

# Math 245

James Yu

September 28, 2017



# Contents

<b>1</b>	<b>September 8 - September 13</b>	<b>1</b>
1.1	Inner Product . . . . .	1
1.2	Cauchy-Bunyakovsky-Schwarz Inequality . . . . .	2
1.3	Orthogonality . . . . .	3
1.4	Gram-Schmidt Procedure . . . . .	4
1.5	Orthogonal Complement . . . . .	5
<b>2</b>	<b>September 15 - September 20</b>	<b>7</b>
2.1	Orthogonal Complement (continued) . . . . .	7
2.2	Adjoint . . . . .	8
2.3	Least Squares (example) . . . . .	10
<b>3</b>	<b>September 22 - September 27</b>	<b>13</b>
3.1	Normal Operators . . . . .	13
3.2	Self-Adjoint Operators . . . . .	15
3.3	Isometries . . . . .	16



# Week 1

## September 8 - September 13

### 1.1 Inner Product

**Definition 1.1.1** (Inner Product Space). An *inner product space* (over  $\mathbb{C}$ ) is a *vector space*  $V$  and a function:

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$$

satisfying:

- i.  $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$  for all  $\vec{x}, \vec{y}, \vec{z} \in V$
- ii.  $\langle c\vec{x}, \vec{y} \rangle = c \langle \vec{x}, \vec{y} \rangle$  for all  $\vec{x}, \vec{y} \in V$  and  $c \in \mathbb{C}$
- iii.  $\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}$  for all  $\vec{x}, \vec{y} \in V$
- iv.  $\langle \vec{x}, \vec{x} \rangle \in \mathbb{R}_{\geq 0}$  if  $\vec{x} \neq \vec{0}$ ,  $\langle \vec{x}, \vec{x} \rangle = 0$  otherwise

**Example 1.1.1.1** (Standard Complex Inner Product).

$$V = \mathbb{C}^n \quad \langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = x_1 \overline{y_1} + \dots + x_n \overline{y_n}$$

Properties *i*, *ii*, and *iii* clearly hold. For *iv*, for any  $\vec{x} = (a_1 + b_1 i, \dots, a_n + b_n i)$

$$\langle \vec{x}, \vec{x} \rangle = (a_1 + b_1 i) \overline{(a_1 + b_1 i)} + \dots + (a_n + b_n i) \overline{(a_n + b_n i)} \quad (1.1)$$

$$= a_1^2 + b_1^2 + \dots + a_n^2 + b_n^2 \quad (1.2)$$

This is the *standard complex inner product*. If we replace  $\mathbb{C}^n$  with  $\mathbb{R}^n$  then we get the *standard real inner product* (dot product).

**Example 1.1.1.2** ( $L^2$  Inner Product).

$$V = C([0, 1]) \quad \langle f, g \rangle = \int_0^1 f \overline{g}$$

This is called the  $L^2$  inner product on  $V$

**Example 1.1.1.3** (Frobenius Inner Product).

$$V = M_n(\mathbb{C}) \qquad \langle A, B \rangle = \text{tr}(A\overline{B}^\top)$$

This is called the *Frobenius inner product* on  $V$ . It satisfies *iv* because, for  $A = (a_{ij})$ ,  $B = (b_{ij})$  we have:

$$\text{tr}(A\overline{B}^\top) = \sum a_{ij}\overline{b_{ij}}$$

which is equivalent to the *standard complex inner product*.

## 1.2 Cauchy-Bunyakovsky-Schwarz Inequality

**Definition 1.2.1** (Length). If  $\vec{v}$  is a vector in an *inner product space*, the *length* of  $\vec{v}$  is:

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

**Theorem 1.2.2** (Cauchy-Schwarz). *Let  $\vec{x}, \vec{y} \in V$  be vectors in an inner product space, then:*

$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\|$$

*Proof.* If  $\vec{y} = \vec{0}$ , this is trivial. Otherwise, for any  $c \in \mathbb{C}$

$$0 \leq \|\vec{x} - c\vec{y}\|^2 \tag{1.3}$$

$$\leq \|\vec{x}\|^2 - \bar{c}\langle \vec{x}, \vec{y} \rangle - c\langle \vec{y}, \vec{x} \rangle + c\bar{c}\|\vec{y}\|^2 \tag{1.4}$$

So, let  $c = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2}$ :

$$0 \leq \|\vec{x}\|^2 - \frac{\overline{\langle \vec{x}, \vec{y} \rangle}}{\|\vec{y}\|^2} \langle \vec{x}, \vec{y} \rangle - \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2} \langle \vec{y}, \vec{x} \rangle + \frac{\langle \vec{x}, \vec{y} \rangle \overline{\langle \vec{x}, \vec{y} \rangle}}{\|\vec{y}\|^4} \|\vec{y}\|^2 \tag{1.5}$$

$$\leq \|\vec{x}\|^2 - \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2} - \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2} + \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2} \tag{1.6}$$

$$|\langle \vec{x}, \vec{y} \rangle|^2 \leq \|\vec{x}\|^2 \|\vec{y}\|^2 \tag{1.7}$$

$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\| \tag{1.8}$$

□

**Remark.** We can define the angle between  $\vec{x}, \vec{y}$  as  $\cos^{-1} \frac{|\langle \vec{x}, \vec{y} \rangle|}{\|\vec{x}\| \|\vec{y}\|}$ .

## 1.3 Orthogonality

**Definition 1.3.1** (Orthogonality). Two vectors  $\vec{x}, \vec{y}$  are *orthogonal* if and only if  $\langle \vec{x}, \vec{y} \rangle = 0$ .

