MATH 235
Personal Notes

Marcus Chan

UW '25

Taught by Nick Rollick

Class 8: Examples of Matrix Representations, Introduction to Inner Product Spaces

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

Let V be a vector space over $\mathbb{F}$.
Then, we say the function $\langle\cdot,\cdot\rangle : V^2 \to \mathbb{F}$
is an "inner product" if

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

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

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

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

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 $\langle v,v\rangle \in \mathbb{R}$ $\forall v \in V$.

ORTHOGONAL [VECTORS] (D8.3)

Let V be an IPS.
Then, we say $v,w \in V$ are "orthogonal" if $\langle v,w\rangle = 0$.

ORTHOGONAL [SETS] (D8.3)

Let $S \subseteq V$, where V is an IPS.
Then, we say S is "orthogonal" if
$\langle v,w\rangle = 0$ $\forall v,w \in S$.

EXAMPLES OF IPS: PART I

The vector space $V = \mathbb{F}^n$ with inner product
$$\left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle = v_1\overline{w_1} + \cdots + v_n\overline{w_n}$$
is an inner product space. (E8.3)

The vector space $V = P_n(\mathbb{F})$ with inner product
$$\langle p,q\rangle = p(0)\overline{q(0)} + \cdots + p(n)\overline{q(n)}$$
is an inner product space. (E8.4)

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

Let $A = (a_{ij}) \in M_{m\times n}(\mathbb{F})$.
Then, the "conjugate" of A, denoted by "$\bar{A}$", is equal to

$$\bar{A} = (\overline{a_{ij}}) \in M_{m\times n}(\mathbb{F}).$$

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

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

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

STANDARD INNER PRODUCT ON $M_{m\times n}(\mathbb{F})$: $\langle A,B\rangle = tr(AB^*)$ (E8.5)

Let $V = M_{m\times n}(\mathbb{F})$.
Then, the "standard inner product" on V is given by
$$\langle A,B\rangle = tr(AB^*),$$
where $tr(A) = \sum_{i=1}^{m} a_{ii}$ for $A \in M_{m\times m}(\mathbb{F})$.

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 $B^* = (\overline{b_{ji}})$.
Then
$$(AB^*)_{ii} = \sum_{k=1}^{n} a_{ik}(B^*)_{ki} = \sum_{k=1}^{n} a_{ik}\overline{b_{ik}}.$$
Hence
$$tr(AB^*) = \sum_{i=1}^{m} (AB^*)_{ii} = \sum_{i=1}^{m}\sum_{k=1}^{n} a_{ik}\overline{b_{ik}}.$$
In particular, this is "similar" to if we wrote the entries of A & B in $\mathbb{F}^{mn}$, and took the standard inner product of these vectors.
It trivially follows that this gives an inner product on V. □

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

We can show $C[a,b]$ with the function
$$\langle f(x), g(x)\rangle = \int_a^b f(x)\overline{g(x)}\,dx$$
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. □

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

Let $w \in V$, and let $T_w : V \to \mathbb{F}$ by $T_w(v) = \langle v,w\rangle$ $\forall v \in V$.
Then necessarily T_w is linear.

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

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 $= \ker(T_w)$. □

$\|v\| \geq 0 \quad \forall v \in V, \qquad v = 0 \iff \|v\| = 0$ (T8·3(1))

Let V be an IPS.
Then necessarily $\|v\| \geq 0$ $\forall v \in V$, 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 $c \in \mathbb{F}$.
Then necessarily $\|cv\| = |c| \cdot \|v\|$.

Proof. This also arises from properties of inner products.

$|\langle v,w\rangle| \leq \|v\| \cdot \|w\|$; $|\langle v,w\rangle| = \|v\|\|w\| \iff v$ & 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):

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

$$\therefore 0 \leq \|v\|^2 - \frac{|\langle v,w\rangle|^2}{\|w\|^2},$$

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

Then, note that

$\|v-cw\| > 0 \iff v - cw \neq 0$ for some $c \in \mathbb{F}$
$\iff v \neq cw$ ''
$\iff v$ & w are lin ind,

and so $\|v-cw\| = 0 \iff v$ & w are lin dep. ∎

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

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

Proof. Note that

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

and the proof follows. ∎

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

ORTHOGONAL & ORTHONORMAL BASIS (D9.1)

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.

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

$S \subseteq V$ IS ORTHOGONAL & HAS NO ZERO VECTORS $\Rightarrow$ S IS LINEARLY INDEPENDENT (T9.1)

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 $c_1, \dots, c_n \in \mathbb{F}$, $v_1, \dots, v_n \in S$ s.t

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

Taking the inner product of each side with v_1, we see that

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

$$\therefore \; 0 = c_1 \|v_1\|^2,$$

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. ☒

V HAS ORTHOGONAL ORDERED BASIS $B = \{v_1, \dots, v_n\} \Rightarrow w = \sum_{i=1}^{n} \frac{\langle w, v_i \rangle}{\|v_i\|^2} v_i$ (T9.2)

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, $\exists c_1, \dots, c_n \ni$

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

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

$$\langle w, v_i \rangle = c_i \|v_i\|^2 \quad \forall 1 \leq i \leq n.$$

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

V HAS ORTHONORMAL ORDERED BASIS $B = \{v_1, \dots, v_n\} \Rightarrow w = \sum_{i=1}^{n} \langle w, v_i \rangle v_i$ (C9.1)

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; $v_i = v_i - \sum_{j=1}^{i-1} \frac{\langle w_i, v_j \rangle}{\|v_j\|^2} v_j \Rightarrow \{v_1, \dots, v_n\}$ IS ORTHOGONAL & $\{v_1, \dots, v_k\}$ IS AN ORTHOGONAL BASIS FOR $\mathrm{Span}\{w_1, \dots, w_k\}$ (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

$$v_i = w_i - \frac{\langle w_i, v_1 \rangle}{\|v_1\|^2} v_1 - \dots - \frac{\langle w_i, v_{i-1} \rangle}{\|v_{i-1}\|^2} v_{i-1} \quad \forall 1 < i \leq n.$$

Then

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

② $\{v_1, \dots, v_k\}$ is an orthogonal basis for $\mathrm{Span}\{w_1, \dots, w_k\}$ for any $1 \leq k \leq n$.

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 $\mathrm{span}\{v_1\} = \mathrm{span}\{w_1\}$, 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 $\mathrm{Span}\{w_1, \dots, w_k\}$.

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

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

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

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

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

By hypothesis, $\mathrm{Span}\{v_1, \dots, v_k\} = \mathrm{Span}\{w_1, \dots, w_k\}$, 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 $\mathrm{Span}\{v_1, \dots, v_{k+1}\} \subseteq \mathrm{Span}\{v_1, \dots, v_k, w_{k+1}\}$.

Then, $\mathrm{Span}\{v_1, \dots, v_k\} = \mathrm{Span}\{w_1, \dots, w_k\} \Rightarrow$ $\mathrm{Span}\{v_1, \dots, v_k, w_{k+1}\} = \mathrm{Span}\{w_1, \dots, w_k, w_{k+1}\}$, and so $\mathrm{Span}\{v_1, \dots, v_{k+1}\} \subseteq \mathrm{Span}\{w_1, \dots, w_{k+1}\}$.

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 $\mathrm{span}\{w_1, \dots, w_{k+1}\} \subseteq \mathrm{span}\{v_1, \dots, v_{k+1}\}$, and so $\mathrm{Span}\{w_1, \dots, w_{k+1}\} = \mathrm{Span}\{v_1, \dots, v_{k+1}\}$.

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 $\mathrm{Span}\{w_1, \dots, w_{k+1}\}$, 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 $\mathrm{Span}\{w_1, \dots, w_n\} = V$. ∎

$\dim V < \infty \Rightarrow V$ 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. ∎

Class 10: Direct Sums of Subspaces and Orthogonal Projections

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

Let $W_1, \ldots, W_n \leq V$.

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

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

Note the following:

① $W_1 + \cdots + W_k$ is a subspace of V, and the smallest subspace containing $W_1, \ldots, W_k$; and

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

DIRECT SUM OF SUBSPACES (D10.2)

Let V be a vector space, and let $W_1, \ldots, W_k \leq V$.

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

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

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

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

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

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: $\operatorname{proj}_W(v)$ (D10.4)

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

Then, the "orthogonal projection map" of V onto W, denoted as "$\operatorname{proj}_W : V \to V$", is defined by

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

$\operatorname{proj}_W(v)$ IS A LINEAR TRANSFORMATION (T10.4 (1))

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

Let $\operatorname{proj}_W : V \to V$ be the associated orthogonal projection mapping.

Then necessarily proj_W is a linear transformation.

Proof. See that

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

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

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

Let $\operatorname{proj}_W : V \to V$ be the associated orthogonal projection mapping.

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

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

Moreover, note that

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

Proof. Let $1 \leq i \leq n$. See that

$$\begin{aligned}
\langle v - \operatorname{proj}_W(v), w_i \rangle &= \left\langle v - \frac{\langle v, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v, w_2 \rangle}{\|w_2\|^2} w_2 - \cdots - \frac{\langle v, w_k \rangle}{\|w_k\|^2} w_k, w_i \right\rangle \\
&= \langle v, w_i \rangle - \frac{\langle v, w_1 \rangle}{\|w_1\|^2} \langle w_1, w_i \rangle - \cdots - \frac{\langle v, w_k \rangle}{\|w_k\|^2} \langle w_k, w_i \rangle \\
&= \langle v, w_i \rangle - \frac{\langle v, w_i \rangle}{\|w_i\|^2} \|w_i\|^2 \\
&= \langle v, w_i \rangle - \langle v, w_i \rangle \\
&= 0,
\end{aligned}$$

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

Hence

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

$\therefore d(v,w)^2 \geq \|v - \operatorname{proj}_W(v)\|^2$, with equality only when $\operatorname{proj}_W(v) - w = 0$, ie $w = \operatorname{proj}_W(v)$.

