldots + c_{m-1} z^{m-1} + z^m$$

Then $q(T)v = 0$. Furthermore, for every $k \in \mathbb{N}$ we have

$$q(T) T^k v = T^k q(T) v = 0$$

Because $v, Tv, \ldots, T^{m-1}v$ is linear independent, so $\dim \operatorname{null} q(T) \geq m$. Hence

$$\dim \operatorname{range} q(T) \leq n - m$$

Because $\operatorname{range} q(T)$ is invariant under T, we consider the operator $T' = T|_{\operatorname{range} q(T)}$ on the vector space $\operatorname{range} q(T)$. Thus there is a monic polynomial $s \in \mathscr{P}(\mathbb{F})$ such that

$$\deg s \leq n - m \text{ and } s(T') = 0$$

Then $\forall v \in V$ we have

$$(sq)(T)v = s(T)q(T)v = 0$$

Thus sq is a monic polynomial such that $\deg sq \leq \dim V$ and $sq(T) = 0$. This completes the existence part.

For uniqueness, let $p \in \mathscr{P}(\mathbb{F})$ be a monic polynomial of the smallest degree such that $p(T) = 0$. Let r be another monic polynomial of the same degree and $r(T) = 0$. Then $(p - r)T = 0$. Because p has the smallest degree, $p - r = 0$. □

Remark:

To compute the minimal polynomial, we can just solve the linear equation

$$c_0 I + c_1 T + \ldots + c_{m-1} T^{m-1} + T^m = 0$$

trying for smallest m possible. However, most of the time we can just check $m = n$ and see if the equation has a unique solution.

More quickly, if we have a vector $v \neq 0$ such that $v, Tv, \ldots, T^{n-1}v$ are linear independent, then writing

$$c_0 v + c_1 T v + \cdots + c_{n-1} T^{n-1} v + T^n v = 0$$

have a unique solution for $c_0, \ldots, c_{n-1}$, then we are done.

Next we shall see the proof in a more geometric aspect.

1. First of all, we have $U = \operatorname{span}\left\{T^k v : k \in \mathbb{N}\right\}$ is an invariant space under T. (We see this like there is a "cycle" in the list $v, Tv, \ldots$.

2. We find a polynomial that compress U to 0. Therefore, we compress V to the quotient space V/U which is $\operatorname{range} q(T)$.

3. V/U is has less dimension, so we use induction assumptions to further compress it to 0.

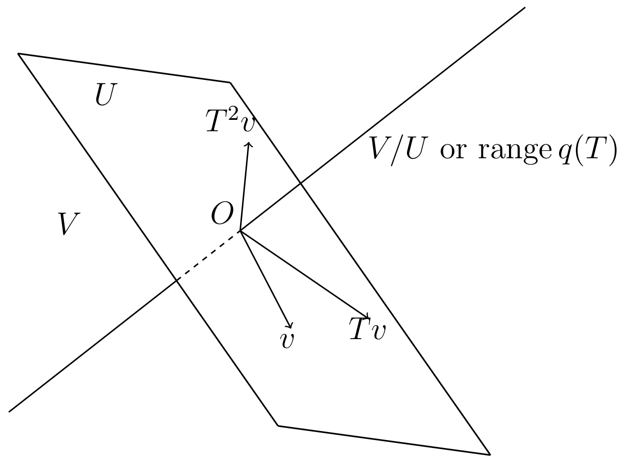

Figure 5.1: Geometric Interpretation of Minimal Polynomials

In this part we see that every operator on a finite dimensional vector space has a minimal polynomial. Conversely, if a polynomial $p(T)$ is given, we can consider its null space: $\operatorname{null} p(T)$. Restricting T to this space, we have $p(T|_{\operatorname{null} p(T)}) = 0$. This is useful to analysis the properties of null spaces of polynomials, as we will see later.

Using the minimal polynomial, we get a stronger version of theorem 5.2.1.

Theorem 5.2.2: Eigenvalues and the Minimal Polynomial

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$.

1. The zeros of the minimal polynomial of T are exactly the eigenvalues of T.
2. If V is a complex vector space, then the minimal polynomial has the form

$$(z - \lambda_1) \cdots (z - \lambda_m)$$

where $\lambda_1, \ldots, \lambda_m$ is a list of all eigenvalues of T, possibly with repetition.

Proof. Let p be the minimal polynomial of T. If λ is a root of p, then

$$p(z) = (z - \lambda)q(z)$$

where $q \in \mathscr{P}(\mathbb{F})$. Because $p(T) = 0$, we have

$$\forall v \in V, (T - \lambda I)q(T)v = 0$$

Because $\deg q < \deg p$, so q is not the minimal polynomial, so $\exists v \in V, q(T)v \neq 0$.

To see that every eigenvalue of T is a zero of p, let λ is an eigenvalue of T. We have $\exists v \in V, v \neq 0$ such that $Tv = \lambda v$, so $T^k v = \lambda^k v$ for all $k \in \mathbb{N}$. Thus,

$$p(T)v = p(\lambda)v = 0$$

Thus $p(\lambda) = 0$. □

Remark:
We shall make it clear that the multiple root of the minimal polynomial is permitted, implying the size of the largest Jordan block of λ.
As the minimal polynomial has degree at most $\dim V$, there can be at most $\dim V$ different eigenvalues, as stated above in corollary 5.1.1.

To see what "minimal" means in another perspective, we have the following

Theorem 5.2.3

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$, $q \in \mathscr{P}(\mathbb{F})$. Then $q(T) = 0$ iff $p \mid q$.

Proof. Let $q = ps + r$ where $\deg r < \deg p$. Then $0 = q(T) = p(T)s(T) + r(T) = r(T)$ so $r = 0$. □

Corollary 5.2.1: Minimal Polynomial of a Restricted Operator

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$, U is a subspace of V that is invariant under T. Let p be the minimal polynomial of T and p' be the minimal polynomial of $T|_U$. Then $p' \mid p$.

Theorem 5.2.4: Equivalent Conditions for Invertibles

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$, p is the minimal polynomial of T. Then the following are equivalent:

1. T is not invertible.
2. 0 is an eigenvalue of T.
3. 0 is a zero of p.
4. the constant term of p is 0.

5.2.3 Eigenvalues on Odd-Dimensional Real Vector Spaces

We shall show that every operator on an odd-dimensional real vector space has an eigenvalue.

Lemma 5.2.1: Even-dimensional Null Space

Suppose $\mathbb{F} = \mathbb{R}$ and V is finite dimensional, $T \in \mathscr{L}(V)$. Let $b, c \in \mathbb{R}$ with $b^2 < 4c$. Then $\dim \operatorname{null}(T^2 + bT + cI)$ is even.

As we are only interested in the properties of null space of the polynomial, feel free to restrict our discussion only to that space. This will make sure that $T^2 + bT + cI = 0$.

Proof. From theorem 5.1.3 we know that $\operatorname{null}(T^2 + bT + cI)$ is invariant under T. Let $U = \operatorname{null}(T^2 + bT + cI)$ and $T' = T|_U$ we have $T'^2 + bT' + cI = 0$. We now restrict our discussion to U.

Suppose $\lambda \in \mathbb{R}$ and $v \in U$ such that $T'v = \lambda v$. Then

$$0 = (T'^2 + bT' + cI)v = (\lambda^2 + b\lambda + c)v = \left(\left(\lambda + \frac{b}{2}\right)^2 + c - \frac{b^2}{4}\right)v$$

as the parentheses > 0, so $v = 0$, implying that T' has no eigenvector.

Let W be a subspace of U with the largest dimension such that:

- W is invariant under T'.
- W has even dimension.

If $W = U$ we're done. If $\exists w \in U, w \notin W$, let $Y = \operatorname{span}(w, T'w)$ then

- Y is invariant under T' for $T(Tw) = -bTw - cw$.
- $\dim Y = 2$ otherwise w is an eigenvector.

We also have $W \cap Y = \{0\}$ because otherwise $W \cap Y$ would be one-dimensional subspace invariant under T', which means T' has an eigenvector.

Therefore, we have

$$\dim(W + Y) = \dim W + \dim Y - \dim(W \cap Y) = \dim W + 2.$$

This contradicts that W has largest dimension. $\square$

Theorem 5.2.5: Eigenvalues of Odd-dimensional Real Vector Spaces

Every operator on an odd-dimensional real vector space has an eigenvalue.

Proof. Let $\mathbb{F} = \mathbb{R}$ and V is finite dimensional, $\dim V = n$ which is odd. Let $T \in \mathscr{L}(V)$. We shall again use induction on n. Note that the result holds for $n = 1$.

Now suppose $n \geq 3$ and the result hold for all cases of fewer odd dimensions. Let p denote the minimal polynomial of T. If p has factor $x - \lambda$ then we are done. So we shall suppose $p(x) = q(x)(x^2 + bx + c)$, where $b^2 < 4c$, thus,

$$0 = p(T) = q(T)(T^2 + bT + cI)$$

This implies that $q(T) = 0$ on $\operatorname{range}(T^2 + bT + cI)$. For $\deg q < \deg p$ and p is the minimal polynomial, we have $\operatorname{range}(T^2 + bT + cI) \neq V$.

As $\dim \operatorname{null}(T^2 + bT + cI)$ is even, then $\dim \operatorname{range}(T^2 + bT + cI)$ is odd, so $T|_{\operatorname{range}(T^2+bT+cI)}$ has an eigenvalue according to our induction hypothesis. So T has an eigenvalue. $\square$

5.2.4 Companion Matrix

Definition 5.2.2: Companion Matrix

Suppose $a_0, \ldots, a_{n-1} \in \mathbb{F}$. Let $T \in \mathscr{L}(\mathbb{F}^n)$ with respect to standard basis whose matrix is:

$$\begin{pmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & \ddots & & & -a_2 \\ & & \ddots & & & \vdots \\ & & & & 0 & -a_{n-2} \\ & & & & 1 & -a_{n-1} \end{pmatrix}$$

The blanks are all 0.
Then the minimal polynomial of the matrix is

$$a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n$$

This shows that every polynomial is the minimal polynomial of some operator.

5.3 Upper Triangle Matrices

When we study operators, we shall assume that we use the same basis for domain and range. A central goal of linear algebra is to show that "given an operator T, there is a reasonably good basis that we can express T is a simple matrix".

Definition 5.3.1: Upper-triangle Matrix

A square matrix in called upper-triangle if all entries below the diagonal are 0.

That is, it has the form

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

The upper triangle matrices fulfills our intuition of "getting least dimension to do the next thing".

Theorem 5.3.1: Conditions for upper-triangle matrices

Suppose $T \in \mathscr{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then the following are equivalent.

1. The matrix of T with respect of $v_1, \ldots, v_n$ is upper triangle.
2. $\text{span}(v_1, \ldots, v_k)$ is invariant under T for each $k = 1, \ldots, n$.

Proof. This is quite straightforward. □

The next result gives a simple equation that the upper triangle matrix satisfies.

Theorem 5.3.2: Equation satisfied by operator with upper-triangle matrix

Suppose $T \in \mathscr{L}(V)$ and V has a basis which T is upper-triangle with diagonal entries $\lambda_1, \ldots, \lambda_n$, then

$$(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0 \tag{5.1}$$

Proof. Our intuition for this is that for any $v \in V$, we have $(T - \lambda_n I)v$ deletes the last dimension of v. Going on, we delete each dimension of v until v become 0.

To see this more closely, let $v_1, \ldots, v_n$ denotes the basis. Then we have $(T - \lambda_i I)v_i \in \operatorname{span}(v_1, \ldots, v_{i-1})$. So we have $(T - \lambda_1 I) \cdots (T - \lambda_n I)v_i = 0$ for all $i = 1, \ldots, n$. □

It is easy to determine the eigenvalues of operators representing as upper-triangle matrices.

Theorem 5.3.3: The Eigenvalues from Upper-triangle Matrices

Suppose $T \in \mathscr{L}(V)$ is represented as upper-triangle matrix in V. Then the eigenvalues are exactly the entries on the diagonal of the matrix.

Proof. Let $v_1, \ldots, v_n$ be the basis, with $\mathscr{M}(T)$ is

$$\mathscr{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

- We have $Tv_1 = \lambda_1 v_1$ so λ_1 is an eigenvalue of T.
- For $k \in \{2, \ldots, n\}$, we see that $(T - \lambda_k I)v_k \in \operatorname{span}(v_1, \ldots, v_{k-1}$. Thus $T - \lambda_k I$ is not injective. (The range has less dimension)

  Therefore $\exists v \in V$ such that $(T - \lambda_k I)v = 0$ so λ_k is an eigenvalue of T.

To prove that T has no other eigenvalues, let $q \in \mathscr{P}(\mathbb{F})$ be $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$, then $q(T) = 0$, thus q is the multiple of the minimal polynomial of T, thus the minimal polynomial has no other roots other then $\lambda_1, \ldots, \lambda_n$. □

This result illustrate the connection of upper-triangle matrix, eigenvalues, and minimal polynomials. We have

Theorem 5.3.4: Equivalent Conditions for having an Upper-triangle Matrix

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$. Then T has an upper-triangle matrix with respect to some basis of V if and only if the minimal polynomial of T equals $(z-\lambda_1) \cdots (z-\lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.

Proof.
- First, if T has an upper-triangle matrix with diagonal entries $\alpha_1, \ldots, \alpha_n$, then $q(z) = (z - \alpha_1) \cdots (z - \alpha_n)$ is the multiple of the minimal polynomial of T.
- To deal with the other side we shall use induction on m.

  If $m = 1$, then $T = \lambda_1 I$ is an upper-triangle matrix. If for any smaller then m is correct, let $U = \operatorname{range}(T - \lambda_m I)$ is invariant under T, Then $T|_U$ is an operator on U.

If $u \in U$, then $u = (T - \lambda_m I)v$ for some $v \in V$ and

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0$$

Hence $(z - \lambda_1) \cdots (z - \lambda_{m-1})$ is a multiple of minimal polynomial of $T|_U$ on U. Therefore, $T|_U$ has an upper triangle matrix for some basis of U, let it be $u_1, \ldots, u_M$. We have

$$Tu_k = (T|_U)u_k \in \operatorname{span}(u_1, \ldots, u_k), \forall k \in \{1, \ldots, M\}$$

Extend our basis to V: $u_1, \ldots, u_M, v_1, \ldots, v_N$. For each $k \in \{1, \ldots, N\}$ we have

$$Tv_k = (T - \lambda_m I)v_k + \lambda_m v_k$$

Thus we have

$$Tv_k \in \operatorname{span}(u_1, \ldots, u_m, v_1, \ldots, v_k)$$

Then T is an upper-triangle matrix with respect to $u_1, \ldots, u_M, v_1, \ldots, v_N$.

□

Corollary 5.3.1: Triangle Matrix of $\mathbb{C}$

Suppose V is a finite dimensional vector space. And $\mathbb{F} = \mathbb{C}$. For every $T \in \mathscr{L}(V)$, there is a basis such that T is an upper-triangle matrix.

Remark:
Note that the upper-triangle matrix here has nothing to do with the one we get from row echelon forms.

Proposition: **Operations of Triangle Matrices**

Suppose A, B are upper-triangle matrices of the same size, with $\alpha_1, \ldots, \alpha_n$ on the diagonal of A and $\beta_1, \ldots, \beta_n$ on the doagonal of B.

1. $A + B$ is an upper-triangle matrix with $\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n$ on the diagonal.

2. AB is also an upper-triangle matrix with $\alpha_1\beta_1, \ldots, \alpha_n\beta_n$ on the diagonal.

If $T \in \mathscr{L}(V)$ is invertible and $v_1, \ldots, v_n$ is a basis of V that makes T an upper-triangle matrix, with $\lambda_1, \ldots, \lambda_n$ on the diagonal. Then the matrix of T^{-1} is also upper-triangle with respect to the basis, with

$$\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}$$

on the diagonal.

Proof. The first two are rather simple matters, for the third one, we have $TT^{-1} = I$. As T is invertible, none of $\lambda_1, \ldots, \lambda_n$ is 0. Let $T^{-1} = M$. We have

$$M_{1,j}T_{\cdot,1} + \ldots + M_{n,j}T_{\cdot,n} = e_j$$

having the form as

$$M_{1,j}\begin{pmatrix}T_{1,1}\\0\\\vdots\\0\\0\end{pmatrix}+M_{2,j}\begin{pmatrix}T_{1,2}\\T_{2,2}\\\vdots\\0\\0\end{pmatrix}+\ldots+M_{n-1,j}\begin{pmatrix}T_{1,n-1}\\T_{2,n-1}\\\vdots\\T_{n-1,n-1}\\0\end{pmatrix}+M_{n,j}\begin{pmatrix}T_{1,n}\\T_{2,n}\\\vdots\\T_{n-1,n}\\T_{n,n}\end{pmatrix}=e_j$$

Solving the equation, we have $M_{j+1,j}=\ldots=M_{n,j}=0$, so M is upper-triangle, The entries on the diagonal are easy to check. □

5.3.1 Diagonalizable Operators

5.3.2 Diagonal Matrices

Definition 5.3.2: Diagonal Matrix

A *diagonal matrix* is a square matrix that is 0 everywhere except possibly on the diagonal. (Well, the diagonal can have 0 really)

Definition 5.3.3: Diagonalizable

An operator on finite dimensional V is called *diagonalizable* if the operator has a diagonal matrix with respect to some basis of V.

Remark:
A diagonal operator means that we can find $\dim V$ directions that the operator is just scaling on these dimensions. And yes, the scaling parameter are just the eigenvalues.