**Definition 1.3.2** (Unit vector). A *unit vector* is a vector of length 1.

**Definition 1.3.3** (Orthogonal Set). An *orthogonal set* is a set  $S$  where  $\langle \vec{x}, \vec{y} \rangle = 0$  for all  $\vec{x}, \vec{y} \in S, \vec{x} \neq \vec{y}$ .

**Definition 1.3.4** (Orthonormal Set). An *orthonormal set* is an *orthogonal set* in which each vector is a *unit vector*.

**Example 1.3.4.1** (Standard Basis). The standard basis in  $\mathbb{R}^n$  is *orthonormal*

**Theorem 1.3.5** (Orthonormal Coordinates). Let  $\{\vec{v}_1, \dots, \vec{v}_n\}$  be an orthogonal basis of an inner product space  $V$ . Then for any  $\vec{x} \in V$  we have:

$$\vec{x} = \sum_{i=1}^n \frac{\langle \vec{x}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$$

*Proof.* Write  $\vec{x} = a_1 \vec{v}_1 + \dots + a_n \vec{v}_n$ . Then, for any  $i$ :

$$\langle \vec{x}, \vec{v}_i \rangle = \langle a_1 \vec{v}_1 + \dots + a_n \vec{v}_n, \vec{v}_i \rangle \quad (1.9)$$

$$= a_1 \langle \vec{v}_1, \vec{v}_i \rangle + \dots + a_n \langle \vec{v}_n, \vec{v}_i \rangle \quad (1.10)$$

$$= a_i \langle \vec{v}_i, \vec{v}_i \rangle \quad (1.11)$$

$$a_i = \frac{\langle \vec{x}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \quad (1.12)$$

□

**Remark.** If  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is *orthonormal*, then  $\vec{x} = \sum_{i=1}^n \langle \vec{x}, \vec{v}_i \rangle \vec{v}_i$

**Remark.** The  $\vec{v}_i$  coordinate of  $\vec{x}$  depends only  $\vec{x}$  and  $\vec{v}_i$ . It does not depend on any other vectors in the basis.

**Remark.** In finite dimensions, *inner product spaces* always have *orthonormal bases*.

**Theorem 1.3.6** (Orthogonal  $\implies$  Linear Independence). Let  $S$  be an orthogonal set of non-zero vectors, then  $S$  is linearly independent.

*Proof.* For any  $\vec{v}_1, \dots, \vec{v}_n \in S$ , set  $a_1 \vec{v}_1 + \dots + a_n \vec{v}_n = \vec{0}$ . By similar construction as 1.3.5, we can show for any  $i$  that:

$$a_i \langle \vec{v}_i, \vec{v}_i \rangle = 0$$

Since  $\vec{v}_i \neq \vec{0}$  by assumption,  $a_i = 0$  for all  $i$ . □

## 1.4 Gram-Schmidt Procedure

Given a basis  $\{\vec{w}_1, \dots, \vec{w}_n\}$  for a (finite dimensional) *inner product space*  $V$ , the Gram-Schmidt gives an *orthogonal basis* for  $V$  as follows:

Step ① Set  $\vec{v}_1 = \vec{w}_1$

Step ② Set  $\vec{v}_2 = \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1$

...

Step ① Set  $\vec{v}_i = \vec{w}_i - \sum_{j=1}^{i-1} \frac{\langle \vec{w}_i, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \vec{v}_j$

**Claim 1.4.1.**  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is an *orthogonal basis* of  $V$

*Proof.* We first check that  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is *orthogonal*.

We proceed by induction on  $i$ . If  $n = 1$ , we are vacuously done.

Otherwise, assume that  $\{\vec{v}_1, \dots, \vec{v}_i\}$  is *orthogonal*. For any  $1 \leq j \leq i$  we have:

$$\langle \vec{v}_{i+1}, \vec{v}_j \rangle = \left\langle \vec{w}_{i+1} - \sum_{k=1}^i \frac{\langle \vec{w}_{i+1}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \vec{v}_k, \vec{v}_j \right\rangle \quad (1.13)$$

$$= \left\langle \vec{w}_{i+1} - \frac{\langle \vec{w}_{i+1}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \vec{v}_j, \vec{v}_j \right\rangle \quad (1.14)$$

$$= \langle \vec{w}_{i+1}, \vec{v}_j \rangle - \frac{\langle \vec{w}_{i+1}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \|\vec{v}_j\|^2 \quad (1.15)$$

$$= \langle \vec{w}_{i+1}, \vec{v}_j \rangle - \langle \vec{w}_{i+1}, \vec{v}_j \rangle \quad (1.16)$$

$$= 0 \quad (1.17)$$

Furthermore,  $\vec{v}_i \neq \vec{0}$ . For  $i = 1$  we have  $\vec{v}_1 = \vec{w}_1 \neq \vec{0}$  by assumption. Otherwise, we have  $\vec{v}_i = \vec{w}_i - \vec{x}$  for  $x \in \text{span}\{\vec{w}_1, \dots, \vec{w}_{i-1}\}$ . Thus  $\vec{v}_i$  is a nonzero linear combination of  $\{\vec{w}_1, \dots, \vec{w}_i\}$  and is therefore non-zero.

Thus  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is a set of  $n$  vectors in an  $n$ -dimensional space that are orthogonal and nonzero. By 1.3.6 they are linearly independent, and thus a basis of  $V$ .  $\square$

**Remark.** To obtain an *orthonormal basis* of  $V$ , simply divide each  $\vec{v}_i$  by its length. This is called *normalizing*.