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

proj_W IS INDEPENDENT OF ORTHOGONAL BASIS FOR W (T10.4 (4))

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

Let $\operatorname{proj}_W : V \to V$ be the associated orthogonal projection mapping.

Let $\{x_1, \ldots, x_n\}$ be another orthogonal basis for W, with associated orthogonal projection proj_W'.

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

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

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

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

Class 11: Orthogonal Complements and Polynomial Interpolation

ORTHOGONAL COMPLEMENTS: $S^\perp$ (D11.1)

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

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

$\dim W < \infty \Rightarrow W \oplus W^\perp = V$ (T11.1(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 def^n, so $w = 0$ necessarily.
To show $V = W + W^\perp$, consider 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 T10.4(2).
Hence $V = W + W^\perp$, as needed. #

$\dim V < \infty$, B_1, B_2 ARE ORTHOGONAL BASES FOR $W, W^\perp$ RESP $\Rightarrow B_1 \cup B_2$ IS AN ORTHOGONAL BASIS FOR V; $\dim W + \dim W^\perp = \dim V$ (T11.1(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 T9.1.
To show this, we just need to show $\langle w_i, x_j \rangle = 0$, as B_1 & 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 $\dim V = \dim W + \dim W^\perp$. ∎

$\text{Span}(S) = W \Rightarrow S^\perp = W^\perp$ (T11.1(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

$$\begin{aligned} v \in W^\perp &\Rightarrow v \text{ is ortho to all vectors in } W \\ &\Rightarrow v \text{ is ortho to all vectors in } S \text{ (as } S \subseteq W) \\ &\Rightarrow v \in S^\perp, \end{aligned}$$

so $W^\perp \subseteq S^\perp$.
Then, let $v \in S^\perp$. We want to show $\langle v, w \rangle = 0$ $\forall w \in W$.
By def^n of S, $\exists w_1, \dots, w_k \in S$, $c_1, \dots, c_k \in \mathbb{F}$ $\ni$

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

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

$$\begin{aligned} \langle v, w \rangle &= \langle v, c_1 w_1 + \dots + c_k w_k \rangle \\ &= \overline{c_1} \langle v, w_1 \rangle + \dots + \overline{c_k} \langle v, w_k \rangle \\ &= \overline{c_1}(0) + \dots + \overline{c_k}(0) = 0, \end{aligned}$$

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

$\dim V < \infty \Rightarrow (W^\perp)^\perp = W$; $S \subseteq V \Rightarrow (S^\perp)^\perp = \text{Span } S$ (T11.1(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 def^n, $\langle w, x \rangle = 0$ $\forall x \in W^\perp$, and so $w \in (W^\perp)^\perp$. Thus $W \subseteq (W^\perp)^\perp$.
By T11.1(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 T11.1(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. ∎

$\ker \text{proj}_W = W^\perp$, $\text{ran } \text{proj}_W = W$ (T11.2)

Let V be an IPS, let W be a finite-dimensional subspace of V, and let $\text{proj}_W : V \to V$ 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 $\text{proj}_W(v) \in W$ $\forall v \in V$, thus $\text{ran } \text{proj}_W \subseteq W$. Now, let $w \in W$. We wish to show $w = \text{proj}_W(w)$.
By T10.4, $d(\text{proj}_W(w), w)$ is minimal. But $d(w, w) = 0$, showing that w is that "minimal element", so that $w = \text{proj}_W(w)$, as needed. ∎

LAGRANGE INTERPOLATION; FINDING A POLYNOMIAL TO APPROXIMATE DATA (T11.3)

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

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

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

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

Proof. See that

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

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

$$\begin{aligned} p = y_1 \hat{p}_1 + \dots + y_{m+1} \hat{p}_{m+1} \iff p(x_i) &= y_1 \hat{p}_1(x_1) + \dots + y_i \hat{p}_i(x_i) + \dots + y_{m+1} \hat{p}_{m+1}(x_{m+1}) \\ \iff p(x_i) &= y_1(0) + \dots + y_i(1) + \dots + y_{m+1}(0) \\ \iff p(x_i) &= y_i. \end{aligned}$$

To show $\hat{B}$ is a basis for $P_m(\mathbb{R})$, we need only show it's 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

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

$\therefore 0 = c_i$,

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

COLUMN SPACE [OF A MATRIX]: Col(A) (D11.2)

1. Let $A \in M_{m \times n}(\mathbb{F})$.
Then, the "column space" of A, denoted as "Col(A)", is the set of vectors in $\mathbb{F}^m$ of the form Ax, where $x \in \mathbb{F}^n$.
Equivalently, Col(A) is the set of all linear combinations of columns of A.

2. In particular, Col(A) is a subspace of $\mathbb{F}^m$, and the columns of A span Col(A). (T11.4)

$A \in M_{m \times n}(\mathbb{R})$; $\mathrm{Col}(A) = \mathrm{Null}(A^T)$ (L11.1)

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

Proof. Since Col(A) is spanned by A's columns, by T11.1(3) we know $y \in (\text{columns of } A)^\perp \Rightarrow y \in \mathrm{Col}(A)^\perp$.
Let $y \in \mathrm{Null}(A^T)$, so $A^T y = 0$. In particular,

$$A^T y = \begin{pmatrix} - a_1 - \\ \vdots \\ - a_n - \end{pmatrix} y = \begin{pmatrix} \langle a_1, y \rangle \\ \vdots \\ \langle a_n, y \rangle \end{pmatrix},$$

std inner product (ie dot product)

So $A^T y = 0 \Rightarrow \langle a_i, y \rangle = 0 \Rightarrow y$ is ortho to each a_i,
so $y \in \mathrm{Col}(A)^\perp$, so $\mathrm{Null}(A^T) \subseteq \mathrm{Col}(A)^\perp$.
Conversely, let $y \in \mathrm{Col}(A)^\perp$, so $\langle a_i, y \rangle = 0 \; \forall i$. By the computation above, $A^T y = 0$, so $y \in \mathrm{Null}(A^T)$. Hence $\mathrm{Col}(A)^\perp \subseteq \mathrm{Null}(A^T)$, and so $\mathrm{Col}(A)^\perp = \mathrm{Null}(A^T)$, as needed. ☒

$x \in \mathbb{R}^n$ MINIMIZES $\|Ax - b\| \iff A^T A x = A^T b$ (T11.5)

Let $A \in M_{m \times n}(\mathbb{R})$ and $b \in \mathbb{R}^m$.
Then $x \in \mathbb{R}^m$ minimizes $\|Ax - b\|$ if and only if $A^T A x = A^T b$.

Proof. See that

$$\begin{aligned} x \text{ minimizes } \|Ax - b\| &\Rightarrow Ax \in \mathrm{proj}_{\mathrm{Col}(A)}(b) && \text{(since } Ax \in \mathrm{Col}(A) \text{ and by T10.4(2))} \\ &\Rightarrow A^T(b - Ax) = 0 && \text{(by L11.1)} \\ &\Rightarrow A^T A x = A^T b, \end{aligned}$$

and

$$\begin{aligned} A^T A x = A^T b &\Rightarrow A^T(b - Ax) = 0 \\ &\Rightarrow b - Ax \in \mathrm{Null}(A^T) = \mathrm{Col}(A)^\perp && \text{(by L11.1)} \\ &\Rightarrow Ax = \mathrm{proj}_{\mathrm{Col}(A)}(b) && \text{(see reasoning on Pensall)} \\ &\Rightarrow \|Ax - b\| \text{ is minimized} \\ &\Rightarrow x \text{ minimizes } \|Ax - b\|, \end{aligned}$$

as needed. ☒

Class 12: Linear Transformations on an Inner Product Space

[T:V→W] PRESERVES INNER PRODUCTS (D12.1)

1. Let $(V, \langle\cdot,\cdot\rangle)$ & $(W, [\cdot,\cdot])$ be IPS, and let $T: V\to W$ be linear.
Then, we say T "preserves inner products" if
$$[T(v_1), T(v_2)] = T(\langle v_1, v_2\rangle) \quad \forall v_1, v_2 \in V.$$

2. In particular, we say T is an "isomorphism" of inner product spaces if T is also an isomorphism.

POLARIZATION IDENTITIES

The polarization identities state that

① V over $\mathbb{R}$ $\Rightarrow$ $\langle x,y\rangle = \frac{1}{4}\|x+y\|^2 + \frac{1}{4}\|x-y\|^2$; &

② V over $\mathbb{C}$ $\Rightarrow$ $\langle x,y\rangle = \frac{1}{4}\|x+y\|^2 + \frac{i}{4}\|x+iy\|^2 - \frac{1}{4}\|x-y\|^2 - \frac{i}{4}\|x-iy\|^2$.

Proof. This can be verified by expanding the norms in the RHS in terms of IPs. ■

T PRESERVES INNER PRODUCTS ⟺ T PRESERVES NORMS (T12.1(1))

Let V, W be IPS, and let $T: V\to W$ be linear.
Then T preserves inner products iff T preserves norms;
ie $\|T(x)\|_W = \|x\|_V \ \forall x \in V$.