Definition 5.3.4: Eigenspace, $E(\lambda, T)$

Suppose $T \in \mathscr{L}(V)$ and $\lambda \in \mathbb{F}$ is an eigenvalue of T. The eigenspace of T corresponding to λ is the subspace $E(\lambda, T) \subseteq V$ defined by

$$E(\lambda, T) = \operatorname{null}(T - \lambda I) = \{v \in V : Tv = \lambda v\}$$

Hence $E(\lambda, T)$ is the set of all eigenvectors of T corresponding to λ.

It is easy to notice that the operator T restricted to $E(\lambda, T)$ is just the operator $\cdot\lambda$.

Theorem 5.3.5: Direct Sum of Eigenspaces

Suppose $T \in \mathscr{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are eigenvalues of T, then

$$E(\lambda_1, T) + \ldots + E(\lambda_m, T)$$

is a direct sum. If V is finite dimensional, we have

$$\dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T) \leq \dim V$$

Proof. Using the linear independence of eigenvectors corresponding to distinct eigenvalues would suffice. □

5.3.3 Conditions for Diagonalizability

Theorem 5.3.6: Equivalent Conditions of Diagonalizability

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denotes the distinct eigenvalues of T. Then the following are equivalent:

1. T is diagonalizable.
2. V has a basis consisting of eigenvectors of T.
3. $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.
4. $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

Proof.

- An operator $T \in \mathscr{L}(V)$ has a diagonal matrix

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

with respect to a basis $v_1, \ldots, v_n$ of V iff $T_k v_k = \lambda_k v_k$. So 1 and 2 are equivalent.

- Suppose 2 holds, so V has a basis of eigenvectors. So

$$V = E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

Using theorem 5.3.5 would imply 3.

- 4 follows 3 is obvious.
- Suppose 4 holds, Choose a basis of each $E(\lambda_i, T)$ and put them together, forming a basis of V. That will imply 2.

□

Remark:
We already know that every operator on a finite dimensional complex vector space has an eigenvalue, but there exists operators on it that is not diagonalizable, such as

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Corollary 5.3.2: Enough Eigenvalues implies Diagonalizability

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$ has $\dim V$ eigenvalues, then T is diagonalizable.

We can use diagonalizability to easily calculate powers. For instance, if $T \in \mathscr{L}(V)$ is diagonalizable, changing to the basis in which the matrix of T is diagonal is S, we have

$$M = SDS^{-1} \tag{5.2}$$

thus

$$M^k = SD^kS^{-1}$$

where D^k is just taking powers from the diagonal.

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad D^k = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix}$$

The next result is also showing the power of minimal polynomial, and implying what "minimal" means also.

Theorem 5.3.7: Diagonalizability and Minimal Polynomial

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$. Then T is diagonalizable if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for distinct $\lambda_1, \ldots, \lambda_m$.

Proof. First suppose T is diagonalizable. Thus there is a basis $v_1, \ldots, v_n$ of V consisting of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues corresponding to v_i, that is, $\forall j \exists \lambda_k, (T - \lambda_k I)v_j = 0$. So

$$\forall j \in \{1, \ldots, n\}, (T - \lambda_1 I) \cdots (T - \lambda_m I)v_j = 0$$

so $p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$ is a multiple of the minimal polynomial of T.

For the other direction, we do it by induction on m. For $m = 1$ we have $T = \lambda_1 I$ which is diagonalizable.

Suppose $m > 1$ and the result hold for smaller m, we have range$(T - \lambda_m I)$ is invariant under T. Restrict our discussion to range$(T - \lambda_m I)$.

If $u \in$ range$(T - \lambda_m I)$, then $u = (T - \lambda_m I)v$ for some $v \in V$, so

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0$$

Hence $(z - \lambda_1) \cdots (z - \lambda_{m-1})$ is a multiple of minimal polynomial of T restricted to range$(T - \lambda_m I)$. So there is a basis of range$(T - \lambda_m I)$ consisting of eigenvectors of T.

Suppose that $u \in$ range$(T - \lambda_m I) \cap$ null$(T - \lambda_m I)$. Then $Tu = \lambda_m u$. Now we have

$$\begin{aligned} 0 &= (T - \lambda_1 I) \cdots (T - \lambda_m I)u \\ &= (\lambda_m - \lambda_1) \cdots (\lambda_m - \lambda_{m-1})u. \end{aligned}$$

Because $\lambda_1, \ldots, \lambda_m$ are distinct, this implies $u = 0$, therefore range$(T - \lambda_m I) \cap$null$(T - \lambda_m I) = \{0\}$, making

$$V = \text{range}(T - \lambda_m I) \oplus \text{null}(T - \lambda_m I)$$

Joining the basis we get earlier with a basis of null$(T - \lambda_m I)$ we get a basis of V consisting of eigenvectors of T, completing the proof. □

The next result is quite straightforward.

Proposition: **Restriction of Diagonalizable Operators to Invariant Subspaces**

Suppose $T \in \mathscr{L}(V)$ is diagonalizable and U is a subspace of V that is invariant under T. Then $T|_U$ is a diagonalizable operator on U.

Proof. Because T is diagonalizable over V, then the minimal polynomial has the form $(z - \lambda_1)\cdots(z-\lambda_m)$ with distinct $\lambda_1, \ldots, \lambda_m$. Also the minimal polynomial is a multiple of the minimal polynomial of $T|_U$ so we get our result. □

5.3.4 Gershgorin Disk Theorem

Definition 5.3.5: Gershgorin Disks

Suppose $T \in \mathscr{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Let A denote the matrix of T with respect to the basis. A *Gershgorin disk* of T with respect to the basis is a set:

$$\left\{ z \in \mathbb{F} : |z - A_{j,j}| \leq \sum_{1 \leq k \leq n, k \neq j} |A_{j,k}| \right\}, \tag{5.3}$$

where $j \in \{1, \ldots, n\}$. (Well T has n Gershgorin disks)

The Gershgorin disks are a small range near the diagonal entry, indicating the tiny shift from diagonal. As we can see, a diagonal matrix has zero radius Geometric disks.

Theorem 5.3.8: Gershgorin Disk Theorem

Suppose $T \in \mathscr{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then each eigenvalue of T is contained in some Gershgorin disk of T with respect to the basis $v_1, \ldots, v_n$.

Proof. Suppose $\lambda \in \mathbb{F}$ is an eigenvalue of T. Let $w \in V$ ba a corresponding eigenvector. There exists $c_1, \ldots, c_n \in \mathbb{F}$ such that

$$w = c_1 v_1 + \ldots + c_n v_n$$

Let A denote the matrix of T for v_i, applying T to both sides we get

$$\lambda w = \sum_{k=1}^{n} c_k T v_k = \sum_{k=1}^{n} c_k \sum_{j=1}^{n} A_{j,k} v_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} A_{j,k} c_k \right) v_j.$$

Let $j \in \{1, \ldots, n\}$ be such that

$$|c_j| = \max\{|c_1|, \ldots, |c_n|\}$$

Therefore we have

$$\lambda c_j = \sum_{k=1}^{n} A_{j,k} c_k$$

To subtract $A_{j,j}$ we have

$$|\lambda - A_{j,j}| = \left| \sum_{k=1, k\neq j}^{n} A_{j,k} \frac{c_k}{c_j} \right| \leq \sum_{k=1, k\neq j}^{n} |A_{j,k}|.$$

□

In the proof c_j is the major contribution to λw. It is quite intuitive that the major contribution would not differ much from the actual eigenvalue.

5.4 Commuting Operators

Commuting operators, like $AB = BA$, is very rare in common. They often indicates some similarity of operators.

Theorem 5.4.1: Eigenspace is Invariant Under Commuting Operators

Suppose $S, T \in \mathscr{L}(V)$ commute, then $E(\lambda, S)$ is invariant under T.

Proof. Suppose $v \in E(\lambda, S)$, then

$$S(Tv) = (ST)v = (TS)v = T(Sv) = \lambda Tv$$

So $Tv \in E(\lambda, S)$. □

Remark:
In our common sense, commutativeness comes from "stretching at the same direction". A simple image is diagonal operators commute. Changing directions would cause some non-commuting behavior, like rotating in different axis.

Theorem 5.4.2: Commutativity and Simultaneous Diagonalizability

Two diagonalizable operators on the same vector space have diagonal matrices with respect to the same basis if and only if the two operators commute.

Proof. First suppose that S, T have diagonal matrices for the same basis V, then for this basis we have S and T commute.

The other side we now suppose $S, T \in \mathscr{L}(V)$ are diagonalizable operators that commute. Let $\lambda_1, \ldots, \lambda_m$ denote distinct eigenvalues of S. Then we have

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S)$$

Then for each $k = 1, \ldots, m$ the subspace $E(\lambda_k, S)$ is invariant under T. Therefore by proposition 5.3.3 we have $T|_{E(\lambda_k, S)}$ is diagonalizable. So there is a basis of $E(\lambda_k, S)$ consisting the eigenvalues of T. Forming together we get a basis consisting of eigenvectors for both S and T. □

If we cancel out the diagonalizable assumption, we have a result for the complex space:

Proposition: **Common Eigenvector for Commuting Operators**

Every pair of commuting operators on a finite dimensional nonzero complex vector space has a common eigenvector.

Proof. Let λ be an eigenvalue of S, (which indeed exists for complex vector spaces) Thus $E(\lambda, S) \neq \{0\}$ and is invariant under T. Thus $T|_{E(\lambda, S)}$ has an eigenvector, completing the proof. □

Not surprisingly, we reduced to upper triangle matrices for generalization of non-diagonal sizable matrices.

Theorem 5.4.3: Commuting Operators are Simultaneously upper triangle

Suppose V is finite dimensional complex vector space and S, T are commuting operators on V. Then $\exists$ a basis of V for which S and T are both upper triangle.

Proof. Let $n = \dim V$, We shall use induction on n. For $n = 1$ the result holds. Now suppose $n > 1$ and the result holds for complex vector spaces that has dimension $n - 1$.

Let v_1 be any common eigenvector of S and T. Then $Sv_1 \in \operatorname{span}(v_1)$ and $Tv_1 \in \operatorname{span}(v_1)$. Let W be a subspace of V such that

$$V = \operatorname{span}(v_1) \oplus W.$$

We collapse V to W by $P : V \to W, P(av_1 + w) = w$. For each $\alpha \in \mathbb{C}$ and $w \in W$, define $\hat{S}, \hat{T} \in \mathscr{L}(W)$ by

$$\forall w \in W, \quad \hat{S}w = P(Sw) \text{ and } \hat{T}w = P(Tw)$$

We shall see $\hat{S}, \hat{T}$ commute:

$$(\hat{S}\hat{T})w = \hat{S}(P(Tw)) = \hat{S}(Tw - av_1) = P(S(Tw - av_1)) = P((ST)w). \quad \text{for } P(S(v_1)) = 0$$

Therefore, we state that there exists a basis $v_2, \ldots, v_n \in W$ such that $\hat{S}$ and $\hat{T}$ has upper-triangle matrices. The list $v_1, \ldots, v_n$ is a basis of V.

For $k \in \{2, \ldots, n\}$ there exists $a_k, b_k \in \mathbb{C}$ such that

$$Sv_k = a_k v_1 + \hat{S}v_k \quad \text{and} \quad Tv_k = b_k v_1 + \hat{T}v_k$$

So $\hat{S}v_k \in \operatorname{span}(v_1, \ldots, v_k)$ and $\hat{T}v_k \in \operatorname{span}(v_1, \ldots, v_k)$ as usual. So we get our result. $\square$

Remark:
We've seen many proof using the technique of collapsing a dimension and then using induction.
In general, it is not possible to determine the eigenvalues of sum and product of two operators. However, there are something happening with comuuting operators.

Theorem 5.4.4: Eigenvalues of sum and product of Commuting Operators

Suppose V is finite dimensional complex vector space and S, T are commuting operators on V. Then

- every eigenvalue of $S + T$ is an eigenvalue of S + an eigenvalue of T.
- every eigenvalue of ST is an eigenvalue of S $\times$ an eigenvalue of T.

Proof. Using the basis that makes moth S and T upper-triangular, we have

$$\mathscr{M}(S + T) = \mathscr{M}(S) + \mathscr{M}(T) \text{ and } \mathscr{M}(ST) = \mathscr{M}(S)\mathscr{M}(T)$$

$\square$

Chapter 6

Inner Product Space

In this part, we add geometric features (angles, distances) to vector spaces via inner products.

We assume $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.

6.1 Inner Products and Norms

6.1.1 Inner Products

The intuition of inner products comes form the dot product we defined on $\mathbb{R}^n$, with $\boldsymbol{x} \cdot \boldsymbol{y} = x_1y_1 + \ldots + x_ny_n$, which indicates the Euclidean metric $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$. For complex spaces however, we have $|\lambda|^2 = \lambda\overline{\lambda}$. So we incline to use conjugate for complex vector spaces, such that $\|z\|^2 = w_1\overline{z_1} + \ldots + w_n\overline{z_n}$.

Definition 6.1.1: Inner Product

An inner product on V is a function $V \times V \to \mathbb{F}$, denoted $\langle u, v\rangle$ such that:

- *Positivity:* $\forall v \in V, \langle v, v\rangle \geq 0$.
- *Definiteness:* $\langle v, v\rangle = 0 \leftrightarrow v = 0$.
- *Additivity in first slot:* $\forall u, v, w \in \mathbb{F}, \langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
- *Homogeneity in first slot:* $\forall u, v \in V, \forall\lambda \in \mathbb{F}, \langle \lambda u, v\rangle = \lambda \langle u, v\rangle$.
- *Conjugate Symmetry:* $\forall u, v \in V, \langle u, v\rangle = \overline{\langle v, u\rangle}$.

An inner product space is a vector space V along with an inner product on V.

The most common inner product is the Euclidean product on $\mathbb{F}^n$ given by:

$$\forall u, v \in \mathbb{F}^n, \langle u, v\rangle = \sum_{i=1}^{n} u_i\overline{v_i}$$

We can give a inner product for $C[-1, 1]$ by

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\mathrm{d}x$$

Theorem 6.1.1: Basic Properties of Inner Products

- For each $v \in V$, the map $f : V \to \mathbb{F}, v \mapsto \langle u, v \rangle$ is a linear map.
- $\forall v \in V, \langle v, 0 \rangle = \langle 0, v \rangle = 0$.
- $\forall u, v, w \in V, \langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$.
- $\forall \lambda \in \mathbb{F}, u, v \in V, \langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle$.

6.1.2 Norms

Definition 6.1.2: Norms

$\forall v \in V$, the norm of v, denoted by $\|v\|$, is defined by

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Proposition: **Properties of Norm**

Suppose $v \in V$, then

1. $\|v\| = 0 \leftrightarrow v = 0$.
2. $\forall \lambda \in \mathbb{F}, \|\lambda v\| = |\lambda| \|v\|$.

Definition 6.1.3: Orthogonal

$u, v \in V$ are called orthogonal if $\langle u, v \rangle = 0$.

The order of u, v does not matter for $\langle u, v \rangle = 0 \leftrightarrow \langle v, u \rangle = 0$.

Proposition: **Pythagorean Theorem**

Suppose $v, w \in V$, If v, u are orthogonal, then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Next we formalize the idea of projection. We want to write a $u = cv + w$ where $\langle v, w \rangle = 0$.

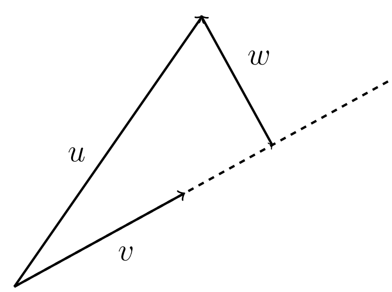

Figure 6.1: Projection

We have $u = cv + (u - cv)$. Therefore we have

$$0 = \langle u - cv, v\rangle = \langle u, v\rangle - c\|v\|^2$$

Then we write,

$$u = \frac{\langle u, v\rangle}{\|v\|^2}v + \left(u - \frac{\langle u, v\rangle}{\|v\|^2}v\right)$$

This is called an orthogonal decomposition.

Theorem 6.1.2: Cauchy-Schwartz Inequality

Suppose $u, v \in V$, then

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

This equality holds iff u, v are linear dependent.

Proof.

$$\|u\|^2 = \|\frac{\langle u, v\rangle}{\|v\|^2}v\|^2 + \|w\|^2 = \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}$$

□

The next result is triangle inequality.

Theorem 6.1.3: Triangle Inequality

Suppose $u, v \in V$, then

$$\|u + v\| \leq \|u\| + \|v\|$$

Proposition: **Parallelogram Equality**

Suppose $u, v \in V$, then

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$$

6.1.3 Extension of Inner Product

- Suppose $V_1, \ldots, V_m$ are inner product space. Then
$$\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m) \rangle = \langle u_1, v_1 \rangle + \ldots + \langle u_m, v_m \rangle$$
Is an inner product on $V_1 \times \cdots \times V_m$.
- Suppose V is a real vector space, for $u, v, w, x \in V$, then define
$$\langle u+iv, w+ix \rangle_{\mathbb{C}} = \langle u, w \rangle + \langle v, x \rangle + (\langle v, w \rangle - \langle u, x \rangle)\, i$$
makes $V_{\mathbb{C}}$ into a complex inner product space. (Well, as you can see, it has the same structure of $(a+bi)\cdot(c+di) = (a+bi)(c-di)$, the standard definition of inner product on $\mathbb{C}^n$)

6.2 Orthonormal Bases

Definition 6.2.1: Orthonormal

A list of vectors is called orthonormal if each vector in the list has norm 1 and is orthogonal to other vectors in the list. That is,

$$\langle v_i, v_j \rangle = \delta_{i,j} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

An orthonormal basis is a orthonormal list of vectors that forms a basis.