## 1.5 Orthogonal Complement

**Definition 1.5.1** (Orthogonal Complement). Let  $V$  be an *inner product space* and  $W \subset V$  a subspace. The orthogonal complement of  $W$  is:

$$W^\perp = \{\vec{v} \in V : \langle \vec{v}, \vec{w} \rangle = 0 \text{ for all } \vec{w} \in W\}$$

**Example 1.5.1.1** (Orthogonal complement of entire space is empty).

$$\begin{aligned} W = V \qquad W^\perp &= \{\vec{v} \in V : \langle \vec{v}, \vec{w} \rangle = 0 \text{ for all } \vec{w} \in V\} = \{\vec{0}\} \\ &\text{because } \langle \vec{v}, \vec{v} \rangle = 0 \implies \vec{v} = \vec{0} \end{aligned}$$

**Example 1.5.1.2** (Orthogonal complement of empty subspace is the entire space).

$$W = \{\vec{0}\} \qquad W^\perp = V$$

**Example 1.5.1.3** (Orthogonal complement of a line is a plane).

$$V = \mathbb{R}^3 \qquad W = \{(0, 0, z) : z \in \mathbb{R}\} \qquad W^\perp = \{(x, y, 0) : x, y \in \mathbb{R}\}$$



## Week 2

# September 15 - September 20

### 2.1 Orthogonal Complement (continued)

**Theorem 2.1.1.** *Let  $V$  be a finite-dimensional inner product space, and  $W \subset V$  be a subspace, then:*

$$V \simeq W \oplus W^\perp$$

via the transformation  $T : W \oplus W^\perp \rightarrow V$  given by:

$$T(\vec{w}, \vec{w}') = \vec{w} + \vec{w}'$$

*Proof.* We prove the theorem by writing an inverse for  $T$ . Let  $\{\vec{w}_1, \dots, \vec{w}_k\}$  be an orthonormal basis of  $W$  and define:

$$\Psi : V \rightarrow W \oplus W^\perp \quad (2.1)$$

$$\Psi(\vec{v}) = \left( \sum_{i=1}^k \langle \vec{v}, \vec{w}_i \rangle \vec{w}_i, \vec{v} - \sum_{i=1}^k \langle \vec{v}, \vec{w}_i \rangle \vec{w}_i \right) \quad (2.2)$$

$\Psi$  is well defined since the first entry is in  $W$  by being a linear combination of  $\vec{w}_i$ , and the right entry is in  $W^\perp$  because it is orthogonal to each  $\vec{w}_i$  in our basis. It clear that  $T \circ \Psi = \text{id}_V$ , so it remains to be shown that  $\Psi \circ T = \text{id}_{W \oplus W^\perp}$ :

$$\Psi(T(\vec{w}, \vec{w}')) = \Psi(\vec{w} + \vec{w}') \quad (2.3)$$

$$= \left( \sum_{i=1}^k \langle \vec{w} + \vec{w}', \vec{w}_i \rangle \vec{w}_i, \vec{w} + \vec{w}' - \sum_{i=1}^k \langle \vec{w} + \vec{w}', \vec{w}_i \rangle \vec{w}_i \right) \quad (2.4)$$

$$= \left( \sum_{i=1}^k \langle \vec{w}, \vec{w}_i \rangle \vec{w}_i, \vec{w} + \vec{w}' - \sum_{i=1}^k \langle \vec{w}, \vec{w}_i \rangle \vec{w}_i \right) \quad (2.5)$$

$$= (\vec{w}, \vec{w}') \quad (2.6)$$

Thus  $T$  and  $\Psi$  are inverses. Since  $T$  and  $\Psi$  are linear transformations,  $T$  is an isomorphism.  $\square$

**Corollary 2.1.1.1** (Extension of orthonormal basis). *Let  $\{\vec{w}_1, \dots, \vec{w}_k\}$  be an orthonormal basis of a subspace  $W$ . One can extend this to an orthonormal basis of the entire space:*

$$\{\vec{w}_1, \dots, \vec{w}_k, \vec{v}_{k+1}, \dots, \vec{v}_n\}$$

where  $\{\vec{v}_{k+1}, \dots, \vec{v}_n\}$  is an orthonormal basis of  $W^\perp$ .

**Corollary 2.1.1.2** (Dimension of orthogonal complement).

$$\dim V = \dim W + \dim W^\perp$$

**Corollary 2.1.1.3** (Duality of orthogonal complement).

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

**Corollary 2.1.1.4** (Intersection of subspace and orthogonal complement).

$$W \cap W^\perp = \{0\}$$

**Definition 2.1.2** (Projection onto a subspace). Let  $W \subset V$  be a subspace and  $\vec{v} \in V$ . Then for  $\Psi : V \rightarrow W \oplus W^\perp$  as defined in 2.1.1, we define the *projection of  $\vec{v}$  onto  $W$*  to be the first coordinate  $\Psi(\vec{v})$ , denoted:

$$\text{proj}_W(\vec{v})$$

**Remark.** If  $\{\vec{w}_1, \dots, \vec{w}_k\}$  is an orthonormal basis of  $W$ , then:

$$\text{proj}_W(\vec{v}) = \sum_{i=1}^k \langle \vec{v}, \vec{w}_i \rangle \vec{w}_i$$

