# MATH 245 by Alexandru Nica

Eason Li

2024 S

Will be talking about vector spaces, linear transformations, eigenvalues, and diagonalization.

Remark: We will be using $\mathbb{R}$ and $\mathbb{C}$ as fields of scalars.

## Contents

# 1 Inner Product Space over $\mathbb{R}$

**Definition 1.1: Inner Product and Inner Product Space**

Let $V$ be a vector space over $\mathbb{R}$. An **inner product** on $V$ is a *bilinear, symmetric* and *positive definite* function "$\langle \cdot, \cdot \rangle$" from $V \times V$ to $\mathbb{R}$. The couple $(V, \langle \cdot, \cdot \rangle)$ is called an **inner product space**.

Explanation of the terms:

1. Function $\langle \cdot, \cdot \rangle$ from $V \times V$ to $\mathbb{R}$ is a function associate every $x, y \in V$ to a number $\langle x, y \rangle \in \mathbb{R}$, called the inner product of $x$ and $y$.
2. **Bilinarity of** $\langle \cdot, \cdot \rangle$, we have
   (a) $\langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle = \langle \alpha_1 x_1, y \rangle + \langle \alpha_2 x_2, y \rangle$ for all $x_1, x_2, y \in V$ and $\alpha_1, \alpha_2 \in \mathbb{R}$.
   (b) $\langle x, \beta_1 y_1 + \beta_2 y_2 \rangle = \langle x, \beta_1 y_1 \rangle + \langle x, \beta_2 y_2 \rangle$ for all $x, y_1, y_2 \in V$ and $\beta_1, \beta_2 \in \mathbb{R}$.
3. **Symmetry**, we have
$$\langle x, y \rangle = \langle y, x \rangle$$
for all $x, y \in V$.
4. **Positive Definite**, we have
$$\langle x, x \rangle \geq 0$$
for all $x \in V$, and equality holds if and only if $x = 0_V$.

**Discovery 1.1**

Notice we can find that

1. Bilinearity implies that
$$\langle 0_V, y \rangle = 0 \qquad \forall\, y \in V$$
*Proof.* We have
$$\langle 0_V, y \rangle = \langle 2 \cdot 0_V, y \rangle \quad \Rightarrow \quad \langle 0_V, y \rangle = 0$$
□
2. Note, in connection with positive definite, we indeed have $\langle 0_V, 0_V \rangle = 0$. The point of positive definite is
$$\begin{cases} x \in V \\ \langle x, x \rangle = 0 \end{cases} \quad \Rightarrow \quad x = 0_V$$
so we can identify if two vectors are the same.

**Example 1.1: Standard Inner Product**

Suppose

$$\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{R}\}$$

Then the standard inner product is defined as follows:
For $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, we have

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n$$

**Example 1.2**

Let $a < b \in \mathbb{R}$ and let

$$\begin{aligned} V &= \mathcal{C}([a,b], \mathbb{R}) \\ &= \{f : [a,b] \to \mathbb{R},\ f \text{ continuous}\} \end{aligned}$$

operations defined pointwise

$$\begin{aligned} f + g &: [a,b] \to \mathbb{R} \\ (f+g)(t) &= f(t) + g(t) \qquad \forall\, t \in [a,b] \end{aligned}$$

Now, for $f, g \in \mathcal{C}([a,b], \mathbb{R})$, we define

$$\langle f, g \rangle = \int_a^b f(t) g(t)\ dt$$

which can be shown to be an inner product. **Exercise:** For $f \in \mathcal{C}([a,b], \mathbb{R})$ such that

$$\int_a^b [f(t)]^2\ dt = 0$$

show that $f = 0$.

**Definition 1.2: Norm**

Suppose $(V, \langle \cdot, \cdot \rangle)$ is an inner product space. For $x \in V$, denote $\|x\|$, called the **norm** of $x$, is defined as

$$\|x\| = \sqrt{\langle x, x \rangle} \ \in\ [0, \infty)$$

**Example 1.3**

For example 1.1, we simply have

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2} \qquad \text{“Pythagoras”}$$

**Example 1.4**

For example 1.2, we have

$$\|f\| = \sqrt{\langle f, f\rangle} = \sqrt{\int_a^b [f(t)]^2 \, dt}$$

## 1.1 Cauchy Schwarz

**Proposition 1.1: Cauchy-Schwarz**

Suppose $(V, \langle \cdot, \cdot \rangle)$ is an inner product space, for every $x, y \in V$, we have

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$$

and the equality holds if and only if $x$ and $y$ are linearly dependent.

**Result 1.1: "Angle between $x$ and $y$"**

Suppose we have $(V, \langle \cdot, \cdot \rangle)$ and $x, y \in (V, \langle \cdot, \cdot \rangle)\backslash\{0\}$, then

$$\frac{\langle x, y\rangle}{\|x\| \cdot \|y\|} \in [-1, 1]$$

and there exists a unique $\theta \in [0, \pi]$ such that

$$\cos\theta = \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|}$$

this is called the **angle between $x$ and $y$**.

Lecture 2 - Monday, May 8

*Proof.* This is the proof for Cauchy Schwatz 1.1.
Pick $x, y \in V$ for which we verify Cauchy schwarz. If $x$ or $y = 0_V$, then it holds with equality since

$$\langle x, y\rangle = 0 = \|x\| \cdot \|y\|$$

Therefore WLOG assume $x \neq 0_V \neq y$. Consider the function

$$\begin{aligned} q \ : \ &\mathbb{R} \to \mathbb{R} \\ &t \to \langle x + ty, x + ty\rangle \end{aligned}$$

Observe that $q(t) \geq 0$ for all $t \in \mathbb{R}$. Moreover,

$$\begin{aligned} a(t) &= \langle x, x\rangle + \langle x, ty\rangle + \langle ty, x\rangle + \langle ty, ty\rangle \\ &= \|y\|^2 t^2 + 2\langle x, y\rangle t + \|x\|^2 \end{aligned}$$

which has at most one root, thus

$$\begin{aligned} \Delta = 4\langle x, y\rangle^2 - 4\|y\|^2\|x\|^2 \leq 0 \\ \Rightarrow \quad |\langle x, y\rangle| \leq \|x\|\,\|y\| \end{aligned}$$

□

**Definition 1.3: Angle between $x$ and $y$**

In $\mathbb{R}^N$, Cauchy-schwarz inequality 1.1 implies that for $x, y \neq 0$, then

$$\frac{\langle x, y\rangle}{\|x\|\,\|y\|} \in [-1, 1]$$

we can find a unique $\theta \in [0, \pi]$ such that

$$\cos\theta = \frac{\langle x, y\rangle}{\|x\|\,\|y\|}$$

we define $\theta$ to be the angle between $x$ and $y$.

**Proposition 1.2**

Suppose $(V, \langle\cdot,\cdot\rangle)$ an inner product space. Consider the map $V \to \mathbb{R}$ defined as $x \to \|x\|$ known as the norm function. The norm function as the following properties:

1. $\|x\| \geq 0$ for all $x \in V$
2. $\|x\| = 0$ if and only if $x = 0_V$
3. $\|\alpha x\| = |\alpha|\,\|x\|$ for all $\alpha \in \mathbb{R}$
4. $\|x + y\| \leq \|x\| + \|y\|$
5. Remark from the fourth one we can obtain that

$$\|x - y\| \geq |\,\|x\| - \|y\|\,|$$

**Discovery 1.2: Normed Vector Space**

Properties 1-4 can be used as a *list of axioms* defining the notion of **normed vector space**, denoted as $(V, \|\cdot\|)$.

**Example 1.5: Exercise**

Let $V = \mathbb{R}^m$ for some $m \geq 2$. For $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, put

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_m|\}$$

1. Prove that this is a normed vector space by proving it satisfies properties 1-4.
2. Prove that it is not possible to find an inner product of $\mathbb{R}^m$, "$\langle\cdot,\cdot\rangle_\infty$", such that

$$\|x\|_\infty = \sqrt{\langle x, x\rangle_\infty} \qquad \forall\, x \in \mathbb{R}^m$$

*Proof.* It is easy to obtain the first statement, we will be focusing on the second statement in this proof. Suppose we have $x = (1, 0, \ldots, 0) = e_1 \in \mathbb{R}^m$ and $y = (0, 1, \ldots, 0) = e_2 \in \mathbb{R}^m$. Thus we have

$$\begin{aligned} 1 = \|x + y\| &= \sqrt{\langle x + y, x + y\rangle} \\ \Rightarrow \quad 1 &= \langle x + y, x + y\rangle \\ \Rightarrow \quad 1 &= \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \end{aligned}$$

since we know that $\langle x, x\rangle = \|x\|^2 = 1 = \langle y, y\rangle$, we can obtain that $\langle x, y\rangle = -1/2$. Similarly, we also have

$$\begin{aligned} 2 = \|x + 2y\| &= \sqrt{\langle x + 2y, x + 2y\rangle} \\ \Rightarrow \quad 4 &= \langle x, x\rangle + 4\langle x, y\rangle + 4\langle y, y\rangle \end{aligned}$$

where we can find that $\langle x, y\rangle = -1/4$, which is a contradiction. $\square$

## 2 Orthogonal and Orthonormal Systems

**Definition 2.1: Orthogonal**

For $(V, \langle\cdot,\cdot\rangle)$ ips. Let $x, y \in V$. If $\langle x, y\rangle = 0$, we say that $x$ and $y$ are **orthogonal** to each other, denoted as $x \perp y$.

**Result 2.1**

Notice that $\forall\, x \in V$, $0_V \perp x$.
Also notice that $x \perp y$ if and only if $\theta_{x,y} = \pi/2$.

**Definition 2.2: Orthogonal System / Orthonormal System**

Suppose $(V, \langle\cdot,\cdot\rangle)$ ips, let $(x_i)_{i\in I}$ be a set of vectors in $V$.

1. We say that $(x_i)_{i\in I}$ is an **orthogonal system** if
   (a) $x_i \perp x_j$, $\forall\, i, j \in I$ such that $i \neq j$
   (b) $x_i \neq 0$, $\forall\, i \in I$
2. We say that $(x_i)_{i\in I}$ is an **orthonormal system** if
   (a) $x_i \perp x_j$, $\forall\, i, j \in I$ such that $i \neq j$
   (b) $\|x_i\| = 1$, $\forall\, i \in I$

**Discovery 2.1**

1. Orthonormal implies orthogonal
2. Conversely, we can manipulate an orthogonal system to construct an orthonormal by **re-norming** (**normalizing**) the $x_i$'s.

**Proposition 2.1**

For $(V, \langle\cdot,\cdot\rangle)$ an ips. Let $x_1, \dots, x_m$ be an orthogonal system of vectors in $V$, then $x_1, \dots, x_m$ are linearly independent.

*Proof.* Let $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ be such that

$$\alpha_1 x_1 + \cdots + \alpha_m x_m = 0$$

STP $\alpha_1 = \cdots = \alpha_m = 0$. We simply inner product the equation with $x_1, \dots, x_m$ on both sides one by one. $\square$

Lecture 4 - Monday, May 13

**Discovery 2.2**

For $(V, \langle\cdot,\cdot\rangle)$ being an inner product space and let $(x_i)_{i\in I}$ be an orthogonal system where the index set $I$ is infinite. We can still conclude that $(x_i)_{i\in I}$ forms a linearly independent set.

## 2.1 Gram-Schmidt Orthogonalization Procedure

**Definition 2.3: What is** span

Question: What is $\text{span}(x_1, \dots, x_n)$ for $x_1, \dots, x_n \in V$ which is a vectors space over $\mathbb{R}$.

1. $\text{span}(x_1, \dots, x_n) = \{\alpha_1 x_1 + \cdots + \alpha_n x_n \ : \ \alpha_i \in \mathbb{R} \text{ for } i = 1, \dots, n\}$
2. It is the smallest possible linear subspace of $V$ which contains $x_1, \dots, x_n$.

**Remark:**

1. $\text{span}(x_1, \dots, x_n)$ is a linear subspace of $V$ which containes $x_1, \dots, x_n$.
2. Whenever $W \subseteq V$ is a linear subspace containing $x_1, \dots, x_n$, it follows that $\text{span}(x_1, \dots, x_n) \subseteq W$.

**Definition 2.4: Increasing chain**

Vecror space $V$ over $\mathbb{R}$, let $x_1, \dots, x_n$ be a linearly independent set of vectors in $V$. Look at the linear subspaces

$$V_1 = \text{span}(x_1) \quad V_2 = \text{span}(x_1, x_2) \quad \dots \quad V_n = \text{span}(x_1, \dots, x_n)$$

Then

$$V_1 \ \subseteq \ V_2 \ \subseteq \ \cdots \ \subseteq \ V_n$$

is an **increasing chain** of linear subspaces of $V_i$ with

$$\dim(V_i) = i \qquad \forall\, i \in \{1, \dots, n\}$$

### 2.1.1 Gram-Schmidt

**Theorem 2.1: Gram-Schmidt**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space. Let $n \in \mathbb{N}$ and let $x_1, \dots, x_n$ be a linearly independent family of vectors in $V$. Then one can find an orthogonal system $y_1, \dots, y_n \in V$ such that

$$\begin{aligned} \operatorname{span}(y_1) &= \operatorname{span}(x_1) \\ \operatorname{span}(y_1, y_2) &= \operatorname{span}(x_1, x_2) \\ &\vdots \\ \operatorname{span}(y_1, \dots, y_n) &= \operatorname{span}(x_1, \dots, x_n) \end{aligned}$$

*Proof.* We prove this using induction. For the base case, $n = 1$, we can simply take $y_1 = x_1 \neq 0$. Therefore, suppose the statement holds for $n$, we want to show that it still holds for $n+1$. Let $x_1, \dots, x_n, x_{n+1}$ be linearly independent, hence by the induction hypothesis, $y_1, \dots, y_n$ form an orthogonal system such that

$$\operatorname{span}(y_1, \dots, y_i) = \operatorname{span}(x_1, \dots, x_i) \qquad \forall\, 1 \leq i \leq n$$

Our goal is to find $y_{n+1}$ such that

1. $y_{n+1} \neq 0$ and $y_{n+1} \perp y_i$ for all $1 \leq i \leq n$
2. $\operatorname{span}(y_1, \dots, y_{n+1}) = \operatorname{span}(x_1, \dots, x_{n+1})$

We find $y_{n+1}$ by the following formula:

$$\boxed{y_{n+1} = x_{n+1} - \left[\frac{\langle x_{n+1}, y_1\rangle}{\|y_1\|^2} y_1 + \cdots + \frac{\langle x_{n+1}, y_n\rangle}{\|y_n\|^2} y_n\right]}$$

It is easy to find that $y_{n+1} \neq 0$ since we have $\{x_{n+1}, y_1, \dots, y_n\}$ is linearly independent. Now we want to show that $y_{m+1} \perp y_i$ for all $1 \leq i \leq n$: We pick an arbitrary $i$ within that range and compute

$$\begin{aligned} \langle y_{n+1}, y_i\rangle &= \langle x_{n+1} - [\beta_1 y_1 + \cdots + \beta_n y_n], y_i\rangle \\ &= \langle x_{n+1}, y_i\rangle - \beta_1 \langle y_1, y_i\rangle - \cdots - \beta_n \langle y_n, y_i\rangle \\ &= \langle x_{n+1}, y_i\rangle - \beta_i \langle y_i, y_i\rangle \\ &= \langle x_{n+1}, y_i\rangle - \frac{\langle x_{n+1}, y_i\rangle}{\|y_i\|^2} \langle y_i, y_i\rangle \\ &= 0 \end{aligned}$$

Now it leaves us to show that $\operatorname{span}(y_1, \dots, y_{n+1}) = \operatorname{span}(x_1, \dots, x_{n+1})$, we prove it by proving inclusion in both direction. First of all, for $1 \leq i \leq n$, we have that

$$x_i \in \operatorname{span}(x_1, \dots, x_n) = \operatorname{span}(y_1, \dots, y_n) \subseteq \operatorname{span}(y_1, \dots, y_n, y_{n+1})$$

Moreover, we also have

$$x_{n+1} = y_{n+1} + \left[\frac{\langle x_{n+1}, y_1\rangle}{\|y_1\|^2}y_1 + \cdots + \frac{\langle x_{n+1}, y_n\rangle}{\|y_n\|^2}y_n\right]$$
$$\subseteq \operatorname{span}(y_1, \ldots, y_n, y_{n+1})$$

Therefore, we have

$$\operatorname{span}(x_1, \ldots, x_{n+1}) \subseteq \operatorname{span}(y_1, \ldots, y_n, y_{n+1})$$

The direction,

$$\operatorname{span}(y_1, \ldots, y_n, y_{n+1}) \subseteq \operatorname{span}(x_1, \ldots, x_n, x_{n+1})$$

is left as an exercise. □

Lecture 5 - Wednesday, May 15

## 2.2 Orthonormal Basis

**Definition 2.5: Orthonormal Basis**

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. We say that $z_1, \ldots, z_n \in V$ form an **orthonormal basis** for $V$ if $z_1, \ldots, z_n$ is an orthonormal system, and

$$\operatorname{span}(z_1, \ldots, z_n) = V$$

**Corollary 2.1**

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space of finite dimension, $\dim(V) = n \quad (n \neq 0)$. Then we can find an **orthonormal basis** for $V$.

*Proof.* Start with any basis, $x_1, \ldots, x_n$, for $V$. We apply the Gram-Schmidt Prodecure to find an orthogonal system such that

$$\operatorname{span}(y_1, \ldots, y_n) = \operatorname{span}(x_1, \ldots, x_n)$$

and then we normalize each of the elements in the orthogonal system to construct an orthonormal basis. □

Lecture 6 - Friday, May 17

**Exercise:** What does Gram-Schmidt do if the given $x_1, \ldots, x_n$ are already an orthogonal system?
Easy induction. $x_1 = y_1$ and since we are subtracting off $\frac{\langle x_{k+1}, y_i\rangle}{||y_i||^2}y_i = 0$ it changes nothing.

**Corollary 2.2**

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space with $\dim(V) = n \ll \infty$. Let $W \subseteq V$ be a linear sunspace with $\dim(W) = m$. Assume $1 \leq m < n$, let $z_1, \ldots, z_m$ be an orthonormal basis for $W$. Then, one can find $z_{m+1}, \ldots, z_n \in V$ such that $z_1, \ldots, z_n$ is an orthonormal basis for $V$.