We have many properties of orthonormal lists. For example,

$$\|a_1 e_1 + \ldots + a_n e_n\|^2 = \|a_1\|^2 + \ldots + \|a_n\|^2$$

Proposition: **Linear Independence of Orthonormal**

Every orthonormal list of vectors in V is linear independent.

Proof.

$$a_1 e_1 + \ldots + a_n e_n = 0$$

Therefore we have $\|a_1\|^2 + \ldots + \|a_n\|^2 = 0$. □

Writing a vector in orthonormal basis $e_1, \ldots, e_n$, we have

- $v = \langle v, e_1 \rangle e_1 + \ldots + \langle v, e_n \rangle e_n$
- $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \ldots + |\langle v, e_n \rangle|^2$
- $\langle u, v \rangle = \langle u, e_1 \rangle \overline{\langle v, e_1 \rangle} + \ldots + \langle u, e_n \rangle \overline{\langle v, e_n \rangle}$

6.2.1 Gram Schmidt Procedure

We now give a way to construct an orthogonal basis form an arbitrary basis of V. The intuition of this procedure is projecting the vectors onto the orthogonal space of previous vectors.

Suppose $v_1, \ldots, v_m$ is a linear independent list of vectors in V, Let $f_1 = v_1$, define f_k inductively as:

$$f_k = v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \ldots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}$$

Then $f_1, \ldots, f_n$ is an orthogonal basis. Let $e_i = \dfrac{f_i}{\|f_i\|}$, then $e_1, \ldots, e_n$ is an orthonormal basis.

Now we have

Corollary 6.2.1: Existence of Orthonormal Basis

Every inner product space has an orthonormal basis.
Every orthonormal list extends to an orthonormal basis.

It is natural to think that there is some connection between the upper-triangle matrix.

Theorem 6.2.1: Upper-triangle matrices on Orthonormal Basis

Suppose V is finite dimensional and $T \in \mathscr{L}(V)$. Then T has an upper-triangle form then T has an upper-triangle form of some orthonormal basis.

Proof. Suppose T has an upper triangle matrix for some basis $v_1, \ldots, v_n$, applying the Gram-Schmidt Procedure to $e_1, \ldots, e_n$ we have

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k), \forall k$$

$\square$

Applying it to the complex vector space, we have

Theorem 6.2.2: Schur's Theorem

Every operator on a finite dimensional complex inner product space has an upper-triangle matrix with respect to some orthonormal basis.

6.2.2 Linear Fuunctionals on Inner Product Space

We have seen that $\forall v \in V$, the function $u \mapsto \langle v, u \rangle$ is a linear functional on V. We shall show that every linear functional has this form.

Theorem 6.2.3: Riesz Representation Theorem

Suppose V is finite dimensional and φ is a linear functional on V. Then there is a unique $v \in V$ such that

$$\forall u \in V, \varphi(u) = \langle u, v\rangle$$

Well, we can say that $\forall v \in V$, define $\varphi_v \in V'$ by

$$\varphi_v(u) = \langle u, v\rangle$$

then $v \mapsto \varphi_v$ is a bijection of $V \to V'$. (Note that this map may not be linear)

This is not surprising as we view linear functionals as row vectors, and dot product has the same form.

Proof. We first construct a vector. Let $e_1, \ldots, e_n$ be an orthonormal basis of V, then

$$\begin{aligned}\varphi(u) &= \varphi\left(\langle u, e_1\rangle e_1 + \ldots + \langle u_n, e_n\rangle e_n\right)\\ &= \langle u, e_1\rangle \varphi(e_1) + \ldots + \langle u, e_n\rangle \varphi(e_n)\\ &= \left\langle u, \overline{\varphi(e_1)}e_1 + \ldots + \overline{\varphi(e_n)}e_n\right\rangle\end{aligned}$$

Let

$$v = \overline{\varphi(e_1)}e_1 + \ldots + \overline{\varphi(e_n)}e_n \tag{6.1}$$

would do. For uniqueness let

$$\forall u \in V, \varphi(u) = \langle u, v_1\rangle = \langle u, v_2\rangle$$

then $0 = \langle u, v_1\rangle - \langle u, v_2\rangle = \langle u, v_1 - v_2\rangle$, then $v_1 = v_2$. □

Remark:
The Riesz representation theorem implies that an inner product gives a homomorphism from V' to V with $\varphi \Leftrightarrow u$.

6.3 Orthogonal Complements and Minimization Problems

6.3.1 Orthogonal Complements

Definition 6.3.1: Orthogonal Complements

If U is a subsset of V, then the orthogoanl complements of U, denoted $U^{\perp}$, is defined by

$$U^{\perp} = \{v \in V : \forall u \in U, \langle u, v\rangle = 0\} \tag{6.2}$$

There is a very geometric image of the . It is the "verticle set" of U. And the following are some obvious consequences.

Proposition: **Properties of Orthogonal Complement**

1. $\forall U \subseteq V$, $U^\perp$ is a subspace of V.

   Proof. $\langle \lambda u_1 + \mu u_2, v\rangle = \lambda \langle u_1, v\rangle + \mu \langle u_2, v\rangle$. □

2. $\forall U \subseteq V, U \cap U^\perp \subseteq \{0\}$.
3. $\forall G \subseteq H \subseteq V, H^\perp \subseteq G^\perp$.

Usually we care more about subsets U that are subspaces of V.

Proposition: **Direct Sum and its Orthogonal Complement**

Suppose U is a finite dimensional subspace of V, then

$$V = U \oplus U^\perp. \tag{6.3}$$

Proof. First we show that $V = U + U^\perp$. Let $e_1, \dots, e_m$ be an orthonormal basis of U. Then $\forall v \in V$, we have

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - (\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m)}_{w}$$

then we have $\langle w, e_k\rangle = \langle v, e_k\rangle - \langle v, e_k\rangle = 0$. So that $w \in U^\perp$. Also we know that $U \cap U^\perp = \{0\}$, so $V = U \oplus U^\perp$. □

A simple corollary is that if V is a finite dimensional vector space, we have

$$\dim U^\perp = \dim V - \dim U$$

Theorem 6.3.1: Orthogonal Complement of Orthogonal Complement

Suppose U is finite dimensional subspace of V, then

$$(U^\perp)^\perp = U \tag{6.4}$$

Proof.
- First we show $U \subseteq (U^\perp)^\perp$. Suppose $u \in U$, then $\forall w \in U^\perp, \langle u, w\rangle = 0$. We have $u \in (U^\perp)^\perp$.
- The other side we suppose $v \in (U^\perp)^\perp$. Then we can write $v = u + w$ where $u \in U \subseteq (U^\perp)^\perp$ and $w \in U^\perp$. Hence we have $v - u = w \in U^\perp$, and also $v - u \in (U^\perp)^\perp$ so $v - u = w = 0$. Thus $v \in U$.

□

Next we give the formalization of orthogonal projection, a generalization of the projection onto a one-dimensional space above.

Definition 6.3.2: Orthogonal Projection

Suppose U is a finite dimensional subspace of V. The orthogonal projection of V onto U is the operator $P_U \in \mathscr{L}(V)$:

$$\forall v \in V, \text{ let } v = u + w, u \in U, w \in U^{\perp}, \text{ then } P_U v = u \tag{6.5}$$

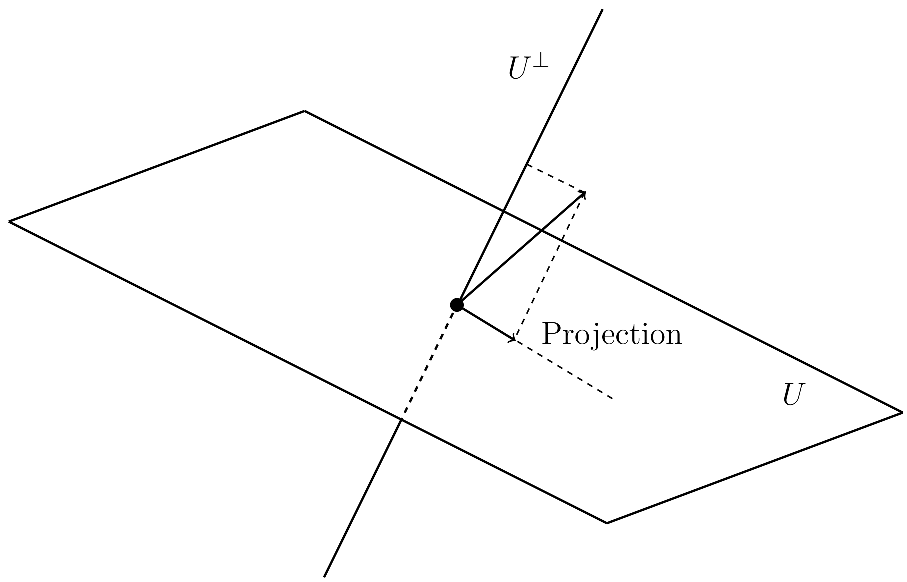

Figure 6.2: Orthogonal Projection

Proposition: **Properties of Orthogonal Projection**

Suppose U is a finite dimensional subspace of V, then

- $\forall u \in U, P_U u = u$ and $\forall v \in U^{\perp}, P_U v = 0$.
- range $P_U = U$ and null $P_U = U^{\perp}$.
- $\forall v \in V, v - P_U v \in U^{\perp}$.
- $P_U^2 = P_U$.
- If $e_1, \ldots, e_m$ is an orthonormal basis of U, then

$$P_U v = \langle v, e_1 \rangle e_1 + \ldots + \langle v, e_m \rangle e_m$$

Now we have a more clear review of the Riesz representation theorem. As $\varphi_v(u) = \langle u, v \rangle$, then $v \in (\operatorname{null} \varphi)^{\perp}$ to make $\langle u, v \rangle = 0$ whenever $\varphi_v(u) = 0$. And $(\operatorname{null} \varphi)^{\perp}$ has only one dimension unless $\varphi = 0$.

6.3.2 Minimization Problems

We take about problems that includes finding $\|v-u\|$ as small as possible when $v \in V$ and $u \in U$.

Theorem 6.3.2: Minimizing Distance

Suppose U is a finite dimensional subspace of V. Let $v \in V$, and $u \in U$. Then

$$\|v - P_U v\| \leq \|v - u\| \tag{6.6}$$

where equality holds iff $u = P_U v$.

6.3.3 Pseudoinverse

Suppose $T \in \mathscr{L}(V, W)$ and $b \in W$, consider finding $x \in V$ such that

$$Tx = b$$

If T is invertible, the unique solution would be $x = T^{-1}b$. Even if the equation has no solution, we can still manage to find x such that $\|Tx - b\|$ is as small as possible. There is where Pseudoinverse comes in.

Restriction of a linear map to $(\operatorname{null} T)^\perp$ would get what we want, as $(\operatorname{null} T)^\perp$ is the U in figure 6.2.

Proposition: **Restriction of a linear map**

Suppose V is finite dimensional and $T \in \mathscr{L}(V, W)$. Then $T|_{(\operatorname{null} T)^\perp}$ is a bijection of $(\operatorname{null} T)^\perp \to \operatorname{range} T$.

Proof. Obviously, $(\operatorname{null} T)^\perp \cong V/\operatorname{null} T \cong \operatorname{range} T$. □

Now we define the pseudoinverse $T^\dagger$. First we want $Tx = P_{\operatorname{range} T} b$ for minimal norm. Then as x may have multiple choice, we restrict our discussion to $(\operatorname{null} T)^\perp$.

Definition 6.3.3: Pseudoinverse

Suppose V is finite dimensional and $T \in \mathscr{L}(V, W)$. The pseudoinverse $T^\dagger \in \mathscr{L}(W, V)$ is defined

$$\forall w \in W, T^\dagger w = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T} w \tag{6.7}$$

We shall see that the pseudoinverse act like an inverse.

Proposition: **Properties of Pseudoinverse**

- If T is invertible, then $T^{-1} = T^\dagger$.
- $TT^\dagger = P_{\operatorname{range} T}$.
- $T^\dagger T = P_{(\operatorname{null} T)^\perp}$.

Remark:
To solve the equation $Tx = b$, taking $x = T^\dagger b$ gives the best fit to the equation so that $\|Tx - b\|$ is least possible. Also, for all vectors that makes $\|Tx - b\|$ equally small, $T^\dagger b$ has the smallest norm, as we take $(\operatorname{null} T)^\perp$.

Theorem 6.3.3: Pseudoinverse provides the best approximate solution

Suppose V is finite dimensional, $T \in \mathscr{L}(V, W)$ and $b \in W$.

- If $x \in V$, then
$$\|T(T^\dagger b) - b\| \le \|Tx - b\|$$
with equality holds iff $x \in T^\dagger b + \operatorname{null} T$.

- If $x \in T^\dagger b + \operatorname{null} T$, then
$$\|T^\dagger b\| \le \|x\|$$
with equality iff $x = T^\dagger b$.

Chapter 7

Operators on Inner Product Spaces

Standing Assumptions:

- $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.
- V, W are nonzero finite dimensional vector spaces.

7.1 Self-Adjoint and Normal Operators

7.1.1 Adjoints

Definition 7.1.1: Adjoint, T^*

Suppose $T \in \mathscr{L}(V, w)$. Then the adjoint of T is the function $T^* : W \to V$ such that

$$\langle Tv, w \rangle = \langle v, T^*w \rangle \tag{7.1}$$

To see why the definition males sense. Let $w \in W$. Consider the linear functional $v \mapsto \langle Tv, w \rangle$ on V. Then by Riesz representation theorem, there is a unique $u \in V$, such that

$$\langle Tv, w \rangle = \langle v, u \rangle$$

We let $u = T^*w$, would give us

$$\langle Tv, w \rangle = \langle v, T^*w \rangle$$

as desired.

We can also understand it as

$$\langle Tv, w \rangle = (Tv)^t \overline{w} = v^t T^t \overline{w} = v^t \overline{T^H w} = \langle v, T^*w \rangle .$$

The matrix form implies that the adjoint is a linear map too.

Theorem 7.1.1: Adjoint of a linear map is a linear map

If $T \in \mathscr{L}(V, W)$, then $T^* \in \mathscr{L}(W, V)$.

Proof.

$$\langle Tv, w_1 + w_2 \rangle = \langle v, T^* w_1 + T^* w_2 \rangle$$

and

$$\langle Tv, \lambda w \rangle = \overline{\lambda} \langle Tv, w \rangle = \langle v, \lambda T^* w \rangle$$

making T^* a linear map. $\square$

Proposition: **Properties of Adjoint**

Suppose $T \in \mathscr{L}(V, W)$, then

- $(S + T)^* = S^* + T^*$, for $\forall S \in \mathscr{L}(V, W)$.
- $(\lambda T)^* = \overline{\lambda} T^*$.
- $(T^*)^* = T$.
- $(ST)^* = T^* S^*$.
- $(T^{-1})^* = (T^*)^{-1}$.

Remark:
Keep in mind that adjoints are complex conjugate transpose.

Theorem 7.1.2: Null spaces and range of T^*

Suppose $T \in \mathscr{L}(V, W)$. Then

- $\operatorname{null} T^* = (\operatorname{range} T)^\perp$. And $\operatorname{null} T = (\operatorname{range} T^*)^\perp$.
- $\operatorname{range} T^* = (\operatorname{null} T)^\perp$. And $\operatorname{range} T = (\operatorname{null} T^*)^\perp$.

Proof.

$$w \in \operatorname{range} T^* \iff T^* w = 0 \iff \forall v \in V, \langle v, T^* w \rangle = \langle Tv, w \rangle = 0 \iff w \in (\operatorname{range} T)^\perp$$

$\square$

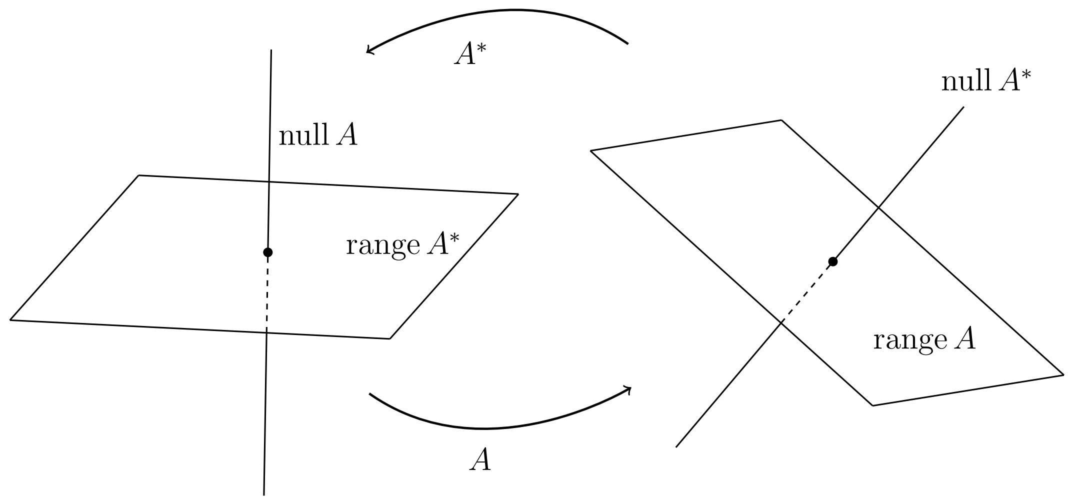

Figure 7.1: Geometric Interpretation of null and range of Adjoints

We can formally link adjoints to conjugate transposes.

