

# MATH 235

## Personal Notes

---

Marcus Chan

UW '25

Taught by Nick Rollie



# Class 8:

## Examples of Matrix Representations, Introduction to Inner Product Spaces

### INNER PRODUCT & INNER PRODUCT SPACES: $\langle v, w \rangle$ (D8.1)

$\exists_1$  Let  $V$  be a vector space over  $\mathbb{F}$ . Then, we say the function  $\langle \cdot, \cdot \rangle: V^2 \rightarrow \mathbb{F}$

is an "inner product" if

- ①  $\langle v, v \rangle \in \mathbb{R}$  &  $\langle v, v \rangle \geq 0 \quad \forall v \in V;$  } Positivity
- ②  $\langle v, v \rangle = 0 \iff v = 0 \quad \forall v \in V;$
- ③  $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle \quad \forall v_1, v_2, w \in V;$  } Linearity
- ④  $\langle cv, w \rangle = c\langle v, w \rangle \quad \forall c \in \mathbb{F}, v, w \in V;$  and } Conjugate Symmetry
- ⑤  $\langle w, v \rangle = \overline{\langle v, w \rangle} \quad \forall v, w \in V.$

$\exists_2$  In this case, we call  $\langle v, w \rangle$  the "inner product" of  $v$  &  $w$ .

$\exists_3$  We refer to  $V$  together with  $\langle \cdot, \cdot \rangle$  as an "inner product space".

### LENGTH [OF A VECTOR]: $\|v\|$ (D8.2)

$\exists_1$  Let  $V$  be an inner product space, and let  $v \in V$ .

Then, the "length" of  $v$ , denoted by " $\|v\|$ ", is defined to be equal to

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

We can do this because  $\langle v, v \rangle \in \mathbb{R} \quad \forall v \in V$ .

### ORTHOGONAL [VECTORS] (D8.3)

$\exists_1$  Let  $V$  be an IPS.

Then, we say  $v, w \in V$  are "orthogonal"

if  $\langle v, w \rangle = 0$ .

### ORTHOGONAL [SETS] (D8.3)

$\exists_1$  Let  $S \subseteq V$ , where  $V$  is an IPS.

Then, we say  $S$  is "orthogonal" if

$$\langle v, w \rangle = 0 \quad \forall v, w \in S.$$

### EXAMPLES OF IPS: PART 1

$\exists_1$  The vector space  $V = \mathbb{F}^n$  with inner product

$$\langle \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \rangle = v_1 \bar{w}_1 + \dots + v_n \bar{w}_n$$

is an inner product space. (E8.3)

$\exists_2$  The vector space  $V = P_n(\mathbb{F})$  with inner product

$$\langle p, q \rangle = p(0) \bar{q}(0) + \dots + p(n) \bar{q}(n)$$

is an inner product space. (E8.4)

### CONJUGATE MATRIX: $\bar{A}$ (D8.4)

$\exists_1$  Let  $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$ .

Then, the "conjugate" of  $A$ , denoted by " $\bar{A}$ ", is equal to

$$\bar{A} = (\bar{a}_{ij}) \in M_{m,n}(\mathbb{F}).$$

### CONJUGATE TRANSPOSE MATRIX: $A^* = \bar{A}^T$ (D8.4)

$\exists_1$  Then, the "conjugate transpose" of  $A$  is defined to be the matrix

$$A^* = \bar{A}^T \in M_{n,m}(\mathbb{F}).$$

### STANDARD INNER PRODUCT ON $M_{m,n}(\mathbb{F})$ :

$$\langle A, B \rangle = \text{tr}(AB^*) \quad (\text{E8.5})$$

$\exists_1$  Let  $V = M_{m,n}(\mathbb{F})$ .

Then, the "standard inner product" on  $V$  is given by

$$\langle A, B \rangle = \text{tr}(AB^*),$$

where  $\text{tr}(A) = \sum_{i=1}^m a_{ii}$  for  $A \in M_{m,n}(\mathbb{F})$ .

$\exists_2$  We can prove this is indeed an IPS.

Proof. Linearity is trivial (arises from fact that trace & matrix multiplication is linear).

Note that for  $A = (a_{ij})$  &  $B = (b_{ij})$ , then  $B^* = (\bar{b}_{ji})$ .

Then

$$(AB^*)_{ii} = \sum_{k=1}^n a_{ik} (B^*)_{ki} = \sum_{k=1}^n a_{ik} \bar{b}_{ik}.$$

Hence

$$\text{tr}(AB^*) = \sum_{i=1}^m (AB^*)_{ii} = \sum_{i=1}^m \sum_{k=1}^n a_{ik} \bar{b}_{ik}.$$

In particular, this is "similar" to if we wrote the entries of  $A$  &  $B$  in  $\mathbb{F}^m$ , and took the standard inner product of these vectors.

It trivially follows that this gives an inner product on  $V$ .  $\square$

### INNER PRODUCT ON $C[a,b]$ : $\langle f(x), g(x) \rangle = \int_a^b f(x)g(x) dx$ (E8.6)

$\exists_1$  We can show  $C[a,b]$  with the function

$$\langle f(x), g(x) \rangle = \int_a^b f(x)g(x) dx$$

is an inner product space.

Proof. Linearity & conjugate symmetry (ie "normal" symmetry, since field =  $\mathbb{R}$ ) follow pretty easily.