*Proof.* By JPBell we can extend to a basis. We use Gram-Schmidt on this basis, and by the same reasons as the above exercise, the first $m$ vectors will not be changed. □

# 3 Orthogonal complements, orthogonal projection onto a linear subspace

**Definition 3.1: Orthogonal Complement**

Suppose $(V, \langle \cdot, \cdot \rangle)$ an inner product space and $W \subseteq V$ is a linear subspace. The orthogonal complement of $W$ denoted as $W^\perp$ is

$$\begin{aligned} W^\perp &= \{x \in V : x \perp w, \ \forall\, w \in W\} \\ &= \{x \in V : \langle x, w \rangle = 0, \ \forall\, w \in W\} \end{aligned}$$

**Discovery 3.1**

Suppose $(V, \langle \cdot, \cdot \rangle)$ an inner product space and $W \subseteq V$ is a linear subspace same as above. Then $W^\perp$, too, is a linear subspace of $V$. (Hence $W \rightsquigarrow W^\perp$ is an operation with linear subspaces.)

*Proof.* Here is the verification that $W^\perp$ is a linear subspace:

1. Do we have $0_V \in W^\perp$?
   Indeed, we have $0_V \perp x, \forall\, x \in V$, hence $0_V$ is indeed in $W^\perp$
2. Do we have $\left[x_1, x_2 \in W^\perp\right] \Rightarrow \left[x_1 + x_2 \in W^\perp\right]$?
   Indeed, for every $w \in W$, we have
   $$\langle x_1 + x_2, w \rangle = \langle x_1, w \rangle + \langle x_2, w \rangle = 0 + 0 = 0$$
3. Do we have $\left[x \in W^\perp, \alpha \in \mathbb{R}\right] \Rightarrow \left[\alpha x \in W^\perp\right]$?
   Indeed.

$\square$

**Discovery 3.2**

Suppose an $(V, \langle \cdot, \cdot \rangle)$ is an inner product space and $W \subseteq V$ is a linear subspace, consider the new linear subspace $W^\perp$.

1. Suppose $w_1, \ldots, w_k \in W$ are such that $\operatorname{span}(w_1, \ldots, w_k) = W$. Then for $x \in V$, we have $x \in W^\perp$ if and only if $x \perp w_i$ for all $i = 1, \ldots, k$.

Lecture 7 - Tuesday, May 21

**Result 3.1**

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and $W \subseteq V$ be a linear subspace. Then $W \cap W^\perp = \{0_V\}$.

*Proof.* The "$\supseteq$" direction is easy. For "$\subseteq$", we know that suppose $x \in W \cap W^\perp$, thus $x \in W$ and $x \in W^\perp$, which means that $x \perp x$, which suggests that $x = 0_V$. □

**Example 3.1**

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Take $W_1 = \{0_V\}$, then $W_1^\perp = V$. Take $W_2 = V$, then $W_2^\perp = \{0_V\}$ because it again, is perpendicular to itself.

## 3.1 The "x = w + q" Decomposition

**Theorem 3.1: The "x = w + q" Decomposition**

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space with $\dim(V) = n \ll \infty$. Let $W \subseteq V$ be a linear subspace of $V$. Then every $x \in V$ can be decomposed as a sum $x = w + q$ with $w \in W$ and $q \in W^\perp$. Moreover, this decomposition is unique.

*Proof.*

1. (*Existence*)
   We would assume that $W$ is a proper subspace of $V$ with $\dim(W) = m$ such that $1 \leq m < n$ (The two special cases are shown in the example already). Pick an orthonormal basis, $z_1, \ldots, z_m$ for $W$ and extend it to an orthonormal basis for $V$ represented as $z_1, \ldots, z_m, z_{m+1}, \ldots, z_n$. We can observe that $z_{m+1}, \ldots, z_n \in W^\perp$. So now take an $x \in V$, which we wan to decompose as $x = w + q$ for $w \in W$ and $q \in W^\perp$. Write $x$ in terms of the orthonormal basis, we get

   $$x = \underbrace{\alpha_1 z_1 + \cdots \alpha_m z_m}_{w \in W} + \underbrace{\alpha_{m+1} z_{m+1} + \cdots + \alpha_n z_n}_{q \in W^\perp}$$

   thus we are done proving the existence.

2. (*Uniqueness*)
   Let $x \in V$ and can be written as $x = w + q = w' + q'$ with $w, w' \in W$ and $q, q' \in W^\perp$. We have

   $$w - w' = q' - q := u$$

   which suggests that $u = 0_V$ since $u \in W$ and $u \in W^\perp$. Therefore, it must be the case that $w = w'$ and $q = q'$.

□

## 3.2 Orthogonal Projection

**Definition 3.2: Orthogonal Projection**

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space with $\dim(V) = n \ll \infty$. Let $W \subseteq V$ be a linear subspace and $x \in V$. In the unique decomposition $x = w + q$ with $w \in W$ and $q \in W^\perp$, the vector $w \in W$ is called the **orthogonal projection** of $x$ onto $W$. Denoted as $P_W(x)$. In short, the projection has the following property.

$$P_W(x) \in W \qquad \text{and} \qquad x - P_W(x) \in W^\perp$$

**Proposition 3.1**

There is another way to characterize what $P_W(x)$ is. Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space with $\dim(V) = n \ll \infty$. Let $W \subseteq V$ be a linear subspace and $x \in V$. Let $w_0 = P_W(x) \in W$, then

$$\|x - w_0\| \leq \|x - w\| \qquad \forall\, w \in W$$

In fact, we have

$$\|x - w_0\| < \|x - w\| \qquad \forall\, w \in W, w \neq w_0$$

Lecture 8 - Wednesday, May 22

*Proof.* STP the second inequality. Fix $w \neq w_0$ in $W$, write

$$x - w = (x - w_0) + (w_0 - w)$$

We observe that

$$\underbrace{x - w_0}_{\in W^\perp} \perp \underbrace{w_0 - w}_{\in W}$$

and this implies that

$$\begin{aligned} &\|x - w\|^2 = \|x - w_0\|^2 + \underbrace{\|w_0 - w\|^2}_{\neq 0} \\ \Rightarrow\quad &\|x - w\|^2 > \|x - w_0\|^2 \\ \Rightarrow\quad &\|x - w\| > \|x - w_0\| \end{aligned}$$

□

# 4 Orthogonal Projection as a Linear Operator

**Definition 4.1: Orthogonal Projection onto Subspace**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space with $\dim(V) = n \ll \infty$ and $W \subseteq V$ is a linear subspace. The function

$$\begin{aligned} P_W : V &\to V \\ x &\mapsto P_W(x) \end{aligned}$$

is called the **orthogonal projection** onto the subspace $W$.
**Remark:** Note the target space of $P_W$ is taken to be all of $W$.

### 4.0.1 Linearity of the Orthogonal Projection Map $P_W$

**Proposition 4.1: Linearity of the Orthogonal Projection Map $P_W$**

For $(V, \langle \cdot, \cdot \rangle)$ and $W \subseteq V$ described as above. The map $P_W$ is linear (hence is a linear operator on $V$).

*Proof.* **Claim 1**: $P_W(x_1 + x_2) = P_W(x_1) + P_W(x_2)$ for all $x_1, x_2 \in V$.

Indeed, we have $P_W(x_1), P_W(x_2) \in W$, and therefore

$$(x_1 + x_2) - (P_W(x_1) + P_W(x_2)) = (x_1 - P_W(x_1)) + (x_2 - P_W(x_2)) \in W^\perp$$

Hence $x_1 + x_2$ has the property which determines uniquely what $P_W(x_1 + x_2)$ is. In particular, we have $P_W(x_1 + x_2) = P_W(x_1) + P_W(x_2)$.

**Claim 2**: $P_W(\alpha x) = \alpha P_W(x)$ for $\alpha \in \mathbb{R}$ and all $x \in V$.

Exercise. □

## 4.1 Subspaces related to $P_W$

**Definition 4.2: Null and Range (Review from Math146)**

Let $L : V_1 \to V_2$ be a linear map between two vector spaces. Then define

$$\begin{aligned} \text{Null}(L) &= \{x \in V_1 : L(x) = 0_{V_2}\} \\ &\subseteq V_1 \text{ linear subcpace} \\ \text{Ran}(L) &= \{y \in V_2 : \exists\, x \in V_1 \text{ such that } L(x) = y\} \end{aligned}$$

**Proposition 4.2**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space with $\dim(V) = n \ll \infty$ and $W \subseteq V$ is a linear subspace. Consider the linear operator $P_W : V \to V$. Then

1. $\text{Null}(P_W) = W^\perp$
2. $\text{Ran}(P_W) = W$

We can also review $W$ from $P_W$, in particular,

$$W = \{x \in V : P_W(x) = x\}$$

*Proof. Part one*:

It is easy to find that $W \subseteq \{x \in V : P_W(x) = x\}$. Now we pick any $x \in W^\perp$, then we can write $x$ as

$$x = \underbrace{0_V}_{\in W} + \underbrace{x}_{\in W^\perp}$$

which gives the unique decomposition pf $x = w + q$ with $w \in W$ and $q \in W^\perp$. Hence $P_W(x) = 0_V$. This suggests that $x$ is indeed in the nullspace of $P_W$.

*Part two and three*:

We check that

$$W \subseteq \{x \in V : P_W(x) = x\} \subseteq \text{Ran}(P_W) \subseteq W$$

$\text{Ran}(P_W) \subseteq W$ is trivial from the definition. Additionally, $\{x \in V : P_W(x) = x\} \subseteq \text{Ran}(P_W)$ is clear as well. Therefore, essentially, we want to show that $W \subseteq \{x \in V : P_W(x) = x\}$. Pick $x \in W$, we have

$$x = \underbrace{x}_{\in W} + \underbrace{0_V}_{\in W^\perp}$$

which gives the unique decomposition. Hence $P_W(x) = x$, which is indeed in the set $\{x \in V : P_W(x) = x\}$. ☐

Lecture 9 - Friday, May 24

## 4.2 Pop Up Quiz

**Exercise:** For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{R}$. For any subset $A \subseteq V$, define

$$A^\perp := \{x \in V : x \perp a, \forall\, a \in A\} = \{x \in V : \langle x, a \rangle = 0, \forall\, a \in A\}$$

1. Let $A_1, A_2 \subseteq V$ such that $A_1 \subseteq A_2$, prove that $A_1^\perp \supseteq A_2^\perp$.
2. Prove that for every subset $A \subseteq B$, we have $A^\perp = (\text{span}(A))^\perp$

*Proof.* 1. We consider an arbitrary $x \in A_2^\perp$, thus we have

$$\langle x, a_2 \rangle = 0 \qquad \forall\, a_2 \in A_2$$

Since we are given that $A_1 \subseteq A_2$, thus we have

$$\langle x, a_1 \rangle = 0 \qquad \forall\, a_1 \in A_1 \subseteq A_2$$

which suggests that $x \in A_1^\perp$, thus we have

$$A_1^\perp \supseteq A_2^\perp$$

2. From part (a), we know that $A \subseteq \text{span}(A)$, thus we have $A^\perp \supseteq (\text{span}(A))^\perp$. Moreover, consider any $x \in A^\perp$, we know that $x \perp a$ for all $a \in A$, thus by the bilinearity of inner product, we know that $x$ is orthogonal to any linear combination of all the elements $a \in A$. Therefore, we can deduce that $A^\perp \subseteq (\text{span}(A))^\perp$. Thus we can conclude that

$$A^\perp = (\text{span}(A))^\perp$$

☐

### 4.2.1 $P_W + P_{W^\perp} = I$

**Proposition 4.3:** $P_W + P_{W^\perp} = I$

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{R}$ with $\dim(V) \ll \infty$. Let $W \subseteq V$ be a linear subspace, and consider the linear operatorss $P_W : V \to V$ and $P_{W^\perp} : V \to V$. then

$$P_W(x) + P_{W^\perp}(x) = x \qquad \forall\, x \in V$$

In other words, we have

$$P_W + P_{W^\perp} = I$$

where $I$ is the identity operator.

**Corollary 4.1**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{R}$ with $\dim(V) \ll \infty$. Let $W \subseteq V$ be a linear subspace. Then

$$(W^\perp)^\perp = W$$

*Proof.* Observe that

$$P_{W^\perp} + P_{(W^\perp)^\perp} = I = P_W + P_{W^\perp}$$

Then

$$P_{(W^\perp)^\perp} = \mathrm{Ran}(P_{(W^\perp)^\perp}) = P_W = W$$

Done. $\square$

Lecture 10 - Monday, May 27

*Proof.* of Proposition

Recall the description of $P_W(x)$: it is the '$w$' part in the unique decomposition $x = w + q$ with $w \in W$ and $q \in W^\perp$. For convenience, denote $W^\perp = Y$, then similarly, $P_Y(x)$ is the '$y$' component in $x = y + z$ with $y \in Y$ and $z \in Y^\perp$.

1. *Claim 1*: $Y^\perp \supseteq W$
   This is clear since we know that $Y^\perp = (W^\perp)^\perp \supseteq W$.

2. *Claim 2*: Let $x \in V$ and consider the unique decomposition $x = w + q$ with $w \in W$ and $q \in W^\perp$. Then writing $x = q + w$ gives the unique decomposition $x = y + z$ with $y \in Y$ and $z \in Y^\perp$.
   In the writing $x = q + w$, we indeed have $q \in W^\perp$, so $w \in W$ implies that $w \in Y^\perp$ (using claim 1).

3. *Claim 3*: For every $x \in V$, we have

   $$x = P_W(x) + P_Y(x)$$

   Consider the decomposition $x = w + q$, say that $P_W(x) = w$. Now write the same decomposition as in claim 2, $x = q + w = y + z$ in connection to $Y$, saying $P_Y(x) = q$. Hence $x = w + q = P_W(x) + P_Y(x)$ as claimed.

$\square$

**Discovery 4.1**

Where did we really use the fact that $\dim(V) \ll \infty$?

**Solution:** We did that when we prove the unique decomposition of every $x \in V$ as $x = w + q$ with $w \in W$ and $q \in W^\perp$. $\square$

# 5 The Algebra $\mathcal{L}(V)$ of Linear Operators on $V$, and its *-operation

**Definition 5.1: Algebra**

An **algebra** over $\mathbb{R}$ is a vector space $\mathcal{A}$ over $\mathbb{R}$ which is endowed with an operation of multiplication: have $a_1a_2 \in \mathcal{A}$ defined for every $a_1, a_2 \in \mathcal{A}$ such that a list of axioms is satisfied:

1. $(a_1a_2)a_3 = a_1(a_2a_3)$ for all $a_1, a_2, a_3 \in \mathcal{A}$;
2. $a_1(a_2 + a_3) = a_1a_2 + a_1a_3$;
3. $(a_1 + a_2)a_3 = a_1a_3 + a_2a_3$;
4. $(\alpha a_1)a_2 = a_1(\alpha a_2) = \alpha a_1a_2$ for all $\alpha \in \mathbb{R}$.

Lecture 11 - Wednesday, May 29

**Definition 5.2: Commutative**

Let $\mathcal{A}$ be an algebra as above. If $a_1a_2 = a_2a_1$ for all $a_1, a_2 \in \mathcal{A}$, then $\mathcal{A}$ is **commutative**.

**Definition 5.3: Unit and Unital**

Let $\mathcal{A}$ be an algebra as above. An element $u \in \mathcal{A}$ is said to be **unit** if

$$ua = a = au \qquad \forall\, a \in \mathcal{A}$$

If such unit exists, then we say $\mathcal{A}$ is a **unital** algebra.

**Discovery 5.1**

If $\mathcal{A}$ is a unital algebra, then its unit $u$ is uniquely determined.

*Proof.* Suppose we have two units, just multiply them together. □

**Example 5.1: of Algebra**

The set

$$A := \{f : [0,1] \to \mathbb{R} : f \text{ is continuous}\}$$

is an algebra, where the operations are defined pointwise. e.g. multiplication: for $f, g \in A$

$$f \cdot g : [0,1] \to \mathbb{R}$$
$$(f \cdot g)(x) = f(x) \cdot g(x) \qquad \forall\, x \in [0,1]$$

This algebra is unital and commutative, where the unit is the function

$$\begin{aligned} 1_A &: [0,1] \to \mathbb{R} \\ 1_A(x) &= 1 \qquad \forall\, x \in [0,1] \end{aligned}$$

**Example 5.2: of Algebra**

The set

$$B := \{f : [0,1] \to \mathbb{R} : f \text{ is continuous}, f(0) = 0\}$$

is an algebra. This is known for "Algebra of Brownion Paths". This algebra is commutative, but not unital

**Exercise:** Show that $B$ is not unital.

**Example 5.3: of Algebra**

Pick an $n \in \mathbb{N}$, let $A = \mathcal{M}_n(\mathbb{R})$. This is an algebra over $\mathbb{R}$ with operations seen in MATH146. This algebra is unital with

$$1_{\mathcal{M}_n(\mathbb{R})} = I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

but it is not commutative.

**Definition 5.4: Linear Operators?**

Let $V$ be a vector space over $\mathbb{R}$ ($\dim(V) \ll \infty$ is not required). Let

$$L(V) := \{T : V \to V : T \text{ is linear}\}$$

We have the operations on $L(V)$:

1. *Addition*:
   For $T_1, T_2 \in L(V)$, define

   $$\begin{aligned} T_1 + T_2 &: V \to V \\ (T_1 + T_2)(x) &= T_1(x) + T_2(x) \end{aligned}$$

2. *Scalar Multiplication*:
   For $T \in L(V)$ and $\alpha \in \mathbb{R}$, define

   $$\begin{aligned} \alpha T &: V \to V \\ (\alpha T)(x) &= \alpha T(x) \end{aligned}$$

3. *Multiplication*:

For $T_1, T_2 \in L(V)$, define

$$T_1T_2 : V \to V$$
$$(T_1T_2)(x) = T_1(T_2(x))$$

**Exercise:** $L(V)$ is an algebra over $\mathbb{R}$.

**Definition 5.5: (Unital) Algebra Homomorphism**

Let $A, B$ be algebras over $\mathbb{R}$. A function $\varphi : A \to B$ is said to be an **algebra homomorphism** when it respects the 3 operations:

1. $\varphi(a_1 + a_2) = \varphi(a_1) + \varphi(a_2)$
2. $\varphi(\alpha a) = \alpha\varphi(a)$
3. $\varphi(a_1a_2) = \varphi(a_1)\varphi(a_2)$

In addition to the above, if both $A$ and $B$ are unital and if $\varphi(1_A) = 1_B$, then $\varphi$ is a **unital algebra homomorphism**.