## 2.2 Adjoints

**Definition 2.2.1** (Conjugate Transpose). For any matrix  $B$ , we define  $B^*$  to be the *conjugate transpose* given by taking the conjugate of each entry in  $B^\top$ , that is:

$$B^* = \overline{B^\top}$$

**Lemma 2.2.2** (Unique inner product form of a linear transformation). *Let  $\mathcal{U} : V \rightarrow \mathbb{F}$  be a linear transformation, then there exists some unique  $z \in V$  such that:*

$$\mathcal{U}(\vec{v}) = \langle \vec{v}, \vec{z} \rangle \text{ for all } \vec{v} \in V$$

*Proof.* Let  $\{\vec{v}_1, \dots, \vec{v}_n\}$  be an orthonormal basis of  $V$  and define  $\vec{z} \in V$  to be:

$$\vec{z} = \overline{\mathcal{U}(\vec{v}_1)}\vec{v}_1 + \dots + \overline{\mathcal{U}(\vec{v}_n)}\vec{v}_n$$

Then we check that  $\mathcal{U}(\vec{v}) = \langle \vec{v}, \vec{z} \rangle$  for all  $\vec{v} \in V$ :

$$\mathcal{U}(\vec{v}) = \mathcal{U}(a_1 \vec{v}_1 + \dots + a_n \vec{v}_n) \quad (2.7)$$

$$= a_1 \mathcal{U}(\vec{v}_1) + \dots + a_n \mathcal{U}(\vec{v}_n) \quad (2.8)$$

$$\langle \vec{v}, \vec{z} \rangle = \left\langle \vec{v}, \overline{\mathcal{U}(\vec{v}_1)} \vec{v}_1 + \dots + \overline{\mathcal{U}(\vec{v}_n)} \vec{v}_n \right\rangle \quad (2.9)$$

$$= \mathcal{U}(\vec{v}_1) \langle \vec{v}, \vec{v}_1 \rangle + \dots + \mathcal{U}(\vec{v}_n) \langle \vec{v}, \vec{v}_n \rangle \quad (2.10)$$

$$= a_1 \mathcal{U}(\vec{v}_1) + \dots + a_n \mathcal{U}(\vec{v}_n) \quad (2.11)$$

$$= \mathcal{U}(\vec{v}) \quad (2.12)$$

To show that  $\vec{z}$  is unique, suppose that  $\langle \vec{v}, \vec{z} \rangle = \langle \vec{v}, \vec{z}' \rangle$  for all  $\vec{v} \in V$ , then:

$$0 = \langle \vec{v}, \vec{z} \rangle - \langle \vec{v}, \vec{z}' \rangle = \langle \vec{v}, \vec{z} - \vec{z}' \rangle$$

Since this holds for all  $\vec{v}$ , we must have  $\vec{z} - \vec{z}' = 0$  (indeed,  $V^\perp = (0)$ ), we have  $\vec{z}' = \vec{z}$  as required.  $\square$

**Theorem 2.2.3** (Existence of unique adjoint). *Let  $T : V \rightarrow V$  be a linear transformation on an inner product space  $V$ . There exists a unique linear transformation  $T^* : V \rightarrow V$  satisfying:*

$$\langle T\vec{x}, \vec{y} \rangle = \langle \vec{x}, T^*\vec{y} \rangle \text{ for all } \vec{x}, \vec{y} \in V$$

*This  $T^*$  is called the adjoint of  $T$ .*

*Proof.* For any  $\vec{y} \in V$ , define  $g_{\vec{y}} : V \rightarrow \mathbb{F}$  (where  $\mathbb{F}$  is  $\mathbb{C}$  or  $\mathbb{R}$ ), by:

$$g_{\vec{y}}(\vec{v}) = \langle T\vec{v}, \vec{y} \rangle$$

Then  $g_{\vec{y}}$  is a linear transformation because:

$$g_{\vec{y}}(\vec{v} + \vec{w}) = \langle T(\vec{v} + \vec{w}), \vec{y} \rangle \quad (2.13)$$

$$= \langle T\vec{v} + T\vec{w}, \vec{y} \rangle \quad (2.14)$$

$$= \langle T\vec{v}, \vec{y} \rangle + \langle T\vec{w}, \vec{y} \rangle \quad (2.15)$$

$$= g_{\vec{y}}(\vec{v}) + g_{\vec{y}}(\vec{w}) \quad (2.16)$$

$$g_{\vec{y}}(c\vec{v}) = \langle T(c\vec{v}), \vec{y} \rangle \quad (2.17)$$

$$= c \langle T\vec{v}, \vec{y} \rangle \quad (2.18)$$

$$= c g_{\vec{y}}(\vec{v}) \quad (2.19)$$

Then we can define  $T^* : V \rightarrow V$  by the map from  $\vec{y} \in V$  to the unique  $\vec{z}$  generated by 2.2.2 for  $g_{\vec{y}}$ . Then, by definition of  $\vec{z}$  we have:

$$\langle T\vec{x}, \vec{y} \rangle = g_{\vec{y}}(\vec{x}) = \langle \vec{x}, \vec{z} \rangle = \langle \vec{x}, T^*\vec{y} \rangle$$