For positivity, note that

$$\langle f(x), f(x) \rangle = \int_a^b f(x)^2 dx \geq \int_a^b 0 dx = 0.$$

If  $f \neq 0$ , then trivially  $\int_a^b f(x)^2 dx > 0$  by EVT, completing the proof.  $\square$

### $T_w: V \rightarrow \mathbb{F}$ BY $T_w(v) = \langle v, w \rangle$ IS LINEAR (T8.2(1))

$\exists_1$  Let  $w \in V$ , and let  $T_w: V \rightarrow \mathbb{F}$  by  $T_w(v) = \langle v, w \rangle \quad \forall v \in V$ .

Then necessarily  $T_w$  is linear.

### SET OF VECTORS ORTHOGONAL TO $w$ IS A SUBSPACE OF $V$ (T8.2(2))

$\exists_1$  Let  $w \in V$ .

Then the set of vectors orthogonal to  $w$  is a subspace of  $V$ .

Proof. This follows from the fact that the set =  $\ker(T_w)$ .  $\square$

$\|v\| \geq 0 \quad \forall v \in V, \quad v=0 \Leftrightarrow \|v\|=0$  (T8-3(1))

Let  $V$  be an IPS.  
Then necessarily  $\|v\| \geq 0 \quad \forall v \in V$ , and  $\|v\|=0$  if and only if  $v=0$ .

Proof. This arises from properties of inner products.

$\|cv\| = |c| \cdot \|v\|$  (T8-3(2))

Let  $V$  be an IPS, and let  $c \in F$ .  
Then necessarily  $\|cv\| = |c| \cdot \|v\|$ .

Proof. This also arises from properties of inner products.

$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| ; \quad |\langle v, w \rangle| = \|v\| \cdot \|w\| \Leftrightarrow v \& w$

ARE LINEARLY DEPENDENT

(THE CAUCHY-SCHWARTZ INEQUALITY) (T8-3(3))

Let  $v, w \in V$ .  
Then necessarily  $|\langle v, w \rangle| \leq \|v\| \cdot \|w\|$ , and equality holds iff  $v$  and  $w$  are linearly dependent.

Proof. We show  $|\langle v, w \rangle|^2 \leq \|v\|^2 \cdot \|w\|^2$ .  
First, if  $w=0$ , the result is trivial. Otherwise, assume

$w \neq 0$ , and let  $c = \frac{\langle v, w \rangle}{\|w\|^2}$ .

By T8-3(1):

$$\begin{aligned} 0 &\leq \|v - cw\|^2 \\ &= \langle v - cw, v - cw \rangle \\ &= \langle v, v - cw \rangle - c \langle w, v - cw \rangle \\ &= \langle v, v \rangle - \overline{c} \langle v, w \rangle - c \langle w, v \rangle + c\overline{c} \langle w, w \rangle \\ &= \|v\|^2 - \overline{c} \langle v, w \rangle - c \langle w, v \rangle + |c|^2 \|w\|^2 \\ &= \|v\|^2 - \frac{\langle v, w \rangle}{\|w\|^2} \langle v, w \rangle - \frac{\langle v, w \rangle}{\|w\|^2} \langle w, v \rangle + \frac{|\langle v, w \rangle|^2}{\|w\|^4} \|w\|^2 \\ &= \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2} - \frac{|\langle v, w \rangle|^2}{\|w\|^2} + \frac{|\langle v, w \rangle|^2}{\|w\|^2} \\ \therefore 0 &\leq \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2}, \end{aligned}$$

and so  $|\langle v, w \rangle|^2 \leq \|v\|^2 \|w\|^2$ , as needed.

Then, note that

$$\begin{aligned} \|v - cw\| > 0 &\Leftrightarrow v - cw \neq 0 \text{ for some } c \in F \\ &\Leftrightarrow v \neq cw \\ &\Leftrightarrow v \& w \text{ are lin ind,} \end{aligned}$$

and so  $\|v - cw\| = 0 \Leftrightarrow v \& w \text{ are lin dep.}$   $\square$

$\|v+w\| \leq \|v\| + \|w\| \quad \forall v, w \in V$  (TRIANGLE INEQUALITY) (T8-3(4))

Let  $v, w \in V$ .  
Then necessarily  $\|v+w\| \leq \|v\| + \|w\|$ .

Proof. Note that

$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w \rangle \\ &= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \|v\|^2 + \langle v, w \rangle + \overline{\langle v, w \rangle} + \|w\|^2 \\ &= \|v\|^2 + \|w\|^2 + 2\operatorname{Re} \langle v, w \rangle \\ &\leq \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle \\ &\leq \|v\|^2 + \|w\|^2 + 2\|v\| \|w\| \quad (\text{by CSE}) \\ &\leq \|v\|^2 + 2\|v\| \|w\| + \|w\|^2, \end{aligned}$$

and the proof follows.  $\square$

# Class 9: Orthogonal and Orthonormal Bases; The Gram-Schmidt Procedure

## ORTHOGONAL & ORTHONORMAL BASIS (D9.1)

$\exists$ : Let  $V$  be an IPS, and let  $B \subseteq V$ .

Then, we say  $B$  is an "orthogonal basis"

for  $V$  if:

- ①  $B$  is a basis for  $V$ ; and
- ②  $B$  is an orthogonal set of vectors.

$\exists$ : We say  $B$  is an "orthonormal basis" for  $V$  if the above conditions are satisfied and  $\|v\|=1 \forall v \in B$ .

$S \subseteq V$  IS ORTHOGONAL & HAS NO ZERO VECTORS  $\Rightarrow$

$S$  IS LINEARLY INDEPENDENT (T9.1)

$\exists$ : Let  $V$  be an IPS, and let  $S \subseteq V$  be orthogonal and have no zero vectors.

Then necessarily  $S$  is linearly independent.