Definition 7.1.2: Conjugate Transpose

Suppose $A \in \mathbb{F}^{m\times n}$, then the conjugate transpose $A^* \in \mathbb{F}^{n\times m}$ is defined by

$$\forall j,k,\, (A^*)_{j,k} = \overline{A_{k,j}}$$

Theorem 7.1.3: Adjoints and Conjugate Transpose

Let $T \in \mathscr{L}(V,W)$. Suppose $e_1,\dots,e_n$ is an orthonormal basis of V and $f_1,\dots,f_m$ is an orthonormal basis of W. Then $\mathscr{M}(T^*,(f_1,\dots,f_m),(e_1,\dots,e_n))$ is the conjugate transpose of $\mathscr{M}(T,(e_1,\dots,e_n),(f_1,\dots,f_m))$. That is,

$$\mathscr{M}(T^*) = \mathscr{M}^*(T) \tag{7.2}$$

Proof. Writing $Te_k = \langle Te_k, f_1\rangle f_1 + \ldots + \langle Te_k, f_m\rangle f_m$. We have $\mathscr{M}(T)_{j,k} = \langle Te_k, f_j\rangle$. Also $\mathscr{M}(T^*)_{j,k} = \langle T^* f_k, e_j\rangle = \overline{\langle e_j, T^* f_k\rangle} = \overline{\mathscr{M}(T)_{k,j}}$. $\square$

Remark:
The Riesz representation theorem provides an equivalent way of dealing with dual spaces. For any linear functional in $f \in V'$ corresponds to a vector $u \in V$ such that $\forall v \in V, f(v) = \langle v, w\rangle$. In this case, the orthogonal complement $U^\perp$ corresponds to the annihilator U^0. Also, the adjoint map T^* corresponds to the dual map T'.

Proposition: **Eigenvalues of Adjoints**

Suppose $T \in \mathscr{L}(V)$ and $\lambda \in \mathbb{F}$. then

$$\lambda \text{ is an eigenvalue of } T \Longleftrightarrow \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

Proof. If $Tv = \lambda v$, then $\forall u \in V$, we have

$$\left\langle v, T^*u - \overline{\lambda}u \right\rangle = \langle Tv, u \rangle - \lambda \langle v, u \rangle = 0$$

This implies range$(T^* - \overline{\lambda}I) \subseteq v^\perp$, Thus is not injective. □

Proposition: **Invariant Subspaces of Adjoints**

Suppose $T \in \mathscr{L}(V)$ and U is a subspace of V, then

$$U \text{ is invariant under } T \Leftrightarrow U^\perp \text{ is invariant under } T^*.$$

Proof. $\forall u \in U, Tu \in U$, then $\forall v \in U^\perp$, we have

$$\forall w \in U, \langle w, T^*v \rangle = \langle Tw, v \rangle = 0$$

Then $T^*v \in U^\perp$ as expected. □

7.1.2 Self-Adjoint Operators

Definition 7.1.3: Self-Adjoint

An operator $\mathcal{T} \in \mathscr{L}(V)$ is called self-adjoint if $T = T^*$.

Theorem 7.1.4: Eigenvalues of Self-adjoint Operators

Every eigenvalue of a self-adjoint operator is real.

Proof. If $Tv = \lambda v, v \neq 0$, then

$$\lambda \|v\|^2 = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \overline{\lambda}\|v\|^2$$

so λ is real. □

The following are some results relevant to self-adjoint operators.

Theorem 7.1.5

Suppose V is a complex inner product space and $T \in \mathscr{L}(V)$, then

$$\forall v \in V, \langle Tv, v \rangle = 0 \Longleftrightarrow T = 0$$

Proof. If $u, w \in V$, then we have

$$\langle Tu, w \rangle = \frac{1}{4}\left(\langle T(u+w), u+w \rangle - \langle T(u-w), u-w \rangle\right) + \frac{1}{4}\left(\langle T(u+iw), u+iw \rangle - \langle T(u-iw), u-iw \rangle\right) i$$

If $\forall v \in V, \langle Tv, v\rangle = 0$, then $\forall u, w \in V, \langle Tu, w\rangle = 0$ as above. Letting $w = Tu$ we have $Tu = 0, \forall u$, that is, $T = 0$. $\square$

Remark:
The above equality can be thought of

$$a\bar{b} = \frac{1}{4}\left(\|a+b\|^2 - \|a-b\|^2\right) + \frac{1}{4}\left(\|a+ib\|^2 - \|a-ib\|^2\right)i$$

Using i^2 we get a -1 to eliminate the $b\overline{a}$.
This result is false for real inner product spaces, as a mere $\frac{\pi}{2}$ rotation would break it.

Theorem 7.1.6

Suppose V is a complex inner product space and $T \in \mathscr{L}(V)$. Then

$$T \text{ is self-adjoint } \iff \forall v \in V, \langle Tv, v\rangle \in \mathbb{R}$$

Proof. If $v \in V$, then

$$\langle T^*v, v\rangle = \overline{\langle v, T^*v\rangle} = \overline{\langle Tv, v\rangle}$$

So we have

$$\begin{aligned} T \text{ is self-adjoint } &\iff T - T^* = 0 \\ &\iff \forall v \in V, \langle (T-T^*)v, v\rangle = 0 \\ &\iff \forall v \in V, \langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \\ &\iff \forall v \in V, \langle Tv, v\rangle \in \mathbb{R} \end{aligned}$$

$\square$

On a real inner product space, theorem 7.1.5 is true for self-adjoint operators.

Theorem 7.1.7

Suppose T is a self-adjoint operator on V. Then

$$\forall v \in V, \langle Tv, v\rangle = 0 \iff T = 0.$$

Proof. Using

$$\langle Tu, w\rangle = \frac{1}{4}\left(\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle\right)$$

We have what we desired like theorem 7.1.5. $\square$

Remark:
In this sense, the self-adjoint operator act like reals.

7.1.3 Normal Operators

Definition 7.1.4: Normal

In an inner product space V, $T \in \mathscr{L}(V)$ is normal if $TT^* = T^*T$.

Obviously, every self-adjoint operator is normal.

Theorem 7.1.8: Normality and Equal Norms

Suppose $T \in \mathscr{L}(V)$, then

$$T \text{ is normal } \iff \forall v \in V, \|Tv\| = \|T^*v\|$$

Proof. We have

$$\begin{aligned}
T \text{ is normal } &\iff T^*T - TT^* = 0 \\
(\text{for } T^*T - TT^* \text{ is self-adjoint}) &\iff \forall v \in V, \langle (T^*T - TT^*)v, v\rangle = 0 \\
&\iff \forall v \in V, \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle \\
&\iff \forall v \in V, \langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle \\
&\iff \forall v \in V, \|Tv\| = \|T^*v\|
\end{aligned}$$

□

Remark:
This gives a picture of the term "normal" as not changing the norm of a vector. Like rotation etc.

Proposition: **Properties of Normal Operators**

Suppose $T \in \mathscr{L}(V)$ is normal, then

- $\operatorname{null} T = \operatorname{null} T^*$ and $\operatorname{range} T = \operatorname{range} T^*$.
- $V = \operatorname{null} T \oplus \operatorname{range} T$.
- $\forall \lambda \in \mathbb{F}, T - \lambda I$ is normal.
- If $v \in V, \lambda \in \mathbb{F}$, then $Tv = \lambda v \iff T^*v = \overline{\lambda} v$.

Proof. 1. $v \in \operatorname{null} T \iff \|Tv\| = 0 \iff \|T^*v\| = 0 \iff v \in \operatorname{null} T^*$. And range = null$^\perp$.

2. $V = (\operatorname{null} T) \oplus (\operatorname{null} T)^\perp = \operatorname{null} T \oplus \operatorname{range} T^* = \operatorname{null} T \oplus \operatorname{range} T$.

3. Suppose $\lambda \in \mathbb{F}$, then

$$\begin{aligned}
(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \overline{\lambda} I) \\
&= TT^* - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\
&= T^*T - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\
&= (T - \lambda I)^*(T - \lambda I)
\end{aligned}$$

4. We have

$$\|(T - \lambda I)v\| = \|(T^* - \overline{\lambda} I)v\|.$$

Then $\|(T - \lambda I)v\| = 0 \iff \|(T^* - \overline{\lambda} I)v\| = 0$.

□

Theorem 7.1.9: Orthogonal Eigenvectors for Normal Operators

Suppose $T \in \mathscr{L}(V)$ is normal, then the eigenvectors corresponding to different eigenvalues are orthogonal.

Proof. Suppose α, β are different eigenvalues of T, with eigenvectors u, v. Then $Tu = \alpha u$ and $Tv = \beta v$. Also $T^* v = \overline{\beta} v$, thus

$$(\alpha - \beta)\langle u, v\rangle = \langle \alpha u, v\rangle - \left\langle u, \overline{\beta} v\right\rangle = \langle Tu, v\rangle - \langle u, T^* v\rangle = 0.$$

If $\alpha \neq \beta$ the equation implies $u \perp v$. □

Theorem 7.1.10: Normal and Conmutating Real and Imaginary Parts

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, then T is normal iff $\exists$ conmutating self-adjoint operators A and B such that $T = A + iB$.

Proof.

- Suppose T is normal, then let

$$A = \frac{T + T^*}{2} \quad \text{and} \quad B = \frac{T - T^*}{2i}$$

Then A and B are self-adjoint and

$$AB - BA = \frac{T^*T - TT^*}{2i} = 0$$

Let $T = A + iB$ then $T^* = A - iB$. $T^*T - TT^* = 0$ as desired. □

We can decompose a normal operator in another way as follows:

Definition 7.1.5: Skew

An operator $B \in \mathscr{L}(V)$ is called skew iff

$$B^* = -B$$

Theorem 7.1.11: Self-adjoint – Skew decomposition

Suppose $T \in \mathscr{L}(V)$. Then T is normal iff $\exists$ commuting operators A, B such that A is self-adjoint and B skew, and $T = A + B$.

Proof. Letting $A = (T + T^*)/2$ and $B = (T - T^*)/2$. □

7.2 Spectral Theorem

The nicest operators on V are those for which there is an orthonormal eigenvectors basis of V. Then for this basis T is diagonal. We shall characterize these operators as the self-adjoint operators when $\mathbb{F} = \mathbb{R}$ and normal operators when $\mathbb{F} = \mathbb{C}$. (The weaker assumption for $\mathbb{C}$ is intuitive for the generalization of the field, in $\mathbb{R}$ we must have eigenvalues in $\mathbb{R}$).

7.2.1 Real Spectral Theorem

First we need some preliminary results:

Proposition: **Invertible quadratic expressions**

Suppose $T \in \mathscr{L}(V)$ is self-adjoint and $b, c \in \mathbb{R}$ are such that $b^2 < 4c$, then $T^2 + bT + cI$ is an invertible operator.

Thinking about the quadratic expression with $\mathbb{R}$ coefficient, this implies $x^2 + bx + c > 0$.

Proof. Let v be a nonzero vector in V, then

$$\begin{aligned}
\left\langle (T^2 + bT + cI)v, v \right\rangle &= \left\langle T^2 v, v \right\rangle + b \left\langle Tv, v \right\rangle + c \left\langle v, v \right\rangle \\
&= \left\langle Tv, Tv \right\rangle + n \left\langle Tv, v \right\rangle + c\|v\|^2 \\
&\geq \|Tv\|^2 - |b| \, \|Tv\| \|v\| + c\|v\|^2 \\
&= \left(\|Tv\| - \frac{|b| \, \|v\|}{2} \right)^2 + \left(c - \frac{b^2}{4} \right) \|v\|^2 \\
&> 0.
\end{aligned}$$

This implies that $\forall v \neq 0, (T^2 + bT + cI)v \neq 0$, completing the proof. ☐

Lemma 7.2.1: Minimal Polynomial of self-adjoint operator

Suppose $T \in \mathscr{L}(V)$ is self-adjoint, then the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m) \in \mathbb{R}$.

Proof.

- Firts suppose $\mathbb{F} = \mathbb{C}$ then the minimal polynomial roots are the eigenvalues of T which are real.
- Now suppose $\mathbb{F} = \mathbb{R}$, by factorization of a polynomial over $\mathbb{R}$ we have

  $$p(T) = (T - \lambda_1) \cdots (T - \lambda_m)(T^2 + b_1 T + c_1 I) \cdots (T^2 + b_N T + c_N I) = 0$$

  where either m or N could be 0. If $N > 0$, for $T^2 + b_i T + c_i I$ is invertible, we have

  $$(T - \lambda_1) \cdots (T - \lambda_m) = 0$$

  which has less degree, contradicts.

☐

This is sufficient to say that a self-adjoint operator has a upper-triangular form.

Theorem 7.2.1: Real Spetral Theorem

Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathscr{L}(V)$, then the following are equivalent.

1. T is self-adjoint.
2. T has a diagonal matrix to some orthonormal basis of V.
3. V has an orthonormal eigenvector-of-T basis.

Proof.
- First if T is self-adjoint, then T has an orthonormal basis to which T has a upper-triangular form (recall Gram-Schmidt). However, $T^* = T$ so the matrix is diagonal. (Conjugate Transpose)
- The other part is obvious.

□

Remark:
We can understand the diagonalizability on an orthonormal basis as stretching on orthogonal directions. And being self-adjoint is linked to this idea.

7.2.2 Complex Spectral Theorem

Theorem 7.2.2: Complex Spectral Theorem

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$. Then the following are equivalent.

1. T is normal.
2. T has a diagonal matrix to some orthonormal basis of V.
3. V has an orthonormal eigenvector-of-T basis.

Proof.
- First suppose T is normal, then from Schur's theorem T is an orthonormal basis that makes it an upper-triangular matrix. Thus we can write

$$\mathscr{M}(T, (e_1, \ldots, e_n)) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix}$$

We will show that this matrix is a diagonal matrix. By

$$\|Te_1\|^2 = |a_{11}|^2$$

$$\|T^*e_1\|^2 = |a_{11}|^2 + \ldots + |a_{1n}|^2$$

And $\|Te_1\| = \|T^*e_1\|$ so we have $a_{12} = \ldots = a_{1n} = 0$. Thus, $\mathscr{M}$ being diagonal.

The other part is obvious. □

7.2.3 Other Properties

Theorem 7.2.3: Self-adjoint and Normal

A normal operator on $\mathbb{C}$ vector space is self-adjoint iff all its eigenvalues are real.

Proof. Diagonalizing it would do. □

Proposition: **Skew and Normal**

A normal operator on a $\mathbb{C}$ vector space is skew iff all its eigenvalues are pure imaginary.

Proposition: **Conditions for Normality**

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$.

- T is normal iff every eigenvector of T is also an eigenvector of T^*.
- T is normal iff there is a polynomial $p \in \mathscr{P}(\mathbb{C})$ such that $T^* = p(T)$.

Proposition: **Square and Cube Roots**

- Suppose $\mathbb{F} = \mathbb{C}$, every normal operator has a square root, i.e. $\exists S \in \mathscr{L}(V), S^2 = T$.
- For every $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, every self-adjoint operator has a cube root.

7.3 Positive Operators

Definition 7.3.1: Positive Operators

An operator $T \in \mathscr{L}(V)$ is called positive if T is self-adjoint and

$$\forall v \in V, \langle Tv, v \rangle \geq 0 \tag{7.3}$$

Note that if $\mathbb{F} = \mathbb{C}$, then the requirement of T being self-adjoint can be dropped due to theorem 7.1.6.

Theorem 7.3.1: Characterization of Positive Operators

Let $T \in \mathscr{L}(V)$. Then the following are equivalent.

1. T is a positive operator.
2. T is self-adjoint and all eigenvalues ≥ 0.
3. T has a positive square root.
4. T has a self-adjoint square root.
5. $T = R^*R$ for some $R \in \mathscr{L}(V)$.

Proof.

- Suppose 1 holds, so that T is positive, so T is self-adjoint, and T can be diagonalized to orthonormal basis. Let v be an eigenvector to λ, then

$$0 \leq \langle Tv, v \rangle = \lambda \langle v, v \rangle$$

making $\lambda \geq 0$. The the diagonal of the diagonalized matrix is all non-negative numbers.

- Followed by the last result, T has a positive, self-adjoint square root, that is, taking the square root of the diagonal entries of the diagonal matrix. Then $1 \to 2 \to 3 \to 4 \to 5$.

- If $T = R^*R$, then $T^* = T$ and $\langle Tv, v\rangle = \langle Rv, Rv\rangle \geq 0$.

$\square$

Corollary 7.3.1: Uniqueness of Positive Square Root

Every operator on V has a unique square root.

Proof. Let $e_1, \ldots, e_n$ be an eigenvector-of-R orthogonal basis of V. Then let $Re_k = \sqrt{\lambda_k} e_k$. Let

$$v = a_1 e_1 + \ldots + a_n e_n$$

be an eigenvector of T such that $Tv = \lambda v$. Thus,

$$Rv = a_1\sqrt{\lambda_1} e_1 + \ldots + a_n \sqrt{\lambda_n} e_n.$$

$$\lambda v = R^2 v = a_1 \lambda_1 e_1 + \ldots + a_n \lambda_n e_n.$$

Thus, $a_k(\lambda - \lambda_k) = 0$ for $k = 1, \ldots, n$. Hence,

$$v = \sum_{\{k : \lambda_k = \lambda\}} a_k e_k$$

$$Rv = \sum_{\{k : \lambda_k = \lambda\}} a_k \sqrt{\lambda} e_k = \sqrt{\lambda} v.$$

As T has an eigenvalue-basis, then R is uniquely determined. $\square$

Notation 7.3.1: $\sqrt{T}$

For a positive T, $\sqrt{T}$ denotes the unique square root of T.

The next result does not explicitly involve a square root.