Proof. If T preserves inner products, by defn of the norm, it also preserves norms.
Now, suppose T preserves norms. If V, W are over $\mathbb{R}$, then by the polarization identities:
$$\begin{aligned}[T(x), T(y)] &= \tfrac{1}{4}\|T(x)+T(y)\|_W^2 + \tfrac{1}{4}\|T(x)-T(y)\|_W^2 \\ &= \tfrac{1}{4}\|T(x+y)\|_W^2 + \tfrac{1}{4}\|T(x-y)\|_W^2 \\ &= \tfrac{1}{4}\|x+y\|_V^2 + \tfrac{1}{4}\|x-y\|_V^2 \\ &= \langle x,y\rangle,\end{aligned}$$
and the case where V, W are over $\mathbb{C}$ is similar, showing T preserves IPs as well. ■

T PRESERVES INNER PRODUCTS ⇒ T IS 1-1 (T12.1(2))

Let $T: V\to W$ preserve inner products.
Then necessarily T is 1-1.

Proof. We will show $\ker T = \{0\}$.
If $v \in V \ni T(v)=0$, then trivially
$$[T(v), T(v)] = [0,0] = 0.$$
Since T preserves IPs, we have that
$[T(v), T(v)] = \langle v, v\rangle = 0$, so $v=0$ if $T(v)=0$.
Thus $\ker T = \{0\}$, as needed. ■

T IS AN ISOMORPHISM OF IPS, $\{v_1, \dots, v_n\}$ IS AN ORTHOGONAL/ORTHONORMAL BASIS FOR V ⇒ $\{T(v_1), \dots, T(v_n)\}$ IS AN ORTHOGONAL/ORTHONORMAL BASIS FOR W (T12.1(3))

Let $T: V\to W$ be an isomorphism of inner product spaces, and let $\{v_1, \dots, v_n\}$ be an orthogonal (or orthonormal) basis for V.
Then necessarily $\{T(v_1), \dots, T(v_n)\}$ is an orthogonal (or orthonormal) basis for W.

Proof. We first show $\{T(v_1), \dots, T(v_n)\}$ is orthogonal.
Since $\{v_1, \dots, v_n\}$ is orthogonal, thus $\langle v_i, v_j\rangle = 0 \ \forall i \neq j$. As T preserves IPs, thus $[T(v_i), T(v_j)] = 0 \ \forall i\neq j$, which shows $\{T(v_1), \dots, T(v_n)\}$ is orthogonal.
Then, as T is an isomorphism, thus T is 1-1, so $T(v_i) \neq 0 \ \forall 1 \leq i \leq n$ and $\ker T = \{0\}$.
In particular, $\{T(v_1), \dots, T(v_n)\}$ is lin ind by T9.1. Since $T(V) = W$, thus $\dim V = \dim W$, and so it follows that $\{T(v_1), \dots, T(v_n)\}$ is a basis for W, which is what we wanted to prove. ■

$\{v_1, \dots, v_n\}$ IS AN ORTHONORMAL BASIS FOR V, $\{T(v_1), \dots, T(v_n)\}$ IS AN ORTHONORMAL BASIS FOR W ⇒ T IS AN ISOMORPHISM OF IPS (T12.1(4))

Let $T: V\to W$ be linear, and let $\{v_1, \dots, v_n\}$ be an orthonormal basis for V such that $\{T(v_1), \dots, T(v_n)\}$ is an orthonormal basis for W.
Then necessarily T is an isomorphism of inner product spaces.

Proof. First, we show T preserves IPs.
Let $x_1, x_2 \in V$, say
$$x_1 = c_1v_1 + \dots + c_nv_n \quad \& \quad x_2 = d_1v_1 + \dots + d_nv_n.$$
Then
$$T(x_1) = c_1T(v_1) + \dots + c_nT(v_n) \quad \& \quad T(x_2) = d_1T(v_1) + \dots + d_nT(v_n).$$
See that
$$\begin{aligned}\langle x_1, x_2\rangle &= \langle c_1v_1 + \dots + c_nv_n, d_1v_1 + \dots + d_nv_n\rangle \\ &= \sum_{i=1}^n\sum_{j=1}^n c_i\overline{d_j}\langle v_i, v_j\rangle \\ &= \sum_{i=1}^n c_i\overline{d_i}\langle v_i, v_i\rangle \\ &= \sum_{i=1}^n c_i\overline{d_i}(1) \qquad (\text{as } \|v_i\| = 1)\end{aligned}$$
$$\therefore \langle x_1, x_2\rangle = \sum_{i=1}^n c_i\overline{d_i},$$
and
$$\begin{aligned}[T(x_1), T(x_2)] &= [T(c_1v_1 + \dots + c_nv_n), T(d_1v_1 + \dots + d_nv_n)] \\ &= [c_1T(v_1) + \dots + c_nT(v_n), d_1T(v_1) + \dots + d_nT(v_n)] \\ &= \sum_{i=1}^n\sum_{j=1}^n c_i\overline{d_j}[T(v_i), T(v_j)] \\ &= \sum_{i=1}^n c_i\overline{d_j}[T(v_i), T(v_i)] \\ &= \sum_{i=1}^n c_i\overline{d_j}(1) \qquad (\text{as } \|T(v_i)\| = 1) \\ &= \sum_{i=1}^n c_i\overline{d_j},\end{aligned}$$
showing T preserves IPs.
By T12.1(2), thus T is 1-1. In particular, since $\dim V = \dim W$, thus T is also an isomorphism, and we're done. ■

$B = (v_1, \dots, v_n)$ IS AN ORTHONORMAL BASIS FOR V ⇒ $[T]_B = (\langle T(v_j), v_i\rangle)_{ij} \in M_{n\times n}(\mathbb{F})$ (T12.2)

Let V be a finite-dimensional IPS, and in particular, let $B = (v_1, \dots, v_n)$ be an ordered orthonormal basis for V.
Let $A = [T]_B$. Then necessarily
$$A_{ij} = \langle T(v_j), v_i\rangle \quad \forall 1 \leq i,j \leq n.$$

Proof. By C9.1, we have that
$$[T(v_j)]_B = \begin{pmatrix}\langle T(v_j), v_1\rangle \\ \vdots \\ \langle T(v_j), v_n\rangle\end{pmatrix} \quad \forall 1\leq j\leq n.$$
Since $[T(v_j)]_B$ is the j^{th} column in $[T]_B$, it follows that the entry in the i^{th} row & j^{th} column of $[T]_B$ is $\langle T(v_j), v_i\rangle$, as needed. ■

$B = (v_1, \dots, v_n)$ IS AN ORDERED ORTHONORMAL BASIS FOR $V \Rightarrow \langle x,y\rangle = [y]_B^* [x]_B$ (L12.1)

Let V be a finite-dimensional IPS, and in particular, let $B = (v_1, \dots, v_n)$ be an ordered orthonormal basis for V.

Let $x, y \in V$. Then necessarily

$$\langle x,y\rangle = [y]_B^* [x]_B.$$

Proof. Let $[x]_B = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ & $[y]_B = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$. Then

$$\begin{aligned}\langle x,y\rangle &= \langle c_1 v_1 + \dots + c_n v_n, d_1 v_1 + \dots + d_n v_n\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n c_i \overline{d_j} \langle v_i, v_j\rangle \\ &= \sum_{i=1}^n c_i \overline{d_i} \\ &= (\overline{d_1} \cdots \overline{d_n}) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \\ &= [y]_B^* [x]_B. \quad \square\end{aligned}$$

$B = (v_1, \dots, v_n)$ IS AN ARBITRARY ORDERED ORTHONORMAL BASIS FOR V; $T: V \to V$ IS AN IPS ISOMORPHISM $\Leftrightarrow [T]_B^* = [T]_B^{-1}$ (T12.3)

Let V be a finite-dimensional IPS, let $T: V \to V$ be linear, and let $B = (v_1, \dots, v_n)$ be an ordered orthonormal basis for V.

Then T is an inner product space isomorphism iff $[T]_B^* = [T]_B^{-1}$.

Proof. ($\Rightarrow$) By defn., $\langle T(x), T(y)\rangle = \langle x,y\rangle$ $\forall x,y \in V$.

By L12.1, $\langle x,y\rangle = [y]_B^* [x]_B$ & $\langle T(x), T(y)\rangle = [T(y)]_B^* [T(x)]_B$.

We also know

$$[T(x)]_B = [T]_B [x]_B \quad \& \quad [T(y)]_B = [T]_B [y]_B.$$

Hence

$$\begin{aligned}\langle T(x), T(y)\rangle &= [T(y)]_B^* [T(x)]_B \\ &= ([T]_B [y]_B)^* ([T]_B [x]_B) \\ &= [y]_B^* ([T]_B^* [T]_B) [x]_B \\ &(= \langle x,y\rangle = [y]_B^* [x]_B).\end{aligned}$$

We claim this implies $[T]_B^* [T]_B = I_n$.

Indeed, let $[y]_B = e_i$ & $[x]_B = e_j$.

Then $([T]_B^* [T]_B)[x]_B$ picks out the j^{th} column of $[T]_B^* [T]_B$, & taking the product on the left with $[y]_B^*$ gives the (i,j) entry of $[T]_B^* [T]_B$.

On the other hand, $[y]^* [x]_B = e_i^* e_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$, which suffices to show $[T]_B^* [T]_B = I_n$, and so $[T]_B^* = ([T]_B)^{-1}$, as needed. $\square$

($\Leftarrow$) We first show T preserves IPs. See that for $x, y \in V$, we have that

$$\begin{aligned}\langle T(x), T(y)\rangle &= [T(y)]_B^* [T(x)]_B \\ &= [y]_B^* ([T]_B^* [T]_B) [x]_B \\ &= [y]_B^* (I_n) [x]_B \quad \text{(by assumption)} \\ &= [y]_B^* [x]_B = \langle x,y\rangle \text{ by L12.1.}\end{aligned}$$

Hence, by T12.1(2), T is 1-1. Since $T: V \to V$, thus T is an isomorphism of IPS, as needed. $\square$

UNITARY MATRICES: $A^* = A^{-1}$ (D12.2)

Let $A \in M_{n\times n}(\mathbb{F})$.