Proof: Let  $c_1, \dots, c_n \in F$ ,  $v_1, \dots, v_n \in S$  s.t.

$$c_1 v_1 + \dots + c_n v_n = 0.$$

Taking the inner product of each side with  $v_i$ , we see that

$$\begin{aligned} 0 &= \langle 0, v_i \rangle \\ &= \langle c_1 v_1 + \dots + c_n v_n, v_i \rangle \\ &= c_1 \langle v_1, v_i \rangle + \dots + c_n \langle v_n, v_i \rangle \\ &= c_1 \|v_1\|^2 + c_2(0) + \dots + c_n(0) \quad (\text{since } S \text{ is orthogonal}) \\ \therefore 0 &= c_1 \|v_1\|^2, \end{aligned}$$

and so since  $v_1 \neq 0$  it follows that  $c_1 = 0$ .

Repeating this argument by taking inner product with  $v_2, \dots, v_n$  gives us that  $c_1 = \dots = c_n = 0$ , showing that the vectors are linearly independent.  $\square$

$V$  HAS ORTHOGONAL ORDERED BASIS  $B = \{v_1, \dots, v_n\} \Rightarrow$

$w = \sum_{i=1}^n \frac{\langle w, v_i \rangle}{\|v_i\|^2} v_i \quad (T9.2)$

$\exists$ : Let  $V$  be a finite-dimensional IPS, and let  $V$  have an orthogonal ordered basis  $B = \{v_1, \dots, v_n\}$ .

Let  $w \in V$  be arbitrary.

Then necessarily

$$w = \frac{\langle w, v_1 \rangle}{\|v_1\|^2} v_1 + \dots + \frac{\langle w, v_n \rangle}{\|v_n\|^2} v_n.$$

In other words,

$$[w]_B = \begin{pmatrix} \frac{\langle w, v_1 \rangle}{\|v_1\|^2} \\ \vdots \\ \frac{\langle w, v_n \rangle}{\|v_n\|^2} \end{pmatrix}$$

Proof: Since  $B$  is a basis,  $\exists c_1, \dots, c_n \in F$  s.t.

$$w = c_1 v_1 + \dots + c_n v_n.$$

Taking IP of both sides w/  $v_1$  yields that

$$\begin{aligned} \langle w, v_1 \rangle &= c_1 \langle v_1, v_1 \rangle + \dots + c_n \langle v_n, v_1 \rangle \\ &= c_1 \|v_1\|^2 + c_2(0) + \dots + c_n(0) \\ &= c_1 \|v_1\|^2, \end{aligned}$$

and doing similarly for  $v_2, \dots, v_n$  yields that

$$\langle w, v_i \rangle = c_i \|v_i\|^2 \quad \forall i \in \{1, \dots, n\}.$$

Thus  $c_i = \frac{\langle w, v_i \rangle}{\|v_i\|^2}$ , which suffices to prove the claim.  $\square$

$V$  HAS ORTHONORMAL ORDERED BASIS  $B = \{v_1, \dots, v_n\} \Rightarrow$

$w = \sum_{i=1}^n \langle w, v_i \rangle v_i \quad (C9.1)$

$\exists$ : Let  $V$  be a finite-dimensional IPS, and let  $V$  have an orthonormal ordered basis  $B = \{v_1, \dots, v_n\}$ .

Let  $w \in V$  be arbitrary.

Then necessarily

$$w = \langle w, v_1 \rangle v_1 + \dots + \langle w, v_n \rangle v_n.$$

In other words,

$$[w]_B = \begin{pmatrix} \langle w, v_1 \rangle \\ \vdots \\ \langle w, v_n \rangle \end{pmatrix}$$

Proof: This follows almost immediately from T9.2.

$S = \{w_1, \dots, w_n\}$  IS LINEARLY INDEPENDENT;

$v_i = v_i - \sum_{j=1}^{i-1} \frac{\langle w_j, v_i \rangle}{\|v_j\|^2} v_j \Rightarrow \{v_1, \dots, v_n\}$  IS ORTHOGONAL &  $\{v_1, \dots, v_n\}$  IS AN

ORTHOGONAL BASIS FOR  $\text{Span}\{w_1, \dots, w_n\}$

(THE GRAM-SCHMIDT PROCEDURE) (L9.1)

💡 Let  $V$  be an IPS, and let  $S = \{w_1, \dots, w_n\} \subseteq V$  be linearly independent.

Define  $\{v_1, \dots, v_n\}$  recursively by  $v_1 = w_1$ , and

$$v_i = w_i - \frac{\langle w_i, v_1 \rangle}{\|v_1\|^2} v_1 - \dots - \frac{\langle w_i, v_{i-1} \rangle}{\|v_{i-1}\|^2} v_{i-1} \quad \forall i \in \mathbb{N}.$$

Then

①  $\{v_1, \dots, v_n\}$  is orthogonal and

②  $\{v_1, \dots, v_n\}$  is an orthogonal basis for  $\text{Span}\{w_1, \dots, w_n\}$  for any taken.

Proof. We prove this by induction.

( $n=1$ ) Since  $w_1 \neq 0$  (as  $S$  is lin ind), hence  $\{v_1\}$  is orthogonal, and since  $v_1 = w_1$ , so  $\text{Span}\{v_1\} = \text{Span}\{w_1\}$ , so the conclusions trivially follow.

(Inductive) Suppose the claim is true for  $1 \leq k < n$ .

So  $\{v_1, \dots, v_k\}$  is an orthogonal basis for  $\text{Span}\{w_1, \dots, w_k\}$ .