By uniqueness of  $\vec{z}$ , this mapping  $T^*$  is unique. Thus it remains only to show that  $T^*$  is linear. For all  $\vec{x}, \vec{y}, \vec{z} \in V$  and  $c \in \mathbb{F}$ :

$$\langle \vec{x}, T^*(c\vec{y}) \rangle = \langle T\vec{x}, c\vec{y} \rangle \quad (2.20)$$

$$= \bar{c} \langle T\vec{x}, \vec{y} \rangle \quad (2.21)$$

$$= \bar{c} \langle \vec{x}, T^*\vec{y} \rangle \quad (2.22)$$

$$= \langle \vec{x}, cT^*\vec{y} \rangle \quad (2.23)$$

By the same orthogonal complement dimension argument as in 2.2.2, since this holds for all  $\vec{x}$ ,  $T^*(c\vec{y}) = cT^*\vec{y}$  as required. Similarly:

$$\langle \vec{x}, T^*(\vec{y} + \vec{z}) \rangle = \langle T\vec{x}, \vec{y} + \vec{z} \rangle \quad (2.24)$$

$$= \langle T\vec{x}, \vec{y} \rangle + \langle T\vec{x}, \vec{z} \rangle \quad (2.25)$$

$$= \langle \vec{x}, T^*\vec{y} \rangle + \langle \vec{x}, T^*\vec{z} \rangle \quad (2.26)$$

$$= \langle \vec{x}, T^*\vec{y} + T^*\vec{z} \rangle \quad (2.27)$$

Again, by the argument used in 2.2.2, since this holds for all  $\vec{x}$ , we have  $T^*(\vec{y} + \vec{z}) = T^*\vec{y} + T^*\vec{z}$  as required. Thus  $T^*$  is unique and linear as required.  $\square$

**Theorem 2.2.4** (Equivalence of conjugate transpose and adjoint). *If  $B$  is an orthonormal basis of  $V$ , then:*

$$[T]^*_B = [T^*]_B$$

*Proof.* Let  $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ ,  $[T]_B = (a_{ij})$  and  $[T^*]_B = (b_{ji})$ . Then for any  $i, j$ :

$$b_{ij} = \langle T^*\vec{v}_j, \vec{v}_i \rangle = \langle \vec{v}_j, T\vec{v}_i \rangle = \overline{\langle T\vec{v}_i, \vec{v}_j \rangle} = \overline{a_{ji}}$$

$\square$

## 2.3 Least Squares (example)

Say  $\{(x_1, y_1), \dots, (x_m, y_m)\}$  is a set of points in  $\mathbb{R}^2$  and we want to find the line that best fits the data. More precisely, we want to find  $a, b \in \mathbb{R}$  such that the line  $y = ax + b$  minimizes the quantity:

$$E = \sum_{i=1}^m |y_i - (ax_i + b)|^2$$

Define the following:

$$A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} \quad \vec{x} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Then we can rewrite the error  $E$  as:

$$E = \|A\vec{x} - \vec{y}\|^2$$

This is minimized when  $A\vec{x} = \text{proj}_{\text{im } A}(\vec{y})$ , so we just need to find  $\vec{x}$  given  $A\vec{x}$ .

**Remark** (Author's Note). In the following section we will take the adjoint of  $A$  even though  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^n$  and we only defined the adjoint for endofunctions. You can justify this to yourself by imagining the extensions of  $A$  and  $\vec{x}$  given by:

$$A' = \begin{bmatrix} x_1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_m & 1 & 0 & \dots & 0 \end{bmatrix} \quad \vec{x} = \begin{bmatrix} a \\ b \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

And convincing yourself that we have not fundamentally changed the problem. In general though, there is something called the Hermitian adjoint which properly describes this property for  $A : H_1 \rightarrow H_2$  where  $H_1$  and  $H_2$  are Hilbert spaces that have an appropriate inner product. I don't know anything about Hilbert spaces, this just happens to be a thing that exists.

We know that  $\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^*\vec{y} \rangle$ . Thus, if  $A^*A\vec{x} = \vec{0}$  we have:

$$A^*A\vec{x} = \vec{0} \implies \langle A^*A\vec{x}, \vec{x} \rangle = 0 \quad (2.28)$$

$$\implies \langle A\vec{x}, A\vec{x} \rangle = 0 \quad (2.29)$$

$$\implies A\vec{x} = \vec{0} \quad (2.30)$$

This tells us that if  $\ker A = (\vec{0})$ , then  $\ker(A^*A) = (\vec{0})$  meaning  $A^*A$  is invertible. In any practical case  $\ker A = (\vec{0})$  since, otherwise, that would mean all of our  $x_i$ s are equal, so our line doesn't represent anything interesting. Thus, if  $\vec{x}$  is our best fit line, then:

$$A\vec{x} - \vec{y} \in (\text{im } A)^\perp \quad (2.31)$$

$$\implies \langle A\vec{z}, A\vec{x} - \vec{y} \rangle = 0 \text{ for all } \vec{z} \in \mathbb{R}^2 \quad (2.32)$$

$$\implies \langle \vec{z}, A^*(A\vec{x} - \vec{y}) \rangle = 0 \quad (2.33)$$

$$\implies A^*(A\vec{x} - \vec{y}) = 0 \quad (2.34)$$

$$\implies A^*A\vec{x} = A^*\vec{y} \quad (2.35)$$

$$\implies \vec{x} = (A^*A)^{-1} A^*\vec{y} \quad (2.36)$$





# Week 3

## September 22 - September 27

### 3.1 Normal Operators

**Definition 3.1.1.** Let  $T : V \rightarrow V$  be a linear transformation on an inner product space  $V$ . We say  $T$  is *normal* if:

$$T^*T = TT^*$$

**Remark.** If there exists an orthonormal basis  $B$  such that  $[T]_B$  is diagonal, then  $[T^*]_B = [T]_B^*$  is also diagonal thus:

$$[T]_B^*[T]_B = [T]_B[T]^*B \quad (3.1)$$

$$T^*T = TT^* \quad (3.2)$$

So  $T$  is *normal*.