Lecture 12 - Friday, May 31

**Discovery 5.2**

If $\dim(V) = n \ll \infty$, then $\mathcal{L}(V) \cong \mathcal{M}_n(\mathbb{R})$, where if $\xi_1, \ldots, \xi_n$ is a basis, then for any $x_1, \ldots, x_n \in V$, there exists an operator $T \in \mathcal{L}(V)$, uniquely determined, such that

$$T(\xi_1) = x_1, \; \ldots \; , T(\xi_n) = x_n$$

which means that we have a bijection

$$\begin{aligned} \mathcal{L}(V) &\longleftrightarrow V^n \\ T &\longleftrightarrow (T(\xi_1), \ldots, T(\xi_n)) \end{aligned}$$

Go one step further, for $T \in \mathcal{L}(V)$, write the coordinate

$$\begin{aligned} T(\xi_1) &= \alpha_{11}\xi_1 + \alpha_{21}\xi_2 + \cdots + \alpha_{n1}\xi_n \\ T(\xi_2) &= \alpha_{12}\xi_1 + \alpha_{22}\xi_2 + \cdots + \alpha_{n2}\xi_n \\ &\vdots \\ T(\xi_n) &= \alpha_{1n}\xi_1 + \alpha_{2n}\xi_2 + \cdots + \alpha_{nn}\xi_n \end{aligned}$$

This creates a matrix

$$A_T = [\alpha_{ij}]_{1 \le i,j \le n} \in \mathcal{M}_n(\mathbb{R})$$

where the coordinates of $T(\xi_j)$ gives the column $j$ of $A_T$

**Result 5.1**

Notice we have

$$T(\xi_j) = \sum_i \alpha_{ij}\xi_i$$

The above creates $A_T \in \mathcal{M}_n(\mathbb{R})$, called **the matrix of** $T$ in the basis $\xi_1, \ldots, \xi_n$. Thus re-shapes the bijection into a bijection defined as following:

$$\varphi : \mathcal{L}(V) \to \mathcal{M}_n(\mathbb{R})$$
$$\varphi(T) = A_T$$

which is a unital algebra homomorphism.

**Example 5.4**

Suppose $T(x, y) = (2x + 3y, 5x - y)$, thus we have

$$\begin{cases} T(1,0) & = (2,5) \\ T(0,1) & = (3,-1) \end{cases} \quad \Rightarrow \quad \begin{bmatrix} 2 & 3 \\ 5 & -1 \end{bmatrix}$$

Thus we have the result $\begin{bmatrix} 2 & 3 \\ 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = T(x, y)$.

Now move to the setting where $V$ also has an inner product (i.p.s.).

Remark: in same notation as for preceding remark, assume now that $V$ (with $\dim(V) = n \ll \infty$) is an inner product space (with inner product $\langle \cdot, \cdot \rangle$) and assume that the fixed basis $\xi_1, \ldots, \xi_n$ is an orthonormal basis for $V$. Observe a specific formula for the entries of $A_T$. More precisely, let $T \in \mathcal{L}(V)$ and let $A_T = [\alpha_{ij}]_{1 \le i,j \le n}$ be the matrix associated to $T$ in the basis $\xi_1, \ldots, \xi_n$, then we have

$$\alpha_{ij} = \langle T(\xi_j), \xi_i \rangle \qquad \forall\, 1 \le i, j \le j$$

*Proof.* We know that

$$T(\xi_j) = \alpha_{1j}\xi_1 + \alpha_{2j}\xi_2 + \cdots + \alpha_{nj}\xi_n$$

Take inner product with $\xi_i$ on both sides

$$\begin{aligned} \langle T(\xi_j), \xi_i \rangle &= \langle \alpha_{1j}\xi_1 + \alpha_{2j}\xi_2 + \cdots + \alpha_{nj}\xi_n, \xi_i \rangle \\ &= \alpha_{1j}\langle \xi_1, \xi_i \rangle + \alpha_{2j}\langle \xi_2, \xi_i \rangle + \cdots + \alpha_{nj}\langle \xi_n, \xi_i \rangle \\ &= \alpha_{ij} \end{aligned}$$

□

## 5.1 Adjoint Always Exists for Finite Dimension and is Uniquely Determined

### 5.1.1 Adjoint for an operator $T \in \mathcal{L}(V)$

**Proposition 5.1: Adjoint for an operator $T \in \mathcal{L}(V)$**

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space where $\dim(V) = n \lll \infty$. Let $T \in \mathcal{L}(V)$. There exists $S \in \mathcal{L}(V)$, uniquely determined, such that $\langle T(x), y \rangle = \langle x, S(y) \rangle$ for all $x, y \in V$.

*Proof.* Pick an orthonormal basis $\xi_1, \dots, \xi_n$ for $V$.

1. *Claim 1*: For an operator $S \in \mathcal{L}(V)$, we have that for all $x, y \in V$,

$$\langle T(x), y \rangle = \langle x, S(y) \rangle \iff \langle T(\xi_i), \xi_j \rangle = \langle \xi_i, S(\xi_j) \rangle \quad \forall\, i \leq i, j \leq n$$

   *Verification of Claim 1*: Indeed, we have the forward direction because the right hand side is just a special case for the left hand side. For the backward direction, for $x, y \in V$, write

$$x = \sum_{i=1}^{n} a_i \xi_i \quad \text{and} \quad y = \sum_{j=1}^{n} b_j \xi_j$$

   with $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$. Note that

$$T(x) = T\left(\sum_{i=1}^{n} a_i \xi_i\right) = \sum_{i=1}^{n} a_i T(\xi_i)$$

   Then

$$\begin{aligned}
\langle T(x), y \rangle &= \langle \sum_{i=1}^{n} a_i T(\xi_i), \sum_{j=1}^{n} b_j \xi_j \rangle = \sum_{i,j} \langle T(\xi_i), \xi_j \rangle = \sum_{i,j} \langle \xi_i, S(\xi_j) \rangle \\
&= \langle \sum_{i=1}^{n} a_i \xi_i, \sum_{j=1}^{n} b_j S(\xi_j) \rangle = \langle \sum_{i=1}^{n} a_i \xi_i, S\left(\sum_{j=1}^{n} b_j \xi_j\right) \rangle \\
&= \langle x, S(y) \rangle
\end{aligned}$$

2. *Claim 2*: Let $A_T = [\alpha_{ij}]_{i \leq i, j \leq n}$ be the matrix associated to $T$ with respect to $\xi_1, \dots, \xi_n$. Let $S$ be some operator in $\mathcal{L}(V)$, and let $A_S = [\beta_{ij}]_{1 \leq i, j \leq n}$ be the matrix of $S$. Then

$$S \text{ satisfies above RHS} \iff A_S = A_T^T$$

   *Verification of Claim 2*: Recll that

$$\alpha_{ij} = \langle T(\xi_j), \xi_i \rangle, \qquad \beta_{ij} = \langle S(\xi_j), \xi_i \rangle$$

   from the remark on the previous page, so then

$$\begin{aligned}
S \text{ satisfies above RHS} &\iff \underbrace{\langle T(\xi_i), \xi_j \rangle}_{\alpha_{ji}} = \underbrace{\langle \xi_i, S(\xi_j) \rangle}_{\beta_{ij}} \quad \forall\, 1 \leq i, j \leq n \\
&\iff A_S = A_T^T
\end{aligned}$$

Conclusion: Recall that the map $\mathcal{L}(V) \to \mathcal{M}_n(\mathbb{R}), X \mapsto A_X$ is a bijection. Using this bijection, we see that there exists $S \in \mathcal{L}(V)$, uniquely determined, such that $A_S = A_T^T$. Thus $S$ has the property stated in the Proposition and is uniquely determined. $\square$

Lecture 13 - Monday, Jun 3

**Discovery 5.3**

Proof of Proposition (5.1) shows that when we fix an orthonormal basis $\xi_1, \ldots, \xi_n$ for $V$, then the operation

$$T \to T^* \text{ on } \mathcal{L}(V)$$

becomes

$$A_T \to A_T^T \text{ on } \mathcal{M}_n(\mathbb{R})$$

**Definition 5.6: Adjoint**

Suppose $(V, \langle \cdot, \cdot \rangle)$ an inner product space with $\dim(V) = n \ll \infty$ and $T \in \mathcal{L}(V)$ as in the above Proposition. The uniquely determined $S \in \mathcal{L}(V)$ is called the **adjoint of** $T$, and is denoted as $T^*$.

*Proof.* We verify that $S$ is uniquely determined: Suppose $S, S' \in \mathcal{L}(V)$ both satisfy the equation, hence for any $x, y \in V$, we have

$$\langle T(x), y \rangle = \langle x, S(y) \rangle = \langle x, S'(y) \rangle$$

Let $D = S - S' \in \mathcal{L}(V)$, then we have for every $x, y \in V$, we have

$$\langle x, D(y) \rangle = \langle x, (S - S')(y) \rangle = \langle x, S(y) \rangle - \langle x, S'(y) \rangle = 0$$

Fix for the moment a vector $y \in V$ we have

$$\langle x, D(y) \rangle = 0 \qquad \forall\, x \in V$$

In particular, the latter equality holds for $x = D(y)$, thus we have

$$\langle D(y), D(y) \rangle = 0 \implies D(y) = 0_V$$

Now unfix $y$, we got $D(y) = 0$ for all $y \in V$, thus $S = S'$. $\square$

Lecture 14 - Wednesday, Jun 5

**Result 5.2**

It is easy to find examples that if $\dim(V) = \infty$, then there exists $T \in \mathcal{L}(V)$ such that $T$ does not have adjoint.

**Proposition 5.2**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space (no assumption for the dimension). Let

$$\mathcal{A} = \{T \in \mathcal{L}(V) : T \text{ has an adjoint}\}$$

Then $\mathcal{A}$ has the following stability properties:

1. $T_1, T_2 \in \mathcal{A}$, then $T_1 + T_2 \in \mathcal{A}$ and $(T_1 + T_2)^* = T_1^* + T_2^*$.
2. $T \in \mathcal{A}$ and $\alpha \in \mathbb{R}$, then $\alpha T \in \mathcal{A}$ and $(\alpha T)^* = \alpha T^*$.

3. $T_1, T_2 \in \mathcal{A}$, then $T_1T_2 \in \mathcal{A}$ and $(T_1T_2)^* = T_2^*T_1^*$.

4. $T \in \mathcal{A}$, then $T^* \in \mathcal{A}$ and $(T^*)^* = T$.

*Proof.* Verifications for them are similar to each other, so here we will only check one of them, namely the third one. Pick $T_1, T_2 \in \mathcal{A}$, and let

$$S := T_2^*T_1^* \in \mathcal{L}(V)$$

Observe that for every $x, y \in V$, we have

$$\langle (T_1T_2)(x), y\rangle = \langle T_1(T_2(x)), y\rangle = \langle T_2(x), T_1^*(y)\rangle = \langle x, T_2^*(T_1^*(y))\rangle = \langle x, S(y)\rangle$$

gg. □

**Discovery 5.4**

Notice that

$$\mathcal{A} = \{T \in \mathcal{L}(V) : T \text{ has an adjoint}\}$$

is a **subalgebra** of $\mathcal{L}(V)$. It is a **unital** subalgebra, meaning that $I \in \mathcal{A}$, where $I$ is the unit of $\mathcal{A}$. It is immediate to check that $I^*$ exists, and $I^* = I$.

**Definition 5.7: Normal**

$T \in \mathcal{L}(V)$ is called **normal** if $TT^* = T^*T$.

**Example 5.5: An infinite dimensional analogue for $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$.**

Look at

$$\mathbb{R}^\infty_{\text{fin}} = \{(x_1, x_2, \ldots, x_n, \ldots) : x_n \in \mathbb{R}, \forall\, n \in \mathbb{N}, \exists\, n_0 \in \mathbb{N} \text{ s.t. } x_n = 0 \forall\, n > n_0\}$$

For $x = (x_1, x_2, \ldots, x_n, \ldots)$ and $y = (y_1, y_2, \ldots, y_n, \ldots)$ and $\alpha \in \mathbb{R}$, we put

$$\begin{aligned} x + y &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots) \\ \alpha x &= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n, \ldots) \end{aligned}$$

As the inner product for $x, y$ as above is defined as

$$\langle x, y\rangle = \sum_{n}^{\infty} x_n y_n = \sum_{n=1}^{n_0} x_n y_n$$

In this way, we have created an inner product space, $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ (exercise).

Useful Remark: Consider linear subspaces $V_1, V_2, \dots, V_n, \dots$ of $\mathbb{R}^n$ described as

$$
\begin{aligned}
V_1 &= \{(x_1, 0, 0, \dots, 0, \dots) : x_1 \in \mathbb{R}\} \\
V_2 &= \{(x_1, x_2, 0, \dots, 0, \dots) : x_1, x_2 \in \mathbb{R}\} \\
&\vdots \\
V_n &= \{(x_1, x_2, x_3, \dots, x_n, \dots) : x_1, \dots, x_n \in \mathbb{R}\}
\end{aligned}
$$

Therefore we have

$$V_1 \subseteq V_2 \subseteq \cdots \subseteq V_n \subseteq \cdots \quad \text{and} \quad \bigcup_{n=1}^{\infty} V_n = \mathbb{R}^\infty_{\text{fin}}$$

Moreover, for every $i \in \mathbb{N}$, let

$$e_i = (0, \dots, 0, \underset{|\ i^{th} \text{ position}}{1}, 0, \dots, 0)$$

Then $\{e_i\}_{i \in \mathbb{N}}$ form a linear basis for $\mathbb{R}^\infty_{\text{fin}}$ with

$$\text{span}\{e_1, \dots, e_n\} = V_n \qquad \forall\, n \in \mathbb{N}$$

**Result 5.3**

For any sequence $(x_n)$ in $\mathbb{R}^\infty_{\text{fin}}$, there exists $T \in \mathcal{L}(\mathbb{R}^\infty_{\text{fin}})$ such that $T(e_n) = x_n \quad \forall\, n \in \mathbb{N}$.

**Exercise:** Let $T \in \mathcal{L}(\mathbb{R}^\infty_{\text{fin}})$ be uniquely determined by the requirements that $T(e_n) = e_1 + e_2 + \cdots + e_n$ for all $n \in \mathbb{N}$. Prove that $T$ does not have an adjoint $T^*$.

Lecture 15 - Friday, Jun 7

# 6 Some Special Classes of Linear Operators

**Definition 6.1: Self-adjoint, Normal, Projection, Unitary and Isometry (Co-isometry)**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{R}$. Let $\mathcal{A} = \{T \in \mathcal{L}(V) : T \text{ has an adjoint}\}$. For $T \in \mathcal{A}$, we say that

1. $T$ is **self-adjoint** if $T = T^*$;
2. $T$ is **normal** if $T^*T = TT^*$;
3. $T$ is a **projection** if $T = T^* = T^2$ (idemponent);
4. $T$ is a **unitary** if $T^*T = I = TT^*$;
5. $T$ is a **isometry** if $T^*T = I$ (without the requiring the other equality);
6. $T$ is a **co-isometry** if $TT^* = I$ (without the requiring the other equality)

**Discovery 6.1: Unitaries and Isometries**

We notice that

$$\begin{aligned} T \text{ is unitary} &\iff T \text{ is invetible with } T^{-1} = T^* \\ T \text{ is an isometry} &\iff T \text{ is an inverse on left for } T \\ T \text{ is a co-isometry} &\iff T \text{ is an inverse on right for } T \end{aligned}$$

**Theorem 6.1**

In the case when $\dim(V) = n \ll \infty$, we know from Math146 that $T$ is isometry (or co-isometry) implies that $T$ is unitary. In fact, we know that $\mathcal{A} = \mathcal{L}(V)$.

*Proof.* We fix an orthonormal basis $\xi_1, \dots, \xi_n$ for $V$ and we use the isomorphism of unital algebras:

$$\begin{aligned} \mathcal{L}(V) &\to \mathcal{M}_n(\mathbb{R}) \\ T &\to A_T \end{aligned}$$

so then if $T$ is isometry, this implies that $T^*T = I$, which further suggests that $A_{T^*} \cdot A_T = A_{T^*T} = A_I = I_n$ with $I_n$ being the identatity matrix $\in \mathcal{M}_n(\mathbb{R})$. Math146 then tells us that we also have

$$A_T \cdot A_{T^*} = I_n \quad \Rightarrow \quad A_{TT^*} = A_T \cdot A_{T^*} = I_n$$

This implies that $TT^* = I$. We can now conclude that $T$ is a co-isometry as well. $\square$

## 6.1 Isometry, but not Unitary

**Theorem 6.2**

The above result does not hold for the case when $\dim(V) = \infty$. We may have non-unitary isometries!

**Example 6.1**

Let $v = \mathbb{R}^\infty_{\text{fin}}$. Recall that

$$\mathbb{R}^\infty_{\text{fin}} = \{(x_1, x_2, \dots, x_n, \dots) : x_n \in \mathbb{R}, \forall\, n \in \mathbb{N}, \exists\, n_0 \in \mathbb{N} \text{ s.t. } x_n = 0\, \forall\, n > n_0\}$$

Recall that we have a orthonormal basis $e_1, e_2, \dots, e_n, \dots$, then given any sequence of vectors $v_1, v_2, \dots, v_n, \dots \in \mathbb{R}^\infty_{\text{fin}}$, there exists an operator $T \in \mathcal{L}(\mathbb{R}^\infty_{\text{fin}})$, uniquely determined, such that $T(e_1) = v_1, \dots, T(e_n) = v_n, \dots$.