Then, we say A is "unitary" if $A^* = A^{-1}$.

ORTHOGONAL MATRICES: $A^T = A^{-1}$ (D12.2)

Let $A \in M_{n\times n}(\mathbb{F})$.

Then, we say A is "orthogonal" if $A^T = A^{-1}$.

Class 13: Diagonalization Review

$A=\begin{pmatrix}-a_1-\\ \vdots \\ -a_n-\end{pmatrix} \Rightarrow \exists D:(\mathbb{F}^n)^n\to\mathbb{F} \ni D(e_1,\dots,e_n)=1$ & $D(a_1,\dots,a_{i-1},a_i+cb_i,a_{i+1},\dots,a_n)=D(a_1,\dots,a_n)+D(a_1,\dots,a_{i-1},b_i,a_{i+1},\dots,a_n)$ & $a_i=a_j \Rightarrow D(a_1,\dots,a_i,\dots,a_j,\dots,a_n)=0$ (T13.1)

1: There exists a unique function $D:(\mathbb{F}^n)^n\to\mathbb{F}$ such that:

① If $e_1,\dots,e_n$ are the standard basis vectors, then $D(e_1,\dots,e_n)=1$;

② D is "multilinear"; ie if we fix $1\le i\le n$, $c\in\mathbb{F}$ & $b_i\in\mathbb{F}^n$, we have that
$$D(a_1,\dots,a_{i-1},a_i+b_i,a_{i+1},\dots,a_n)=D(a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n)+D(a_1,\dots,a_{i-1},b_i,a_{i+1},\dots,a_n);$$ &

③ D is an "alternating function"; ie if $a_i=a_j$ for $i\neq j$, then
$D(a_1,\dots,a_i,\dots,a_j,\dots,a_n)=0$.

2: In particular, $D(a_1,\dots,a_n)$ is the "determinant" of $A=\begin{pmatrix}-a_1-\\ \vdots \\ -a_n-\end{pmatrix}$.

* this is an alternative definition of determinants (aside from the cofactor definition).

DIAGONALIZABLE [LINEAR OPERATOR] (D13.2)

Let V be finite-dimensional, and let $T:V\to V$ be linear. Then, we say T is "diagonalizable" if there exists an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix.

DIAGONALIZABLE [MATRIX] (D13.2)

Let $A\in M_{n\times n}(\mathbb{F})$. Then, we say A is "diagonalizable" if the matrix multiplication operator $T_A:\mathbb{F}^n\to\mathbb{F}^n$ by $T_A(x)=Ax$ $\forall x\in\mathbb{F}^n$ is diagonalizable.

EIGENVECTOR & EIGENVALUE [OF LINEAR OPERATORS] (D13.3)

1: Let $T:V\to V$ be linear. Then, we say $0\neq v\in V$ is an "eigenvector" of T if there exists some $c\in\mathbb{F}$ such that $T(v)=cv$.

2: In particular, we call $c\in\mathbb{F}$ an "eigenvalue" of T.

EIGENVECTOR & EIGENVALUE [OF MATRICES] (D13.3)

1: Let $A\in M_{n\times n}(\mathbb{F})$. Then, the eigenvectors and eigenvalues of A are just the corresponding eigenvectors and eigenvalues of T_A.

2: In other words, $0\neq x\in\mathbb{F}^n$ is an eigenvector of A if $Ax=cx$, and in this case say that $c\in\mathbb{F}$ is the associated eigenvalue.

EIGENSPACE (T13.3)

1: Let $T:V\to V$ be linear, and let $c\in\mathbb{F}$ be an eigenvalue of T. Then, the "eigenspace" associated to c, denoted as "E_c", is defined to be the set
$$E_c=\{v\in V:\ T(v)=cv\}.$$

2: Indeed, E_c is a subspace of V.

T IS INVERTIBLE <=> $[T]_\beta$ IS INVERTIBLE FOR ANY ORDERED BASIS β OF V (L13.1)

Let $T:V\to V$ be linear on a finite-dimensional vector space V. Then necessarily T is invertible iff $[T]_\beta$ is invertible for every ordered basis β of V.

Proof. ($\Rightarrow$) Let β be an arbitrary basis of V.
We know $\exists T^{-1}:V\to V \ni T^{-1}\circ T=T\circ T^{-1}=I$.
Hence
$$I_n=[I]_\beta=[T\circ T^{-1}]_\beta=[T]_\beta[T^{-1}]_\beta,$$
showing $[T]_\beta$ is invertible.

($\Leftarrow$) Suppose T is not invertible. In particular T is not 1-1, so $\exists 0\neq v\in V \ni T(v)=0$.
Take any ordered basis β. Translating into matrices, this
$$[T(v)]_\beta=[T]_\beta[v]_\beta=[0]_\beta=0,$$
showing that $[v]_\beta\in \mathrm{Null}([T]_\beta)$.
Since $[v]_\beta\neq 0$, thus $[T]_\beta$ is not invertible. ☒

$c\in\mathbb{F}$ IS AN EIGENVALUE OF T <=> $\det[T-cI]_\beta=0$ FOR SOME BASIS β OF V (L13.2)

Let $\dim V<\infty$, and let $T:V\to V$ be linear. Then, $c\in\mathbb{F}$ is an eigenvalue of T iff $\det[T-cI]_\beta=0$ for some choice of ordered basis β of V.

Proof. ($\Rightarrow$) Let $c\in\mathbb{F}$ be an eigenvalue of T, w/ eigenvector $v\in V$.
So $T(v)=cv$. and so $(T-cI)(v)=0$.
In particular $c\in\ker(T-cI)$, and as $c\neq 0$, thus $T-cI$ is not 1-1.
Hence $[T-cI]_\beta$ is not invertible by L13.1 and so $\det[T-cI]_\beta=0$.

($\Leftarrow$) If $\det[T-cI]_\beta=0$, then by L13.1 $T-cI$ is not invertible.
In particular, there exists a $0\neq v\in V \ni (T-cI)(v)=0$.
Hence
$$T(v)=cI(v)=cv,$$
showing that c is an eigenvalue of V. ☒

CHARACTERISTIC POLYNOMIAL [OF $T:V\to V$]: $C(t)$ (D13.4)

1: Let $\dim V<\infty$, and let $T:V\to V$ be linear. Then, the "characteristic polynomial" of T is the polynomial
$$C(t)=\det[T-tI]_\beta,$$
where β is any ordered basis for V.

2: In particular, if β' is another ordered basis for V, then necessarily
$$\det([T-tI]_\beta)=\det([T-tI]_{\beta'}).\qquad \text{(L13.2)}$$

Proof. We know
$$[T]_{\beta'}={}_{\beta'}[I]_\beta\cdot[T]_\beta\cdot({}_{\beta'}[I]_\beta)^{-1}.$$
Let $P={}_{\beta'}[I]_\beta$. Note $[I]_\beta=[I]_{\beta'}=n=\dim V$.
Then, see that
$$\begin{aligned}\det[T-tI]_{\beta'} &= \det([T]_{\beta'}-t[I]_{\beta'})\\ &= \det([T]_{\beta'}-tI_n)\\ &= \det(P[T]_\beta P^{-1}-t(PI_nP^{-1}))\\ &= \det(P([T]_\beta-tI_n)P^{-1})\\ &= \det(P)\det([T]_\beta-tI_n)\det(P^{-1})\\ &= \det([T]_\beta-tI_n)\\ &= \det[T-tI_n]_\beta\end{aligned}$$
as needed. ☒

ALGEBRAIC MULTIPLICITY [OF AN EIGENVALUE]: a_c (D13.5)

Let $\dim V < \infty$, and let $T: V \to V$ be linear, and let $c \in \mathbb{F}$ be an eigenvalue of T.
Then, the "algebraic multiplicity" of c, denoted "a_c", is the largest positive integer $k \in \mathbb{Z}^+$ such that $(t-c)^k$ is a factor of the characteristic polynomial $C(t)$.

GEOMETRIC MULTIPLICITY [OF AN EIGENVALUE]: g_c (D13.5)

Let $\dim V < \infty$, and let $T: V \to V$ be linear, and let $c \in \mathbb{F}$ be an eigenvalue of T.
Then, the "geometric multiplicity" of c, denoted by "g_c", is defined to be equal to

$$g_c = \dim E_c.$$

$1 \le g_c \le a_c$ (T13.4(1))

Let $T: V \to V$ be linear, and let $c \in \mathbb{F}$ be an eigenvalue of T.
Then necessarily $1 \le g_c \le a_c$.

T IS DIAGONALIZABLE $\iff$ $g_c = a_c$ $\forall c$ (T13.4(2))

Let $T: V \to V$ be linear.
Then T is diagonalizable iff $g_c = a_c$ for any eigenvalues $c \in \mathbb{F}$ of T.

$c_1, \dots, c_k$ ARE THE DISTINCT ROOTS OF $C(t)$ $\Rightarrow$ $E_{c_1} \oplus \cdots \oplus E_{c_k}$; T IS DIAGONALIZABLE $\iff$ $E_{c_1} \oplus \cdots \oplus E_{c_k} = V$ (T13.4(3))

Let $T: V \to V$ be linear, with characteristic polynomial $C(t)$.
Let $c_1, \dots, c_k$ be the distinct roots of $C(t)$.
Then necessarily
① $E_{c_1} \oplus \cdots \oplus E_{c_k}$ (ie sum is direct); and
② T is diagonalizable iff $E_{c_1} \oplus \cdots \oplus E_{c_k} = V$.