Theorem 7.3.2

Suppose T is a positive operator on V, then

$$\langle Tv, v\rangle = 0 \to Tv = 0.$$

Proof. We have

$$0 = \langle Tv, v\rangle = \left\langle \sqrt{T}\sqrt{T} v, v\right\rangle = \left\langle \sqrt{T} v, \sqrt{T} v\right\rangle = \|\sqrt{T} v\|^2$$

So $\sqrt{T} v = 0$, then $Tv = 0$. $\square$

7.4 Isometries, Unitary Operators and Matrix Factorization

7.4.1 Isometries

Linear maps that proserve norm is an isometry. (This is just the one the is defined in topology, form the Greek word *isos metron*).

Definition 7.4.1: Isometry

A linear map $S \in \mathscr{L}(V, W)$ is called an isometry if

$$\forall v \in V, \|Sv\| = \|v\|$$

It is obvious that every isometry is injective for

$$\|Sv\| = 0 \to \|v\| = 0$$

Example: **Isometry**

An orthonormal basis to an orthonormal list is an isometry.

Theorem 7.4.1: Characterization for Isometries

Suppose $S \in \mathscr{L}(V, W)$, and $e_1, \ldots, e_n$ being an orthonormal basis for V and $f_1, \ldots, f_m$ is an orthonormal basis for W. Then the following are equivalent.

1. S is an isometry.
2. $S^*S = I$.
3. $\forall u, v \in V, \langle Su, Sv\rangle = \langle u, v\rangle$
4. $Se_1, \ldots, Se_n$ is an orthonormal list in W.
5. The columns of $\mathscr{M}(S, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ form an orthonormal list in $\mathbb{F}^m$ with Euclidean inner product.

Proof.

- First suppose S is an isometry, then $\forall v \in V$,

$$\langle (I - S^*S)v, v\rangle = \langle v, v\rangle - \langle Sv, Sv\rangle = 0.$$

Therefore, the self-adjoint $I - S^*S = 0$.

- If $S^*S = I$, then $\forall u, v \in V$,

$$\langle Su, Sv\rangle = \langle S^*Su, v\rangle = \langle u, v\rangle .$$

- $2 \to 3$ we have $\langle Se_i, Se_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}$. $4 \to 5 \to 1$ is obvious.

$\square$

7.4.2 Unitary Operators

A unitary operator is an isometry.

Definition 7.4.2: Unitary Operators

An operator $S \in \mathscr{L}(V)$ is called unitary if S is an invertible isometry.

Well the "invertible" is not necessary for every injective operator on finite dimensional vector space is invertible.

Theorem 7.4.2: Characterization of Unitary Operators

Suppose $S \in \mathscr{L}(V)$, and $e_1, \dots, e_n$ is an orthonormal basis of V. Then the following are equivalent:

1. S is a unitary operator.
2. $S^*S = SS^* = I$.
3. S is invertible and $S^{-1} = S^*$.
4. $Se_1, \dots, Se_n$ is an orthonormal basis of V.
5. The rows of $\mathscr{M}(S, (e_1, \dots, e_n))$ forms an orthonormal basis of $\mathbb{F}^n$ in Euclidean inner product.
6. S^* is a unitary operator.

Remark:
Analogous to complex numbers, with z corresponding to S and $\overline{z}$ corresponding to S^*. The real numbers $z = \overline{z}$ corresponds to self-adjoint operators, and the nonnegative numbers corresponds to positive operators. The elements on the unit circle $|z| = 1$ corresponds to unitary operators.

Theorem 7.4.3: Eigenvalues for Unitary Operators

Suppose λ is an eigenvalue for a unitary operator U, then $|\lambda| = 1$.

Proof.

$$|\lambda| \, \|v\| = \|\lambda v\| = \|Sv\| = \|v\|$$

□

Theorem 7.4.4: Characterization of Unitary Operators on $\mathbb{C}$ vector spaces

Suppose $\mathbb{F} = \mathbb{C}$ and $S \in \mathscr{L}(V)$. Then the following are equivalent:

1. S is a unitary operator.
2. There is an orthonormal basis of V consisting of eigenvectors of S whose eigenvalues all has $|\lambda| = 1$.

7.4.3 QR Factorization

We shift our attention to matrices.

Definition 7.4.3: Unitary Matrices

An n-by-n matrix is unitary iff its columns forms an orthonormal list in $\mathbb{F}^n$.

Remark:
As is shown in theorem 7.4.2, if $S \in \mathscr{L}(V)$ and $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ are orthonormal basis of V, then S is a unitary operator iff $\mathscr{M}(S, (e_1, \ldots, e_n), (f_1, \ldots, f_n))$ is a unitary matrix.

Theorem 7.4.5: QR Factorization

Suppose A is a square matrix with lineaely independent columns. Then $\exists$ unique matrices Q, R such that Q is unitary, R is upper-triangular (with only positive numbers on the diagonal), and $A = QR$.

Proof. The existence is by reducing the columns can be reduced to an orthonormal basis by Gram-Schmidt procedure, which creates the Q, R desired.

To show that Q and R are unique, let $A = \hat{Q}\hat{R}$, then we have $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(q_1, \ldots, q_k)$ and $\langle v_k, q_k \rangle > 0$ by induction on k, which means $v_k = q_k$ for all $k = 1, \ldots, n$, saying that $Q = \hat{Q}$ and thus $R = \hat{R}$. □

Remark:
This shows that a unitary matrix could be generated by Gram-Schmidt procedure.
If QR factorization is available, then it is easy to solve a corresponding system of linear equations without Gaussian elimination process. By $Ax = b$ we have

$$Rx = Q^*b$$

And R^{-1} is very easy to obtain for it is upper-triangular.

7.4.4 Cholesky Factorization

We now introduce positive invertible (positive definite) operators.

Theorem 7.4.6: Positive Invertible

A self-adjoint operator is positive invertible iff $\forall v \neq 0, \langle Tv, v\rangle > 0$.

Definition 7.4.4: Positive Definite

A matrix $B \in \mathbb{F}^{n,n}$ is called positive definite if $B^* = B$ and

$$\forall x \in \mathbb{F}^n, x \neq 0, \langle Bx, x\rangle > 0$$

Theorem 7.4.7: Cholesky Factorization

Suppose B is a positive definite matrix. Then there exists a unique upper-triangular matrix R with only positive numbers on the diagonal such that

$$B = R^*R$$

Proof. Because B is positive definite, then there is a invertible A such that $B = A^*A$, then let $A = QR$ be the QR factorization of A and then

$$B = A^*A = R^*Q^*QR = R^*R$$

as desired.

To prove the uniqueness, let $B = S^*S$ also, then S is invertible for B is invertible, then we have $BS^{-1} = S^*$, we shall prove $S = R$. We have

$$(AS^{-1})^*(AS^{-1}) = (S^*)^{-1}A^*AS^{-1} = (S^*)^{-1}AS^{-1} = I$$

Hence, AS^{-1} is unitary. As $A = (AS^{-1})S$ is a QR factorization then by the uniqueness $S = R$. (This is a way that "comparing" the difference of A and S, as we want to reverse the R^*Q^*QR process to get A). □

7.5 Singular Value Decomposition

7.5.1 Singular Values

Theorem 7.5.1: Properties of T^*T

Suppose $T \in \mathscr{L}(V, W)$. Then

- T^*T is a positive operator on V.
- $\operatorname{null} T^*T = \operatorname{null} T$.
- $\operatorname{range} T^*T = \operatorname{range} T^*$.
- $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

Proof.

- We have

$$(T^*T)^* = T^*T$$

So T^*T is self-adjoint. If $v \in V$, then

$$\langle (T^*T)v, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0$$

So T^*T is positive.

- Suppose $v \in \operatorname{null} T^*T$, then

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^*Tv, v\rangle = \langle 0, v\rangle = 0.$$

- As T^*T is self-adjoint, then

$$\operatorname{range} T^*T = (\operatorname{null} T^*T)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*.$$

- It's just the rank.

$\square$

Definition 7.5.1: Singular Values

Suppose $T \in \mathscr{L}(V, W)$. Let $\lambda_1, \dots, \lambda_m$ be eigenvalues of T^*T, then the singular values of T are $\sigma_i = \sqrt{\lambda_i}$, as $\lambda_i \geq 0$. The σ_i are listed in decreasing order, including many times as the dimension of $E(\lambda_i, T^*T)$.

Theorem 7.5.2: Roles of Sigular Values

Suppose $T \in \mathscr{L}(V, W)$, then

- T is injective $\Leftrightarrow$ 0 is not a singular value of T.
- The number of positive singular values of T is $\dim \operatorname{range} T$.
- T is surjective $\Leftrightarrow$ number of positive singular values of T is $\dim W$.

Proof.

- We have

$$T \text{ is injective } \Leftrightarrow \operatorname{null} T = \operatorname{null} T^*T = \{0\} \Leftrightarrow 0 \text{ is not a eigenvalue of } T^*T.$$

- According to the spectral theorem for T^*T, we have $\dim \operatorname{range} T^*T$ equals the number of positive eigenvalues (counting repetitions). The equality of repetitions of eigenvalues and eigenspaces is due to diagonalizability.

$\square$

Remark:
Note that when an operator is diagonalizable, the algebraic multiplicity and geometric multiplicity of every eigenvalue is the same (this is straightly from the fact that there is an eigenvector basis).

List of Eigenvalues	List of Singular Values
On vector spaces	On inner product spaces
Defined only on $\mathscr{L}(V)$	Defined on $\mathscr{L}(V, W)$
includes $0 \Leftrightarrow$ operator is not invertible	includes $0 \Leftrightarrow$ linear map is not injective
No standard order	Always listed in decreasing order

Theorem 7.5.3: Isometries and Singular Values

Suppose $S \in \mathscr{L}(V, W)$. Then

$$S \text{ is an isometry } \Leftrightarrow \text{ all sigular values of } S \text{ equals } 1\ .$$

Proof. We have

$$\begin{aligned} S \text{ is an isometry } &\Leftrightarrow S^*S = I \\ &\Leftrightarrow \text{ all eigenvalues of } S^*S \text{ equals } 1 \\ &\Leftrightarrow \text{ all singular values of } S \text{ equals } 1\ . \end{aligned}$$

□

7.5.2 SVD for Linear Maps and for Matrices

The SVD Decomposition presents a remarkable result that shows every linear map form V to W can be represented as orthonormal basis and singular value.

Theorem 7.5.4: Singular Value Decomposition

Suppose $T \in \mathscr{L}(V, W)$ and the positive singular values of T are $s_1, \ldots, s_m$. Then there exist orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

Proof. Let $s_1, \ldots, s_n$ denote the singular values of T, $(\dim V = n\ )$. For T^*T is a positive operator, then the spectral theorem guarantee that there is an orthonormal basis $e_1, \ldots, e_n$ of V with

$$T^*Te_k = s_k^2 e_k$$

for $k = 1, \ldots, n$. For each $k = 1, \ldots, m$ let

$$f_k = \frac{Te_k}{s_k}$$

We shall show that f_i is orthonormal.

$$\langle f_j, f_k \rangle = \frac{1}{s_j s_k} \langle Te_j, Te_k \rangle = \frac{1}{s_j s_k} \langle e_j, T^*Te_k \rangle = \frac{s_k}{s_j} \langle e_j, e_k \rangle = \delta_{jk}$$

For those $k \in \{1, \ldots, n\}, k > m$, we have $s_k = 0$ then $T^*Te_k = 0$, by 7.3.2 we have $Te_k = 0$. Then

$$\begin{aligned} Tv &= T\left(\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n\right) \\ &= s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m \end{aligned}$$

□

Remark:
The proof here gives us new insights of what the operator T^*T means: the eigenvectors of T^*T forms a basis of $(\operatorname{null} T)^\perp$. The SVD can be understood as every operator can be seen as a compression of coordinates to specific orthonormal basis.
To compute SVD, remember that:

- e_i is the orthonormal diagonal basis of T^*T.
- $f_i = Te_i/s_i$.

Theorem 7.5.5: SVD of Adjoints and Psudoinverse

Suppose $T \in \mathscr{L}(V,W)$ and the positive singular values of T are $s_1, \dots, s_m$. Suppose $e_1, \dots, e_m$ and $f_1, \dots, f_m$ are orthonormal lists in V and W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle e, v_m \rangle f_m$$

Then

$$T^*w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m$$

and

$$T^\dagger w = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} e_m$$

for $\forall w \in W$.

Proof.

- If $v \in V, w \in W$ then

$$\begin{aligned}\langle Tv, w \rangle &= s_1 \langle v_1, e_1 \rangle \langle f_1, w \rangle + \cdots + s_m \langle v, e_m \rangle \langle f_m, w \rangle \\ &= \langle v, s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m \rangle\end{aligned}$$

- Suppose $w \in W$ and

$$v = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} e_m$$

$$\begin{aligned}Tv &= \langle w, f_1 \rangle f_1 + \cdots + \langle w, f_m \rangle f_m \\ &= P_{\operatorname{range} T} w\end{aligned}$$

As $e_k \in \operatorname{range} T^* = (\operatorname{null} T)^\perp$, we know $v = T^\dagger w$.

□

Theorem 7.5.6: Matrix Version of SVD

Suppose $A \in \mathbb{F}^{m \times n}$ has rank $r \geq 1$, then there exists a $B \in \mathbb{F}^{m \times r}$ with orthonormal columns, $D \in \mathbb{F}^{r \times r}$ being diagonal, and $C \in \mathbb{F}^{n \times m}$ with orthonormal columns such that

$$A = BDC^* \tag{7.4}$$

Or more generally, there exists orthonormal U, V and diagonal D such that

$$A = UDV^* \tag{7.5}$$

Theorem 7.5.7: Singular Values for Adjoints

Suppose $T \in \mathscr{L}(V, W)$, then T^* and T have the same positive singular values

7.6 Consequences of SVD

7.6.1 Norms of Linear Maps

Theorem 7.6.1: Upper Bounds for $\|Tv\|$

Suppose $T \in \mathscr{L}(V, W)$. Let s_1 be the largest singular value of T, then

$$\forall v \in V, \|Tv\| \leq s_1 \|v\|$$

Proof.

$$\begin{aligned}
\|Tv\|^2 &= s_1^2 \left|\langle v, e_1\rangle\right|^2 + \cdots + s_m^2 \left|\langle v, e_m\rangle\right|^2 \\
&\leq s_1^2 \left(\left|\langle v, e_1\rangle\right|^2 + \cdots + \left|\langle v, e_m\rangle\right|^2\right) \\
&\leq s_1^2 \|v\|^2
\end{aligned}$$

□

If $\|v\| \leq 1$, then we have $\|Tv\| \leq s_1$, and $\|Te_1\| = s_1$ so

$$\max\left\{\|Tv\| : v \in V \wedge \|v\| \leq 1\right\} = s_1$$

Definition 7.6.1: Norm of a linear map

Suppose $T \in \mathscr{L}(V, W)$. Then the norm of T denoted by $\|T\|$, is defined as

$$\|T\| = \max\left\{\|Tv\| : v \in V \wedge \|v\| < 1\right\}$$

Remark:
The norm of a linear map represents the largest stretching of the linear transform, the longest semi-axis of the ellipsoid from a unit sphere.

We shall see that the norm of a linear transform is indeed a norm on $\mathbb{F}^{m \times n}$.

Theorem 7.6.2: Basic Properties of norms of Linear Maps

Suppose $T \in \mathscr{L}(V, W)$. Then

- $\|T\| \geq 0$
- $\|T\| = 0 \Leftrightarrow T = 0$
- $\forall \lambda \in \mathbb{F}, \|\lambda T\| = |\lambda| \, \|T\|$
- $\forall S \in \mathscr{L}(V, W), \|S + T\| \leq \|S\| + \|T\|$

Proof. The last one, using $\exists v \in V, \|S+T\| = \|(S+T)v\|$, so

$$\|S+T\| = \|(S+T)v\| \leq \|Sv\| + \|Tv\| \leq \|S\| + \|T\|$$

□

We usually call $\|S-T\|$ be the distance between S and T. And $\|S-T\|$ is a small number means that S and T are close together. We will show that $\forall T \in \mathscr{L}(V)$ there is an invertible operator as closed to T as we wish.

Theorem 7.6.3: Alternative formulas for $\|T\|$

Suppose $T \in \mathscr{L}(V, W)$, then

1. $\|T\|$ is the largest singular value of T.
2. $\|T\| = \max\{\|Tv\| : v \in V \wedge \|v\| = 1\}$.
3. $\|T\|$ is the smallest number c such that $\forall v \in V, \|Tv\| \leq c\|v\|$.

Theorem 7.6.4: Norm of the Adjoint

Suppose $T \in \mathscr{L}(V, W)$, then $\|T\| = \|T^*\|$.

Proof. Suppose $w \in W$, then

$$\|T^*w\|^2 = \langle T^*w, T^*w\rangle = \langle TT^*w, w\rangle \leq \|TT^*w\|\|w\| \leq \|T\|\|T^*w\|\|w\|$$

thus $\|T^*w\| \leq \|T\|\|w\|$, which means $\|T^*\| \leq \|T\|$.

Also, this is quite clear from theorem 7.5.7. □

7.6.2 Approximation of Linear Transforms

Theorem 7.6.5: Best Approximation for Linear Maps

Suppose $T \in \mathscr{L}(V, W)$ and $s_1 \geq \cdots \geq s_m$ are the positive singular values of T. Suppose $1 \leq k \leq m$. Then

$$\min\{\|T-S\| : S \in \mathscr{L}(V, W) \wedge \dim \operatorname{range} S \leq k\} = s_{k+1}$$

Furthermore, if

$$Tv = s_1 \langle v, e_1\rangle f_1 + \cdots + s_m \langle v, e_m\rangle f_m$$

is a singular value decomposition of T and $T_k \in \mathscr{L}(V, W)$ is defined by

$$T_k v = s_1 \langle v, e_1\rangle f_1 + \cdots + s_k \langle v, e_k\rangle f_k$$

then $\dim \operatorname{range} T_k = k$ and $\|T - T_k\| = s_{k+1}$.