Warning:

$$\mathcal{A} = \{T \in \mathcal{L}(\mathbb{R}^\infty_{\text{fin}}) : T \text{ has an adjoint}\} \subset \mathcal{L}(\mathbb{R}^\infty_{\text{fin}})$$

Let $S \in \mathcal{L}(\mathbb{R}^\infty_{\text{fin}})$ be defined via requirements that $S(e_1) = e_2, S(e_2) = e_3, \dots, S(e_n) = e_{n+1}, \dots$.
Let $R \in \mathcal{L}(\mathbb{R}^\infty_{\text{fin}})$ be defined via requirements that $R(e_1) = 0, R(e_2) = e_1, \dots, R(e_n) = e_{n-1}, \dots$.
Here we have several claims:

1. *Claim 1:* $\langle S(e_i), e_j\rangle = \langle e_i, R(e_j)\rangle$ for all $i, j \in \mathbb{N}$;
2. *Claim 2:* $R, S \in \mathcal{A}$ and $R = S^*$;
3. *Claim 3:* $S^*S = I \neq SS^*$

Lecture 16 - Monday, Jun 10

1. *Verification of claim 1:*
   If $j = 1$, then both sides are 0:

$$LHS = \langle S(e_i), e_j\rangle = \langle e_{i+1}, e_1\rangle = 0 = \langle e_i, R(e_j)\rangle = \langle e_i, R(e_1)\rangle = RHS$$

   If $j \neq 1$, then

$$LHS = \langle e_{i+1}, e_j\rangle = \begin{cases} 1 & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases} \quad RHS = \langle e_i, e_{j-1}\rangle = \begin{cases} 1 & \text{if } i = j-1 \\ 0 & \text{otherwise} \end{cases}$$

2. *Verification of claim 2:*
   Write $x = \sum_{i=1}^n \alpha_i e_i$ and $y = \sum_{j=1}^m \beta_j e_j$ for some $n, m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m \in \mathbb{R}$. Then

$$S(x) = \sum_{i=1}^n \alpha_i S(e_i) \qquad R(y) = \sum_{j=1}^m \beta_j R(e_j)$$

Then

$$\begin{aligned}
\langle S(x), y\rangle &= \langle \sum_{i=1}^{n} \alpha_i S(e_i), \sum_{j=1}^{m} \beta_j e_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i\beta_j \langle S(e_i), e_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i\beta_j \langle e_i, R(e_j)\rangle \\
&= \langle \sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{m} \beta_j R(e_j)\rangle = \langle x, R(y)\rangle
\end{aligned}$$

3. *Verification of claim 3:*
   STP that $RS$ and $I$ agree on the basis, or $e_i$ for all $i \in \mathbb{N}$. Indeed, we have

   $$(RS)(e_i) = R(S(e_i)) = R(e_{i+1}) = e_i = I(e_i)$$

   Now we wish to show that $SS^* \neq I$:

   $$(SR)(e_i) = S(R(e_i)) = S(0) = 0 \neq I(e_i)$$

Therfore, we can conclude that $S$ is a non-unitary isometry, while $R = S^*$ is a non-unitary co-isometry. □

**Discovery 6.2**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{R}$. Let $T \in \mathcal{L}(V)$ be an isometry, we have

1. $\|T(x)\| = \|x\|, \forall\, x \in V$;

   $$\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*(T(x))\rangle = \langle x, x\rangle = \|x\|^2 \quad \square$$

2. For the above $T$ (we know $T^*T = I$, but may have $TT^* \neq I$), we observe that we can be sure that $TT^*$ is a projection.

   $$(TT^*)^* = (T^*)^*T^* = TT^* \text{ and } (TT^*)(TT^*) = T(T^*T)T^* = TT^*$$

**Algorithm 6.1: Tricks for checking equalities of operators**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space and $Z \in \mathcal{L}(V)$, suppose we have

$$\langle Z(x), y\rangle = 0 \quad \forall\, x, y \in V$$

Then $Z$ is the zero-operator on $V$, or $(Z(x) = 0_V, \forall\, x \in V)$.
Suppose we know that $Z$ is self-adjoint, $Z = Z^*$, and that

$$\langle Z(x), x\rangle = 0, \forall\, x \in V$$

then $Z$ is the zero-operator on $V$.

Lecture 17 - Wednesday, Jun 12

The reason why we can conclude $Z(x) = 0_V$ for all $x \in V$ if we have $\langle Z(x), x\rangle = 0$ for all $x \in V$ is that we have

$$\begin{aligned}
0 &= \langle Z(x+y), x+y\rangle \\
&= \langle Z(x)+Z(y), x+y\rangle \\
&= \underbrace{\langle Z(x), x\rangle}_{=0} + \langle Z(x), y\rangle + \langle Z(y), x\rangle + \underbrace{\langle Z(y), y\rangle}_{=0} \\
&= \langle Z(x), y\rangle + \underbrace{\langle y, Z^*(x)\rangle}_{=\langle y, Z(x)\rangle} \\
&= 2\langle Z(x), y\rangle
\end{aligned}$$

Then use the above result, we know that $Z$ has to be the 0-operator.

**Result 6.1**

If we do not have $Z = Z^*$, then the second result does not hold, here is an counterexample for the result:
Say $V = \mathbb{R}^2$ with standard inner product. Let $Z : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $Z(x_1, x_2) = (-x_2, x_1)$.

**Proposition 6.1**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{R}$, $T \in \mathcal{L}(V)$, TFAE:

1. $T^*T = I$ (isometry);
2. $\|T(x)\| = \|x\|$ for all $x \in V$.

*Proof.* 1. $(\Longrightarrow)$
$\|T(x)\| = \|x\|$, $\forall\, x \in V$;

$$\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*(T(x))\rangle = \langle x, x\rangle = \|x\|^2$$

2. $(\Longleftarrow)$
Let $Z = T^*T - I \in \mathcal{L}(V)$, notice that thus we have

$$Z^* = (T^*T - I)^* = T^*T - I^* = T^*T - I$$

Now it suffices to prove that $\langle Z(x), x\rangle = 0$ for all $x \in V$, indeed, we have

$$\begin{aligned}
\langle Z(x), x\rangle &= \langle (T^*T)(x) - I(x), x\rangle \\
&= \langle T^*(T(x)) - x, x\rangle \\
&= \langle T^*(T(x)), x\rangle - \langle x, x\rangle \\
&= \langle T(x), T(x)\rangle - \|x\|^2 \\
&= \|T(x)\|^2 - \|x\|^2 = 0
\end{aligned}$$

Therefore, $Z$ is the 0-operator, which suggests that $T^*T = I$. $\square$

## 6.2 Interpretation of Isometry, Normal, and Projection

**Discovery 6.3**

Proposition (6.1) gives two faces of the notion of isometry:

1. It is the sheer algebra in $\mathcal{L}(V)$: $T^*T = I$;
2. Geometry of vectors: $\|T(x)\| = \|x\|$ for all $x \in V$.

Aside: "iso-metry" is a Greek word meaning "preserves distances":

$$\|T(x) - T(y)\| = \|T(x-y)\| = \|x-y\|$$

One has similar two face descriptions for normal operators and projection operators:

**Corollary 6.1: For normal operators**

1. It is the sheer algebra in $\mathcal{L}(V)$: $T^*T = TT^*$;
2. Geometry of vectors: $\|T(x)\| = \|T^*(x)\|$ for all $x \in V$.

*Proof.* HW4. □

**Corollary 6.2: For projection operators**

1. It is the sheer algebra in $\mathcal{L}(V)$: $T = T^* = T^2$;
2. Geometry of vectors: $T = P_W$ for a suitable linear subspace $W \subseteq V$.

*Proof.* HW4. □

# 7 Eigenvalues/ Eigenspaces for Self-adjoint Operators

This is the last section which we use vactor spaces over $\mathbb{R}$. The goal after this is we want to look at *normal operators* where we have to switch over to $\mathbb{C}$.

Lecture 18 - Friday, Jun 14

**Definition 7.1: Eigenvalues, Eigenspaces, and Eigenvectors**

For $V$ a vector space over $\mathbb{R}$ and $T \in \mathcal{L}(V)$, $\lambda \in \mathbb{R}$. We say that $\lambda$ is an **eigenvalue** for $T$ to mean that there exists $x \neq 0_V$ in $V$ such that $T(x) = \lambda x$. Equivalently, we have

$$\lambda \text{ is eigenvalue for } T \iff \text{Null}(T - \lambda I) \neq \{0_V\}$$

If $\lambda$ is an eigenvalue of $T$, then the linear subspace

$$E_\lambda := \text{Null}(T - \lambda I) \subseteq V$$

is called the **eigenspace** of $T$. The vectors in $E_\lambda \backslash \{0_V\}$ are called the **eigenvectors of** $T$ corresponding to the eigenvalue $\lambda$.

**Discovery 7.1**

Suppose that $0 < \dim(V) \ll \infty$, then we have a longer equivalence:

$$\lambda \text{ is eigenvalue for } T \iff \text{Null}(T - \lambda I) \neq \{0_V\} \iff \text{the operator } T - \lambda I \in \mathcal{L}(V) \text{ is not invertible}$$

Indeed, if $0 < \dim(V) \ll \infty$, then (from MATH146) for $R \in \mathcal{L}(V)$, we have

$$R \text{ invertible} \iff \text{Null}(R) = \{0_V\} \iff \text{Ran}(R) = V$$

**Discovery 7.2**

For general vector space $V$ over $\mathbb{R}$ and $T \in \mathcal{L}(V)$, we define

$$\text{Spectrum}(T) = \{\lambda \in \mathbb{R} : T - \lambda I \text{ not invertible}\}$$

Note that

$$\lambda \text{ is eigenvalue of } T \implies T - \lambda I \text{ not injective} \implies T - \lambda I \text{ not invertible}$$

For $\dim(V) \ll \infty$, the converse also holds (see above discovery), hence for finite dimensional $V$ we have

$$\text{Spectrum}(T) = \{\lambda \in \mathbb{R} : \lambda \text{ eigenvalue of } T\}$$

**Example 7.1**

Let $S \in \mathcal{L}(V)$ be the forward shift operator, then $0 \in \text{Spectrum}(S)$ ($S$ is not surjective, hence not invertible), but 0 is not an eigenvalue of $S$ ($S$ is injective, $\text{Null}(S - 0I) = \{0_V\}$).

**Example 7.2**

Let $V = \mathbb{R}^2$ with standard inner product and let $Z \in \mathcal{L}(V)$ be defined by

$$Z(x_1, x_2) = (-x_2, x_1) \qquad \forall\, (x_1, x_2) \in \mathbb{R}^2$$

then there is no $\lambda \in \mathbb{R}$ being an eigenvalue for $Z$.

*Proof.* Suppose $\lambda \in \mathbb{R}$ is an eigenvalue, and let $x \neq (0,0) \in \mathbb{R}^2$ be such that $Z(x) = \lambda x$. Now we have

$$0 = \langle Z(x), x \rangle = \langle \lambda x, x \rangle = \lambda \left\| x \right\|^2$$

which implies that $x = 0$, which is a contradiction. $\square$

The point of Lecture 7 is to prove the following fact:

**Theorem 7.1**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{R}$ with $0 < \dim(V) \ll \infty$, $T \in \mathcal{L}(V)$ such that $T = T^*$, then $T$ has eigenvalue in $\mathbb{R}$.

**Discovery 7.3**

When proving the "fact" just stated, we may assume that $T$ is invertible. (If $T$ is not, then $\text{Null}(T) \neq \{0_V\}$, hence $\lambda = 0$ will do as eigenvalue).

**Discovery 7.4**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{R}$ with $\dim(V) = n \in \mathbb{N}$, let $T \in \mathcal{L}(V)$. Then the set

$$\{\|T(x)\| : x \in V, \|x\| = 1\} \subseteq [0, \infty)$$

is a bounded subset of $[0, \infty)$, and has a maxmal number $\lambda$ in it. That is, there exists $x_0 \in V$ with $\|x_0\| = 1$ such that $\|T(x_0)\| = \lambda$. So $\lambda$ and $x_0$ are found such that

$$\|T(x)\| \leq \lambda \qquad \forall\, x \in V \text{ with } \|x\| = 1$$

*Proof.* Let $\xi_1, \ldots, \xi_n$ be an orthonormal basis for $V$. For $x \in V$ such that $\|x\| = 1$, we write $x = \sum_{i=1}^n t_i \xi_i$ for $t_1, \ldots, t_n \in \mathbb{R}$ and observe that $\sum_{i=1}^n t_i^2 = \|x\|^2 = 1$, hence $(t_1, \ldots, t_n)$ lies on the unit sphere of $\mathbb{R}^n$. We wish to maximize $\|T(x)\|$ (or equivalently $\|T(x)\|^2$). Note that we have $T(x) = \sum_{i=1}^n t_i T(\xi_i)$, so

$$\begin{aligned}\|T(x)\|^2 = \langle T(x), T(x)\rangle &= \langle \sum_{i=1}^n t_i T(\xi_i), \sum_{j=1}^n t_j T(\xi_j)\rangle \\ &= \sum_{i,j=1}^n t_i t_j \langle T(\xi_i), T(\xi_j)\rangle \\ &= \sum_{i,j=1}^n \alpha_{ij} t_i t_j\end{aligned}$$

where $\alpha_{ij} = \langle T(\xi_i), T(\xi_j)\rangle$. Hence, in coordinates, "maximizes $\{\|T(x)\|^2 : x \in V, \|x\| = 1\}$" becomes "maximizes $\{\sum_{i,j=1}^n \alpha_{ij} t_i t_j : (t_1, \ldots, t_n \in \mathbb{R}^n), \sum_i t_i^2 = 1\}$". The Extreme Value Theorem from calculus says that this is possible: find $(t_1^{(0)}, \ldots, t_n^{(0)}) \in \mathbb{R}^n$ with $[t_1^{(0)}]^2 + \cdots + [t_n^{(0)}]^2 = 1$, which achieves a maximal value of $\sum_{i,j=1}^n \alpha_{ij} [t_i^{(0)}][t_j^{(0)}]$. Then $x_0 := t_1^{(0)}\xi_1 + \cdots + t_n^{(0)}\xi_n$ has $\|x_0\| = 1$ and will achieve

$$\|T(x_0)\|^2 = \max\{\|T(x)\|^2 : x \in V, \|x\| = 1\}$$

hence also

$$\|T(x_0)\| = \max\{\|T(x)\| : x \in V, \|x\| = 1\} := \lambda$$

□

Lecture 19 - Monday, Jun 17

Towards of the proof of the Theorem (7.2), we use two lemmas.

**Lemma 7.1**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{R}$ with $\dim(V) = n \in \mathbb{N}$, let $T \in \mathcal{L}(V)$ such that $T = T^*$, we have

$$\|T(v)\| \leq \lambda \|v\| \qquad \forall\, v \in V$$

*Proof.* If $v = 0$, then $T(v) = 0_V$, and $\|T(v)\| = 0 = \lambda \|v\|$. Thus we may assume that $v \neq 0_V$, then we consider the vector

$$x = \frac{1}{\|v\|} v$$

which has the property that $\|x\| = 1$. Definition of $\lambda$ implies that $\|T(x)\| \leq \lambda$. And

$$T(x) = T\left(\frac{1}{\|v\|} v\right) = \frac{1}{\|v\|} T(v)$$

hence $\|T(xc)\| = \left\|\frac{1}{\|v\|} T(v)\right\| = \frac{1}{\|v\|} \|T(v)\|$, which further suggests that

$$\lambda \geq \frac{1}{\|v\|} \|T(v)\| \;\;\Rightarrow\;\; \|T(v)\| \leq \lambda \|v\|$$

which was desired. □

**Lemma 7.2**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{R}$ with $\dim(V) = n \in \mathbb{N}$, let $T \in \mathcal{L}(V)$ such that $T = T^*$ with

$$\lambda := \max\{\|T(x)\| : x \in V, \|x\| = 1\}$$

and $x_0 \in V$ such that $\|x_0\| = 1$ and $\|T(x_0)\| = \lambda$. Then $T^2(x_0) = \lambda^2 x_0$.

*Proof.* The idea is that $T^2(x_0)$ and $x_0$ will be found to be proportional, which happens because they satisfy Cauchy-Schwarz (1.1) with equality. Compute:

$$\begin{aligned}\langle T^2(x_0), x_0\rangle &= \langle T(T(x_0)), x_0\rangle \\ &= \langle T(x_0), T^*(x_0)\rangle \\ &= \langle T(x_0), T(x_0)\rangle = \|T(x)\|^2 = \lambda^2\end{aligned}$$

On the other hand, $\|x_0\| = 1$ and

$$\left\|T^2(x_0)\right\| = \left\|T(\underbrace{T(x_0)}_{v})\right\| \leq \lambda \|T(x_0)\| = \lambda \cdot \lambda = \lambda^2$$

hence $\left\|T^2(x_0)\right\| \leq \lambda^2$. Putting things together, we find that

$$\left\|T^2(x_0)\right\| \cdot \|x_0\| \leq \lambda^2 \cdot 1 = \langle T^2(x_0), x_0\rangle \leq |\langle T^2(x_0), x_0\rangle| \leq \left\|T^2(x_0)\right\| \cdot \|x_0\|$$

All this inequalities are now forced to be equalities. In particular, Cauchy-Schwarz also holds with equality, and this implies the existence of an $\alpha \in \mathbb{R}$ such that $T^2(x_0) = \alpha x_0$. Now we need to determine what $\alpha$ is. Recall that we have

$$\lambda^2 = \langle T^2(x_0), x_0\rangle = \langle \alpha x_0, x_0\rangle = \alpha \|x_0\|^2 = \alpha$$

which was desired. □

Lecture 20 - Wednesday, Jun 19

## 7.1 Either $\lambda$ or $-\lambda$ is an eigenvalue of $T$.

**Theorem 7.2**

For $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{R}$ with $\dim(V) = n \in \mathbb{N}$, let $T \in \mathcal{L}(V)$ such that $T = T^*$ with

$$\lambda := \max\{\|T(x)\| : x \in V, \|x\| = 1\}$$

Either $\lambda$ or $-\lambda$ is an eigenvalue of $T$.

*Proof.* Let $x_0 \in V$ be such that $\|x_0\| = 1$ and $\|T(x_0)\| = \lambda$. We saw from Lemma (7.2) that $T^2(x_0) = \lambda^2 x_0$. This means that $(T^2 - \lambda^2 I)(x_0) = 0_V$. Observe that we have $T^2 - \lambda^2 = (T + \lambda I)(T - \lambda I)$, hence we have

$$((T + \lambda I)(T - \lambda I))(x_0) = (T + \lambda I)[(T - \lambda I)(x_0)] = 0_V$$

Now we consider two cases:

1. *Case 1:* $(T - \lambda I)(x_0) = 0_V$
   this suggests that $T(x_0) = \lambda x_0$, thus $\lambda$ is an eigenvalue of $T$ with eigenvector $x_0 \neq 0$.

2. *Case 1:* $(T - \lambda I)(x_0) \neq 0_V$
   let $x_1 := (T - \lambda I)(x_0)$, this suggests that $T(x_1) = -\lambda x_1$, thus $-\lambda$ is an eigenvalue of $T$ with eigenvector $x_1 \neq 0$.

Therefore we conclude that either $\lambda$ or $-\lambda$ is an eigenvalue of $T$. □

# 8 Eigenvalues/ Eigenspaces for Self-adjoint Operator II

The goal for this section is to prove a stronger Theorem: for $T = T^* \in \mathcal{L}(V)$ with $\dim(V) = n \in \mathbb{N}$, one can find an orthonormal basis for $V$ consisting of eigenvectors of $T$.