**Definition 3.1.2.** Let  $T : V \rightarrow V$  be a linear transformation on a vector space  $V$ , and let  $W$  be a subspace of  $V$ . We say  $W$  is  $T$ -invariant if, for all  $\vec{w} \in W$ ,  $T\vec{w} \in W$ .

**Lemma 3.1.3** (Schur). *Let  $T : V \rightarrow V$  be a linear transformation on an inner product space  $V$ . If the characteristic polynomial of  $T$  splits completely, then there is an orthonormal basis  $B$  of  $V$  such that  $[T]_B$  is upper triangular.*

*Proof.* We induce on  $\dim V$ . The case  $\dim V = 1$  is trivial since all  $1 \times 1$  matrices are upper triangular. So we assume the lemma holds for all inner product spaces  $W$  with  $\dim W < \dim V$ . Since the characteristic polynomial splits completely, there is some eigenvector  $\vec{v} \in V$  and corresponding eigenvalue  $\lambda$  satisfying:

$$T\vec{v} = \lambda\vec{v}$$

Thus, for any  $\vec{x} \in V$ :

$$0 = \langle (T - \lambda I) \vec{v}, \vec{x} \rangle \quad (3.3)$$

$$= \langle \vec{v}, (T^* - \bar{\lambda}I) \vec{x} \rangle \quad (3.4)$$

Which means that  $\vec{v} \in (\text{im}(T^* - \bar{\lambda}I))^\perp$ . Thus  $(T^* - \bar{\lambda}I)$  is not surjective, so by rank-nullity theorem, there is some nonzero  $\vec{z} \in \ker(T^* - \bar{\lambda}I)$ , giving:

$$(T^* - \bar{\lambda}I) \vec{z} = 0 \quad (3.5)$$

$$T^* \vec{z} = \bar{\lambda} \vec{z} \quad (3.6)$$

Without loss of generality, assume that  $\|\vec{z}\| = 1$ , since the equality holds under scalar multiplication of  $\vec{z}$ . Let  $W = \text{span}\{\vec{z}\}$ , then  $W$  is  $T^*$ -invariant. Then, for all  $\vec{y} \in W^\perp$ :

$$\langle T\vec{y}, c\vec{z} \rangle = \bar{c} \langle \vec{y}, T^* \vec{z} \rangle \quad (3.7)$$

$$= \bar{c} \lambda \langle \vec{y}, \vec{z} \rangle \quad (3.8)$$

$$= 0 \text{ by choice of } \vec{y} \quad (3.9)$$

Thus  $W^\perp$  is  $T$ -invariant. This means  $T|_{W^\perp} : W^\perp \rightarrow W^\perp$  is a linear transformation (whose characteristic polynomial splits completely, proof omitted in this class but this follows from the fact that  $T$  splits), and  $\dim W^\perp = \dim V - 1$ . Thus, by our inductive hypothesis there exists an orthonormal basis  $\beta = \{\vec{v}_1, \dots, \vec{v}_{n-1}\}$  of  $W^\perp$  such that  $[T|_{W^\perp}]_\beta$  is upper triangular. Thus:

$$[T^*]_B = \begin{bmatrix} [T^*|_{W^\perp}]_\beta & 0 \\ * & \bar{\lambda} \end{bmatrix} \quad (3.10)$$

$$[T]_B = \begin{bmatrix} [T|_{W^\perp}]_\beta & * \\ 0 & \lambda \end{bmatrix} \quad (3.11)$$

which is upper triangular. □

**Theorem 3.1.4** (Orthonormal Diagonalizability of Complex Linear Transformations). *If  $T : V \rightarrow V$  is a normal linear transformation on a complex inner product space  $V$ , then there exists an orthonormal basis  $B$  such that  $[T]_B$  is diagonal.*

*Proof.* Since all polynomials split over  $\mathbb{C}$ , by 3.1.3, there is an orthonormal basis:

$$B = \{\vec{v}_1, \dots, \vec{v}_n\}$$

such that  $[T_B]$  is upper triangular. We will show that  $[T]_B$  is also diagonal. Let  $[T]_B = (a_{ij})$ , we will show that  $a_{ij} = 0$  if  $i \neq j$  by induction on  $j$ . If  $j = 1$ , this is immediate from upper triangularity, so if the claim holds for all  $j' < j$ . If  $i < j$  then:

$$0 = \|T\vec{v}_i - \lambda\vec{v}_i\|^2 \text{ for } \lambda = a_{ii} \quad (3.12)$$

$$= \langle T\vec{v}_i - \lambda\vec{v}_i, T\vec{v}_i - \lambda\vec{v}_i \rangle \quad (3.13)$$

$$= \langle (T - \lambda I)\vec{v}_i, (T - \lambda I)\vec{v}_i \rangle \quad (3.14)$$

$$= \langle \vec{v}_i, (T - \lambda I)^*(T - \lambda I)\vec{v}_i \rangle \quad (3.15)$$

$$= \langle \vec{v}_i, (T - \lambda I)(T - \lambda I)^*\vec{v}_i \rangle \quad (3.16)$$

$$= \langle (T^* - \bar{\lambda}I)\vec{v}_i, (T^* - \bar{\lambda}I)\vec{v}_i \rangle \quad (3.17)$$

$$= \|T^*\vec{v}_i - \bar{\lambda}\vec{v}_i\| \quad (3.18)$$

Thus  $T^*\vec{v}_i = \bar{\lambda}\vec{v}_i$ . Then consider:

$$T\vec{v}_j = a_{1j}\vec{v}_1 + \cdots + a_{jj}\vec{v}_j$$

By orthonormality of our basis, it follows that:

$$a_{ij} = \langle T\vec{v}_j, \vec{v}_i \rangle \quad (3.19)$$

$$= \langle \vec{v}_j, T^*\vec{v}_i \rangle \quad (3.20)$$

$$= \langle \vec{v}_j, \bar{\lambda}\vec{v}_i \rangle \quad (3.21)$$

$$= 0 \quad (3.22)$$

As required, each entry  $a_{ij}$  with  $i < j$  is 0, and entries with  $i > j$  follow from upper triangularity.  $\square$