Proof. If $v \in V$ then

$$\begin{aligned}\|(T - T_k)v\|^2 &= \|s_{k+1}\langle v, e_{k+1}\rangle f_{k+1} + \cdots + s_m \langle v, e_m\rangle f_m\|^2 \\ &\leq s_{k+1}^2 \left(|\langle v, e_{k+1}\rangle|^2 + \cdots + s_m^2 |\langle v, e_m\rangle|^2\right) \\ &\leq s_{k+1}^2 \|v\|^2\end{aligned}$$

Thus $\|T - T_k\| \leq s_{k+1}$, and $(T - T_k)e_{k+1} = s_{k+1}f_{k+1}$, so $\|T - T_k\| = s_{k+1}$.

Now we consider an $S \in \mathscr{L}(V, W)$ that $\dim \operatorname{range} S \leq k$, thus $Se_1, \ldots, Se_{k+1}$ is linear dependent. Hence, there exists $a_1, \ldots, a_{k+1} \in \mathbb{F}$ not all 0 such that

$$a_1 Se_1 + \ldots + a_{k+1}Se_{k+1} = 0$$

So $a_1e_1 + \ldots + a_{k+1}e_{k+1} \neq 0$. We have

$$\|(T - S)(a_1e_1 + \ldots + a_{k+1}e_{k+1}\|^2 = \|T(a_1e_1 + \ldots + a_{k+1}e_{k+1}\|^2 \geq s_{k+1}^2\|a_1e_1 + \ldots + a_{k+1}e_{k+1}\|^2$$

Thus $\|T - S\| \geq s_{k+1}$. □

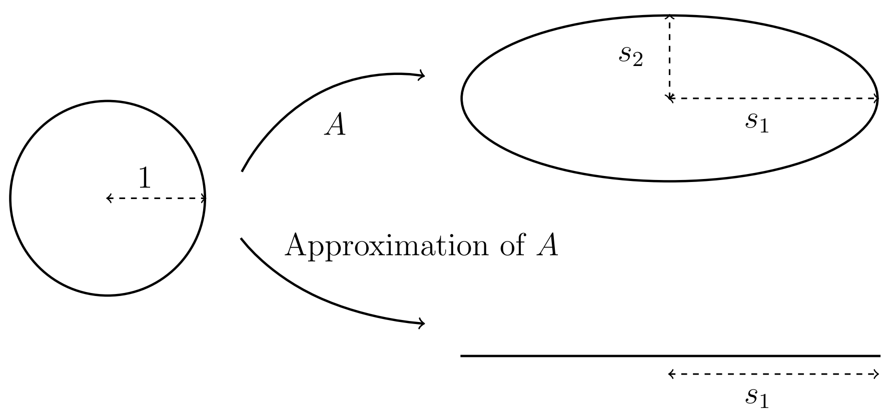

Figure 7.2: The Approximation of Linear Maps

7.6.3 Polar Decomposition

Using our analogy of unitary operators and complex numbers. As every $z \in \mathbb{C}$ can be represented as

$$z = \left(\frac{z}{|z|}\right)|z| = \left(\frac{z}{|z|}\right)\sqrt{\bar{z}z}$$

which is a polar decomposition, with $|z/|z|| = 1$. This lead us to guess that every operator $T \in \mathscr{L}(V)$ can be written as a unitary operator times $\sqrt{T^*T}$. Therefore, we turn any operator with a unitary operator and a positive operator, both we know extremely well.

Specifically, if $\mathbb{F} = \mathbb{C}$ and $T = S\sqrt{T^*T}$. Then there is an orthonormal basis such that S has a diagonal matrix, and there is (another) orthonormal basis such that $\sqrt{T^*T}$ has a diagonal matrix. (There may not be an orthonormal basis that simultaneously do both.)

However, we have

$$\text{There is an orthonormal basis such that } S \text{ and } \sqrt{T^*T} \text{ are both diagonal } \Leftrightarrow T \text{ is normal .}$$

Theorem 7.6.6: Polar Decomposition

Suppose $T \in \mathscr{L}(V)$. Then there exists z unitary operator such that

$$T = S\sqrt{T^*T}$$

This is true for both $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Proof. Let $s_1, \ldots, s_m$ be the positive singular values of T. And let $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ be such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m.$$

Extend then to an orthonormal basis $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$. Define $S \in \mathscr{L}(V)$ by

$$Sv = \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

Then we have

$$\begin{aligned} \|Sv\|^2 &= \| \langle v, e_1 \rangle f_1 + \cdots + \langle v, e_n \rangle f_n\|^2 \\ &= |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 \\ &= \|v\|^2 \end{aligned}$$

Thus S is a unitary operator.

Also by

$$T^*Tv = s_1^2 \langle v, e_1 \rangle e_1 + \cdots + s_m^2 \langle v, e_m \rangle e_m$$

therefore

$$\sqrt{T^*T}v = s_1 \langle v, e_1 \rangle e_1 + \cdots + s_m \langle v, e_m \rangle e_m$$
$$S\sqrt{T^*T}v = Tv$$

As we can verify. ☐

Remark:
The positive operator can be seen as a stretching in some direction, and the unitary operator is a rotation.

7.6.4 Operators Applied to Ellipsoids and Parallelepipeds

In this part we shall give a geometric interpretation of operators. We define a ball in V :

Definition 7.6.2: Ball

The ball in V of radius 1 centered at O, denoted by B, is defined

$$B = \{v \in V : \|v\| < 1\} \tag{7.6}$$

We can view an ellipsoid as a ball that is stretched on some axis.

Definition 7.6.3: Ellipsoid

Suppose that $f_1, \ldots, f_n$ is an orthonormal basis of V and $s_1, \ldots, s_n$ are positive numbers. The ellipsoid $E(s_1 f_1, \ldots, s_n f_n)$ with principle axis $s_1 f_1, \ldots, s_n f_n$ is defined by

$$E(s_1 f_1, \ldots, s_n f_n) = \left\{ v \in V : \frac{|\langle v, f_1 \rangle|^2}{s_1^2} + \cdots + \frac{|\langle v, f_n \rangle|^2}{s_n^2} < 1 \right\} \tag{7.7}$$

The definition here has a very intuitive geometric meaning. The ellipsoid $E(f_1, \ldots, f_n)$ equals the unit ball for all orthonormal basis $f_1, \ldots, f_n$. Also, any invertible operator in $\mathscr{L}(V)$ maps a ball into an ellipsoid with principle axes from the SVD of the operator

Theorem 7.6.7: Invertible Operators takes Ball to Ellipsoid

Suppose $T \in \mathscr{L}(V)$ is invertible, and unit ball B in V. Then $T(B)$ is an ellipsoid in V.

Proof. Suppose T has the singular value decomposition:

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

We can show that $T(B) = E(s_1 f_1, \ldots, s_n f_n)$. □

Corollary 7.6.1: Invertible Operators takes Ellipsoids to Ellipsoids

Suppose $T \in \mathscr{L}(V)$ is invertible, and $\forall$ ellipsoid E in V. Then $T(E)$ is an ellipsoid in V.

Next we consider parallelepipeds: the generalization of parallelogram in $\mathbb{R}^2$ to $\mathbb{F}^n$.

Definition 7.6.4: Parallelepipeds

Suppose $v_1, \ldots, v_n$ is a basis of V. Let

$$P(v_1, \ldots, v_n) = \{a_1 v_1 + \cdots + a_n v_n : a_i \in (0, 1)\}.$$

A parallelepiped is a set of the form $u + P(v_1, \ldots, v_n$ for some $u \in V$. The vectors $v_1, \ldots, v_n$ are called the edges of the parallelepiped.

Theorem 7.6.8: Invertible Operators takes Parallelepipeds to Parallelepipeds

Suppose $u \in V$ and $v_1, \ldots, v_n$ is a basis of V, $T \in \mathscr{L}(V)$ is invertible, then

$$T(u + P(v_1, \ldots, v_n) = Tu + P(Tv_1, \ldots, Tv_n)$$

Just like rectangular, we defined box to be parallelepipeds with orthogonal edges.

Definition 7.6.5: Box

A box in V is the set of the form

$$u + P(r_1 e_1, \ldots, r_n e_n)$$

where $u \in V$ and $r_1, \ldots, r_n$ are positive numbers and $e_1, \ldots, e_n$ are orthonormal basis.

Just like the ellipsoid, every invertible operator takes a box along the principle axis to another box.

Theorem 7.6.9: Invertible Operators takes some boxes to boxes

Suppose $T \in \mathscr{L}(V)$ is invertible and has the singular value decomposition:

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

In this case, T maps $u + P(r_1 e_1, \ldots, r_n e_n)$ to $Tu + P(r_1 s_1 e_1, \ldots, r_n s_n e_n)$.

7.6.5 Volume and Singular Values

We use an intuitive way to define the volume.

Definition 7.6.6: Volume of a Box

Suppose $\mathbb{F} = \mathbb{R}$. If $u \in V$ and $r_1, \ldots, r_n \in \mathbb{R}_+$ and $e_1, \ldots, e_n$ is an orthonormal basis, then

$$\operatorname{volume}(u + P(r_1 e_1, \ldots, r_n e_n)) = r_1 r_2 \cdots r_n. \tag{7.8}$$

The volume of an arbitrary subset $\Omega \subseteq V$ is defined in analysis by integral.

Theorem 7.6.10: Volume Change by Invertibles

Suppose $\mathbb{F} = \mathbb{R}$, $T \in \mathscr{L}(V)$ in invertible, and $\Omega \subseteq V$. Then

$$\operatorname{volume} T(\Omega) = s_1 \cdots s_n \operatorname{volume} \Omega$$

where $s_1, \ldots, s_n$ are singular values of T.

Remark:
This implies that $s_1 \cdots s_n = |\det T|$, as we will see later.

7.7 Summary

Properties of a Normal Operator	**Eigenvalues Range**
Inveritible	$\mathbb{C}\setminus\{0\}$
Self-adjoint	$\mathbb{R}$
Skew	$\{\lambda \in \mathbb{C} : \Re\lambda = 0\}$
Orthogonal Projection	$\{0, 1\}$
Positive	$\mathbb{R}_{\geq 0}$
Unitary	$\{\lambda \in \mathbb{C} : \lvert\lambda\rvert = 1\}$

Chapter 8

Operators on Complex Vector Spaces

Standing Assumption

- $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.
- V denotes finite dimensional vector spaces.

8.1 Generalized Eigenvectors and Nilpotent Operators

8.1.1 Null Spaces of T^k

Proposition: **Sequences of Increasing Null Spaces**

Suppose $T \in \mathscr{L}(V)$, then

$$\{0\} = \operatorname{null} T^0 \subseteq \operatorname{null} T^1 \subseteq \cdots \subseteq \operatorname{null} T^k \subseteq \cdots$$

Proof. If $v \in \operatorname{null} T^k$, then $T^k v = 0$, then $T^{k+1} v = 0$, then $v \in \operatorname{null} T^{k+1}$. □

Proposition: **Equality in the Sequence of Null spaces**

Suppose $T \in \mathscr{L}(V)$ and $m \in \mathbb{N}$ such that

$$\operatorname{null} T^m = \operatorname{null} T^{m+1}$$

Then $\forall k \in \mathbb{N}$, we have

$$\operatorname{null} T^{m+k} = \operatorname{null} T^m$$

Proof. We want to prove $\operatorname{null} T^{m+k} = \operatorname{null} T^{m+k+1}$. We already know that $\operatorname{null} T^{m+k} \subseteq \operatorname{null} T^{m+k+1}$. Suppose $v \in \operatorname{null} T^{m+k+1}$, then

$$T^{m+1}(T^k v) = 0$$

then $T^k v \in \operatorname{null} T^{m+1} = \operatorname{null} T^m$, so $v \in \operatorname{null} T^{m+k}$. □

This result shows that the sequence of null spaces will tend to stable. A direct result is that the null spaces can grow only $\dim V$ times, which increase 1 dimension every step.

Proposition: **Null spaces Stop growing**

Suppose $T \in \mathscr{L}(V)$, then

$$\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1} = \cdots$$

It is not always true that $V = \operatorname{null} T \oplus \operatorname{range} T$ for $\forall T \in \mathscr{L}(V)$, however, the following proposition is a useful substitute.

Theorem 8.1.1: $V = \operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V}$

Suppose $T \in \mathscr{L}(V)$, then

$$V = \operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V}$$

Proof. We already have $\dim V = \dim \operatorname{null} T^{\dim V} + \dim \operatorname{range} T^{\dim V}$. So we only need to prove

$$(\operatorname{null} T^n) \cap (\operatorname{range} T^n) = \{0\}$$

Where $n = \dim V$. Suppose $v \in (\operatorname{null} T^n) \cap (\operatorname{range} T^n)$, then $T^n v = 0$. Also $\exists u \in V, v = T^n u$, so $T^{2n} u = 0$, thus $u \in \operatorname{null} T^{2n} = \operatorname{null} T^n$. Therefore $v = T^n u = 0$. $\square$

Accordingly, we have a similar result for range.

Proposition: **The range of powers**

Suppose $T \in \mathscr{L}(V)$, then

- The range sequence:
$$V = \operatorname{range} T^0 \supseteq \operatorname{range} T^1 \supseteq \cdots$$
- If $m \in \mathbb{N}, \operatorname{range} T^m = \operatorname{range} T^{m+1}$, then $\forall k \in \mathbb{N}$, we have
$$\operatorname{range} T^{m+k} = \operatorname{range} T^m$$
- $\operatorname{range} T^{\dim V} = \operatorname{range} T^{\dim V+1} = \cdots$.

The proof is similar.

Proposition: **Null Range Relation of Powers**

Suppose $T \in \mathscr{L}(V)$, and $n \in \mathbb{N}$, then

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} \Leftrightarrow \operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Proof. Using the same dimension would do. $\square$

8.1.2 Generalized Eigenvectors

Some operator do not have enough eigenvectors to fully describe its behavior. When we try to decompose V into invariant subspaces

$$V = V_1 \oplus \cdots \oplus V_m$$

The simplest approach is to decompose into one-dimensional spaces, which can only be done when there is an eigenvector basis. That is,

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T).$$

where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T. (Some of $E(\lambda_k, T)$ may have larger dimension.)

Definition 8.1.1: Generalized Eigenvector

Suppose $T \in \mathscr{L}(V)$, and λ is an eigenvalue of T. A vector $v \in V$ is called a generalized eigenvector of T corresponding to λ if $v \neq 0$ and

$$(T - \lambda I)^k v = 0$$

for some $k \in \mathbb{Z}_+$.

Well we have v is a generalized eigenvector of T iff $v \neq 0$ and

$$(T - \lambda I)^{\dim V} = 0$$

Theorem 8.1.2: A basis of Generalized Eigenvectors

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.

Proof. Let $n = \dim V$, and use induction on n. First the result holds for $n = 1$. Suppose $n > 1$ and the result holds for all smaller n. Let λ be an eigenvalue of T (using the hypothesis $\mathbb{F} = \mathbb{C}$), then as

$$V = \operatorname{null}(T - \lambda I)^n \oplus \operatorname{range}(T - \lambda I)^n$$

If $\operatorname{null}(T - \lambda I)^n = V$, then every nonzero $v \in V$ is a generalized eigenvector of T.

If $\operatorname{null}(T - \lambda I)^n \neq V$, then $\operatorname{range}(T - \lambda I)^n \neq \{0\}$, and $\operatorname{null}(T - \lambda I)^n \neq \{0\}$ for λ is an eigenvalue. Thus

$$0 < \dim \operatorname{range}(T - \lambda I)^n < n$$

As $\operatorname{range}(T - \lambda I)^n$ is invariant under T. We let $S = \mathscr{L}\left(\operatorname{range}(T - \lambda I)^n\right)$ be the restriction of T onto $\operatorname{range}(T - \lambda I)^n$. Then by induction hypothesis to S, we have there is a generalized eigenvector basis for S (also T) of $\operatorname{range}(T - \lambda I)^n$. Joining the basis of $\operatorname{range}(T - \lambda I)^n$ and a basis of $\operatorname{null}(T - \lambda I)^n$. □

Remark:
This result does not hold for $\mathbb{F} = \mathbb{R}$, as there may not be any eigenvalue initially.

Theorem 8.1.3: A Basis of Generalized Eigenvectors, Revised

Suppose $T \in \mathscr{L}(V)$, then there is a basis of V consisting of generalized eigenvectors of T iff the minimal polynomial of T equals $(z-\lambda_1)\cdots(z-\lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$. Which is equivalent to T has an upper-triangular form.

Proof. The proof of the "if" part is the same as above, for T has an eigenvalue. □

Theorem 8.1.4: Generalized Eigenvector Corresponds to one Eigenvalue

Suppose $T \in \mathscr{L}(V)$, theneach generalized eigenvector of T Corresponds to only one eigenvalue of T.