**Definition 8.1: Invariant**

For $V$ a vector space over $\mathbb{R}$ and $T \in \mathcal{L}(V)$. A linear subspace $W \subseteq V$ is said to be **invariant** for $T$ when it has the property that for $w \in W$, $T(w) \in W$. When $W$ is invariant for $T$, we get to have a restricted operator $\check{T} \in \mathcal{L}(V), \check{T} : W \to W$ defined by

$$\check{T}(w) := T(w) \in W \qquad \forall\, w \in W$$

**Lemma 8.1**

Suppose $(V, \langle \cdot, \cdot \rangle)$ is an inner product space over $\mathbb{R}$, $T \in \mathcal{L}(V)$ such that $T^*$ exists with $T^* = T$. Let $W \subseteq V$ be a linear subspace that is invariant for $T$, then $W^\perp$ is also invariant for $T$.

*Proof.* Pick $y \in W^\perp$, for which we want to check if $T(y) \in W^\perp$. That is, we want to check if $\langle T(y), w \rangle = 0$ for all $w \in W$. And indeed, for every $w \in W$,

$$\langle T(y), x \rangle = \langle y, T(x) \rangle = 0$$

because we have $T(x) \in W$. □

**Discovery 8.1**

For $(V, \langle \cdot, \cdot \rangle)$ and inner product space over $\mathbb{R}$, $T = T^* \in \mathcal{L}(V)$ as in Lemma (8.1). Let $W \subseteq V$ be a linear subspace which is invariant for $T$, and let $\check{T} \in \mathcal{L}(V)$ be the restriction of $T$ to $W$, then $\check{T}^* = \check{T}$. Indeed, for all $x, y \in W$, we have

$$\langle \check{T}(x), y \rangle = \langle T(x), y \rangle = \langle x, T(y) \rangle = \langle x, \check{T}(y) \rangle$$

**Theorem 8.1**

For $n \in \mathbb{N}$, take $(V, \langle\cdot,\cdot\rangle)$ inner product space over $\mathbb{R}$ with $\dim(V) = n$ and $T \in \mathcal{L}(V)$ such that $T^* = T$, then one can find an orthonormal basis $\xi_1, \ldots, \xi_n$ for $V$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$T(\xi_1) = \lambda_1\xi_1, \ \ldots, \ T(\xi_n) = \lambda_n\xi_n$$

*Proof.* Proof is by induction on $n$.

1. *Base Case:*
   $\dim(V) = n = 1$, then pick a vector $\xi_1 \in V$ with $\|\xi_1\| = 1$. Since $\dim(V) = 1$, then we have

   $$V = \{\alpha\xi_1 : \alpha \in \mathbb{R}\}$$

   Sicne $T : V \to V$, there exists $\lambda_1 \in \mathbb{R}$ such that $T(\xi_1) = \lambda_1\xi_1$.

Lecture 21 - Friday, Jun 21

2. *Induction Step:*
   Let $n \geq 2$, suppose the theorem is proved for $1, 2, \ldots, n-1$, we will also prove it for the case $n$. So let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{R}$ with $\dim(V) = n$. Let $T \in \mathcal{L}(V)$ be such that $T^* = T$. Then Theorem (7.2) says that there exists $\lambda \in \mathbb{R}$ which is an eigenvalue for $T$. Let $x \in V$ be an eigenvalue for $T$ corresponding to $\lambda$. That is, $x \neq 0_V$ and $T(x) = \lambda x$. For convenience, we denote $\lambda =: \lambda_n$ and let $\xi_n = \frac{1}{\|x\|}x \in V$ (vector with $\|\xi_n\| = 1$). Observe that

   $$T(\xi_n) = T\left(\frac{1}{\|x\|}x\right) = \frac{1}{\|x\|}T(x) = \frac{1}{\|x\|}\lambda x = \lambda\left(\frac{1}{\|x\|}x\right) = \lambda\xi_n$$

   Hence $\xi_n$ is an eigenvector for $T$, corresponding to the eigenvalue $\lambda = \lambda_n$. Let $W = \{\alpha\xi_n : \alpha \in \mathbb{R}\}$, then $W$ is a linear subspace of $V$, which is invariant for $T$. Recall that by Lemma (8.1), the linear subspace $W^\perp$ is also invariant for $T$. Then Discovery (8.1) says that consider the operator $\check{T} \in \mathcal{L}(W^\perp)$ obtained by restrciting $T$, and we have $\check{T}^* = \check{T}$. Since the dimension of $W^\perp$ is

   $$\dim(W^\perp) = n - \dim(W) = n - 1$$

   the induction hypothesis applies to $\check{T}$, hence we can find an orthonormal basis for $W^\perp$, denoted as $\xi_1, \ldots, \xi_{n-1}$ with corresponding $\lambda_1, \ldots, \lambda_{n-1} \in \mathbb{R}$ such that

   $$\check{T}(\xi_i) = \lambda_i\xi_i \qquad \forall\, i \in \{1, \ldots, n-1\}$$

   Notice that we also have

   $$T(\xi_i) = \check{T}(\xi_i) = \lambda_i\xi_i \qquad \forall\, i \in \{1, \ldots, n-1\}$$

   That is, $\lambda_i$ are also eigenvalues for $T$ with eigenvectors $\xi_i$. The punchline is that $\xi_1, \ldots, \xi_n$ is an orthonormal basis for $V$. Verification: Easy to find that they all have norm of 1, and $\xi_n$ is orthogonal to all other $\xi_i$'s because for the reason that $\xi_n \in W$ and $\xi_i \in W^\perp$.

Finally, the $n$ linearly independent vectors in a space $V$ with $\dim(V) = n$ must be a basis. $\square$

**Definition 8.2**

Let $n \in \mathbb{N}$, and looke at $\mathcal{M}_n(\mathbb{R}) = \{A = [\alpha_{ij}]_{1\leq i,j\leq n} : \alpha_{ij} \in \mathbb{R} \quad \forall\, 1 \leq i,j \leq n\}$. For $A = [\alpha_{ij}]_{1\leq i,j\leq j} \in \mathcal{M}_n(\mathbb{R})$, we denote by $A^{\text{tr}}$ the transpose of $A$, that is

$$A^{\text{tr}} = [\beta_{ij}]_{1\leq i,j\leq n} \quad \text{where } \beta_{ij} = \alpha_{ji} \quad \forall\, 1 \leq i,j \leq n$$

$A \in \mathcal{M}_n(\mathbb{R})$ is said to be

1. *Symmetric*: When it satisfies $A^{\text{tr}} = A$;
2. *Orthogonal*: when it satisfies
$$A^{\text{tr}} \cdot A = I = A \cdot A^{\text{tr}}$$

## 8.1 "Every Symmetrix Matrix is Orthogonally Diagonalizable"

**Theorem 8.2**

Let $n \in \mathbb{N}$, and let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric. Then there exists a diagonal matrix $D \in \mathcal{M}_n(\mathbb{R})$

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \quad \text{with } \lambda_1 \geq \cdots \geq \lambda_n$$

and an orthogonal matrix $U \in \mathcal{M}_n(\mathbb{R})$ such that

$$A = UDU^{\text{tr}}$$

**Discovery 8.2**

We have that

$$A = UDU^{\text{tr}} \equiv A = UDU^{-1}$$

In the framework of MATH146, this says that $A$ is **similar** to the diagonal matrix $D$, where $U$ serves as the similarity matrix. Hence $A$ is diagonalizable with orthogonal similarity matrix.

*Proof.* We will obtain that as a consequence of Theorem (8.1). So let $n \in \mathbb{N}$ and let $A \in \mathcal{M}_n(\mathbb{R})$ be such that $A = A^{\text{tr}}$. Let $V \in \mathbb{R}^n$ ne endowed with standard inner product, and let $T = T_A \in \mathcal{L}(\mathbb{R}^n)$ be the linear operator of multiplication with $A$ on the left: if $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, then $T(x) = y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ is defined via the requirement:

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

□

Lecture 22 - Monday, Jun 24

*Proof.* Consider the inner product space $(V, \langle\cdot,\cdot\rangle)$ where $V = \mathbb{R}^n$ endowed with standard inner product. The given matrix $A \in \mathcal{M}_n(\mathbb{R})$ has an associated operator $T_A \in \mathcal{L}(\mathbb{R}^n)$ defined via matrix multiplication: for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we put $T_A(x) = y = (y_1, \ldots, y_n)$, where

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Now we proceed in steps:

1. *Step 1*: Observe that $T_A = T_A^*$
   Why? We know (HW1Q4) that

   $$\langle T_A(x), x' \rangle = \langle x, T_{A^{\mathrm{tr}}}(x') \rangle \qquad \forall\, x, x' \in \mathbb{R}^n$$

   This means that $T_A = T_{A^{\mathrm{tr}}} = T_A^*$.

2. *Step 2*: Applying Theorem (8.1)
   Thus the Theorem gives an orthonormal basis $\xi_1, \ldots, \xi_n$ for $\mathbb{R}^n$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

   $$T_A(\xi_1) = \lambda_1 \xi_1,\ \ldots,\ T_A(\xi_n) = \lambda_n \xi_n$$

   By re-labelling the $\lambda_i$'s we may assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

3. *Step 3*: Create the required matrices $D$ and $U$
   How? Let

   $$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} | & & | \\ \xi_1 & \cdots & \xi_n \\ | & & | \end{bmatrix}$$

   We know that $U$ is an orthogonal matrix by HW1Q3.

4. *Step 4*: Use $U$ to connect $\xi_1, \ldots, \xi_n$ to the "standard" orthonormal basis $e_1, \ldots, e_n$ for $\mathbb{R}^n$
   Indeed, we have

   $$T_U(e_i) = \xi_i,\ \ldots,\ T_U(e_n) = \xi_n$$

   because

   $$U \cdot e_i = \begin{bmatrix} | & & | \\ \xi_1 & \cdots & \xi_n \\ | & & | \end{bmatrix} \begin{bmatrix} 0 \\ 1_i \\ 0 \end{bmatrix} = [\xi_i]$$

   gives exactly the $i^{th}$ column of $U$. Therefore, we obtain the following two formualas:

   (a) $T_U(e_i) = \xi_i$;

   (b) $T_{U^{\mathrm{tr}}}(\xi_i) = T_{U^{-1}}(\xi_i) = {T_U}^{-1}(\xi_i) = e_i$.

5. *Step 5*: Let $B = UDU^{-1}$. Check: what is $T_B(\xi_i)$?
   We compute,

   $$T_B(\xi_i) = T_{UDU^{-1}}(\xi_i) = (T_U T_D T_{U^{-1}})(\xi_i) = T_U(T_D(T_{U^{-1}}(\xi_i))) = \lambda_i \xi_i$$

6. *Step 6*: With $D$ and $U$ defined in step 3 and $B$ defined in step 5, we have $B = A$.
   For every $1 \leq i \leq n$, we have
$$T_B(\xi_i) = \lambda_i \xi_i = T_A(\xi_i)$$
   so have $T_A, T_B \in \mathcal{L}(\mathbb{R}^n)$, they agree on a basis, which implies that $T_A = T_B$. Finally, we must have that $A = B$.

$\square$

#### 8.1.1 Example: Adjacency Matrix

**Example 8.1**

Make $n = 3$, and consider $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. So we have the graph ${}_2\overset{1}{\triangle}_3$ and the adjacency matrix:

$$A = [\alpha_{ij}]_{1 \leq i,j \leq 3} \quad \alpha_{ij} = \begin{cases} 1 & \text{if \{i, j\} is on edge} \\ 0 & \text{otherwise} \end{cases}$$

The question is: what is the orthogonal diagonalization of $A$?

*Proof.* Observe that $A = M - I$ where $M = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. What is interesting about $M$ is that $M^2 = 3M$. Put $P = 1/3M$ and observe that $P^2 = 1/3M = P = P^{\text{tr}}$ (we call $P$ a projection matrix). Let $Q = I - P \in \mathcal{M}_3(\mathbb{R})$. Easy to verify that $Q$ is also a projection matrix. Also observe that

$$PQ = P(I - P) = P \cdot I - P^2 = P - P = \mathcal{O}$$

We write this as $P \perp Q$. Back to the given matrix $A$, have

$$A = M - I = 3P - I = 3P - (P + Q) = 2P - Q$$

Hence we have a linear combination $A = 2P + (-1)Q$ with $P$, $Q$ are projection matrices with $P \perp Q$.

**Result 8.1: Rule of Thumb (Verified in L9)**

We can read the eigenvalues of $A$ as 2 and $-1$.

We can also easily find that $\text{Ran}(P) = \{\alpha(1,1,1) : \alpha \in \{z \in \mathbb{C} : z - \overline{z} = \overline{z} - z\}\}$, so we pick $\xi_1 = \frac{1}{\sqrt{3}}(1,1,1)$. Note also

$$\text{Ran}(Q) = \text{Ran}(P)^\perp = \text{Null}(P) = \text{span}\{(1,-1,0),(0,1,-1)\} \overset{GS}{=} \text{span}\{(1,-1,0),(1,1,-2)\}$$

so we take $\xi_2 = \frac{1}{\sqrt{2}}(1,-1,0)$ and $\xi_3 = \frac{1}{\sqrt{6}}(1,1,-2)$. Now we have

$$A = UDU^{-1} = UDU^{\text{tr}} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix}^{\text{tr}}$$

$\square$

# 9 Spectral Theorem and Functional Calculus for a Self-adjoint Operator

Fix for this lecture, $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{R}$ with finite dimension $\dim(V) = n \in \mathbb{N}$.

**Definition 9.1**

Suppose $W_1, W_2 \subseteq V$ two linear subspaces. Write $W_1 \perp W_2$ to mean that $w_1 \perp w_2$ for all $w_1 \in W_1$ and $w_2 \in W_2$. Note that $W_1 \perp W_2$ is equivalent to $W_1 \subseteq W_2^{\perp}$ (or $W_2 \subseteq W_1^{\perp}$).
Let $P_1, P_2 \in \mathcal{L}(V)$ be projection operators (for $i = 1, 2$, we have $P_i = P_i^2 = P_i^*$). We write $P_1 \perp P_2$ to mean that $P_1 P_2 = \mathcal{O}$ (recall HW4Q4).

**Proposition 9.1**

Let $W_1, W_2 \subseteq V$ be two linear subspaces. For $i = 1, 2$, let $P_i \in \mathcal{L}(V)$ be the operator of orthogonal projection onto $W_i$. Then

$$W_1 \perp W_2 \quad \Longleftrightarrow \quad P_1 \perp P_2$$

*Proof.* 1. ($\Longrightarrow$):
Pick an $x \in V$, look at $(P_1P_2)(x)$. We have $P_1(P_2(x))$. Observe that

$$P_2(x) \in \mathrm{Ran}(P_2) = W_2 \subseteq W_1^{\perp} = \mathrm{Null}(P_1)$$

by Proposition (4.2). So we have $P_2(x) \in \mathrm{Null}(P_1)$, hence $(P_1(P_2(x))) = 0$.

2. ($\Longleftarrow$):
Pick $w \in W_2$ and write $0_V = (P_1P_2)(w) = P_1(P_2(w)) = P_1(w)$. We found that $P_1(w) = 0_V$, which suggests that $w \in \mathrm{Null}(P_1) = W_1^{\perp}$. Thus we have $W_2 \subseteq W_1^{\perp}$, which further tells us that $W_1 \perp W_2$.

□

We will now set some notations concerning an operator $T = T^* \in \mathcal{L}(V)$, which will end in writing the operator as a linear combination:

$$T = \gamma_1 P_1 + \cdots + \gamma_m P_m$$

with $P_1, \ldots, P_m$ are projector operators such that $P_i \perp P_j$ for $i \neq j$ and such that $P_1 + \cdots + P_m = I$.

Lecture 24 - Friday, Jun 28

### 9.0.1 Spectrum

**Definition 9.2: Spectrum**

Pick $T = T^* \in \mathcal{L}(V)$ and let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ and $\xi_1, \ldots, \xi_n \in V$ be as in Theorem (8.1). That is, $T(\xi_i) = \lambda_i \xi_i$ for $1 \leq i \leq n$. Relabel the $\lambda_i$'s and the $\xi_i$'s to keep track of possible repetitions. Define $\lambda_1^{(1)} = \cdots = \lambda_{d_1}^{(1)} =: \gamma_1, \ldots, \lambda_1^{(m)} = \cdots = \lambda_{d_m}^{(m)} =: \gamma_m$ where $\gamma_i \neq \gamma_j$ for $i \neq j$, then the **spectrum of** $T$ is

$$\mathrm{Spec}(T) = \{\gamma_1, \ldots, \gamma_m\}$$

and for every $1 \leq j \leq m$, $d_j$ is the **multiplicity** of $\gamma_j$ in the list of eigenvalues $\lambda_1, \ldots, \lambda_n$.

**Discovery 9.1**

We can also perform the parallel relabelling on the $\xi_i$'s: having

$$\xi_1^{(i)}, \ldots, \xi_{d_i}^{(i)} \quad \text{– the part of the onb wrt } \gamma_i$$

For $1 \leq j \leq m$, we let

$$E_j = \operatorname{span}\{\xi_1^{(j)}, \ldots, \xi_{d_j}^{(j)}\} \subseteq V$$

as a linear subspace of $V$ with $\dim(E_j) = d_j$.

**Result 9.1**

Notation as above, we observe three facts about the linear subspace $E_j$:

1. $[x \in E_j] \;\; \Rightarrow \;\; [T(x) = \gamma_j x]$;
2. Every $x \in V$ can be written as a sum $x = x_1 + \cdots x_m$, with $x_1 \in E_1, \ldots, x_m \in E_m$.
3. For $1 \leq i, j \leq m$ with $i \neq j$, we have $E_i \perp E_j$.

*Proof.* Proof for (1):
We have

$$x = \alpha_1 \xi_1^{(j)} + \cdots + \alpha_{d_j} \xi_{d_j}^{(j)}$$

for some $\alpha_1, \ldots, \alpha_{x_j} \in \mathbb{R}$, thus

$$\begin{aligned} T(x) &= \alpha_1 T(\xi_1^{(j)}) + \cdots + \alpha_{d_j} T(\xi_{d_j}^{(j)}) \\ &= \alpha_1(\gamma_1 \xi_1^{(j)}) + \cdots + \alpha_{d_j}(\gamma_{d_j} \xi_{d_j}^{(j)}) \\ &= \gamma_j \left(\alpha_1 \xi_1^{(j)} + \cdots + \alpha_{d_j} \xi_{d_j}^{(j)}\right) = \gamma_j x \end{aligned}$$

Proof for (2):
Write $x$ in terms of the re-labelled orthonormal basis $\xi_1, \ldots, \xi_n$, then we have

$$x = \underbrace{\left(\alpha_1^{(1)} \xi_1^{(x)} + \cdots + \alpha_{d_1}^{(1)} \xi_{d_1}^{(1)}\right)}_{\in E_1} + \cdots + \underbrace{\left(\alpha_1^{(m)} \xi_1^{(x)} + \cdots + \alpha_{d_m}^{(m)} \xi_{d_m}^{(m)}\right)}_{\in E_m}$$

Proof for (3):
We know that $\xi_k^{(i)} \perp \xi_l^{(j)}$ for all $1 \leq k \leq d_i$ and all $1 \leq l \leq d_j$ because they are distinct vectors in an orthonormal basis. Therefore the linear combinations of $\xi_k^{(i)}$'s and $\xi_l^{(j)}$'s are still perpendicular to each other.