Proof Let's first show ①. We do this by showing

$$E_{c_i} \cap (E_{c_1} + \cdots + E_{c_k}) = \{0\} \quad \forall 1 \le i \le k.$$

Let $v \in E_{c_i} \cap (E_{c_1} + \cdots + E_{c_k})$. In particular,

$$T(v) = c_i v \quad \& \quad v = w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_k$$

for some $w_j \in E_{c_j}$ for each j.
Suppose we chose $w_1, \dots, w_k$ so that the # of non-zero vectors is as minimal as possible.
If $\neg(w_1 = \cdots = w_k = 0)$, WLOG assume $w_k \neq 0$.
Then,

$$\begin{aligned} c_i v = T(v) &= T(w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_k) \\ &= T(w_1) + \cdots + T(w_{i-1}) + T(w_{i+1}) + \cdots + T(w_k) \\ &= c_1 w_1 + \cdots + c_{i-1} w_{i-1} + c_i w_i + \cdots + c_k w_k, \end{aligned}$$

but as $v = w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_k$, thus

$$c_k v = c_k w_1 + \cdots + c_k w_{i-1} + c_k w_{i+1} + \cdots + c_k w_k.$$

Hence

$$(c_i - c_k) v = (c_1 - c_k) w_1 + \cdots + (c_{i-1} - c_k) w_{i-1} + (c_{i+1} - c_k) w_{i+1} + \cdots + 0 w_k.$$

Dividing by $c_i - c_k$, we see that we have written v as a sum of fewer non-zero vectors from the other subspaces than we had before — a contradiction to our initial assumption.
Thus $w_1 = \cdots = w_{i-1} = w_{i+1} = \cdots = w_k = 0$, so $v = 0$, completing the proof that $E_{c_1} + \cdots + E_{c_k}$ is direct.

Next, since a basis for $E_{c_1} \oplus E_{c_2} \oplus \cdots \oplus E_{c_k}$ is built by combining bases from $E_{c_1}, \dots, E_{c_k}$, if $V = E_{c_1} \oplus \cdots \oplus E_{c_k}$, then we can obtain a basis for V by the above.
In particular, we can find a basis of eigenvectors of T for V, showing that T is diagonalizable.
Conversely, if T is diagonalizable, then there is a basis for V consisting of eigenvectors of T.
Each of these eigenvectors belongs to one of the spaces E_{c_i}, and so putting together bases for $E_{c_1}, \dots, E_{c_k}$ yields a basis of V, which suffices to show $E_{c_1} \oplus \cdots \oplus E_{c_k} = V$. ◻

Class 14: Orthogonal Diagonalization

$A \in M_{n\times n}(\mathbb{R})$ IS ORTHOGONAL $\Leftrightarrow$ COLUMNS OF A FORMS AN ORTHONORMAL BASIS FOR $\mathbb{R}^n$ WRT STD INNER PRODUCT (L14.1)

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

Then necessarily A is orthogonal iff the columns of A forms an orthonormal basis for $\mathbb{R}^n$ (with respect to the standard inner product).

Proof. Recall A is orthogonal $\Leftrightarrow A^T = A^{-1}$, ie

$$I_n = A^T A.$$

Then, the (i,j) entry of A^TA is given by

$$\sum_{k=1}^{n} (A^T)_{ik} A_{kj} = \sum_{k=1}^{n} A_{ki} A_{kj} = \langle a_i, a_j \rangle,$$

where a_i is the i^{th} column of A.

Then, $A^TA = I_n$ iff $(A^TA)_{ii} = 1$ & $(A^TA)_{ij} = 0$ $\forall i \neq j$, and so in particular,

$$\langle a_i, a_i \rangle = 1 \quad \& \quad \langle a_i, a_j \rangle = 0 \quad \forall i \neq j.$$

Hence the a_i forms an orthonormal set for $\mathbb{R}^n$, and as $|\{a_1, \dots, a_n\}| = n = \dim \mathbb{R}^n$, this set is also a basis for $\mathbb{R}^n$, as needed. ☒

$A \in M_{n\times n}(\mathbb{C})$ IS UNITARY $\Leftrightarrow$ COLUMNS OF A FORMS AN ORTHONORMAL BASIS FOR $\mathbb{C}^n$ WRT STD INNER PRODUCT (L14.1)

Let $A \in M_{n\times n}(\mathbb{C})$.

Then, A is unitary iff the columns of A form an orthonormal basis for $\mathbb{C}^n$ with respect to its standard inner product.

Proof. See that A is unitary $\Leftrightarrow A^* = A^{-1}$, ie $A^*A = I_n$, or in other words

$$\sum_{k=1}^{n} (A^*)_{ik} A_{kj} = \sum_{k=1}^{n} \overline{A_{ki}} A_{kj},$$

which is exactly $\mathbb{C}^n$'s IP. The rest of the proof is like the real case. ☒

T_A IS DIAGONALIZABLE WRT ORDERED BASIS $B \Rightarrow \exists$ UNITARY (ORTHOGONAL) $P \ni D = P^{-1}[T_A]_B P$; $P = (b_1 \cdots b_n)$

Let $A \in M_{n\times n}(\mathbb{F})$.

Suppose T_A is diagonalizable with respect to some orthonormal ordered basis B, so that $D = [T_A]_B$ is diagonal.

Then necessarily there exists an unitary (orthogonal if $\mathbb{F} = \mathbb{R}$) matrix $P \in M_{n\times n}(\mathbb{F})$ such that

$$D = P^{-1}AP,$$

and if $B = (v_1, \dots, v_n)$, then

$$P = (v_1 \cdots v_n) \in M_{n\times n}(\mathbb{F}).$$

Proof. Let S be the std ord basis of $\mathbb{F}^n$, so that

$$[T_A]_S = A.$$

Then

$$[T_A]_B = ({}_S[I_V]_B)^{-1} [T_A]_S ({}_S[I_V]_B), \quad \text{— } (*)$$

where $V = \mathbb{F}^n$. Then, see that

$${}_S[I_V]_B = ([v_1]_S \cdots [v_n]_S),$$

and as B is orthogonal & $[v_i]_S$ are just the "standard" representations of v_i for each i, by L14.1 ${}_S[I_V]_B$ is unitary (ortho if $\mathbb{F} = \mathbb{R}$). Letting $P = {}_S[I_V]_B$ and subbing back into $(*)$, we see that

$$D = [T_A]_B = P^{-1}[T_A]_S P = P^{-1}AP, \quad \text{(since } S \text{ is the std ord basis of } \mathbb{F}^n)$$

as needed. ☒

ORTHOGONALLY SIMILAR (D14.1)

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

Then, we say A is "orthogonally similar" to B if there exists an orthogonal matrix $P \in M_{n\times n}(\mathbb{R})$ such that

$$B = P^{-1}AP = P^TAP \quad (\text{since } P \text{ is orthogonal}).$$

ORTHOGONALLY DIAGONALIZABLE (D14.1)

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

Then, we say A is "orthogonally diagonalizable" if A is orthogonally similar to some diagonal matrix $D \in M_{n\times n}(\mathbb{R})$.

UNITARILY SIMILAR (D14.1)

Let $A, B \in M_{n\times n}(\mathbb{C})$.

Then, we say A is "unitarily similar" to B if there exists an unitary matrix $P \in M_{n\times n}(\mathbb{C})$ such that

$$B = P^{-1}AP = P^*AP \quad (\text{since } P \text{ is unitary}).$$

UNITARILY DIAGONALIZABLE (D14.1)

Let $A \in M_{n\times n}(\mathbb{C})$.

Then, we say A is "unitarily diagonalizable" if A is unitarily similar to a diagonal matrix $D \in M_{n\times n}(\mathbb{C})$.

$A \in M_{n\times n}(\mathbb{R})$ IS ORTHOGONALLY DIAGONALIZABLE $\Rightarrow A$ IS SYMMETRIC (L14.2)

Let $A \in M_{n\times n}(\mathbb{R})$, and suppose A is orthogonally diagonalizable.

Then necessarily A is symmetric (ie $A^T = A$).

Proof. Since A is ortho diag, $\exists$ diag matrix $D \in M_{n\times n}(\mathbb{R})$ such that A is ortho similar to D; ie $\exists$ ortho $P \in M_{n\times n}(\mathbb{F}) \ni$

$$D = P^{-1}AP.$$

Hence

$$A = PDP^{-1} = PDP^T.$$

Since D is diagonal, it is also symmetric. Taking transposes of both sides yields that

$$A^T = (PDP^T)^T = PD^TP^T = PDP^T = A,$$

showing that $A^T = A$, as needed.

$A \in M_{n\times n}(\mathbb{R})$ IS SYMMETRIC $\Rightarrow$ EVERY EIGENVALUE OF A IS A REAL NUMBER (T14.1)

Let $A \in M_{n\times n}(\mathbb{R})$ be symmetric.
Then necessarily every eigenvalue of A is a real number.