Proof. Suppose $v \in V$ is a generalized eigenvector corresponds to λ and α of T. Let m be the smallest number such that $(T-\alpha I)^m v = 0$, and $n = \dim V$. Then

$$\begin{aligned} 0 &= (T-\lambda I)^n v \\ &= \left((T-\alpha I) + (\alpha-\lambda)I\right)^n v \\ &= \sum_{k=0}^{n} b_k(\alpha-\lambda)^{n-k}(T-\alpha I)^k v \end{aligned}$$

Where $b_0 = 1$. Applying $(T-\alpha I)^{m-1}$ to both sides we have

$$0 = (\alpha-\lambda)^n (T-\alpha I)^{m-1} v$$

Thus $\alpha = \lambda$. □

Theorem 8.1.5: Linear Independent Generalized Eigenvectors

Suppose $T \in \mathscr{L}(V)$, then every list of generalized eigenvectors corresponding to distinct eigenvalues are linear independent.

Proof. Similar to the way we prove eigenvectors are linear independent. Let m be the smaller number that $v_1, \ldots, v_m$ corresponding to distinct $\lambda_1, \ldots, \lambda_m$ are linear dependent. We have $m \geq 2$ and $a_1, \ldots, a_m \in \mathbb{F}$ such that

$$a_1 v_1 + \ldots + a_m v_m = 0.$$

None of $a_i = 0$ for the minimality of m. Applying $(T-\lambda_m I)^n$ to both sides we have

$$a_1(T-\lambda_m)^n v_1 + \cdots + a_{m-1}(T-\lambda_m I)^n v_{m-1} = 0$$

As $(T-\lambda_m I)^n v_k \neq 0$ according to theorem 8.1.4, then $(T-\lambda_m)^n v_k$ is a smaller linear independent list, contradicts. □

8.1.3 Nilpotent Operators

Definition 8.1.2: Nilpotent Operators

An operator $T \in \mathscr{L}(V)$ is a nilpotent operator if $\exists k \in \mathbb{N}, T^k = 0$.

Thus, an operator is nilpotent iff every nonzero vector in V is a generalized eigenvector of T corresponding to the eigenvalue 0.

Proposition: **An upper limit of the power of a nilpotent operator**

Suppose $T \in \mathscr{L}(V)$ is nilpotent, then $T^{\dim V} = 0$.

Proof. If $\exists k \in \mathbb{Z}_+, T^k = 0$, then $\operatorname{null} T^k = V$, then $T^{\dim V} = V$, then $T^{\dim V} = 0$. □

As all generalized eigenvector corresponds to eigenvalue 0, then we say that there are no eigenvalue other than 0 for a nilpotent operator.

Theorem 8.1.6: Eigenvalues of Nilpotent Operators

Suppose $T \in \mathscr{L}(V)$, then

- If T is nilpotent, then 0 is the only eigenvalue of T.
- If $\mathbb{F} = \mathbb{C}$ and 0 is the only eigenvalue of T, then T is nilpotent.

Proof.
- As $T^m = 0$ for some $m \in \mathbb{Z}_+$, then T is not injective, thus 0 is an eigenvalue of T. If $Tv = \lambda v$, then $T^m v = \lambda^m v = 0$, thus $\lambda = 0$.
- Suppose $\mathbb{F} = \mathbb{C}$ and 0 is the only eigenvalue of T, then the minimal polynomial of T equals z^m, thus $T^m = 0$ and T is nilpotent.

□

Remark:
We already know a class of nilpotent operators has the matrix form

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

We shall prove that all nilpotent operators has this form.

Theorem 8.1.7: Minimal Polynomial and Matrix of Nilpotent Operators

Suppose $T \in \mathscr{L}(V)$, then the following are equivalent.

- T is nilpotent.
- The minimal polynomial of T is z^m for some $m \geq 0$.
- There is a basis of V such that the matrix of T is

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

Proof.
- Suppose T is nilpotent, then $T^n = 0$ for some $n \in \mathbb{Z}_+$, then the minimal polynomial is a factor of z^n.

- By the minimal polynomial, we have a upper-triangular form, also the diagonal is the roots of the minimal polynomial, then we get the result.

□

8.2 Generalized Eigenspace Decomposition

8.2.1 Generalized Eigenspaces

Definition 8.2.1: Generalized Eigenspaces

Suppose $T \in \mathscr{L}(V)$, and $\lambda \in \mathbb{F}$, then the generalized eigenspace of T corresponding to λ is defined

$$G(\lambda, T) = \left\{v \in V : \exists k \in \mathbb{Z}_+, (T - \lambda I)^k v = 0\right\} \tag{8.1}$$

Or equivalently,

$$G(\lambda, T) = \operatorname{null}(T - \lambda I)^{\dim V}.$$

It is easy to see that $E(\lambda, T) \subseteq G(\lambda, T)$.

Theorem 8.2.1: Generalized Eigenspace Decomposition

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T. Then

1. (G, λ_k, T) is invariant under T.
2. $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent.
3. $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$.

Proof. 1. As $G(\lambda_k, T) = \operatorname{null}(T - \lambda_k I)^{\dim V}$.

2. Obvious.

3. As generalized eigenvectors corresponding to different eigenvalues are linear independent, the $\bigoplus G(\lambda_k, T)$ is a direst sum. Also there is a generalized eigenvector basis.

□

8.2.2 Multiplicity of an Eigenvalue

Definition 8.2.2: Multiplicity

Suppose $T \in \mathscr{L}(V)$, the multiplicity of an eigenvalue λ of T is defined to be $\dim G(\lambda, T)$.

We have a direct result:

Proposition: **Sum of Multiplicity**

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, then the sum of multiplicity of all eigenvalues is $\dim V$,

Proof. This follows directly from the generalized eigenvector decomposition. □

Remark:
Algebraic Multiplicity and Geometric Multiplicity
Traditional definition of algebraic multiplicity involves the use of determinants. But now we see that it also has a geometric meaning.

- **Algebraic Multiplicity**: $\dim G(\lambda, T) = \dim \operatorname{null}(T - \lambda I)^{\dim V}$.
- **Geometric Multiplicity**: $\dim E(\lambda, T) = \dim \operatorname{null}(T - \lambda I)$.

If V is an inner product space, $T \in \mathscr{L}(V)$ is normal, and λ is an eigenvalue of T, then the algebraic multiplicity equals the geometric multiplicity. (As $T - \lambda I$ is normal, write it as a diagonal matrix).

Definition 8.2.3: Characteristic Polynomial

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues with multiplicities $d_1, \ldots, d_m$, then we define characteristic polynomial

$$(z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m} \tag{8.2}$$

Theorem 8.2.2: Degree and Zeros of Characteristic Polynomial

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, p is the characteristic polynomial. Then

- $\deg p = \dim V$.
- The zeros of p are the eigenvalues of T.

Theorem 8.2.3: Cayley-Hamilton Theorem

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, q is the characteristic polynomial of T. Then $q(T) = 0$.

Proof. We have

$$q(T) = (T - \lambda_1 I)^{d_1} \cdots (T - \lambda_m I)^{d_m}.$$

As $(T - \lambda_k)^{d_k}|_{G(\lambda_k, T} = 0$, then $q(T)|_{G(\lambda_k, T)} = 0$. As $V = \bigoplus G(\lambda_k, I)$ then $q(T) = 0$. □

Proposition: **Characteristic Polynomial and the Minimal Polynomial**

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, then the characteristic polynomial is a multiple of the minimal polynomial.

Proof. This follows immediately from the Cayley-Hamilton Theorem. □

Theorem 8.2.4: Multiplicity and the Diagonal

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$. $v_1, \ldots, v_n$ is a basis which T has an upper-triangular matrix $\mathscr{M}$. Then the multiplicity of λ is the number of times λ appears on the diagonal of $\mathscr{M}$.

Proof. As all the columns that $\lambda_k \neq 0$ are linear independent (obvious), then let d be the number of $\lambda_k = 0$. We have

$$\dim \operatorname{range} T \geq n - d$$

Thus

$$\dim \operatorname{null} T \leq d$$

As T^n has diagonal $\lambda_1^n, \dots, \lambda_n^n$, then

$$\dim \operatorname{null} T^n \leq d$$

Let m_λ be the multiplicity of λ and d_λ be the number of λ appears on the diagonal. Replacing T with $T - \lambda I$ we have

$$m_\lambda \leq d_\lambda$$

As $\sum m_\lambda = \sum d_\lambda = n$, so $m_\lambda = d_\lambda$. □

8.2.3 Block Diagonal Matrices

To see that decomposing a vector space into direct sums of invariant subspaces would simplify matter, we see that the matrix form has a *block-diagonal* pattern:

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

Where A_k are square matrices. This is obvious since each A_k is the restraint of T onto one of the invariant subspaces, mapping a part of the coordinated onto exactly the same part.

Theorem 8.2.5: Block Diagonal Matrices with Upper-triangular Blocks

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, let $\lambda_1, \dots, \lambda_m$ be the eigenvalues of T, with multiplicities d_k. Then there is a basis of V such that T has a block-diagonal form

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

With A_k a $d_k \times d_k$ upper-triangular matrix.

Proof. Combining the upper-triangular-basis of $(T - \lambda_k I)|_{G(\lambda_k, T)}$ would do. □

8.3 Consequences of General Eigenspace Decomposition

8.3.1 Square Root of Operators

We have seen that every positive operator has a square root. This is not true for all operators on $\mathbb{C}$.

Example: **Operators that does not have a Square Root**

$T(z_1, z_2, z_3) = (z_2, z_3, 0)$ does not have a square root.

Proposition: **Identity + Nilpotent has a square root**

Suppose $T \in \mathscr{L}(V)$ is nilpotent. Then $I + T$ has a square root.

Our motivation of the proof comes from the Taylor series of the function $\sqrt{1+x}$.

Proof. We let the Taylor series of $\sqrt{1+x}$ to be

$$\sqrt{1+x} = 1 + a_1 x + a_2 x^2 + \cdots$$

Let m be the smallest integer that $T_m = 0$, consider the polynomial

$$p(T) = I + a_1 T + \cdots + a_{m-1} T^{m-1}$$

And we want to show that $p(T)^2 = I + T$. Just solving each a_k would do. □

Theorem 8.3.1: Invertible Operators over $\mathbb{C}$ have Square Roots

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, then T has a square root.

Proof. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalue of T. As $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent, denote it $T_k \in \mathscr{L}(G(\lambda_k, T))$, thus $T|_{G(\lambda_k,T)} = \lambda_k I + T_k$.

Because T is invertible, $\lambda_k \neq 0$ for all k. Thus, we write

$$T|_{G(\lambda_k,T)} = \lambda_k \left(1 + \frac{T_k}{\lambda_k}\right)$$

We have $1 + T_k/\lambda_k$ has a square root and so do λ_k, thus $T|_{G(\lambda_k,T)}$ has a square root. Changing the diagonal block to its square root we get a square root of T.

(Mathematically, for each $v = u_1 + \ldots + u_m$ where $u_k \in G(\lambda_k, T)$, we let $Rv = R_1 u_1 + \ldots + R_m u_m$ would do.) □

Corollary 8.3.1: k^{th} root of Invertibles over $\mathbb{C}$

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, then T has a k^{th} root for all $k \in \mathbb{Z}_+$.

8.3.2 Jordan Form

We can do even better with the block-diagonal form. We shall prove that V has a basis to which the matrix of T has nonzero elements only on the diagonal and the line above the diagonal.

In the last section, we know that we should focus on the operator $T|_{G(\lambda_k,T)}$. This operator is nilpotent. If we can prove that every nilpotent operator has a basis to which the matrix form has nonzero elements only at the line above the diagonal, we're done!

Definition 8.3.1: Jordan Basis

Suppose $T \in \mathscr{L}(V)$, then a basis of V is called a *Jordan Basis* for T if the matrix form of T looks like

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

in which each A_k is an upper-triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}$$

!!Different A_k may have the same λ_k.

Theorem 8.3.2: Every Nilpotent Operators has a Jordan Basis

Suppose $T \in \mathscr{L}(V)$ is nilpotent, then there is a Jordan basis of V for T.

Proof. We'll prove this result by induction on $\dim V$. If $\dim V = 1$, the result holds, obviously. Assuming that $\dim V > 1$ and the result holds for all smaller dimension. Let m be the smallest number that $T^m = 0$, thus there exists $u \in V$, such that $T^{m-1}u \neq 0$. Let

$$U = \operatorname{span}(u, Tu, \ldots, T^{m-1}u)$$

The list $u, Tu, \ldots, T^{m-1}u$ is linear independent. (If $T^k u = a_1 u + \ldots + a_{k-1}T^{k-1}$, multiply the equation by $T^{m-k}, \ldots, T^m$ we have $a_1 = \cdots = a_{k-1} = 0$, contradicting to the minimality of m.) If $U = V$, writing the list in reverse order gives a Jordan basis.

Assume $U \neq V$, then we have a Jordan basis for $T|_U$, we shall find a W that is invariant under T and $V = U \oplus W$.

Geometrically, a nilpotent T acting on a Jordan basis is like: moving the next basis to the previous one, and deleting the first. A good way to find W is to construct a space that is perpendicular to all $T^k v$.

Let $\varphi \in V'$ be such that $\varphi(T^{m-1}u) \neq 0$, and let

$$W = \left\{ v \in V : \varphi(T^k v) = 0, \forall k = 0, \ldots, m-1 \right\}$$

Then:

- W is a subspace that is invariant under T. (For $\varphi(T^k(Tv)) = 0$.)
- $U + W$ is a direct sum. Suppose $v \in U \cap W, v \neq 0$, Because $v \in U$, there is $c_0, \ldots, c_{m-1}$ such that

  $$v = c_0 u + c_1 Tu + \cdots + c_{m-1}T^{m-1}u.$$

  Let j be the smallest that $c_j \neq 0$, then

  $$T^{m-j-1}v = c_j T^{m-1}u$$

  Applying φ to both sides gives a contradiction.

- $U \oplus W = V$, our intuition is that W is perpendicular to m linear independent vectors, so has at most $\dim V - m$ dimension. Let $Sv = (\varphi(v), \ldots, \varphi(T^{m-1}v))$, then $\operatorname{null} S = W$. Hence $\dim W = \dim V - \dim \operatorname{range} S \geq \dim V - m$. So we have $\dim(U \oplus W) \geq \dim V$.

□

Theorem 8.3.3: Jordan Form

Suppose $\mathbb{F} = \mathbb{C}$, $T \in \mathscr{L}(V)$, then there is a Jordan basis of V for T.

Proof. Easily proved by the previous theorem and the generalized eigenspace decomposition. □

8.4 Trace

Definition 8.4.1: Trace of a Square Matrix

Suppose A is a square matrix in $\mathbb{F}^{n\times n}$. The trace of A is defined to be the sum of the diagonal of A.

Theorem 8.4.1: Trace of AB and BA

Suppose $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times m}$, then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Proof. Expanding the terms would do:

$$\operatorname{tr}(AB) = \operatorname{tr}(BA) = \sum_{j=1}^{m}\sum_{k=1}^{n} A_{j,k}B_{k,j}$$

□

Theorem 8.4.2: Trace do not Depend on Basis

Suppose $T \in \mathscr{L}(V)$, suppose u_k and v_k are basis, then

$$\operatorname{tr} \mathscr{M}(T, u_k) = \operatorname{tr} \mathscr{M}(T, v_k)$$

Proof.

$$\operatorname{tr} A = \operatorname{tr}(C^{-1}BC) = \operatorname{tr}(CC^{-1}B) = \operatorname{tr} B$$

□

This gives us confidence in defining the trace of an operator.

Definition 8.4.2: Trace of an Operator

The trace of an Operator is the trace of any of its matrix

Proposition: **Properties of Trace**

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathscr{L}(V)$, then

- $\operatorname{tr} T$ equals the sum of eigenvalues of T including multiplicities.

$$\operatorname{tr} T = d_1\lambda_1 + \cdots + d_m\lambda_m$$

- $\operatorname{tr} T$ is the negative of the z^{n-1} term coefficient of the characteristic polynomial.

Proposition: **Trace on Inner Product Space**

Suppose V is an inner product space, and $e_1, \ldots, e_n$ is an orthogonal basis. Then

$$\operatorname{tr} T = \langle Te_1, e_1 \rangle + \cdots + \langle Te_n, e_n \rangle .$$

Theorem 8.4.3: The Algebraic Characterization of Trace

The trace is the only linear functional $\tau : \mathscr{L}(V) \to \mathbb{F}$ such that

- $\tau(ST) = \tau(TS)$, for all $S, T \in \mathscr{L}(V)$.
- $\tau(I) = \dim V$.

Proof. We only need to consider $\mathbb{F}^{n\times n} \cong \mathscr{L}(V)$. Find a basis of $\mathbb{F}^{n\times n}$:

$$P_{i,j} = \text{ the } (i, j) \text{ entry is 1, other is 0.}$$

$$P_{i,j}P_{k,l} = \begin{cases} P_{i,l}, & j = k \\ 0, & j \neq k \end{cases}$$

So if $i \neq l$, we have $\operatorname{tr}(P_{i,l}) = \operatorname{tr}(P_{k,l}P_{i,j}) = 0$. Also $\operatorname{tr} P_{i,i} = \operatorname{tr}(P_{i,j}P_{j,i}) = \operatorname{tr} P_{j,j}$, and $\operatorname{tr} I = \operatorname{tr} P_{11} + \cdots + \operatorname{tr} P_{nn} = n$, so $\operatorname{tr} P_{i,i} = 1$. ☐

Example: $AB - BA = I$

In a finite dimensional vector space, there is no operators such that $ST - TS = I$.

Chapter 9

Multilinear Algebra and Determinants

We investigate the bilinear forms and quadratic forms on a vector space.

Standing Assumptions:

- $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.
- V, W are finite dimensional nonzero vector spaces over $\mathbb{F}$.

9.1 Bilinear Forms and Quadratic Forms

9.1.1 Bilinear Forms

Well actually a bilinear form should be phrased as "bilinear functional".