$\square$

## 9.1 Spectral Theorem

**Theorem 9.1: Spectral Theorem**

Let $T = T^* \in \mathcal{L}(V)$ and, in connection to it, consider the setting of the definition of Spectrum, and the above Discovery and Result. For every $1 \leq j \leq m$, let $P_j \in \mathcal{L}(V)$ be the operator of orthogonal projection onto $E_j$, then:

1. $P_i \perp P_j$ for $i \neq j$ (that is $P_i P_j = \mathcal{O}$);
2. $P_1 + \cdots + P_m = I$;
3. $\gamma_1 P_1 + \cdots + \gamma_m P_m = T$.

*Proof.* Proof of (1):
This result follows the Proposition (9.1)

Proof of (2):
Pick $x \in V$, by the above result, we know that we can write $x = x_1 + \cdots + x_m$ with $x_1 \in E_1$ and $x_m \in E_m$. For every $1 \leq j \leq m$, we then get

$$P_j(x) = P_j(x_1) + \cdots + P_j(x_m)$$

Observe that $P_j(x_j) = x_j$ because $x_j \in E_j$ and $P_j$ is the projection onto $E_j$. In addition to that, we also have for every $k \neq j$, we have $P_j(x_k) = 0_V$ because $x_k \in E_k \subseteq E_j^{\perp}$, where the result follows naturally. Hence we have

$$(P_1 + \cdots + P_m)(x) = P_1(x) + \cdots + P_m(x) = x_1 + \cdots + x_m = x$$

Proof of (3):
Write again $x = x_1 + \cdots + x_m$, we have

$$T(x) = T(x_1) + \cdots T(x_m) = \gamma_1 x_1 + \cdots \gamma_m x_m$$

□

Lecture 25 - Wednesday, Jul 3

**Discovery 9.2**

From Theorem (9.1), we get right away that

$$T^k = \gamma_1^k P_1 + \cdots + \gamma_m^k P_m \qquad \forall\ k \in \mathbb{N}$$

*Proof.* We check the above result by induction on $k$. For $k = 1$, it is proved in the proof for Theorem (9.1). Suppose we have the statement holds for $k \in \mathbb{N}$. STP its validity for $k + 1 \in \mathbb{N}$:

$$\begin{aligned} T^{k+1} = T^k T &= (\gamma_1^k P_1 + \cdots + \gamma_m^k P_m)(\gamma_1 P_1 + \cdots + \gamma_m P_m) \\ &= \sum_{i,j=1}^{m} (\gamma_i^k P_i)(\gamma_j P_j) \\ &= \sum_{i,j=1}^{m} (\gamma_i^k \gamma_j)(P_i P_j) \end{aligned}$$

$\square$

**Definition 9.3**

For $T \in \mathcal{L}(V)$. Consider a polynomial function $q : \mathbb{R} \to \mathbb{R}$, $q(t) = a_0 + a_1 t + \cdots + a_k t^k$ for $t \in \mathbb{R}$ for some $k \in \mathbb{N} \cup \{0\}$ and $a_0, \ldots, a_k \in \mathbb{R}$. Denote

$$q(T) = a_0 I + a_1 T + \cdots + a_k T^k \in \mathcal{L}(V)$$

**Corollary 9.1**

Framework as in Theorem (9.1) (with $T = T^* \in \mathcal{L}(V)$ and with $\gamma_1, \ldots, \gamma_m$ and $P_1, \ldots, P_m$ as in the Theorem). Let $q : \mathbb{R} \to \mathbb{R}$ be a polynomial function, then we have

$$q(T) = q(\gamma_1)P_1 + \cdots + q(\gamma_m)P_m$$

*Proof.* Write $q(t) = a_0 + a_1 t + \cdots + a_k t^k$ for $t \in \mathbb{R}$. Consider formulas with $k = 1, \ldots, k$. Theorem (9.1) tells

$$\begin{aligned} I &= P_1 + \cdots + P_m \\ T &= \gamma_1 P_1 + \cdots + \gamma_m P_m \\ &\vdots \\ T^k &= \gamma_1^k P_1 + \cdots + \gamma_m^k P_m \end{aligned}$$

Multiply the first equation by $a_0$, the second by $a_1$, .... Then add the resulting equations will yield us the desired result. $\square$

Lecture 26 - Friday, Jul 5

There is a uniqueness bit which holds in connection to Spectral Theorem (9.1), see more in A6-Q1.

**Corollary 9.2**

Framework as in Theorem (9.1), let $q, r : \mathbb{R} \to \mathbb{R}$ be polynomial functions such that $q(\gamma_j) = r(\gamma_j)$ for all $1 \leq j \leq m$. In words, $q$ and $r$ agree on $\mathrm{Spec}(T)$. Then we have $q(T) = r(T)$.

*Proof.* We have shown that $q(T) = q(\gamma_1)P_1 + \cdots + q(\gamma_m)P_m$, so

$$\begin{aligned} q(T) &= q(\gamma_1)P_1 + \cdots + q(\gamma_m)P_m \\ &= r(\gamma_1)P_1 + \cdots + r(\gamma_m)P_m \\ &= r(T) \end{aligned}$$

$\square$

**Definition 9.4**

Let $T = T^* \in \mathcal{L}(V)$, consider $\mathrm{Spec}(T) = \{\gamma_1, \ldots, \gamma_m\} \subseteq \mathbb{R}$, and let $f : \mathrm{Spec}(T) \to \mathbb{R}$ be a function. We define an operator $f(T) \in \mathcal{L}(V)$ as following: Pick a polynomial function $q : \mathbb{R} \to \mathbb{R}$ such that $q(\gamma_j) = f(\gamma_j)$ for all $1 \leq j \leq m$ and define $f(T) := q(T)$.

**Discovery 9.3**

The existence of $q$ is guaranteed (by Lagrange Interpolation), but its existence is not unique. However, the definition of $f(T)$ makes sense as a result of Corollary (9.2).

**Result 9.2: Functional Calculus**

There is an alternative formula for $f(T)$. Consider the $P_1, \dots, P_m$ from Theorem (9.1) and observe that we have

$$f(T) = f(\gamma_1)P_1 + \cdots f(\gamma_m)P_m$$

Indeed, with $q$ as above, we get

$$f(T) = q(T) = q(\gamma_1)P_1 + \cdots + q(\gamma_m)P_m = f(\gamma_1)P_1 + \cdots f(\gamma_m)P_m$$

The assignment

$$f \mapsto f(T)$$
$$\text{functions from } \mathrm{Spec}(T) \text{ to } \mathbb{R} \rightsquigarrow \text{operator in } \mathcal{L}(V)$$

is called **functional Calculus** for the operator $T$.

**Example 9.1**

For every $T = T^* \in \mathcal{L}(V)$, we can define an operator $e^T$. Possible hw question: given $S = S^*$ and $T = T^*$ in $\mathcal{L}(V)$, do we have $e^{S+T} = e^S \cdot e^T$
The answer is YES if we have $ST = TS$ (A6-Q4 can be useful here).

**Example 9.2**

For every $T = T^* \in \mathcal{L}(V)$, we can define $\sqrt[3]{T} \in \mathcal{L}(V)$. Possible hw question: suppose $R = R^*$ and $S = S^* \in \mathcal{L}(V)$ are such that $R^3 = T = S^3$, can we conclude that $R = S$?
The answer is positive.

**Proposition 9.2**

Functional Calculus respects the algebra operations: For $f, g : \mathrm{Spec}(T) \to \mathbb{R}$ and $\alpha \in \mathbb{R}$, we haven

1. $(f + g)(T) = f(T) + g(T)$;
2. $(\alpha f)(T) = \alpha f(T)$;
3. $(f \cdot g)(T) = f(T) \cdot g(T)$.

Moreover, note that

$$\iota(T) = T$$

where $\iota : \text{Spec}(T) \to \mathbb{R}$ is the identify function $\iota(\gamma) = \gamma$ for all $\gamma \in \text{Spec}(T)$ and

$$\mathbb{1}(T) = I$$

where $\mathbb{1} : \text{Spec}(T) \to \mathbb{R}$ is constantly equal to 1.

Lecture 27 - Monday, Jul 8

**Discovery 9.4**

Notice that we also have $[f(T)]^* = f(T)$.

*Proof.* All required formulas follow easily from the result above (Functional Calculus). Check for instance the formula for $(f \cdot g)(T)$: we have

$$f(T) = \sum_{j=1}^{m} f(\gamma_j) P_j \quad \text{and} \quad g(T) = \sum_{j=1}^{m} g(\gamma_j) P_j$$

Hence

$$\begin{aligned} f(T) \cdot g(T) &= \left[ \sum_{i=1}^{m} f(\gamma_i) P_i \right] \cdot \left[ \sum_{j=1}^{m} g(\gamma_j) P_j \right] \\ &= \sum_{i,j=1}^{m} f(\gamma_i) g(\gamma_j) P_i P_j \\ &= \sum_{i}^{m} f(\gamma_i) g(\gamma_i) {P_i}^2 \\ &= \sum_{i=1}^{m} (f \cdot g)(\gamma_i) P_i = (f \cdot g)(T) \end{aligned}$$

as desired. □

**Proposition 9.3: Spectral Mapping Theorem**

Use the same framework as above, $T = T^* \in \mathcal{L}(V)$ with $\text{Spec}(T) = \{\gamma_1, \ldots, \gamma_m\}$.
Consider $f : \text{Spec}(T) \to \mathbb{R}$ and denote $f(T) =: R \in \mathcal{L}(V)$, then

1. $\text{Spec}(R) = f(\text{Spec}(T))$ (i.e. $\{\rho \in \mathbb{R} : \exists j \in \{1, \ldots, m\} \text{ s.t. } \rho = f(\gamma_j)\}$);
2. Suppose we are given a function $h : \text{Spec}(R) \to \mathbb{R}$, then we have $h(R) = (h \circ f)(T)$.

*Proof.* For part 1:
The idea of the proof is that write

$$R = f(T) = \sum_{j=1}^{m} f(\gamma_j) P_j$$

then group together $i, j$ where $f(\gamma_i) = f(\gamma_j)$.

**Example 9.3**

Look at a "concrete" example, say that $m = 3$ (so $\text{Spec}(T) = \{\gamma_1, \gamma_2, \gamma_3\}$, WLOG assume they are in descending order) and we have $f(\gamma_1) = f(\gamma_2) = \rho_1 \in \mathbb{R}$, $f(\gamma_3) = \rho_2 \in \mathbb{R}$ with $\rho_1, \rho_2$. Then

$$\begin{aligned} R &= f(\gamma_1)P_1 + f(\gamma_2)P_2 + f(\gamma_3)P_3 \\ &= \rho_1 P_1 + \rho_2 P_2 + \rho_3 P_3 \\ &= \rho_1(P_1 + P_2) + \rho_3 P_3 \end{aligned}$$

so we get $R = \rho_1 Q_1 + \rho_2 Q_2$ with $\rho_1 \neq \rho_2$ and $Q_1 \perp Q_2$ are non-zero operators. By HW6Q1 we now have that

$$\text{Spec}(R) = \{\rho_1, \rho_2\}$$

For part 2:

**Example 9.4**

Consider a function $h : \text{Spec}(R) \to \mathbb{R}$, we compute

$$\begin{aligned} h(R) &= h(\rho_1)Q_1 + h(\rho_2)Q_2 \\ &= \sigma_1 Q_1 + \sigma_2 Q_2 \\ &= \sigma_1 P_1 + \sigma_1 P_2 + \Sigma_2 P_3 \\ &= (h \circ f)(\gamma_1)P_1 + (h \circ f)(\gamma_2)P_2 + (h \circ f)(\gamma_3)P_3 = (h \circ f)(T) \end{aligned}$$

$\square$

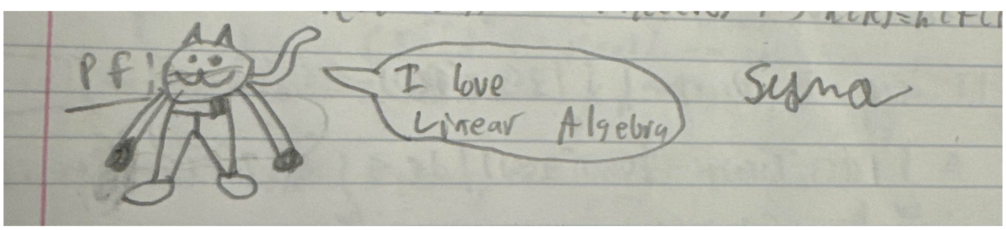


Figure 1: John Spectrum family photo (he's homeless and unmarried)

# 10 Two Applications of Funtional Calculus

## 10.1 Functional Calculus for Symmetric Matrices

In this lecture, we fix $n \in \mathbb{N}$, look at $\mathbb{R}^N$ endowed with standard inner product.

Given $A \in \mathcal{M}_n(\mathbb{R})$, we define $T_A \in \mathcal{L}(\mathbb{R}^N)$, we have

$$T_A{}^* = T_{A^{\text{tr}}}$$

Lecture 28 - Wednesday, Jul 10

We also denote

$$\text{Spec}(A) = \text{Spec}(T_A) = \{\lambda \in \mathbb{R} : \exists x \in V \text{ s.t. } Ax = \lambda x\}$$

which is the ser of distinct eigenvalues of $A \in \mathcal{M}_n(\mathbb{R})$.

**Definition 10.1**

Let $q : \mathbb{R} \to \mathbb{R}$ be a polynomial function, $q(t) = a_0 + a_1 t + \cdots + a_k t^k$ for $k \in \mathbb{N}$ and $a_0, \ldots, a_k \in \mathbb{R}$. Then for every $A \in \mathcal{M}_n(\mathbb{R})$, we denote

$$q(A) = a_0 I + a_1 A + \cdots a_k A^k \in \mathcal{M}_n(\mathbb{R})$$

is the map $A \rightsquigarrow q(A)$ as in $\mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$, which is the matrix-version of the polynomial $q$.

**Lemma 10.1**

Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric ($A = A^{\text{tr}}$). Consider the factorization as in Theorem (8.2)

$$A = UDU^{\text{tr}}$$

where $D$ is diagonal (with eigenvalues of $A$) and $U \in \mathcal{M}_n(\mathbb{R})$ is an orthogonal matrix (with eigenvectors of $A$ forming orthonormal basis). Then for every $k \in \mathbb{N}$, we have

$$A^k = UD^kU^{\text{tr}} = U \begin{bmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{bmatrix} U^{\text{tr}}$$

*Proof.* The proof is by induction on $k$. The base case has been shown in Theorem (8.2). Induction Step: Assume the statement holds for $k$, we wish to show it for $k+1$. We have

$$A^{k+1} = A^k \cdot A = U \begin{bmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{bmatrix} U^{\text{tr}} \cdot U \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} U^{\text{tr}} = U \begin{bmatrix} \lambda_1^{k+1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{k+1} \end{bmatrix} U^{\text{tr}}$$

as desired. □

**Proposition 10.1**

Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric ($A = A^{\text{tr}}$) and write $A = UDU^{\text{tr}}$ as shown above. Let $q : \mathbb{R} \to \mathbb{R}$ be a polynomial function, then

$$q(A) = U \begin{bmatrix} q(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & q(\lambda_n) \end{bmatrix} U^{\text{tr}}$$

*Proof.* Write the polynomial in an explicit way,

$$q(t) = a_0 + a_1 t + \cdots + a_k t^k$$

For every $j \in \{1, \ldots, k\}$, the proceeding Lemma tells us that

$$A^j = U \begin{bmatrix} \lambda_1^j & & 0 \\ & \ddots & \\ 0 & & \lambda_n^j \end{bmatrix} U^{\text{tr}}$$

Note that for $j_0$ we have $a_0 I = U \begin{bmatrix} a_0\lambda_1^0 & & 0 \\ & \ddots & \\ 0 & & a_n\lambda_n^0 \end{bmatrix} U^{\text{tr}}$. Thus amplifying each of the $A^j$ by $a_j$ and sum up the $k+1$ equations we have desired conclusion. □

**Corollary 10.1**

Let $A \in \mathcal{M}_n(\mathbb{R})$ be symmetric ($A = A^{\text{tr}}$). Let $q, r : \mathbb{R} \to \mathbb{R}$ be polynomial functions agreeing on $\text{Spec}(A)$. That is, write $\text{Spec}(A) = \{\gamma_1, \ldots, \gamma_m\}$, we have

$$q(\gamma_j) = r(\gamma_j) \qquad \forall j \in \{1, \ldots, m\}$$

Then we have $q(A) = r(A)$.

*Proof.* Consider the writing $U \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} U^{\text{tr}}$, then we have

$$q(A) = U \begin{bmatrix} q(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & q(\lambda_n) \end{bmatrix} U^{\text{tr}} = \begin{bmatrix} r(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & r(\lambda_n) \end{bmatrix} U^{\text{tr}} = r(A)$$

gg. □

**Definition 10.2**

Let $A = A^{\text{tr}} \in \mathcal{M}_n(\mathbb{R})$ with $\text{Spec}(A) = \{\gamma_1, \dots, \gamma_m\} \in \mathbb{R}$. Let $f : \text{Spec}(A) \to \mathbb{R}$ be a function. We define a new matrix $f(A) \in \mathcal{M}_n(\mathbb{R})$ as following: Let $q$ be a polynomial function such that $q(\gamma_j) = f(\gamma_j)$ for all $j \in \{1, \dots, m\}$, then we define

$$f(A) = q(A) \in \mathcal{M}_n(\mathbb{R})$$

As a result of Corollary (10.1), we know that this is well-defined.