We want to show similarly  $\{v_1, \dots, v_{k+1}\}$  is an orthogonal basis for  $\text{Span}\{w_1, \dots, w_{k+1}\}$ .

Since we know  $\{v_1, \dots, v_k\}$  is orthogonal, we just need to check  $v_{k+1}$  is orthogonal to each  $v_i$  to verify  $\{v_1, \dots, v_{k+1}\}$  is orthogonal.

Observe that

$$\begin{aligned} \langle v_{k+1}, v_i \rangle &= \left\langle w_{k+1} - \frac{\langle w_{k+1}, v_1 \rangle}{\|v_1\|^2} v_1 - \dots - \frac{\langle w_{k+1}, v_k \rangle}{\|v_k\|^2} v_k, v_i \right\rangle \\ &= \langle w_{k+1}, v_i \rangle - \frac{\langle w_{k+1}, v_1 \rangle}{\|v_1\|^2} \langle v_1, v_i \rangle \\ &\quad - \dots - \frac{\langle w_{k+1}, v_k \rangle}{\|v_k\|^2} \langle v_k, v_i \rangle \\ &\quad - \dots - \frac{\langle w_{k+1}, v_k \rangle}{\|v_k\|^2} \langle v_k, v_i \rangle \\ &= \langle w_{k+1}, v_i \rangle - 0 - \dots - \frac{\langle w_{k+1}, v_k \rangle}{\|v_k\|^2} \|v_k\|^2 - \dots - 0 \\ &= \langle w_{k+1}, v_i \rangle - \langle w_{k+1}, v_i \rangle \\ &= 0, \end{aligned}$$

Showing that  $v_{k+1}$  is orthogonal to each  $v_i$ , and so  $\{v_1, \dots, v_{k+1}\}$  is orthogonal.

Next, we show  $\text{Span}\{v_1, \dots, v_{k+1}\} = \text{Span}\{w_1, \dots, w_{k+1}\}$ .

By hypothesis,  $\text{Span}\{v_1, \dots, v_k\} = \text{Span}\{w_1, \dots, w_k\}$ , and since

$$v_{k+1} = w_{k+1} - \frac{\langle w_{k+1}, v_1 \rangle}{\|v_1\|^2} v_1 - \dots - \frac{\langle w_{k+1}, v_k \rangle}{\|v_k\|^2} v_k$$

shows that  $v_{k+1}$  is a lin comb of  $v_1, \dots, v_k, w_{k+1}$ . Since this is also trivially true for  $v_1, \dots, v_k$  as well, thus any lin comb of  $v_1, \dots, v_{k+1}$  is a lin comb of  $v_1, \dots, v_k, w_{k+1}$ , and so  $\text{Span}\{v_1, \dots, v_{k+1}\} \subseteq \text{Span}\{v_1, \dots, v_k, w_{k+1}\}$ .

Then,  $\text{Span}\{v_1, \dots, v_k\} = \text{Span}\{w_1, \dots, w_k\} \Rightarrow$

$\text{Span}\{v_1, \dots, v_k, w_{k+1}\} = \text{Span}\{w_1, \dots, w_k, w_{k+1}\}$ , and so

$\text{Span}\{v_1, \dots, v_{k+1}\} \subseteq \text{Span}\{w_1, \dots, w_{k+1}\}$ .

Conversely, for  $1 \leq i \leq k+1$ , since  $w_i$  is a lin comb of  $v_1, \dots, v_i$  hence any lin comb of  $w_1, \dots, w_{k+1}$  is also a lin comb of  $v_1, \dots, v_{k+1}$ .

so  $\text{Span}\{w_1, \dots, w_{k+1}\} \subseteq \text{Span}\{v_1, \dots, v_{k+1}\}$ , and so

$\text{Span}\{w_1, \dots, w_{k+1}\} = \text{Span}\{v_1, \dots, v_{k+1}\}$ .

Since  $\{w_1, \dots, w_{k+1}\}$  is lin ind, it follows  $\{v_1, \dots, v_{k+1}\}$  is also lin ind, and so  $\{v_1, \dots, v_{k+1}\}$  is an ortho basis for  $\text{Span}\{w_1, \dots, w_{k+1}\}$  completing the inductive step.

$\dim V < \infty \Rightarrow V$  HAS AN ORTHOGONAL BASIS (T9.3)

💡 Let  $V$  be a finite-dimensional IPS.

Then necessarily  $V$  has an orthogonal basis.

Proof. Since  $\dim V < \infty$ ,  $V$  has a finite basis, say  $\{w_1, \dots, w_n\}$ . Then, applying L9.1 to  $\{w_1, \dots, w_n\}$  yields an orthogonal set  $\{v_1, \dots, v_n\}$ , for which  $\{v_1, \dots, v_n\}$  is an orthogonal basis for  $\text{Span}\{w_1, \dots, w_n\} = V$ .  $\square$

$\dim V < \infty \Rightarrow V$  HAS AN ORTHONORMAL BASIS (C9.2)

💡 Let  $V$  be a finite-dimensional IPS.

Then necessarily  $V$  has an orthonormal basis.

Proof. This follows by taking the basis obtained in T9.3 and scaling each vector down by its respective norm.  $\square$

# Class 10:

# Direct Sums of Subspaces and Orthogonal Projections

SUMS OF SUBSPACES:  $W_1 + \dots + W_n$  (D10.1)

Let  $W_1, \dots, W_n \subseteq V$ .