Definition 9.1.1: Bilinear Forms

a bilinear form on V is a function $\beta : V \times V \to \mathbb{F}$ such that

$$v \mapsto \beta(v, u) \quad \text{and} \quad v \mapsto \beta(u, v)$$

are both linear functionals.

This explicitly means that β satisfies:

- $\beta(au + bv, w) = a\beta(u, w) + b\beta(v, w)$.
- $\beta(u, av + bw) = a\beta(u, v) + b\beta(u, w)$.

Note that a bilinear form differs from inner products, it does not require symmetry. Also, a bilinear form is usually not a linear functional from $V \times V \to \mathbb{F}$, for $\beta(u_1+u_2, v_1+v_2) \neq \beta(u_1, v_1)+\beta(u_2, v_2)$.

Definition 9.1.2: $V^{(2)}$

The set of bilinear forms is denoted $V^{(2)}$.

It is easy to see that $V^{(2)}$ is indeed a vector space with the usual addition and scalar multiplication of functions. We usually represent a bilinear form by a matrix.

Definition 9.1.3: Matrix of a Bilinear Form

Suppose β is a bilinear form on V and $e_1, \ldots, e_n$ is a basis of V. The matrix of β to the basis is defined to be $\mathscr{M}(\beta) \in \mathbb{F}^{n \times n}$:

$$\mathscr{M}(\beta)_{j,k} = \beta(e_j, e_k)$$

This can be seen as the bilinear form on $\mathbb{F}^n$, for $V \cong \mathbb{F}^n$, this is quite natural. The result also implies that $V^{(2)} \cong \mathbb{F}^{n \times n}$. Also, in matrix form, if $(v_1, v_2) \in \mathbb{F}^n \times \mathbb{F}^n$, then we have

$$\beta(v_1, v_2) = v_1^T \mathscr{M}(\beta) v_2.$$

Corollary 9.1.1: Dimension of $V^{(2)}$

The map $\beta \to \mathscr{M}(\beta)$ above is an isomorphism of $V^{(2)} \to \mathbb{F}^{n \times n}$.
Therefore, $\dim V^{(2)} = (\dim V)^2$.

Proposition: **Composition of a Bilinear Form and an Operator**

Suppose $\beta \in V^{(2)}$ and $T \in \mathscr{L}(V)$, then defined $\alpha, \rho \in V^{(2)}$

$$\alpha(u, v) = \beta(u, Tv) \qquad \rho(u, v) = \beta(Tu, v)$$

We have

$$\mathscr{M}(\alpha) = \mathscr{M}(\beta)\mathscr{M}(T) \qquad \mathscr{M}(\rho) = \mathscr{M}(T)^t \mathscr{M}(\beta)$$

Proposition: **Change of Basis of Bilinear Forms**

Suppose $\beta \in V^{(2)}$ and e_k, f_k are basis of V. A and B are matrices of β to e_k and f_k. Let $C = \mathscr{M}(I, e_k, f_k)$, then

$$A = C^t BC \tag{9.1}$$

Remark:
We notice that this change of basis formula is different from that of change of basis, which is $A = C^{-1}BC$.
We shall clarify that the nature of bilinear functions and operators are different. The transplantation of previous theorems may sometimes fail. We are confident to say that for $\mathbb{F} = \mathbb{R}$, for a symmetric matrix A, there is a C that $C^t C = I$ and $C^t AC$ is symmetric according to the real spectral theorem. But for an arbitrary $\mathbb{F}$ we need to restate the base-changing condition.

9.1.2 Symmetric Bilinear Forms

Definition 9.1.4: Symmetric Bilinear Forms

A bilinear form $\rho \in V^{(2)}$ is called symmetric if

$$\forall u, w \in V, \rho(u, w) = \rho(w, u)$$

The set of symmetric bilinear forms is denoted $V^{(2)}_{\text{sym}}$.

The next result follows directly form the spectral theorem.

Theorem 9.1.1: Symmetric Bilinear Forms are Diagonalizable

Suppose $\rho \in V^{(2)}$. Then the following are equivalent.

- ρ is a symmetric bilinear form.
- $\mathscr{M}(\rho, e_k)$ is a symmetric matrix for every basis of V.
- $\mathscr{M}(\rho, e_k)$ is a symmetric matrix for some basis of V.
- $\mathscr{M}(\rho, e_k)$ is a diagonal matrix for some basis $e_1, \ldots, e_n$ of V.

Suppose V is a real inner product space, and ρ a symmetric bilinear form on V, then there is an orthonormal basis of V to which ρ has a diagonal matrix.

Proof. Suppose ρ is a symmetric bilinear form. Suppose $e_1, \ldots, e_n$ is a basis of V then $\rho(e_j, e_k) = \rho(e_k, e_j)$ for ρ is symmetric. Thus $\mathscr{M}(\rho, e_i)$ is a symmetric matrix. $2 \to 3$ is obvious and showing $3 \to 1$ requires mere expanding from basis.

To prove that $1 \to 3$ we use induction on $n = \dim V$. If $n = 1$ then the result holds. If the result holds for less dimension then n, then if $\rho = 0$, then the matrix is 0. If $\rho \neq 0$, then $\exists u, v \in V, \rho(u, v) \neq 0$, we have

$$2\rho(u, v) = \rho(u + w, u + w) - \rho(u, u) - \rho(v, v) \neq 0$$

Hence there exists $v \in V, \rho(v, v) \neq 0$. Let $U = \{u \in V : \rho(u, v) = 0\}$, then U is the null space of the linear functional $u \mapsto \rho(u, v)$. We have $U \neq 0$ so $\dim U = n - 1$ which means there is a basis $e_1, \ldots, e_{n-1}$ such that $\rho|_{U \times U}$ has a diagonal matrix. Also, $e_1, \ldots, e_{n-1}, v$ is a basis of V. And $\rho(v, e_k) = \rho(e_k, v) = 0$, so the matrix to this basis is diagonal. □

9.1.3 Alternating Bilinear Forms

Definition 9.1.5: Alternating Bilinear Form

A bilinear form is called alternating if

$$\forall v \in V, \alpha(v, v) = 0$$

The set of alternating bilinear forms is denoted $V^{(2)}_{\text{alt}}$.

There is a more intuitive definition that relate to skew matrices.

Theorem 9.1.2: Characterization of Alternating Bilinear Forms

A bilinear form α on V is alternating iff

$$\forall u, v \in V, \alpha(u, v) = -\alpha(v, u)$$

Proof.

$$\begin{aligned} 0 &= \alpha(u + w, u + w) \\ &= \alpha(u, u) + \alpha(u, w) + \alpha(w, u) + \alpha(w, w) \\ &= \alpha(u, w) + \alpha(w, u) \end{aligned}$$

□

There is an obvious decomposition of bilinear forms into symmetric and alternating forms:

Theorem 9.1.3: Symmetric-Alternating Decomposition

$$V^{(2)} = V^{(2)}_{\text{sym}} + V^{(2)}_{\text{alt}} \tag{9.2}$$

Proof. Setting

$$\rho(u, w) = \frac{1}{2}\left(\beta(u, w) + \beta(w, u)\right) \qquad \alpha(u, w) = \frac{1}{2}\left(\beta(u, w) - \beta(w, u)\right)$$

We'll have $\beta = \rho + \alpha$. □

9.1.4 Quadratic Forms

Definition 9.1.6: Quadratic Form

For $\beta \in V^{(2)}$, define a function $q_\beta : V \to \mathbb{F}$ by setting $q_\beta = \beta(v, v)$.
A function $q : V \to \mathbb{F}$ is called a quadratic form on V if there is a $\beta \in V^{(2)}$ such that $q = q_\beta$.

It is easy to see that $q_\beta = 0$ iff β is alternating.

Proof. By the definition of alternating forms would suffice. □

On $\mathbb{F}^n$, a quadratic form is usually represented as

$$q(x_1, \ldots, x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k. \tag{9.3}$$

The definition of a quadratic form require the mere existence of an associated bilinear form, it is easy to see that the bilinear forms corresponding to the specific quadratic form is not unique.

Theorem 9.1.4: Characterization of Quadratic Forms

Suppose $q : V \to \mathbb{F}$. Then the following are equivalent.

- q is a quadratic form.
- There exists a *unique* symmetric bilinear form $\rho \in V^{(2)}_{\text{sym}}$ such that $q = q_\beta$.
- $\forall \lambda \in \mathbb{F}, v \in V$, we have $q(\lambda v) = \lambda^2 q(v)$, and the function

$$(u, w) \mapsto q(u + w) - q(u) - q(w)$$

  is a symmetric bilinear form on V.
- The previous result hold specifically for $\lambda = 2$. (Or $\lambda = \lambda_0$.)

Proof.

- If q is a quadratic form, we let $q = q_\beta$ where $\beta \in V^{(2)}$. Let $\beta = \rho + \alpha$, where ρ is symmetric. Then $q_\beta = q_\rho$ for $q_\alpha = 0$.
- $q(\lambda v) = \rho(\lambda v, \lambda v) = \lambda^2 \rho(v, v)$. And the next part is obvious.
- Clearly $3 \to 4$.
- Letting

$$\rho(u, w) = \frac{1}{2} \left(q(u + w) - q(u) - q(w) \right)$$

  Then,

$$\rho(v, v) = \frac{4q(v) - 2q(v)}{2} = q(v)$$

□

Remark:
We have the set of quadratic forms $\cong V^{(2)}_{\text{sym}}$.

By the spectral theorem, any symmetric matrix can be diagonized.

Theorem 9.1.5: The Diagonalization of Quadratic Forms

Suppose q is a quadratic form on V.

- There is a basis $e_1, \ldots, e_n \in V$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that:

$$q(x_1 e_1 + \cdots + x_n e_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2.$$

  For all $x_1, \ldots, x_n \in \mathbb{F}$.
- If $\mathbb{F} = \mathbb{R}$ and V is an inner product space, then a basis can be chosen to be orthonormal.

9.2 Alternating Multilinear Forms

9.2.1 Multilinear Forms

We give a similar definition of multilinear forms (functionals), which is a natural generalization of bilinear forms.

Definition 9.2.1: m-linear form, $V^{(m)}$

For any $m \in \mathbb{Z}_+$, an m-linear form on V is a function $\beta : V^m \to \mathbb{F}$, such that for each $k \in \{1, \ldots, m\}$ and $\forall u_1, \ldots, u_m \in V$, the function

$$v \mapsto \beta(u_1, \ldots, u_{k-1}, v, u_{k+1}, \ldots, u_m)$$

is a linear map $V \to \mathbb{F}$.
The set of m-linear forms is denoted $V^{(m)}$.

Definition 9.2.2: Alternating Forms

Suppose $m \in \mathbb{Z}_+$.

- A m-linear form α on V is called alternating if $\alpha(v_1, \ldots, v_m) = 0$ if there is some $v_j = v_k$ for some $j \neq k$.
- The set of alternating m-linear form on V is denoted $V^{(m)}_{\text{alt}}$.

Theorem 9.2.1: Alternating Multilinear Forms and Linear Independence

Suppose $m \in \mathbb{Z}_+$ and $\alpha \in V^{(m)}_{\text{alt}}$ on V. Then if $v_1, \ldots, v_m$ is a linear dependent list on V then

$$\alpha(v_1, \ldots, v_m) = 0$$

Proof. There is v_k that is a linear combination of previous vectors. Let $v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}$. then

$$\begin{aligned}
\alpha(v_1, \ldots, v_m) &= \alpha\left(v_1, \ldots, v_{k-1}, \sum_{=1}^{k-1} b_j v_j, v_{k+1}, \ldots, v_m\right) \\
&= \sum_{j=1}^{k-1} \alpha(v_1, \ldots, v_{k-1}, b_j v_j, v_{k+1}, \ldots, v_m) \\
&= 0
\end{aligned}$$

$\square$

Corollary 9.2.1: An upper limit for nonzero alternating multilinear forms

Suppose $m > \dim V$, then 0 is the only alternating multilinear m-forms.

Theorem 9.2.2: Alternating Multilinear Forms and Permutations

Suppose $m \in \mathbb{Z}_+$ and $\alpha \in V_{\text{alt}}^{(m)}$. Then

$$\alpha(v_{j_1}, \dots, v_{j_m}) = \text{sign}(j_1, \dots, j_m)\alpha(v_1, \dots, v_m).$$

Where sign is the permutation number (odd -1, and even 1).

Proof. We prove this by swapping the first two entries gives a factor -1:

$$0 = \alpha(v_1 + v_2, v_1 + v_2, v_3, \dots, v_m) = \alpha(v_1, v_2, v_3, \dots, v_m) + \alpha(v_2, v_1, v_3, \dots, v_m)$$

□

Theorem 9.2.3: The $(\dim V)$-multilinear alternating form

Let $n = \dim V$. Let $e_1, \dots, e_n$ be a basis of V and $v_1, \dots, v_n \in V$. Let

$$v_k = \sum_{j=1}^{n} b_{k,j} e_j$$

Then

$$\alpha(v_1, \dots, v_n) = \alpha(e_1, \dots, e_n) \sum_{(j_1, \dots, j_n) \in \text{perm}\, n} \text{sign}(j_1, \dots, j_n) b_{1,j_1} \cdots b_{n,j_n}.$$

Therefore, the space $V_{\text{alt}}^{(\dim V)}$ has dimension 1. (We only need to construct a nonzero one).

9.3 Determinants

Notation 9.3.1: α_T

Suppose $m \in \mathbb{Z}_+$ and $T \in \mathcal{L}(V)$. For $\alpha \in V_{\text{alt}}^{(m)}$, define $\alpha_T \in V_{\text{alt}}^{(m)}$ by

$$\alpha_T(v_1, \dots, v_m) = \alpha(Tv_1, \dots, Tv_m)$$

Remark:
We can easily show that the function $\alpha \mapsto \alpha_T$ is a linear operator on $V_{\text{alt}}^{(m)}$. As $\dim V_{\text{alt}}^{(m)} = 1$. Therefore, we define the determinant of the operator to be the multiple coefficient.

Definition 9.3.1: Determinants

Suppose $T \in \mathcal{L}(V)$, Then determinant of T, denoted by $\det T$ is defined to be the unique number in $\mathbb{F}$ such that

$$\alpha_T = (\det T)\alpha$$

for all $\alpha \in V_{\text{alt}}^{(m)}$.
The determinant of a matrix is defined by the determinant of the corresponding operator.

Remark:
We shall see that when $\alpha(v_1, \ldots, v_m)$ stands for the "volume" of the vectors, then the determinant is the "volume change ratio" after the operator T.

Theorem 9.3.1: Determinant is an alternating multilinear form

Suppose $n \in \mathbb{Z}_+$, then the map that takes a list $v_1, \ldots, v_n \in \mathbb{F}^n$ to $\det(v_1, \ldots, v_n)$ is an alternating n-linear form on $\mathbb{F}^n$.

Proof. For the standart basis we have $\det T = \det(v_1, \ldots, v_n)$. Let α be a multilinear form that $\alpha(e_1, \ldots, e_n) = 1$, then

$$\det(v_1, \ldots, v_n) = \det T\alpha(e_1, \ldots, e_n) = \alpha(v_1, \ldots, v_n).$$

□

This result implies several familiar consequences. The formula of determinant of a matrix is

$$\det A = \sum_{(j_1,\ldots,j_n)\in \operatorname{perm} n} \operatorname{sign}(j_1, \ldots, j_n) A_{1,j_1} \cdots A_{n,j_n}. \tag{9.4}$$

For upper-triangular matrices with diagonal entries $\lambda_1, \ldots, \lambda_n$ we have

$$\det A = \lambda_1 \cdots \lambda_n \tag{9.5}$$

9.3.1 Properties of Determinants

Theorem 9.3.2: Determinants are Multiplicative

- Suppose $S, T \in \mathscr{L}(V)$, then $\det(ST) = \det S \det T$.
- For square matrices $A, B \in \mathbb{F}^{n\times n}$, we have $\det(AB) = \det A \det B$.

Proof. Let $\dim V = n$, we have

$$\begin{aligned}\alpha_{ST}(v_1, \ldots, v_n) &= \alpha(STv_1, \ldots, STv_n)\\ &= (\det S)\alpha(Tv_1, \ldots, Tv_n)\\ &= (\det S)(\det T)\alpha(v_1, \ldots, v_n)\end{aligned}$$

Therefore we have $\det(ST) = \det S \det T$. □

Theorem 9.3.3: Invertible and Nonzero Determinant

An operator $T \in \mathscr{L}(V)$ is invertible iff $\det T \neq 0$. Furthermore, $\det(T^{-1}) = 1/\det T$.

Proof. Suppose T is invertible, then $TT^{-1} = I$, then $\det T \det(T^{-1}) = 1$. Thus $\det T \neq 0$.

Suppose $\det T \neq 0$, $\forall v \neq 0$, let $v, e_2, \ldots, e_n$ be a basis, then

$$\alpha(Tv, Te_2, \ldots, Te_n) = (\det T)\alpha(v, e_2, \ldots, e_n) \neq 0$$

Therefore $Tv \neq 0$. □

Theorem 9.3.4: Eigenvalues and Determinants

Suppose $T \in \mathscr{L}(V)$, and $\lambda \in \mathbb{F}$. Then λ is an eigenvalue iff $\det(T - \lambda I) = 0$.

Proof. λ is an eigenvalue iff $T - \lambda I$ is invertible. $\square$

Theorem 9.3.5: Determinant is a Similarity Invariant

Suppose $T \in \mathscr{L}(V)$, and $S : W \to V$ is an invariant linear map, then

$$\det(S^{-1}TS) = \det T$$