Proof. Let $c \in \mathbb{C}$ be an eigenvalue of A, w/ corresponding eigenvector $x \in \mathbb{C}^n$, so that $Ax = cx$.
See that

$$\begin{aligned} c\langle x,x\rangle &= \langle cx, x\rangle \\ &= \langle Ax, x\rangle \\ &= x^*(Ax) \\ &= x^*(A^*x) \quad \text{(since } A\in M_{n\times n}(\mathbb{R}) \text{ \& } A \text{ is symmetric)} \\ &= (x^*A^*)x \\ &= (Ax)^*x \\ &= \langle x, Ax\rangle \\ &= \langle x, cx\rangle \\ &= \bar{c}\langle x,x\rangle, \end{aligned}$$

So in particular, $c = \bar{c}$. As $x \neq 0$, so $\langle x,x\rangle \neq 0$, so $c \in \mathbb{R}$, as needed. ∎

SELF-ADJOINT [MATRIX IN $\mathbb{C}$] (D14.2)

Let $A \in M_{n\times n}(\mathbb{C})$.
Then, we say A is "self-adjoint" if $A^* = A$.

$A \in M_{n\times n}(\mathbb{C})$ IS SELF-ADJOINT $\Rightarrow$ EVERY EIGENVALUE OF A IS A REAL NUMBER (C14.1)

Let $A \in M_{n\times n}(\mathbb{C})$ be self-adjoint.
Then necessarily every eigenvalue of A is a real number.

Proof. Same as T14.1, since $A^* = A$. ∎

A IS SYMMETRIC, (c_1, v_1), (c_2, v_2) ARE EIGENVALUES/VECTORS OF A $\Rightarrow$ v_1 & v_2 ARE ORTHOGONAL (T14.2)

Let $A \in M_{n\times n}(\mathbb{R})$ be symmetric, and let v_1 & v_2 be eigenvectors corresponding to the eigenvalues c_1 & c_2 of A respectively.
Then necessarily v_1 & v_2 are orthogonal with respect to the standard inner product of $\mathbb{R}^n$.

Proof. See that

$$\begin{aligned} c_1\langle v_1, v_2\rangle &= \langle c_1v_1, v_2\rangle \\ &= \langle Av_1, v_2\rangle \\ &= v_2^T(Av_1) \quad \text{(as all entries in } \mathbb{R}) \\ &= v_2^T(A^Tv_1) \quad (A \text{ is symmetric}) \\ &= (Av_2)^Tv_1 \\ &= \langle v_1, Av_2\rangle \\ &= \langle v_1, c_2v_2\rangle \\ &= c_2\langle v_1, v_2\rangle, \end{aligned}$$

and as $c_1 \neq c_2$, the equality holds iff $\langle v_1, v_2\rangle = 0$, showing the claim in question. ∎

A IS SELF-ADJOINT, (c_1, v_1), (c_2, v_2) ARE EIGENVALUES/VECTORS OF A $\Rightarrow$ v_1 & v_2 ARE ORTHOGONAL (C14.2)

Let $A \in M_{n\times n}(\mathbb{C})$ be self-adjoint, and let v_1 & v_2 be eigenvectors associated to the eigenvalues c_1 & c_2 of A respectively.
Then necessarily v_1 & v_2 are orthogonal with respect to the standard inner product on $\mathbb{C}^n$.

Proof. Almost identical to the proof for T14.2.

Class 15: Orthogonal Diagonalization Continued; Unitary Diagonalization

ADJOINT [OF A LINEAR OPERATOR]: T^*; $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$ (T15.1(1))

Let $T: V \to V$ be linear, where $\dim V < \infty$.
Then, the "adjoint of T", denoted as "T^*", is defined to be the unique linear operator such that

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle \quad \forall v, w \in V.$$

Proof. First, assume T^* exists.
Choose an ord orthonorm basis $B = (v_1, \dots, v_n)$ for V.
Then for each v_i, $\exists$ unique $c_{i1}, \dots, c_{in} \in \mathbb{F}$ $\ni$

$$T^*(v_i) = c_{1i}v_1 + \cdots + c_{ni}v_n \quad \text{for each } i.$$

Let $[T]_B = (b_{ij})$. Then we also know

$$T(v_j) = b_{1j}v_1 + \cdots + b_{nj}v_n \quad \text{for each } j.$$

As $\langle T(v_j), v_i\rangle = \langle v_j, T^*(v_i)\rangle$, thus

$$\langle T(v_j), v_i\rangle = \langle b_{1j}v_1 + \cdots + b_{nj}v_n, v_i\rangle = b_{1j}\langle v_1, v_i\rangle + \cdots + b_{nj}\langle v_n, v_i\rangle = b_{ij};$$

and

$$\langle v_j, T^*(v_i)\rangle = \langle v_j, c_{1i}v_1 + \cdots + c_{ni}v_n\rangle = \overline{c_{1i}}\langle v_j, v_1\rangle + \cdots + \overline{c_{ni}}\langle v_j, v_n\rangle = \overline{c_{ji}}.$$

Hence $b_{ij} = \overline{c_{ij}}$ $\forall i, j$, and so $[T^*]_B$ is the conjugate transpose of $[T]_B$; ie $[T^*]_B = [T]_B^*$.

This proves uniqueness of T^*, once we prove it exists.

Then, we know $\exists$ unique $T^*: V \to V$ $\ni$

$$T^*(v_j) = \sum_{i=1}^{n} \overline{b_{ji}}\, v_i \quad \forall 1 \le j \le n.$$

By the above, $[T^*]_B = [T]_B^*$.

By L12.1,

$$\langle T(v), w\rangle = [w]_B^* [T(v)]_B = [w]_B^* [T]_B [v]_B.$$

&

$$\langle v, T^*(w)\rangle = [T^*(w)]_B^* [v]_B = ([T^*]_B [w]_B)^* [v]_B = [w]_B^* [T]_B [v]_B \quad (= \langle T(v), w\rangle),$$

as needed. ☒

$[T^*]_B = [T]_B^*$ (T15.1(2))

Let $T: V \to V$ be linear, and let B be an ordered orthonormal basis for V.
Then necessarily

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

Proof. This was proved in the proof for T15.1(1).

$(aT + bU)^* = \bar{a}T^* + \bar{b}U^*$ (T15.1(3))

Let $T, U: V \to V$ be linear.
Then necessarily

$$(aT + bU)^* = \bar{a}T^* + \bar{b}U^* \quad \forall a, b \in \mathbb{F}.$$

Proof. Follows by defn of T^*. ☒

$(UT)^* = T^*U^*$; $(T^*)^* = T$ (T15.1(3))

Let $T, U: V \to V$ be linear.
Then necessarily
① $(UT)^* = T^*U^*$; and
② $(T^*)^* = T$.

Proof. Follows by defn of T^*. ☒

V OVER $\mathbb{C}$ $\Rightarrow$ $\exists$ ORDERED ORTHONORMAL BASIS FOR V $\ni$ $[T]_B$ IS UPPER TRIANGULAR <<SCHUR'S THEOREM I>> (T15.2(1))

Let V be finite-dimensional and over $\mathbb{C}$, and let $T: V \to V$ be linear.
Then there exists an ordered orthonormal basis B for V such that $[T]_B$ is upper-triangular.

Proof. Proceed by induction on $n = \dim V$.
Consider $n = 1$. Let $0 \neq v \in V$.
Let $w = \frac{v}{\|v\|}$, so that $\|w\| = 1$, and so $B = (w)$ is an ord orthonorm basis for V.
Since $[T]_B$ is 1×1, it is trivially upper triangular.
Now, assume claim is true for all V with $\dim V = n$, and choose V $\ni$ $\dim V = n+1$ & choose some linear $T: V \to V$.
Consider $C(t)$ for T^*. As $C(t) \in P_{n+1}(\mathbb{C})$, thus by Fund Theorem of Alg $C(t)$ has ≥ 1 complex root c, which is an eigenvalue of T^* by defn.
So, we can choose an associated eigenvector $v_{n+1} \in V$ $\ni$

$$T^*(v_{n+1}) = cv_{n+1}.$$

WLOG, as we can divide v_{n+1} by its norm, we may assume that $\|v_{n+1}\| = 1$.
Let $W = \operatorname{span}\{v_{n+1}\}$. As $n+1 = \dim V = \dim W + \dim W^\perp = 1 + \dim W^\perp$,
so $\dim W^\perp = n$.
Now, we claim $W^\perp$ is "fixed under T"; ie $\forall x \in W^\perp$, $T(x) \in W^\perp$.
Let $x \in W^\perp$. We want to show $T(x) \in W^\perp$, ie

$$\langle av_{n+1}, T(x)\rangle = 0 \quad \forall a \in \mathbb{F}.$$

Then see that

$$
\begin{aligned}
\langle av_{n+1}, T(x)\rangle &= \langle T^*(x), x\rangle && \text{(by defn of } T^*) \\
&= \langle aT^*(v_{n+1}), x\rangle \\
&= \langle acv_{n+1}, x\rangle \\
&= ac\langle v_{n+1}, x\rangle \\
&= ac(0) && \text{(as } v_{n+1} \text{ \& } x \text{ are ortho)} \\
&= 0,
\end{aligned}
$$

as needed.
Hence, if we only apply the linear transformation T to the vectors in $W^\perp$, we obtain a "restriction map" $T|_{W^\perp}: W^\perp \to W^\perp$.
As $\dim W^\perp = n$, by IH $\exists$ ord orthonorm basis $C = (v_1, \dots, v_n)$ $\ni$ $[T|_{W^\perp}]_C$ is upper triangular.
Since v_{n+1} is ortho to any vector in $W^\perp$, thus $B = (v_1, \dots, v_n, v_{n+1})$ is orthogonal, and in fact orthonormal (as $\|v_{n+1}\| = 1$).
So $B = (v_1, \dots, v_{n+1})$ is an ord orthonorm basis for V, and indeed, no matter what $T(v_{n+1})$ is, $[T]_B$ is upper triangle.