**Proposition 10.2**

Let $A = A^{\text{tr}} \in \mathcal{M}_n(\mathbb{R})$ written as $A = U \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} U^{\text{tr}}$ as in Theorem (8.2), then for every function $f : \text{Spec}(A) \to \mathbb{R}$, we have $f(A) = U \begin{bmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{bmatrix} U^{\text{tr}}$

Lecture 29 - Friday, Jul 12

*Proof.* Pick a mimicking polynomial $q(t) = a_0 + a_1 t + \cdots + a_k t^k$ which mimicks $f$ on $\text{Spec}(A)$, this means in particular that

$$q(\lambda_i) = f(\lambda_i), \qquad \forall i \in \{1, \dots, n\}$$

It follows that

$$f(A) = q(A) = U \begin{bmatrix} q(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & q(\lambda_n) \end{bmatrix} U^{\text{tr}} = \begin{bmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{bmatrix}$$

as wanted. $\square$

### 10.1.1 Simultaneous Orthogonal Diagonalization

**Exercise:** *(Simultaneous Orthogonal Diagonalization)*

Take two symmetric matrices $A, B \in \mathcal{M}_n(\mathbb{R})$ such that $AB = BA$. One can find $\lambda_1, \dots, \lambda_n, \lambda'_1, \dots, \lambda'_n \in \mathbb{R}$ and $U = \begin{bmatrix} | & & | \\ \xi_1 & \cdots & \xi_n \\ | & & | \end{bmatrix}$ such that

$$A = U \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} U^{\text{tr}} \quad \text{and} \quad B = U \begin{bmatrix} \lambda'_1 & & 0 \\ & \ddots & \\ 0 & & \lambda'_n \end{bmatrix} U^{\text{tr}}$$

### 10.1.2 Pop Up Quiz 2

**Exercise:**

1. What homework question would you use in order to solve Exercise (Simultaneous Orthogonal Diagonalization 10.1.1)?
2. Write two sentences explaining what would be the plan of your solution to Exercise (Simultaneous Orthogonal Diagonalization 10.1.1)

*Proof.* 1. HW6Q4.

2. Since we know that $A$ and $B$ commute, and they are symmetric matrices over $\mathbb{R}$, thus they are self-adjoint. It follows that by HW6Q4, there exists an orthonormal basis consisting eigenvectors for both $A$ and $B$, thus they are simultaneously diagonalizable by such an orthonormal basis.

$\square$

## 10.2 Positive Operators and their Square Roots

In this section, we fix $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{R}$ with $\dim(V) = n \in \mathbb{N}$.

**Lemma 10.2**

Let $T = T^* \in \mathcal{L}(V)$ with $\mathrm{Spec}(T) = \{\gamma_1, \ldots, \gamma_m\}$. Consider the writing $T = \gamma_1 P_1 + \cdots + \gamma_m P_m$, with $P_1, \ldots, P_m$ are projections, $P_i \perp P_j$ for $i \neq j$ and $P_1 + \cdots + P_m = I$. Then for every $x \in V$ we have

$$\langle T(x), x\rangle = \sum_{j=1}^{m} \gamma_j \left\|P_j(x)\right\|^2$$

*Proof.* Pick $x \in V$ and write

$$x = I(x) = P_1(x) + \cdots P_m(x)$$
$$T(x) = \gamma_1 P_1(x) + \cdots + \gamma_m P_m(x)$$

Therefore

$$\begin{aligned}\langle T(x), x\rangle &= \langle \sum_{i=1}^{m} \gamma_i P_i(x), \sum_{j=1}^{m} P_j(x)\rangle \\ &= \sum_{i,j=1}^{m} \gamma_i \langle P_i(x), P_j(x)\rangle \\ &= \sum_{j=1}^{m} \gamma_j \left\|P_j(x)\right\|^2\end{aligned}$$

as desired. $\square$

**Proposition 10.3**

For $T = T^* \in \mathcal{L}(V)$, TFAE:

1. $\mathrm{Spec}(T) \subseteq [0, \infty)$;

2. $\langle T(x), x\rangle \geq 0$, $\forall x \in V$.

*Proof.* ($1 \implies 2$) is obvious by the above Lemma. For the other direction, we pick an arbitrary eigenvector $\lambda \in \mathbb{R}$ for $T$ and we wish to show that $\lambda > 0$. Let $0_v \neq x \in V$ be an eigenvector of $T$ such that $T(x) = \lambda x$. For this $x$ we get that $\langle T(x), x\rangle = \langle \lambda x, x\rangle = \lambda \|x\|^2 \geq 0$, which implies that we must have $\lambda \geq 0$ as wanted. $\square$

Lecture 30 - Monday, Jul 15

**Definition 10.3: Strictly Positive Operator**

For $T = T^* \in \mathcal{L}(V)$, we know by Proposition (10.3) that the following two equivalent conditions holds:

1. $\text{Spec}(T) \subseteq (0, \infty)$;
2. $\langle T(x), x\rangle > 0$, $\forall x \in V$.

We say that $T$ is a **strictly positive operator**.

*Proof.* The proof is similar to the proof of Proposition (10.3) and is left as an **exercise**. $\square$

**Discovery 10.1**

For $T = T^* \in \mathcal{L}(V)$ we have

$$T \text{ is strictly positive } \Leftrightarrow T \text{ is positive and invertible}$$

*Proof.* This is because we know that both are satisfied exactly when $\text{Spec}(T) \subseteq [0, \infty)$ and $0 \notin \text{Spec}(T)$, thus $T$ is strictly positive if and only if $T$ is positive and invertible. $\square$

**Definition 10.4: Square Root of Positive Operator**

Let $T \in \mathcal{L}(V)$ be a positive operator (that is, $T = T^*$ and $\text{Spec}(T) \subseteq [0, \infty)$). Then we an consider the function $f : \text{Spectrum}(T) \to \mathbb{R}$, $f(t) = \sqrt{t}$. Consider the operator $f(T)$ and put

$$\sqrt{T} := f(T)$$

where $\sqrt{T}$ is called the **square root of** $T$.

**Discovery 10.2**

For $T = T^* \in \mathcal{L}(V)$ and $\sqrt{T}$ defined as above, we observe that $\sqrt{T}$ is also a positive operator.

*Proof.* We know that $\sqrt{T} = \left(\sqrt{T}\right)^*$ because we have $f(T) = (f(T))^*$ for any function $f : \text{Spec}(T) \to \mathbb{R}$. Then the Spectral Mapping Theorem (9.3) says that

$$\text{Spec}(\sqrt{T}) = \{\sqrt{\gamma_1}, \ldots, \sqrt{\gamma_m}\}$$

which implies that

$$\mathrm{Spec}(\sqrt{T}) \subseteq [0, \infty)$$

□

**Proposition 10.4**

For $T = T^* \in \mathcal{L}(V)$ a positive operator. Suppose that someone has found a positive operator $R \in \mathcal{L}(V)$ such that $R^2 = T$, then

$$R = \sqrt{T}$$

*Proof.*

1. *Claim 1:* $RT = TR$
$$RT = R \cdot R^2 = R^3 = R^2 \cdot R = TR$$

2. *Claim 2:* $R\sqrt{T} = \sqrt{T}R$
Since $RT = TR$, from Hw6Q3, we know that $f(R)g(T) = g(T)f(R)$ for any two functions $f$ and $g$. Let $f(t) = t$ for $t \in \mathrm{Spec}(R)$ and $g(t) = \sqrt{t}$ for $t \in \mathrm{Spec}(T)$, then we have $f(R) = R$ and $g(T) = \sqrt{T}$, and now we get
$$R\sqrt{T} = \sqrt{T}R$$

3. *Claim 3:* We can find an orthonormal basis $\{\xi_1, \ldots, \xi_n\}$ for $V$ such that $\sqrt{T}(\xi_i) = \alpha_i \xi_i$ and $R(\xi_i) = \beta_i \xi_i$ for all $i = 1, \ldots, n$.
This follows Hw6Q4.

4. *Claim 4:* Let $\alpha_i, \beta_i$ be as introduced in Claim 3, we have $\alpha_i = \beta_i$.
Observe that $\alpha_i$ is an eigenvalue for $\sqrt{T}$, hence $\alpha_i \geq 0$. Likewise, $\beta_i$ is an eigenvalue for $R$, so $\beta_i \geq 0$. Now STP that $\alpha_i^2 = \beta_i^2$:
$$T(\xi_i) = R^2(\xi_i) = R(R(\xi_i)) = R(\beta_i \xi_i) = \beta_i R(\xi_i) = \beta_i^2 \xi_i$$
Likewise, we write
$$T(\xi_i) = \sqrt{T}(\sqrt{T}(\xi_i)) = \sqrt{T}(\alpha_i \xi_i) = \alpha_i \sqrt{T}(\xi_i) = \alpha_i^2 \xi_i$$

5. *Claim 5:* $R = \sqrt{T}$
Since the two operators agree on a linear basis, they must be equal to each other.

GOUWWWWWD PROOF. □

# 11 A Parallel World: Inner Product Space over $\mathbb{C}$

**Definition 11.1: What is an inner product space over $\mathbb{C}$**

Let $V$ be a vector space over $\mathbb{C}$, an inner product on $V$ is an assignment which produces a number $\langle x, y\rangle \in \mathbb{C}$ whenever vectors $x, y \in V$ are given. This assignment must be

1. **sesqui-linear**
2. **conjugate symmetric**
3. **positive definite**

Lecture 31 - Wednesday, Jul 17

**Exercise:** Solve for $f(t)$:

$$\frac{d^2}{dt^2} f(t) = -4f(t)$$

Thinking about what does this have to do with Linear Algebra?

*Proof.* Solve for $f(t)$ we have $f(t) = \sin(2t)$. Notice that this is a Vector Space that is

$$\operatorname{span}\{\sin(2t), \cos(2t)\}$$

Solution with factoring:

$$\begin{aligned} \left(\frac{d^2}{dt^2} + 4I\right) f &= 0 \\ \Rightarrow \left(\frac{d}{dt} + 2iI\right)\left(\frac{d}{dt} - 2iI\right) f &= 0 \end{aligned}$$

Hence we have two cases, which are

$$\begin{cases} \left(\frac{d}{dt} - 2iI\right) f & = 0 \text{ or} \\ \left(\frac{d}{dt} + 2iI\right) f & = 0 \end{cases}$$

which implies that $f(t) = e^{2it}$ or $= e^{-2it}$. As a result, alternatively, the space can also be represented as:

$$\operatorname{span}\{e^{2it}, e^{-2it}\}$$

□

**Discovery 11.1**

Notice that there is an issue. Consider $V := \{f : [0, 2\pi] : f \text{ is continuous}\}$ and define inner product as

$$\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)\, dt$$

As a result, we have

$$\left\|e^{2it}\right\|^2 = \langle e^{2it}, e^{2it}\rangle = \int_0^{2\pi} = 0$$

**Result 11.1**

Therefore, so solve the above issue, we need to **conjugate it**! In particular, we need to define the inner product as:

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\overline{g(t)}\, dt$$

#### 11.0.1 Inner Product Space over $\mathbb{C}$

**Definition 11.2: An inner product space over $\mathbb{C}$**

let $V$ be an vector space over $\mathbb{C}$. The inner product on $V$ must be:

1. **sesqui-linear**

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

2. **conjugate symmetric**

$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$

   Consequently, we have

$$\langle x, \beta y \rangle = \overline{\beta} \langle x, y \rangle \quad \text{and} \quad \langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$$

3. **positive definite**

$$\langle x, x \rangle \geq 0 \text{ with equality only when } x = 0$$

**Example 11.1**

Let $V = \mathbb{C}^n$ and $\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$, this is an inner product space over $\mathbb{C}$.

**Theorem 11.1: Cauchy Schwarz**

We have

$$|\langle x, y \rangle| \leq \|x\| \, \|y\|$$

**Theorem 11.2: Triangle Inequality**

We have

$$\|x + y\| \leq \|x\| + \|y\|$$

#### 11.0.2 Orthogonal System

Just like what we had before,

$$x \perp y \iff \langle x, y \rangle = 0$$

**Definition 11.3: Orthogonal/ Orthonormal System**

An **orthogonal system** is a family of vectors $x_1, \dots, x_n$ such that $x_i \perp x_j$ for all $i \neq j$. Moreover, an orthonormal system is a family of vactors who form an orthogonal system and whose norm are all 1.

**Example 11.2**

Define $X_k(t) = e^{kit}$, notice that

$$X_k \in V := \{f : [0, 2\pi] : f \text{ is continuous}\}$$

with inner product the same as defined above. Then we have for any $n$, the set

$$\{X_1, \dots, X_n\}$$

is an orthogonal system.

*Proof.* For $k \neq j$, we have

$$\begin{aligned}
\langle e^{ikt}, e^{ijt} \rangle &= \int_0^{2\pi} e^{ikt} \overline{e^{ijt}} \, dt \\
&= \int_0^{2\pi} e^{i(k-j)t} \, dt \\
&= \frac{1}{i(k-j)} e^{i(k-j)t} \Big|_0^{2\pi} \\
&= 0
\end{aligned}$$

□

Lecture 32 - Friday, Jul 19

# 12 $\mathcal{L}(V)$ where $V$ is an inner product space over $\mathbb{C}$

## 12.1 Adjoint

In this section, we will be exploring

1. adjoints;
2. eigenvalues.

Throughout this lecture, we fix $(V, \langle\cdot,\cdot\rangle)$ an inner product space over $\mathbb{C}$, with $\dim(V) = n \in \mathbb{N}$.

**Definition 12.1**

We define

$$\mathcal{L}(V) = \{T : V \to V \mid T \text{ is linear}\}$$

where the term **linear** here means that one has

$$T(x) = \alpha x \qquad \forall\, \alpha \in \mathbb{C},\ x \in V$$

On $\mathcal{L}(V)$ we have three operations:

1. addition;
2. scalar multiplication;
3. multiplication.

**Discovery 12.1**

Fix an orthonormal basis, $\zeta_1, \ldots, \zeta_n$ for $V$, then we get a bijection

$$\begin{aligned} \mathcal{L}(V) &\to \mathcal{M}_n(\mathbb{C}) \\ T &\mapsto A_T \end{aligned}$$

where $A_T = [\alpha_{jk}]_{1 \le j,k \le n}$, with $\alpha_{jk} = \langle T(\zeta_k), \zeta_j\rangle$ for $1 \le j, k \le n$. This choice of the $\alpha_{jk}$ is made such that we have

$$T(\zeta_k) = \alpha_{1k}\zeta_1 + \alpha_{2k}\zeta_2 + \cdots + \alpha_{nk}\zeta_n$$

Moreover, the matrix preserves the three operations for all $S, T \in \mathcal{L}(V)$ and $\alpha \in \mathbb{C}$:

1. $A_{S+T} = A_S + A_T$;
2. $A_{\alpha T} = \alpha \cdot A_T$;
3. $A_{ST} = A_S \cdot A_T$ (product of matrices in $\mathcal{M}_n(\mathbb{C})$).

**Theorem 12.1**

Given $T \in \mathcal{L}(V)$, there exists $S \in \mathcal{L}(V)$, uniquely determined, such that

$$\langle T(x), y\rangle = \langle x, S(y)\rangle \qquad \forall\, x, y \in V$$

**Definition 12.2: adjoint**

The $S \in \mathcal{L}(V)$ defined above is called the **adjoint** of $T$ and denoted as $T^*$.

*Proof.* Proof of uniqueness of $S$: **exercise.**
Proof of existence of $S$:
Fix an orthonormal basis $\zeta_1, \dots, \zeta_n$ for $V$. Consider the matrix associated to $T$,

$$A_T = [\alpha_{jk}]_{1 \le j,k \le n} \in \mathcal{M}_n(\mathbb{C})$$

Let

$$B = [\beta_{jk}]_{1 \le j,k \le n} \in \mathcal{M}_n(\mathbb{C})$$

be defined by putting $\beta_{jk} = \overline{\alpha_{kj}}$. Since the map between linear operator and the associated matrix is a bijection, there exists $S \in \mathcal{L}(V)$ such that $A_S = B$, this means that we have

$$\langle S(\zeta_k), \zeta_j\rangle = \beta_{jk}$$

We will show that $S$ satisfies the above equality.

1. *Claim 1*: $\langle T(\zeta_k), \zeta_j\rangle = \langle \zeta_k, S(\zeta_j)\rangle$
   We verify the claim merely by computation:
   $$\langle T(\zeta_k), \zeta_j\rangle = \alpha_{jk} = \overline{\beta_{kj}} = \overline{\langle S(\zeta_j), \zeta_k\rangle} = \langle \zeta_k, S(\zeta_j)\rangle$$

2. *Claim 2*:
   For the $S$ found above, we have $\langle T(x), y\rangle = \langle x, S(y)\rangle$
   To verify, pick $x, y \in V$ and write
   $$x = \sum_{k=1}^{n} a_k \zeta_k \qquad \text{and} \qquad y = \sum_{j=1}^{n} b_j \zeta_j$$
   where $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{C}$. Observe that we have
   $$\begin{aligned}
   \langle T(x), y\rangle &= \langle \sum_{k=1}^{n} a_k T(\zeta_k), \sum_{j=1}^{n} b_j \zeta_j\rangle \\
   &= \sum_{j,k=1}^{n} a_k \overline{b_j} \langle T(\zeta_k), \zeta_j\rangle \\
   &= \sum_{j,k=1}^{n} a_k \overline{b_j} \langle \zeta_k, S(\zeta_j)\rangle \\
   &= \langle \sum_{k=1}^{n} a_k \zeta_k, \sum_{j=1}^{n} b_j S(\zeta_j)\rangle = \langle x, S(y)\rangle
   \end{aligned}$$

as desired. $\square$

**Result 12.1**

The proof of existence of $S$ in the above Theorem allows us to add one more property, in particular, we have

$$A_{T^*} = (A_T)^* \qquad \forall\, T \in \mathcal{L}(V)$$

where for every matrix $M \in \mathcal{M}_n(\mathbb{C})$, $M = [m_{jk}]_{1\leq j,k\leq n}$, we denote

$$M^* = [\overline{m_{kj}}]_{1\leq j,k\leq n}$$

Lecture 33 - Monday, Jul 22

**Discovery 12.2**

We have the following formulas: For all $S, T \in \mathcal{L}(V)$ and $\alpha \in \mathbb{C}$

1. $(S+T)^* = S^* + T^*$;
2. $(\alpha T)^* = \overline{\alpha} T^*$;
3. $(ST)^* = T^* S^*$;
4. $(T^*)^* = T$

*Proof.* Here we verify the second formula. STP that the operator $R := \overline{\alpha} T^*$ has the property that

$$\langle (\alpha T()x), y\rangle = \langle x, R(x)\rangle \qquad \forall\, x, y \in V$$

We process as following:

$$\begin{aligned}\langle (\alpha T()x), y\rangle &= \langle \alpha \cdot T(x), y\rangle \\ &= \alpha\langle T(x), y\rangle \\ &= \alpha\langle x, T^*(y)\rangle \\ &= \langle x, \overline{\alpha} T^*(y)\rangle = \langle x, R(y)\rangle\end{aligned}$$

$\square$

**Definition 12.3: Some Special Classes of Operators**

1. **Self-adjoint**: $T \in \mathcal{L}(V)$ such that $T = T^*$;
2. **Unitary**: $U \in \mathcal{L}(V)$ such that $U^*U = I = UU^*$;
3. **Normal**: $T \in \mathcal{L}(V)$ such that $TT^* = T^*T$;

## 12.2 Eigenvalues and Eigenvectors

**Definition 12.4**

For $T \in \mathcal{L}(V)$, define

1. What it means for $\lambda \in \mathbb{C}$ to be an eigenvalue of $T$:

$$\text{Null}(T - \lambda I) \neq \{0_V\}$$

2. What it means for $x \in V$ to be an eigenvector for $T$, corresponding to eigenvalue $\lambda$:

$$x \neq 0_V \qquad \text{and} \qquad T(x) = \lambda x$$

**Proposition 12.1**

Every $T \in \mathcal{L}(V)$ has eigenvalues.
**Important:** we are working over $\mathbb{C}$.