Then, the "sum" of  $W_1, \dots, W_n$ , denoted as " $W_1 + \dots + W_n$ ", is the set

$$W_1 + \dots + W_n = \{w_1 + \dots + w_n : w_i \in W_i, i=1, \dots, n\}.$$

Note the following:

①  $W_1 + \dots + W_n$  is a subspace of  $V$ , and the smallest subspace containing  $W_1, \dots, W_n$ ; and

② If  $\text{Span}\{S_i\} = W_i$  for each  $i$ , then  $\text{Span}(\bigcup_{i=1}^k S_i) = W_1 + \dots + W_k$ .

DIRECT SUM OF SUBSPACES (D10.2)

Let  $V$  be a vector space, and let  $W_1, \dots, W_k \subseteq V$ .

Then, we say  $V$  is the "direct sum" of  $W_1, \dots, W_k$ ,

if

① there exist unique vectors  $w_i \in W_i$  with  $v = w_1 + \dots + w_k$  for each  $v \in V$ ;

②  $V = W_1 + \dots + W_k$  and  $W_i \cap (W_1 + \dots + W_{i-1} + W_{i+1} + \dots + W_k) = \{0\}$  for each  $i=1, \dots, k$ ;

③ If  $B_i$  is a basis for  $W_i$ , then  $\bigcup_{i=1}^k B_i$  is a basis for  $V$ .

Note the following conditions are equivalent.

In this case, we write that

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_k.$$

DISTANCE [BETWEEN VECTORS]:  $d(v, w)$  (D10.3)

Let  $V$  be an IPS, and let  $v, w \in V$ .

Then, the "distance" between  $v$  to  $w$ , denoted as " $d(v, w)$ ", is equal to

$$d(v, w) = \|v - w\|.$$

The distance function obeys the "usual" properties of distance (T10.3).

ORTHOGONAL PROJECTION MAP:  $\text{proj}_W(v)$  (D10.4)

Let  $V$  be an IPS, and let  $W \subseteq V$  be finite-dimensional, with orthogonal basis  $\{w_1, \dots, w_n\}$ .

Then, the "orthogonal projection map" of  $V$  onto  $W$ ,

denoted as " $\text{proj}_W: V \rightarrow W$ ", is defined by

$$\text{proj}_W(v) = \frac{\langle v, w_1 \rangle}{\|w_1\|^2} w_1 + \dots + \frac{\langle v, w_n \rangle}{\|w_n\|^2} w_n.$$

$\text{proj}_W(v)$  IS A LINEAR TRANSFORMATION

(T10.4(1))

Let  $V$  be an IPS, and let  $W \subseteq V$  be finite-dimensional, with orthogonal basis  $\{w_1, \dots, w_n\}$ .

Let  $\text{proj}_W: V \rightarrow W$  be the associated orthogonal projection mapping.

Then necessarily  $\text{proj}_W$  is a linear transformation.

Proof. See that

$$\begin{aligned} \text{proj}_W(c_1 v_1 + c_2 v_2) &= \frac{\langle c_1 v_1 + c_2 v_2, w_1 \rangle}{\|w_1\|^2} w_1 + \dots + \frac{\langle c_1 v_1 + c_2 v_2, w_n \rangle}{\|w_n\|^2} w_n \\ &= \frac{c_1 \langle v_1, w_1 \rangle + c_2 \langle v_2, w_1 \rangle}{\|w_1\|^2} w_1 + \dots + \frac{c_1 \langle v_1, w_n \rangle + c_2 \langle v_2, w_n \rangle}{\|w_n\|^2} w_n \\ &= c_1 \left[ \frac{\langle v_1, w_1 \rangle}{\|w_1\|^2} w_1 + \dots + \frac{\langle v_1, w_n \rangle}{\|w_n\|^2} w_n \right] + c_2 \left[ \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1 + \dots + \frac{\langle v_2, w_n \rangle}{\|w_n\|^2} w_n \right] \\ &= c_1 \text{proj}_W(v_1) + c_2 \text{proj}_W(v_2). \end{aligned}$$

$d(v, \text{proj}_W(w)) \leq d(v, w) \quad \forall w \in W$  (T10.4(2))

Let  $V$  be an IPS, and let  $W \subseteq V$  be finite-dimensional, with orthogonal basis  $\{w_1, \dots, w_n\}$ .

Let  $\text{proj}_W: V \rightarrow W$  be the associated orthogonal projection mapping.

Then necessarily for any  $v \in V$ , we have that

$$d(v, \text{proj}_W(v)) \leq d(v, w) \quad \forall w \in W.$$

Moreover, note that

$$d(v, w) \text{ is smallest } \Leftrightarrow w = \text{proj}_W(v). \quad (\text{T10.4(3)})$$

Proof. Let  $w \in W$ . See that

$$\begin{aligned} \langle v - \text{proj}_W(v), w \rangle &= \langle v - \frac{\langle v, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v, w_2 \rangle}{\|w_2\|^2} w_2 - \dots - \frac{\langle v, w_n \rangle}{\|w_n\|^2} w_n, w \rangle \\ &= \langle v, w \rangle - \frac{\langle v, w_1 \rangle}{\|w_1\|^2} \langle w_1, w \rangle - \dots - \frac{\langle v, w_n \rangle}{\|w_n\|^2} \langle w_n, w \rangle \\ &= \langle v, w \rangle - \frac{\langle v, w \rangle}{\|w\|^2} \|w\|^2 \\ &= \langle v, w \rangle - \langle v, w \rangle \\ &= 0, \end{aligned}$$

and so  $v - \text{proj}_W(v)$  is orthogonal to all the vectors  $w_i$ . It follows that  $v - \text{proj}_W(v)$  is orthogonal to any  $w \in W$ .