$A \in M_{n\times n}(\mathbb{C}) \Rightarrow \exists$ UNITARY MATRIX $P \in M_{n\times n}(\mathbb{C})$, UPPER-TRIANGULAR MATRIX $U \in M_{n\times n}(\mathbb{C}) \ni U = P^{-1}AP$ « SCHUR'S THEOREM II » (T15.2(2))

Let $A \in M_{n\times n}(\mathbb{C})$.

Then there necessarily exists a unitary matrix $P \in M_{n\times n}(\mathbb{C})$ and upper-triangular matrix $U \in M_{n\times n}(\mathbb{C})$ such that

$U = P^{-1}AP$.

Proof. By T15.2(1), $T_A : \mathbb{C}^n \to \mathbb{C}^n$ is upper-triangularizable, ie $\exists$ orthonorm ord basis B for $\mathbb{C}^n \ni [T_A]_B$ is upper triangular.

Then

$$[T_A]_B = ({}_S[I_{\mathbb{C}^n}]_B)^{-1} [T_A]_S \, {}_S[I_{\mathbb{C}^n}]_B,$$

where S is the standard ord basis for $\mathbb{C}^n$.

Now, the columns of ${}_S[I_{\mathbb{C}^n}]_B$ are given by the standard coordinates of the vectors in B, which is an orthonormal basis, so by L14.1 $P = {}_S[I_{\mathbb{C}^n}]_B$ is unitary.

Since $U = [T_A]_B$ is upper-triangular & $[T_A]_S = A$, hence

$U = P^{-1}AP$, as needed. ∎

$A \in M_{n\times n}(\mathbb{R})$, ALL EIGENVALUES OF A ARE REAL $\Rightarrow \exists$ ORTHOGONAL MATRIX $P \in M_{n\times n}(\mathbb{R})$ & UPPER-TRIANGULAR MATRIX $U \in M_{n\times n}(\mathbb{R}) \ni U = P^{-1}AP$ « SCHUR'S THEOREM III » (T15.2(3))

Let $A \in M_{n\times n}(\mathbb{R})$, and suppose all the eigenvalues of A are real.

Then there necessarily exists an orthogonal matrix $P \in M_{n\times n}(\mathbb{R})$ and upper-triangular matrix $U \in M_{n\times n}(\mathbb{R})$ such that

$U = P^{-1}AP$.

Proof. Similar to proof for T15.2(2). ∎

$A \in M_{n\times n}(\mathbb{R})$; A IS ORTHOGONALLY DIAGONALIZABLE $\Leftrightarrow$ A IS SYMMETRIC (T15.3)

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

Then A is symmetric iff A is orthogonally diagonalizable.

Proof. By L14.2,

A is orthogonally diagonalizable $\Rightarrow$ A is symmetric.

So, let A be symmetric! By T14.1, A only has real eigenvalues.

So, by T15.2(3), $\exists$ an upper-triangular matrix U & orthogonal matrix $P \ni$

$$U = P^{-1}AP = P^TAP.$$

Hence

$$U^T = (P^TAP)^T = P^TAP = U.$$

So U is upper triangular and symmetric, and so U is diagonal!

Hence, as $U = P^TAP$, by defn: this tells us that A is orthogonally diagonalizable. ∎

$A \in M_{n\times n}(\mathbb{C})$ IS SELF-ADJOINT $\Rightarrow$ A IS UNITARILY DIAGONALIZABLE (T15.4)

Let $A \in M_{n\times n}(\mathbb{C})$ be self-adjoint (ie $A^* = A$).

Then necessarily A is unitarily diagonalizable.

Proof. By T15.2(2), $\exists$ unitary $P \in M_{n\times n}(\mathbb{C})$ & upper-tri $U \in M_{n\times n}(\mathbb{C})$

$\ni$

$$U = P^{-1}AP = P^*AP.$$

So

$$U^* = (P^*AP)^* = P^*AP = U,$$

and as U is upper triangular & self-adjoint, it is hence diagonal.

Proof follows.

$A \in M_{n\times n}(\mathbb{C})$ IS UNITARY $\Rightarrow$ A IS UNITARILY DIAGONALIZABLE (T15.5)

Let $A \in M_{n\times n}(\mathbb{C})$ be unitary.

Then necessarily A is unitarily diagonalizable

Proof. By T15.2(2), $\exists$ unitary $P \in M_{n\times n}(\mathbb{C})$ & upper-tri $U \in M_{n\times n}(\mathbb{C}) \ni$

$$U = P^{-1}AP = P^*AP.$$

As A is unitary, thus

$$\begin{aligned} U^* = (P^*AP)^* &= P^*A^*(P^*)^* \\ &= P^{-1}A^{-1}P \\ &= (P^{-1}AP)^{-1} \\ &= U^{-1}, \end{aligned}$$

so U is also unitary.

Since U is upper-triangular, thus U^* is lower-triangular.

But as U is upper-triangular & invertible, thus U^{-1} is also upper triangular.

Since $U^* = U^{-1}$, thus U^* & U^{-1} must be diagonal.

So U is diagonal as well, and the proof follows. ∎

NORMAL [MATRICES] (D15.1)

Let $A \in M_{n\times n}(\mathbb{F})$.

Then, we say A is "normal" if

$A^*A = AA^*$.

NORMAL [LINEAR OPERATORS] (D15.1)

Let $T : V \to V$ be linear, where $\dim V < \infty$.

Then, we say T is "normal" if

$TT^* = T^*T$.

$A \in M_{n\times n}(\mathbb{C})$ IS UNITARILY DIAGONALIZABLE $\Rightarrow$ A IS NORMAL (L15.1)

Let $A \in M_{n\times n}(\mathbb{C})$, and suppose A is unitarily diagonalizable.

Then necessarily A is normal.

Proof. Let A be unit diag. so $\exists$ unitary $P \in M_{n\times n}(\mathbb{C})$ & diag $U \in M_{n\times n}(\mathbb{C})$

$\Rightarrow$

$$D = P^{-1}AP,$$

ie

$$A = PDP^{-1} = PDP^*.$$

We know $A^* = (PDP^*)^* = PD^*P^*$.

Note $DD^* = D^*D$, as D, D^* are both diagonal.

So

$$\begin{aligned} AA^* &= (PDP^*)(PD^*P^*) \\ &= PD(P^*P)D^*P^* \\ &= PDD^*P^* \qquad \text{(as } P \text{ is unitary)} \\ &= PD^*DP^* \\ &= (PD^*P^*)(PDP^*) \\ &= A^*A, \end{aligned}$$

so A is normal as needed. ∎

$A \in M_{n\times n}(\mathbb{C})$ IS UNITARILY DIAGONALIZABLE $\Leftrightarrow$ A IS NORMAL (T15.6)

Let $A \in M_{n\times n}(\mathbb{C})$.

Then A is unitarily diagonalizable iff it is normal.

Proof. The above lemma shows $(\Rightarrow)$.

Let A be normal.

By T15.2(2), $\exists$ unitary $P \in M_{n\times n}(\mathbb{C})$ & upper tri $U \in M_{n\times n}(\mathbb{C})$ such that

$$U = P^{-1}AP = P^*AP.$$

Then

$$\begin{aligned} UU^* &= (P^*AP)(P^*AP)^* \\ &= \cdots \\ &= U^*U, \end{aligned}$$

so U is normal.

Thus, we need only show U is diagonal.

As U is upper-triangular, say

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix},$$

see that

$$(U^*U)_{11} = |u_{11}|^2,$$

but

$$(UU^*)_{11} = |u_{11}|^2 + |u_{12}|^2 + \cdots + |u_{1n}|^2.$$

As $U^*U = U^*U$, then $|u_{12}|^2 + \cdots + |u_{1n}|^2 = 0$,

ie $u_{12} = \cdots = u_{1n} = 0$.

Repeating this argument for each successive row yields that all entries off the diagonal are 0, and so U is diagonal. ∎

$A \in M_{n \times n}(\mathbb{C})$ IS NORMAL, (c, v) IS AN EIGENVALUE/VECTOR OF A $\Rightarrow$ $(\bar{c}, v)$ IS AN EIGENVALUE/VECTOR OF A^* (L15.1)

Let $A \in M_{n \times n}(\mathbb{C})$ be normal, and let $v \in \mathbb{C}^n$ be an eigenvector of A associated with the eigenvalue c.
Then necessarily v is also an eigenvector of A^* associated with the eigenvector $\bar{c}$.

Proof. We wish to show $A^* v = \bar{c} v$.
As $Av = cv$, so $(A - cI)v = 0$, so $\|(A-cI)v\|^2 = 0$.
Thus

$$
\begin{aligned}
0 &= \langle (A-cI)v, (A-cI)v \rangle \\
&= \langle v, (A-cI)^*(A-cI)v \rangle \\
&= \langle v, (A^* - \bar{c}I^*)(A-cI)v \rangle \\
&= \langle v, (A^* - \bar{c}I)(A-cI)v \rangle .
\end{aligned}
$$

Then

$$
\begin{aligned}
(A^* - \bar{c}I)(A-cI) &= A^*A - \bar{c}AI - cA^*I + \bar{c}cI^2 \\
&= AA^* - \bar{c}A - cA^* + \bar{c}cI \quad \text{(by normality of } A\text{)} \\
&= (A-cI)(A^* - \bar{c}I) \\
&= (A-cI)(A-cI)^* .
\end{aligned}
$$

So

$$
\begin{aligned}
0 &= \langle v, (A^* - \bar{c}I)(A-cI)v \rangle \\
&= \langle v, (A-cI)(A-cI)^* v \rangle \\
&= \langle (A-cI)^* v, (A-cI)^* v \rangle ,
\end{aligned}
$$

thus $(A-cI)^* v = 0$. In other words, $(A^* - \bar{c}I)v = 0$, and the proof follows. ☑

$T: V \to V$ IS NORMAL, (c, v) IS AN EIGENVALUE/VECTOR OF T $\Rightarrow$ $(\bar{c}, v)$ IS AN EIGENVALUE/VECTOR OF T^* (L15.2)

Let $T: V \to V$ be normal, where T is linear & $\dim V < \infty$.
Let v be an eigenvector of T corresponding to the eigenvalue c.
Then necessarily v is an eigenvector of T^* corresponding to the eigenvalue $\bar{c}$.