**Corollary 3.1.4.1.** *If  $T : V \rightarrow V$  is a linear transformation on a complex inner product space  $V$ , then there exists an orthonormal basis  $B$  such that  $[T]_B$  is diagonal if and only if  $T$  is normal.*

**Remark.** 3.1.4.1 does not apply to real inner product spaces. Consider  $V : \mathbb{R}^2$  and  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  given by the rotation matrix:

$$\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

Then  $T^*$  describes the opposite rotation, thus  $T^*T = TT^* = I$  so  $T$  is *normal*, however if  $\theta \notin \pi\mathbb{Z}$ ,  $T$  has no real eigenvectors and is thus not diagonalizable.

## 3.2 Self-Adjoint Operators

**Definition 3.2.1.** A linear transformation  $T$  is *self-adjoint* (or Hermitian) if  $T = T^*$ .

**Remark.** If  $T : V \rightarrow V$  is a linear transformation on a real inner product space  $V$ , and there exists an orthonormal basis  $B$  for which  $[T]_B$  is diagonal, then:

$$[T]_B = [T]_B^*$$

So  $T$  is *self-adjoint*

**Remark.** If  $T = T^*$ , then  $T^*T = TT^*$  so  $T$  is normal.

**Theorem 3.2.2** (Orthonormal Diagonalizability of Real Linear Transformations). *If  $T : V \rightarrow V$  is a linear transformation on a real inner product space  $V$ , then  $T$  is self-adjoint if and only if there is an orthonormal basis  $B$  such that  $[T]_B$  is diagonal.*

*Proof.* Note that the characteristic polynomial of  $T$  must split over  $\mathbb{C}$ , so consider any eigenvector  $\vec{x}$  and eigenvalue  $\lambda \in \mathbb{C}$  such that  $T\vec{x} = \lambda\vec{x}$ , then:

$$(T - \lambda I)\vec{x} = \vec{0} \implies (T^* - \bar{\lambda}I)\vec{x} = 0 \text{ (see proof of 3.1.4)} \quad (3.23)$$

$$\implies T^*\vec{x} = \bar{\lambda}\vec{x} \quad (3.24)$$

So if  $T$  is *self-adjoint*:

$$\bar{\lambda}\vec{x} = T^*\vec{x} = T\vec{x} = \lambda\vec{x} \quad (3.25)$$

$$\bar{\lambda} = \lambda \quad (3.26)$$

Thus  $\lambda \in \mathbb{R}$ , so all eigenvalues of  $T$  are real. Thus the characteristic polynomial of  $T$  splits completely over  $\mathbb{R}$ , so invoking 3.1.3, there must exist an orthonormal basis  $B$  such that  $[T]_B$  is upper triangular. However  $[T]_B^* = [T^*]_B$  which must be lower triangular, so  $[T]_B$  is both upper and lower triangular, meaning  $[T]_B$  is diagonal.  $\square$

**Corollary 3.2.2.1** (Orthonormal Diagonalizability of Symmetric Real Matrices). *A real matrix is orthogonally diagonalizable if and only if it's symmetric.*

*Proof.* A real matrix that is *self-adjoint* is just a symmetric matrix, so this follows immediately from 3.2.2.  $\square$

### 3.3 Isometries

**Definition 3.3.1.** A linear transformation  $T : V \rightarrow W$  from an inner product space  $V$  to an inner product space  $W$  is an *isometry* if  $\langle T\vec{x}, T\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$  for all  $\vec{x}, \vec{y} \in V$ .

**Definition 3.3.2.** An *isometry*  $T$  is *unitary* if  $T$  is surjective.

**Definition 3.3.3.** Let  $V, W$  be inner product spaces. If there exists a *unitary isometry*  $T : V \rightarrow W$ , we say  $V$  and  $W$  are *isometric*.

**Remark.** Every *isometry*  $T$  is injective because:

$$T\vec{x} = \vec{0} \implies \langle \vec{x}, \vec{x} \rangle = \langle T\vec{x}, T\vec{x} \rangle = 0 \quad (3.27)$$

$$\implies \vec{x} = 0 \quad (3.28)$$

Thus  $\ker T = \{0\}$ .

**Remark** (Author's Note). Again, in this section, we will use the adjoint of  $T$  even if  $T$  is not an endomorphism. In finite dimensional vector spaces, this exists, and the conjugate transpose of the matrix representation still works, you'll just have to convince yourself.

**Remark.** For every *isometry*  $T : V \rightarrow W$ ,  $T^*T = I$  since, for all  $\vec{x}, \vec{y} \in V$ :

$$\langle T\vec{x}, T\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle \quad (3.29)$$

$$\langle \vec{x}, T^*T\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle \text{ for any } \vec{x} \quad (3.30)$$

**Remark.** If  $T$  is *unitary*,  $T$  is invertible so  $TT^* = I = T^*T$ , so  $T$  is also normal.

**Lemma 3.3.4.** Let  $\mathcal{U} : V \rightarrow V$  be a self-adjoint linear transformation, and  $\langle \vec{x}, \mathcal{U}\vec{x} \rangle = 0$  for all  $\vec{x} \in V$ , then  $\mathcal{U} = 0$ .