Hence

$$\begin{aligned} d(v, w)^2 &= \|v - w\|^2 \\ &= \langle v - w, v - w \rangle \\ &= \langle v - \text{proj}_W(v) + \text{proj}_W(v) - w, v - \text{proj}_W(v) + \text{proj}_W(v) - w \rangle \\ &= \langle v - \text{proj}_W(v), v - \text{proj}_W(v) \rangle + \langle v - \text{proj}_W(v), \text{proj}_W(v) - w \rangle \\ &\quad + \langle \text{proj}_W(v) - w, v - \text{proj}_W(v) \rangle + \langle \text{proj}_W(v) - w, \text{proj}_W(v) - w \rangle \\ &= \langle v - \text{proj}_W(v), v - \text{proj}_W(v) \rangle + 0 + \langle \text{proj}_W(v) - w, \text{proj}_W(v) - w \rangle \\ &= \|v - \text{proj}_W(v)\|^2 + \|(\text{proj}_W(v) - w)\|^2 \\ &\therefore d(v, w)^2 \geq \|v - \text{proj}_W(v)\|^2, \end{aligned}$$

with equality only when  $\text{proj}_W(v) - w = 0$ , ie  $w = \text{proj}_W(v)$ .

Thus  $d(v, w) \geq \|v - \text{proj}_W(v)\| = d(v, \text{proj}_W(v))$ , and the above observation also verifies uniqueness.  $\blacksquare$

$\text{proj}_W$  IS INDEPENDENT OF ORTHOGONAL BASIS FOR  $W$  (T10.4(4))

Let  $V$  be an IPS, and let  $W \subseteq V$  be finite-dimensional, with orthogonal basis  $\{w_1, \dots, w_n\}$ .

Let  $\text{proj}_W: V \rightarrow W$  be the associated orthogonal projection mapping.

Let  $\{x_1, \dots, x_n\}$  be another orthogonal basis for  $W$ , with associated orthogonal projection  $\text{proj}'_W$ .

Then necessarily  $\text{proj}_W = \text{proj}'_W$ .

Proof. By T10.4(2) & (3),  $\text{proj}'_W(v)$  is the unique vector in  $W$  closest to  $v$ .

Since  $\text{proj}_W$  also satisfies this property, thus  $\text{proj}'_W = \text{proj}_W(v)$ .

Since  $v \in V$  was arbitrary, it follows that  $\text{proj}_W = \text{proj}'_W$ , as needed.  $\blacksquare$

# Class III:

## Orthogonal Complements and Polynomial Interpolation

### ORTHOGONAL COMPLEMENTS: $S^\perp$

(TII.1)

Let  $V$  be an IPS, and let  $S \subseteq V$ .

Then, the "orthogonal complement" of  $S$ , denoted as " $S^\perp$ " (read as "S perp") is the set of vectors orthogonal to every vector in  $S$ ; ie

$$S^\perp = \{v \in V \mid \langle v, w \rangle = 0 \quad \forall w \in S\}.$$

$$\dim W < \infty \Rightarrow W \oplus W^\perp = V \quad (\text{TII.1(1)})$$

Let  $W \subseteq V$  be a finite-dimensional subspace.

Then necessarily  $V = W \oplus W^\perp$ .

Proof. See that for any  $w \in W \cap W^\perp$ ,  $\langle w, w \rangle = 0$  by defn, so  $w = 0$  necessarily.

To show  $V = W + W^\perp$ , consider  $\text{proj}_W$ . For any  $v \in V$ ,  $\text{proj}_W(v)$  is a lin comb of vectors in  $W$ , and so belongs to  $W$  itself.

So

$$v = \underbrace{\text{proj}_W(v)}_{\in W} + \underbrace{(v - \text{proj}_W(v))}_{\in W^\perp},$$

since  $v - \text{proj}_W(v) \in W^\perp$  from the proof of TII.1(2).

Hence  $V = W + W^\perp$ , as needed.  $\blacksquare$

$\dim V < \infty$ ,  $B_1, B_2$  ARE ORTHOGONAL BASES FOR

$W, W^\perp$  RESP  $\Rightarrow B_1 \cup B_2$  IS AN ORTHOGONAL BASIS FOR  $V$ ;

$$\dim W + \dim W^\perp = \dim V \quad (\text{TII.1(2)})$$

Let  $V$  be a finite-dimensional IPS, and let  $W \subseteq V$ .

Let  $B_1$  be an orthogonal basis for  $W$ , and let  $B_2$  be an orthogonal basis for  $W^\perp$ .

Then necessarily  $B_1 \cup B_2$  is an orthogonal basis for  $V$ , and in particular,

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

Proof. Let  $B_1 = \{w_1, \dots, w_k\}$  &  $B_2 = \{x_1, \dots, x_\ell\}$ . As  $V = W + W^\perp$ ,

thus  $B_1 \cup B_2$  spans  $V$ .

To show  $B_1 \cup B_2$  is an orthogonal basis, it suffices to show  $B_1 \cup B_2$  is orthogonal, as then  $B_1 \cup B_2$  would be lin ind by TII.1.

To show this, we just need to show  $\langle w_i, x_j \rangle = 0$ , as  $B_1 \cup B_2$  are already orthogonal by construction.

But this follows from the fact that  $w_i \in W$  &  $x_j \in W^\perp$ .

Thus  $B_1 \cup B_2$  is an orthogonal basis for  $V$ , so that

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

$$\text{Span}(S) = W \Rightarrow S^\perp = W^\perp \quad (\text{TII.1(3)})$$

Let  $W \subseteq V$ , and let  $\text{Span}(S) = W$ .