*Proof.* Fix $\zeta_1, \ldots, \zeta_n$ an orthonormal basis for $V$ and look at $A_T \in \mathcal{M}_n(\mathbb{C})$. For $\lambda \in \mathbb{C}$ we have

$$\begin{aligned} \lambda \text{ is not an eigenvalue} &\Leftrightarrow \text{Null}(T - \lambda I) = \{0_V\} \\ &\Leftrightarrow T - \lambda I \text{ is invertible} \\ &\Leftrightarrow A_T - \lambda I_n \text{ is an invertible matrix} \\ &\Leftrightarrow \det(A_T - \lambda I_n) \neq 0 \\ &\Leftrightarrow P(\lambda) \neq 0 \end{aligned}$$

Hence $\lambda$ is an eigenvalue of $T$ if and only if $P(\lambda) = 0$ where $P$ is the characteristic polynomial of $A_T$. By the fundamental theorem of algebra, we know that such $\lambda$'s exist. □

**Example 12.1**

Assume for the moment that our inner product space is $V = \mathbb{C}^2$ endowed with the standard inner product. Let $T \in \mathcal{L}(V)$ be defined by the following:

$$T((z_1, z_2)) = (-z_2, z_1) \qquad \forall\, z_1, z_2 \in \mathbb{C}$$

We wish to find $\lambda \in \mathbb{C}$ and $(z_1, z_2) \neq (0, 0)$ such that

$$T((z_1, z_2)) = \lambda(z_1, z_2)$$

Observe that we need to have $\lambda^2 = 1$, so possible eigenvalues are $\lambda_1 = i$ and $\lambda_2 = -i$. Observe $\zeta_1 = (i, 1)$, then

$$T(\zeta_1) = (-1, i) = i(i, 1)$$

Likewise, we can also see that for $\zeta_2 = (1, i)$:

$$T(\zeta_2) = (-i, 1) = -i(1, i)$$

Lecture 34 - Wednesday, Jul 24

## 13 Spectral Theorem for a Normal Operator over $\mathbb{C}$

In this lecture we consider the following setting: $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{C}$ where $\dim(V) = n \in \mathbb{N}$. We have $T \in \mathcal{L}(V)$ is normal (that is, $TT^* = T^*T$). We will prove that one can find

$$\lambda_1, \ldots, \lambda_n \in \mathbb{C} \quad \text{and o.n.b. } \zeta_1, \ldots, \zeta_n \text{ for } V$$

such that

$$T(\zeta_1) = \lambda_1 \zeta_1, \ \ldots, \ T(\zeta_n) = \lambda_n \zeta_n$$

**Lemma 13.1**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{C}$ with $\dim(V) = n \in \mathbb{N}$. Let $S \in \mathcal{L}(V)$ be normal, then

$$\|S(x)\| = \|S^*(x)\| \qquad \forall\, x \in V$$

*Proof.* Compute

$$\|S^*(x)\|^2 = \langle S^*(x), S^*(x) \rangle = \langle x, S^{**}(S^*(x)) \rangle = \langle x, SS^*(x) \rangle = \langle x, S^*S(x) \rangle = \langle S(x), S(x) \rangle = \|S(x)\|^2$$

□

**Lemma 13.2**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{C}$ with $\dim(V) = n \in \mathbb{N}$. Let $T \in \mathcal{L}(V)$ be normal, then for every $\lambda \in \mathbb{C}$ and $x \in V$ we have

$$\|T(x) - \lambda x\| = \left\|T^*(x) - \overline{\lambda} x\right\|$$

*Proof.* Let $S = T - \lambda I \in \mathcal{L}(V)$, we observe that $S$ is normal:

$$\begin{aligned} SS^* = (T - \lambda I)(T^* - \overline{\lambda} I) &= TT^* - \lambda T^* - \overline{\lambda} T + \lambda \overline{\lambda} I \\ &= T^*T - \lambda T^* - \overline{\lambda} T + \lambda \overline{\lambda} I = (T^* - \overline{\lambda} I)(T - \lambda I) \end{aligned}$$

Then applying the previous Lemma on $S$ yields us the desired result. □

**Proposition 13.1**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{C}$ with $\dim(V) = n \in \mathbb{N}$. Let $T \in \mathcal{L}(V)$ be normal. Suppose $\lambda \in \mathbb{C}$ is an eigenvalue for $T$ with eigenvector $x \in V \backslash \{0_V\}$. Then the same vector $x \in V \backslash \{0_V\}$ is an eigenvector for operator $T^*$ corresponding to to eigenvalue $\overline{\lambda}$.

*Proof.* The proof is immediate after the proceeding Lemma. □

### 13.1 Normal Implies Diagonalizability

**Theorem 13.1**

Let $n \in \mathbb{N}$ and let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{C}$ with dimension $n$. Let $T \in \mathcal{L}(V)$ be normal, then there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and an orthonormal basis $\zeta_1, \ldots, \zeta_n$ for $V$ such that

$$T(\zeta_1) = \lambda_1 \zeta_1, \quad \ldots, \quad T(\zeta_n) = \lambda_n \zeta_n$$

*Proof.* This is a $\mathbb{C}$-analogue for the proof for Theorem (8.1). The proof is by induction on $n$:

1. *Base Case*: $n = 1$,
   We can pick $\zeta_1 \in V$ with $\|\zeta_1\| = 1$. Since $\dim(V) = 1$, we must have

   $$V = \{\lambda \zeta_1 : \lambda \in \mathbb{C}\}$$

   Hence there exists $\lambda_1 \in \mathbb{C}$ such that $T(\zeta_1) = \lambda_1 \zeta_1$.

2. *Induction Step*: $n - 1 \;\Rightarrow\; n$ (with $n \geq 2$)
   Pick $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{C}$ with dimension $n$ and pick $T \in \mathcal{L}(V)$ an normal operator. We will verify some claims about $T$:

   (a) *Claim 1*: we can find $\lambda \in \mathbb{C}$ and $\zeta \in V$ with $\|\zeta\| = 1$ such that $T(\zeta) = \lambda \zeta$.
   This is the fundamental fact that $T$ is sure to have eigenvalues in $\mathbb{C}$ (see Proposition (12.1)).
   We denote the eigenpair as $(\lambda_n, \zeta_n)$. Recall that Proposition (13.1) tells us that

   $$T^*(\zeta_n) = \overline{\lambda_n} \zeta_n$$

   Also denote $\{Y = \{y \in V : \langle y, \zeta_n \rangle = 0\}\}$

   (b) *Claim 2*: $Y$ is a linear subspace of $V$ with $\dim(Y) = n - 1$.
   The claim holds because $Y = W^{\perp}$ for $W = \{\alpha \zeta_n : \alpha \in \mathbb{C}\}$. This implies that $Y$ is a linear subspace with

   $$\dim(Y) = \dim(W^{\perp}) = \dim(V) - \dim(W) = n - 1$$

   Lecture 35 - Friday, Jul 26

   (c) *Claim 3*:

   i. We have

   $$[y \in Y] \quad \Rightarrow \quad [T(y) \in Y]$$

   Hence the operator $T \in \mathcal{L}(V)$ induces (by restriction) an operation $T_0 \in \mathcal{L}(Y)$.

   ii. We have

   $$[y \in Y] \quad \Rightarrow \quad [T^*(y) \in Y]$$

   Hence the operator $T^* \in \mathcal{L}(V)$ induces (by restriction) an operation $S_0 \in \mathcal{L}(Y)$.

   iii. In the inner product space $(Y, \langle \cdot, \cdot \rangle)$, we have $S_0 = T_0^*$, where $S_0$ and $T_0$ are the restrictions of $T^*$ and $T$ respectively.

   We verify each statements in Claim 3:

i. Pick $y \in Y$, look at $T(y)$. We have

$$\langle T(y), \zeta_n \rangle = \langle y, T^*(\zeta_n) \rangle = \langle y, \overline{\lambda_n} \zeta_n \rangle = \lambda_n \langle y, \zeta_n \rangle = 0$$

so indeed $T(y) \in Y$.

ii. Same trick. Pick $y \in Y$, look at $T^*(y)$. We have

$$\langle T^*(y), \zeta_n \rangle = \langle y, T^{**}(\zeta_n) \rangle = \langle y, T(\zeta_n) \rangle = \langle y, \lambda_n \zeta_n \rangle = \overline{\lambda_n} \langle y, \zeta_n \rangle = 0$$

so indeed $T(y) \in Y$.

iii. We have for $y, y' \in Y$: $\langle T_0(y), y' \rangle = \langle T(y), y' \rangle = \langle y, T^*(y') \rangle = \langle y, S_0(y') \rangle$.

(d) *Claim 4*: The opreator $T_0 \in \mathcal{L}(V)$ which appeared in Claim 3 is *normal*.
We must check that $T_0 T_0^* = T_0^* T_0$, i.e. $T_0 S_0 = S_0 T_0$. Pick $y \in Y$ and write

$$(T_0 S_0)(y) = T_0(S_0(y)) = T(S_0(y)) = T(T^*(y)) = (TT^*)(y)$$

Likewise, we also find that

$$(S_0 T_0)(y) = (T^* T)(y)$$

so we have

$$(T_0 S_0)(y) = (TT^*)(y) = (T^* T)(y) = (S_0 T_0)(y)$$

3. Now we apply the induction to the normal operator $T_0 \in \mathcal{L}(V)$. Write $\dim(Y) = n - 1$, this gives us $\lambda_1, \dots, \lambda_{n-1} \in \mathbb{C}$ and an orthonormal basis $\zeta_1, \dots, \zeta_{n-1} \in Y$ such that

$$T_0(\zeta_j) = \lambda_j \zeta_j \qquad \forall\, 1 \le j \le n-1$$

Thus we have $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ and $\zeta_1, \dots, \zeta_n \in V$ satisfy the requirements in the theorem.

□

**Discovery 13.1**

For $(V, \langle \cdot, \cdot \rangle)$ an inner product space over $\mathbb{C}$ with $\dim(V) = n \in \mathbb{N}$ and $T \in \mathcal{L}(V)$ is normal. Denote

$$\operatorname{Spec}(T) = \{\gamma_1, \dots, \gamma_m\} \subset \mathbb{C}$$

For $1 \le j \le m$ let

$$E_j = \{x \in V : T(x) = \gamma_j x\}$$

and let $P_j \in \mathcal{L}(V)$ be the orthogonal projection operator onto $E_j$, then we get

$$P_1 + \cdots + P_m = I$$

with $P_j \perp P_k$ for $j \ne k$. We also get

$$T = \gamma_1 P_1 + \cdots + \gamma_m P_m$$

This allows us to define functional calculus for any $f : \operatorname{Spec}(T) \to \mathbb{C}$.

## 14 The Cayley-Hamilton Theorem (with calculus)

**Framework:** for $n \in \mathbb{N}$, look at $\mathcal{M}_n(\mathbb{R})$

**Definition 14.1**

For $A \in \mathcal{M}_n(\mathbb{R})$, let $Q_A$ be the characteristic polynomial of $A$

$$Q_A(\lambda) = \det(\lambda I_n - A) = c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + \lambda^n$$

for some $c_0, \ldots, c_{n-1} \in \mathbb{R}$.

**Theorem 14.1**

Notation as above, we look at

$$Q_A(A) := c_0 I_n + c_1 A + \cdots + c_{n-1}A^{n-1} + A^n$$

We have $Q_A(A) = \mathcal{O}$, the matrix whose entries are all 0.

Lecture 36 - Monday, Jul 29

**Example 14.1**

Say $n = 2$, and $A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$, then

$$Q_A(\lambda) = \det \begin{bmatrix} \alpha_{11} - \lambda & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - \lambda \end{bmatrix} = \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})$$

Note that $c_0 = \det(A)$ and $c_1 = \operatorname{tr}(A)$. Cayley-Hamilton tells us that

$$c_0 I_2 - c_1 A + A^2 = 0$$

or in other words,

$$A^2 = [\operatorname{tr}(A)]A - [\det(A)]I_2$$

**Definition 14.2**

Let $A \in \mathcal{M}_n(\mathbb{R})$ be a matrix, $A = [\alpha_{ij}]_{1 \le i,j \le n}$, and let $(A_k)_{k=1}^{\infty}$ be a sequence in $\mathcal{M}_n(\mathbb{R})$ where

$$A_k = \left[\alpha_{ij}^{(k)}\right]_{1 \le i,j \le n} \qquad \text{for } k \in \mathbb{N}$$

We say that

$$A_k \longrightarrow_{k \to \infty} A$$

to mean that *for every $i, j \in \{1, \ldots, n\}$ we have $\alpha_{ij}^{(k)} \longrightarrow_{k \to \infty} \alpha_{ij}$ (convergence in $\mathbb{R}$).*

**Discovery 14.1**

Convergence is well-behaved with respect to matrix operations. More precisely, if $A_k \longrightarrow A$, $B_k \longrightarrow B$, and $c_k \longrightarrow c$, then

1. $A_k + B_k \longrightarrow_{k\to\infty} A + B$;
2. $A_k \cdot B_k \longrightarrow_{k\to\infty} A \cdot B$;
3. $c_k \cdot A_k \longrightarrow_{k\to\infty} c \cdot A$;

**Discovery 14.2**

Suppose $A_k \longrightarrow_{k\to\infty} A$ in $\mathcal{M}_n(\mathbb{R})$. Consider the characteristic polynomials

$$\begin{aligned} Q_A(\lambda) &= c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + \lambda^n \\ Q_{A_k}(\lambda) &= c_0^{(k)} + c_1^{(k)}\lambda + \cdots + c_{n-1}^{(k)}\lambda^{n-1} + \lambda^n \end{aligned}$$

Then we have

$$c_i^{(k)} \longrightarrow_{k\to\infty} c_i \qquad \text{for } i \in \{1, \dots, n-1\}$$

This holds because the coefficients $c_j$ and $c_j^{(k)}$ have explicit formulas in terms of the entries of the corresponding matrices.

**Lemma 14.1**

Suppose that $A_k \longrightarrow_{k\to\infty} A$ in $\mathcal{M}_n(\mathbb{R})$, then

$$Q_{A_k}(A_k) \longrightarrow_{k\to\infty} Q_A(A)$$

*Proof.* **Exercise.** (Using the properties shown above that the convergence is well-behaved with respect to matrix operations.) □

**Definition 14.3: Dense**

A set $\mathcal{S} \subset \mathcal{M}_n(\mathbb{R})$ is said to be **dense** in $\mathcal{M}_n(\mathbb{R})$ when the following holds:
For every $A \in \mathcal{M}_n(\mathbb{R})$, there exists $(A_k)_{k=1}^{\infty}$ sequence in $\mathcal{S}$ such that $A_k \longrightarrow_{k\to\infty} A$.

**Proposition 14.1**

Suppose we were able to find $\mathcal{S} \subset \mathcal{M}_n(\mathbb{R})$ such that

1. $Q_A(A) = \mathcal{O}$ for all $A \in \mathcal{S}$;
2. $\mathcal{S}$ is dense in $\mathcal{M}_n(\mathbb{R})$,

then it will follow that $Q_A(A) = \mathcal{O}$ for all $A \in \mathcal{M}_n(\mathbb{R})$, hence the Cayley Hamilton Theorem holds.

*Proof.* Let $A \in \mathcal{M}_n(\mathbb{R})$. Pick $(A_k)_{k=1}^{\infty}$ in $\mathcal{S}$ such that $A_k \longrightarrow_{k\to\infty} A$. The above lemma states that $Q_{A_k}(A_k) \longleftrightarrow_{k\to\infty} Q_A(A)$, which implies that $Q_A(A) = \mathcal{O}$. □

**Proposition 14.2**

Let $\mathcal{S} = \{A \in \mathcal{M}_n(\mathbb{R}) : Q_A \text{ has distinct eigenvalues in } \mathbb{C}\}$, then

1. $Q_A(A) = \mathcal{O}$ for all $A \in \mathcal{S}$;
2. $\mathcal{S}$ is dense in $\mathcal{M}_n(\mathbb{R})$,

*Proof.*

1. From Math146 we know that if $\lambda_1, \ldots, \lambda_n$ are eigenvalues, then $A$ is diagonalizable over $\mathbb{C}$. That is, there exists $S \in \mathcal{M}_n(\mathbb{C})$, invertible, such that

$$A = S \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} S^{-1}$$

then we get

$$A^p = S \begin{bmatrix} \lambda_1^p & & 0 \\ & \ddots & \\ 0 & & \lambda_n^p \end{bmatrix} S^{-1} \qquad \forall\, p \in \mathbb{N}$$

Then take linear combinations we get that for every polynomial $q$, it holds that

$$q(A) = S \begin{bmatrix} q(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & q(\lambda_n) \end{bmatrix} S^{-1}$$

In particular for $q = Q_A$, we get

$$Q_A(A) = S \begin{bmatrix} Q_A(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & Q_A(\lambda_n) \end{bmatrix} S^{-1} = S \begin{bmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{bmatrix} S^{-1} = \mathcal{O}$$

2. Use the notion of *discriminant* of a matrix $A$:

$$\Delta(A) = \prod_{1 \le i,j \le n} (\lambda_i - \lambda_j)^2$$

Observe that

$$\mathcal{S} = \{A \in \mathcal{M}_n(\mathbb{R}) : \Delta(A) \neq 0\}$$

□