*Proof.* Suppose  $\vec{x}$  is an eigenvector of  $\mathcal{U}$  and  $\lambda$  be its corresponding eigenvalue, then:

$$0 = \langle \vec{x}, \mathcal{U}\vec{x} \rangle \quad (3.31)$$

$$= \langle \vec{x}, \lambda\vec{x} \rangle \quad (3.32)$$

$$= \bar{\lambda} \langle \vec{x}, \vec{x} \rangle \quad (3.33)$$

But  $\vec{x} \neq \vec{0}$  by choice of  $\vec{x}$  being an eigenvector, so  $\lambda = 0$ . Since all eigenvalues of  $\mathcal{U}$  are 0 and  $\mathcal{U}$  is diagonalizable (since it is self-adjoint),  $\mathcal{U} = 0$ .  $\square$

**Theorem 3.3.5.** Let  $T : V \rightarrow V$  be a surjective linear transformation on a finite dimensional inner product space  $V$ , then the following are equivalent:

- i.  $TT^* = T^*T = I$
- ii.  $\langle T\vec{x}, T\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$  for all  $\vec{x}, \vec{y} \in V$
- iii. If  $B$  is an orthonormal basis, then so is  $T(B)$
- iv. There exists an orthonormal basis  $B$  such that  $T(B)$  is also orthonormal
- v.  $\|T\vec{x}\| = \|\vec{x}\|$  for all  $\vec{x} \in V$ .

*Proof.* We will prove a ring of implications:

i.  $\implies$  ii. For all  $\vec{x}, \vec{y} \in V$ :

$$\langle T\vec{x}, T\vec{y} \rangle = \langle \vec{x}, T^*T\vec{y} \rangle \quad (3.34)$$

$$= \langle \vec{x}, I\vec{y} \rangle \quad (3.35)$$

$$= \langle \vec{x}, \vec{y} \rangle \quad (3.36)$$

ii.  $\implies$  iii. Let  $B$  be a basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$ , then for any  $\vec{v}_i, \vec{v}_j \in B$ :

$$\langle T\vec{v}_i, T\vec{v}_j \rangle = \langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Thus  $T(B)$  is orthonormal. Recall from 1.3.6 that this is sufficient to show  $T(B)$  is linearly independent and thus a basis.

iii.  $\implies$  iv. Immediate from the fact that  $V$  is finite dimensional so an orthonormal basis exists.

iv.  $\implies$  v. For any  $\vec{x} \in V$  write:

$$\vec{x} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$$

where  $\{\vec{v}_1, \dots, \vec{v}_n\}$  are a subset of the orthonormal basis  $B$  provided by the assumption. Then:

$$\|T\vec{x}\|^2 = \|T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n)\|^2 \quad (3.37)$$

$$= \|a_1T\vec{v}_1 + \cdots + a_nT\vec{v}_n\|^2 \quad (3.38)$$

$$= |a_1|^2 + \cdots + |a_n|^2 \quad (3.39)$$

$$= \|\vec{x}\|^2 \quad (3.40)$$

By non-negativity of the norm, v. holds.

v.  $\implies$  i. From our assumption, for all  $\vec{x}$ :

$$\|\vec{x}\| = \|T\vec{x}\| \quad (3.41)$$

$$= \langle T\vec{x}, T\vec{x} \rangle \quad (3.42)$$

$$= \langle \vec{x}, T^*T\vec{x} \rangle \quad (3.43)$$

We have  $\langle \vec{x}, (T^*T - I)\vec{x} \rangle = 0$  for all  $\vec{x}$ . Note that  $(T^*T - I)$  is self-adjoint because  $(T^*T - I)^* = T^*T - I$ . Thus by 3.3.4,  $T^*T - I = 0$  so  $T^*T = I$ . Thus  $T^*$  is a left inverse of  $T$ , so since  $T$  is an endomorphism,  $T^*T = I = TT^*$ .

□

**Corollary 3.3.5.1.** *Let  $V, W$  be isometric finite dimensional inner product spaces, then  $\dim V = \dim W$ .*

**Corollary 3.3.5.2.** *If  $\dim V = \dim W$  for finite dimensional inner product spaces  $V, W$ , then  $V, W$  are isometric.*

*Proof.* Since  $V$  and  $W$  are finite dimensional, they have orthonormal bases  $\{\vec{v}_1, \dots, \vec{v}_n\}, \{\vec{w}_1, \dots, \vec{w}_n\}$ . Then we can define a linear transformation  $T : V \rightarrow W$  given by  $T(\vec{v}_i) = \vec{w}_i$  for all  $i$ . By 3.3.5,  $T$  is an *isometry*. □

**Corollary 3.3.5.3.** *Any  $n$ -dimensional inner product space is isometric to  $\mathbb{R}^n$  or  $\mathbb{C}^n$  with the standard inner product.*

**Corollary 3.3.5.4.** *If  $T : V \rightarrow W$  is unitary, then its eigenvalues all have absolute value 1.*

*Proof.* For all  $\vec{x} \in T$ :

$$\|\vec{x}\| = \|T\vec{x}\| = \|\lambda\vec{x}\| = |\lambda| \|\vec{x}\|$$

Thus for any eigenvalue  $\lambda$ ,  $|\lambda| = 1$ . □