Then necessarily  $S^\perp = W^\perp$ .

In other words, to check if  $v \in W^\perp$ , it suffices to

Show  $v$  is orthogonal to every vector in  $S$ .

Proof. See that

$$v \in W^\perp \Leftrightarrow v \text{ is ortho to all vectors in } W$$

$$\Leftrightarrow v \text{ is ortho to all vectors in } S \text{ (as } S \subseteq W)$$

$$\Leftrightarrow v \in S^\perp.$$

So  $W^\perp \subseteq S^\perp$ .

Then, let  $v \in S^\perp$ . We want to show  $\langle v, w \rangle = 0 \quad \forall w \in W$ .

By defn of  $S$ ,  $\exists w_1, \dots, w_k \in S$ ,  $c_1, \dots, c_k \in \mathbb{C}$  s.t.

$$c_1 w_1 + \dots + c_k w_k = v.$$

As  $\langle v, w_i \rangle = 0$  for each  $i$  (since  $v \in S^\perp$ ), thus

$$\langle v, w \rangle = \langle v, c_1 w_1 + \dots + c_k w_k \rangle$$

$$= \bar{c}_1 \langle v, w_1 \rangle + \dots + \bar{c}_k \langle v, w_k \rangle$$

$$= \bar{c}_1(0) + \dots + \bar{c}_k(0) = 0,$$

so  $v \in W^\perp$ , so that  $S^\perp \subseteq W^\perp$ , and so  $S^\perp = W^\perp$ , as needed.  $\blacksquare$

$$\dim V < \infty \Rightarrow (W^\perp)^\perp = W; \quad S \subseteq V \Rightarrow (S^\perp)^\perp = \text{Span } S \quad (\text{TII.1(4)})$$

Let  $V$  be a finite-dimensional IPS, and let  $W \subseteq V$  be a subspace. Then necessarily  $(W^\perp)^\perp = W$ .

In general, if  $S \subseteq V$ , then  $(S^\perp)^\perp = \text{Span } S$ .

Proof. Let  $w \in W$ . By defn,  $\langle w, v \rangle = 0 \quad \forall v \in S^\perp$ , and so  $w \in (S^\perp)^\perp$ . Thus  $W \subseteq (S^\perp)^\perp$ .

By TII.1(2),  $\dim W + \dim W^\perp = \dim V = \dim W^\perp + \dim (W^\perp)^\perp$ , so that  $\dim W = \dim (W^\perp)^\perp$ . Hence  $W = (W^\perp)^\perp$ .

Then, let  $S \subseteq V$  and let  $W = \text{Span}(S)$ . By TII.1(3), necessarily  $S^\perp = W^\perp$ .

Taking orthogonal complements on both sides yields

$$(S^\perp)^\perp = (W^\perp)^\perp = W = \text{Span } S,$$

as needed.  $\blacksquare$

$$\ker \text{proj}_W = W^\perp, \quad \text{ran } \text{proj}_W = W \quad (\text{TII.2})$$

Let  $V$  be an IPS, let  $W$  be a finite-dimensional subspace of  $V$ , and let  $\text{proj}_W: V \rightarrow V$  be the orthogonal projection onto  $W$ .

Then necessarily  $\ker \text{proj}_W = W^\perp$  and  $\text{ran } \text{proj}_W = W$ .

Proof. Let  $\{w_1, \dots, w_k\}$  be an ortho basis for  $W$ .

Let  $v \in \ker \text{proj}_W$ . See that

$$\text{proj}_W(v) = \frac{\langle v, w_1 \rangle}{\|w_1\|^2} w_1 + \dots + \frac{\langle v, w_k \rangle}{\|w_k\|^2} w_k = 0.$$

By lin ind of  $\{w_1, \dots, w_k\}$ , hence  $\frac{\langle v, w_i \rangle}{\|w_i\|^2} = 0$ , and so each  $\langle v, w_i \rangle = 0$ , which suffices to show that  $v \in W^\perp$ .

Thus  $\text{proj}_W(v) = 0 \Leftrightarrow v \in W^\perp$ , ie  $\ker \text{proj}_W = W^\perp$ .

As  $\text{proj}_W(v) \in W \quad \forall v \in V$ , thus  $\text{ran } \text{proj}_W \subseteq W$ . Now, let  $w \in W$ . We wish to show  $w = \text{proj}_W(w)$ .

By TII.4,  $d(\text{proj}_W(w), w)$  is minimal. But  $d(w, w) = 0$ , so that showing that  $w$  is that "minimal element", ie  $w = \text{proj}_W(w)$ , as needed.  $\blacksquare$

# LAGRANGE INTERPOLATION; FINDING A POLYNOMIAL TO APPROXIMATE DATA (TII.3)

Let  $m \in \mathbb{N}$ , and let  $(x_1, y_1), \dots, (x_{m+1}, y_{m+1}) \in \mathbb{R}^2$ .

Let  $\hat{B} = \{\hat{p}_1, \dots, \hat{p}_{m+1}\} \subseteq P_m(\mathbb{R})$ , where

$$\hat{p}_i = \frac{\prod_{j \neq i} (x_i - x_j)}{\prod_{j \neq i} (x_i - x_j)}, \quad 1 \leq i \leq m+1.$$

Then  $\hat{B}$  is a basis for  $P_m(\mathbb{R})$ , and

$$p(x_i) = y_i \quad \forall 1 \leq i \leq m+1 \Leftrightarrow p = y_1 \hat{p}_1 + \dots + y_{m+1} \hat{p}_{m+1}.$$

Proof. See that

$$\hat{p}_i(x_k) = \frac{\prod_{j \neq i} (x_k - x_j)}{\prod_{j \neq i} (x_k - x_j)} = (\text{some stuff}) (x_k - x_k) = 0 \quad \text{if } i \neq k,$$

and

$$\hat{p}_i(x_i) = \frac{\prod_{j \neq i} (x_i - x_j)}{\prod_{j \neq i} (x_i - x_j)} = 1.$$

Hence

$$\begin{aligned} p = y_1 \hat{p}_1 + \dots + y_{m+1} \hat{p}_{m+1} &\Leftrightarrow p(x_i) = y_1 \hat{p}_1(x_i) + \dots + y_{m+1} \hat{p}_{m+1}(x_i) \\ &\Leftrightarrow p(x_i) = y_1 c_1 + \dots + y_{m+1} c_{m+1} \\ &\Leftrightarrow p(x_i) = y_i. \end{aligned}$$

To show  $\hat{B}$  is a basis for  $P_m(\mathbb{R})$ , we need only show its lin. ind., as  $|\hat{B}| = m+1 = \dim P_m(\mathbb{R})$ .

Let  $c_1, \dots, c_{m+1} \in \mathbb{R}$  such that

$$c_1 \hat{p}_1 + \dots + c_{m+1} \hat{p}_{m+1} = 0.$$

For any  $1 \leq i \leq m+1$  and evaluating both sides at  $x_i$  yields

$$\begin{aligned} 0 &= c_1 \hat{p}_1(x_i) + \dots + c_i \hat{p}_i(x_i) + \dots + c_{m+1} \hat{p}_{m+1}(x_i) \\ &= 0 + \dots + c_i(1) + \dots + 0 \\ \therefore 0 &= c_i, \end{aligned}$$

and so  $c_1 = \dots = c_{m+1} = 0$ , showing lin. ind., and we're done.  $\blacksquare$

## COLUMN SPACE [OF A MATRIX]: Col(A) (TII.2)

Let  $A \in M_{m,n}(\mathbb{R})$ .

Then, the "column space" of  $A$ , denoted as "Col(A)", is the set of vectors in  $\mathbb{R}^m$  of the form  $Ax$ , where  $x \in \mathbb{R}^n$ .

Equivalently, Col(A) is the set of all linear combinations of columns of  $A$ .

In particular, Col(A) is a subspace of  $\mathbb{R}^m$ , and the columns of  $A$  span Col(A). (TII.4)

$$A \in M_{m,n}(\mathbb{R}); \quad \text{Col}(A) = \text{Null}(A^T) \quad (\text{LII.1})$$

Let  $A \in M_{m,n}(\mathbb{R})$ , and give  $\mathbb{R}^n$  the standard inner product. Then necessarily  $\text{Col}(A)^\perp = \text{Null}(A^T)$ .

Proof. Since Col(A) is spanned by A's columns, by TII.1(2) we know  $y \in (\text{columns of } A)^\perp \Rightarrow y \in \text{Col}(A)^\perp$ .

Let  $y \in \text{Null}(A^T)$ , so  $A^T y = 0$ . In particular,

$$A^T y = \begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix} y = \begin{pmatrix} \langle a_1, y \rangle \\ \vdots \\ \langle a_n, y \rangle \end{pmatrix} \quad \begin{matrix} \text{(std inner product)} \\ \text{(ie dot product)} \end{matrix}$$

So  $A^T y = 0 \Rightarrow \langle a_i, y \rangle = 0 \Rightarrow y$  is ortho to each  $a_i$ ,

so  $y \in \text{Col}(A)^\perp$ , so  $\text{Null}(A^T) \subseteq \text{Col}(A)^\perp$ .

Conversely, let  $y \in \text{Col}(A)^\perp$ , so  $\langle a_i, y \rangle = 0 \quad \forall i$ . By the computation above,  $A^T y = 0$ , so  $y \in \text{Null}(A^T)$ . Hence  $\text{Col}(A)^\perp \subseteq \text{Null}(A^T)$ , and so  $\text{Col}(A)^\perp = \text{Null}(A^T)$ , as needed.  $\blacksquare$

$$x \in \mathbb{R}^n \text{ MINIMIZES } \|Ax - b\| \Leftrightarrow A^T A x = A^T b$$

## (TII.5)

Let  $A \in M_{m,n}(\mathbb{R})$  and  $b \in \mathbb{R}^m$ .

Then  $x \in \mathbb{R}^n$  minimizes  $\|Ax - b\|$  if and only if

$$A^T A x = A^T b.$$

Proof. See that

$$\begin{aligned} x \text{ minimizes } \|Ax - b\| &\Rightarrow Ax \in \text{proj}_{\text{Col}(A)}(b) \quad (\text{since } Ax \in \text{Col}(A) \text{ and} \\ &\quad \text{by TII.4(2)}) \\ &\Rightarrow A^T(b - Ax) = 0 \quad (\text{by LII.1}) \\ &\Rightarrow A^T A x = A^T b, \end{aligned}$$

and

$$\begin{aligned} A^T A x = A^T b &\Rightarrow A^T(b - Ax) = 0 \\ &\Rightarrow b - Ax \in \text{Null}(A^T) = \text{Col}(A)^\perp \quad (\text{by LII.1}) \\ &\Rightarrow Ax = \text{proj}_{\text{Col}(A)}(b) \quad (\text{see reasoning on previous}) \\ &\Rightarrow \|Ax - b\| \text{ is minimized} \\ &\Rightarrow x \text{ minimizes } \|Ax - b\|, \end{aligned}$$

as needed.  $\blacksquare$