Proof. Almost identical to the matrix version. ☑

$A \in M_{n \times n}(\mathbb{C})$ IS NORMAL, (c_1, v_1) & (c_2, v_2) ARE EIGENVALUES/VECTORS OF A, $c_1 \neq c_2$ $\Rightarrow$ v_1 & v_2 ARE ORTHOGONAL (T15.7)

Let $A \in M_{n \times n}(\mathbb{C})$ be normal.
Let $v_1, v_2 \in \mathbb{C}^n$ be eigenvectors of A associated with distinct eigenvalues $c_1, c_2 \in \mathbb{C}$.
Then necessarily v_1 & v_2 are orthogonal.

Proof. We know $(A - c_1 I)v_1 = 0$. So

$$
\begin{aligned}
0 &= \langle (A - c_1 I)v_1, v_2 \rangle \\
&= \langle Av_1, v_2 \rangle - c_1 \langle v_1, v_2 \rangle \\
&= \langle v_1, A^* v_2 \rangle - c_1 \langle v_1, v_2 \rangle \\
&= \langle v_1, \bar{c}_2 v_2 \rangle - c_1 \langle v_1, v_2 \rangle \quad \text{(by L15.2)} \\
&= c_2 \langle v_1, v_2 \rangle - c_1 \langle v_1, v_2 \rangle \\
&= (c_2 - c_1)\langle v_1, v_2 \rangle
\end{aligned}
$$

As $c_2 \neq c_1$, thus $\langle v_1, v_2 \rangle = 0$, as needed. ☑

Class 16: Introduction to Quadratic Forms

QUADRATIC FORMS (D16.1)

A "quadratic form" is a polynomial function $Q : \mathbb{R}^n \to \mathbb{R}$ such that

$$\deg(Q(x)) = 2 \quad \forall x \in \mathbb{R}^n.$$

More explicitly, we have

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,$$

where $a_{ij} \in \mathbb{R} \ \forall 1 \le i, j \le n$ & $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$.

eg $Q(x) = x_1^2 + x_1 x_2 - 2x_2 x_3 + x_3^2$

In particular,

$$Q(x) = x^T A x,$$

where $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

Proof. Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Then

$$\begin{aligned} x^T A x &= (x_1 \cdots x_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= (x_1 \cdots x_n) \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n \end{pmatrix} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \\ &= Q(x). \ \blacksquare \end{aligned}$$

$Q(x)$ IS A QUADRATIC FORM $\Rightarrow \exists$ UNIQUE SYMMETRIC MATRIX $C \in M_{n \times n}(\mathbb{R}) \ni Q(x) = x^T C x$; $C = ((a_{ij} + a_{ji})/2)_{ij}$ (T16.1)

Let $Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$ be a quadratic form.

Then there exists a unique symmetric matrix $C \in M_{n \times n}(\mathbb{R})$ such that $Q(x) = x^T C x$ for each $x \in \mathbb{R}^n$.

Moreover, C is given by

$$c_{ij} = \frac{(a_{ij} + a_{ji})}{2} \quad \forall 1 \le i, j \le n.$$

POSITIVE/NEGATIVE DEFINITE / SEMIDEFINITE; INDEFINITE <<CLASSIFYING QUADRATIC FORMS>> (D16.2)

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form.

Then,

1. We say Q is "positive definite" if $Q(x) \ge 0 \ \forall x \in \mathbb{R}^n$, and $Q(x) = 0 \iff x = 0$;
2. We say Q is "positive semidefinite" if $Q(x) \ge 0 \ \forall x \in \mathbb{R}^n$, but $Q(x) = 0$ for some $x \ne 0$;
3. We say Q is "negative definite" if $Q(x) \le 0 \ \forall x \in \mathbb{R}^n$, and $Q(x) = 0 \iff x = 0$;
4. We say Q is "negative semidefinite" if $Q(x) \le 0 \ \forall x \in \mathbb{R}^n$, but $Q(x) = 0$ for some $x \ne 0$; and
5. We say Q is "indefinite" if $Q(x) > 0$ for some $x \in \mathbb{R}^n$, and $Q(x) < 0$ for some $x \in \mathbb{R}^n$.

DIAGONAL QUADRATIC FORM (D16.3)

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form.

Then, we say Q is "diagonal" if it has the form

$$Q\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1^2 + \cdots + a_n x_n^2$$

for some $a_i \in \mathbb{R}$, $1 \le i \le n$.

Equivalently, Q is of the form

$$Q(x) = x^T D x$$

for some diagonal matrix $D \in M_{n \times n}(\mathbb{R})$.

$Q(x) = x^T D x$; CLASSIFYING Q BASED OFF ENTRIES ON MAIN DIAGONAL (T16.2)

Let $Q(x) = x^T D x$ be a diagonal quadratic form, say $D = \mathrm{diag}(a_1, \ldots, a_n)$.

Then,

1. If $a_i > 0 \ \forall 1 \le i \le n$, then Q is <u>positive definite</u>;
2. If $a_i \ge 0 \ \forall 1 \le i \le n$ & $a_j = 0$ for some $1 \le j \le n$, then Q is <u>positive semidefinite</u>;
3. If $a_i < 0 \ \forall 1 \le i \le n$, then Q is <u>negative definite</u>;
4. If $a_i \le 0 \ \forall 1 \le i \le n$ & $a_j = 0$ for some $1 \le j \le n$, then Q is <u>negative semidefinite</u>; &
5. If $a_i > 0$ & $a_j < 0$ for some $1 \le i, j \le n$, then Q is <u>indefinite</u>.

$Q(x) = x^T C x$; CLASSIFYING Q BASED OFF ITS EIGENVALUES (T16.3)

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form, say $Q(x) = x^T C x \ \forall x \in \mathbb{R}^n$, where C is symmetric.

By T14.1, every eigenvalue of C is real, and by T15.3, since C is symmetric, thus C is orthogonally diagonalizable.

Hence, C has n real eigenvalues, say $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Then,

1. If $\lambda_i > 0 \ \forall 1 \le i \le n$, then Q is <u>positive definite</u>;
2. If $\lambda_i \ge 0 \ \forall 1 \le i \le n$ and $\lambda_j = 0$ for some $1 \le j \le n$, then Q is <u>positive semidefinite</u>;
3. If $\lambda_i > 0 \ \forall 1 \le i \le n$, then Q is <u>negative definite</u>;
4. If $\lambda_i \ge 0 \ \forall 1 \le i \le n$ and $\lambda_j = 0$ for some $1 \le j \le n$, then Q is <u>negative semidefinite</u>; &
5. If $\lambda_i < 0$ & $\lambda_j > 0$ for some $1 \le i, j \le n$, then Q is <u>indefinite</u>.

Class 17:
Optimising Quadratic Forms

$Q(x) = x^T D x$, $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \Rightarrow$
$\max\{Q(x) : \|x\| = 1\} = \max\{\lambda_1, \dots, \lambda_n\}$;
$\min\{Q(x) : \|x\| = 1\} = \min\{\lambda_1, \dots, \lambda_n\}$ (T17.1)

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a diagonal quadratic form,
ie $Q(x) = x^T D x$ for some diagonal matrix
$D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, where $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.
Then necessarily
① $\max\{Q(x) : \|x\| = 1\} = \max\{\lambda_1, \dots, \lambda_n\}$; &
② $\min\{Q(x) : \|x\| = 1\} = \min\{\lambda_1, \dots, \lambda_n\}$.

P IS ORTHOGONAL $\Rightarrow \|Px\| = \|x\|$ (L17.1)

Let $P \in M_{n \times n}(\mathbb{R})$ be orthogonal, and let $x \in \mathbb{R}^n$.
Then necessarily $\|Px\| = \|x\|$.

$Q(x) = x^T C x$, **C IS SYMMETRIC** $\Rightarrow$
$\max\{Q(x) : \|x\| = 1\} = \max\{\lambda_1, \dots, \lambda_n\}$;
$\min\{Q(x) : \|x\| = 1\} = \min\{\lambda_1, \dots, \lambda_n\}$;
$Q(x)$ **MIN** $\Rightarrow$ x **IS AN EV CORRESP TO** $\min\{\lambda_i\}$;
$Q(x)$ **MAX** $\Rightarrow$ x **IS AN EV CORRESP TO** $\max\{\lambda_i\}$
(T17.2)

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form, with $Q(x) = x^T C x \ \forall x \in \mathbb{R}^n$,
where C is symmetric.
By T14.1 & T15.3, C has n real eigenvalues, say
$\lambda_1, \dots, \lambda_n$.
Then, necessarily
① $\max\{Q(x) : \|x\| = 1\} = \max\{\lambda_1, \dots, \lambda_n\}$; &
② $\min\{Q(x) : \|x\| = 1\} = \min\{\lambda_1, \dots, \lambda_n\}$.
Additionally,
$Q(x)$ is min $\Leftrightarrow$ x is a norm-1 eigenvector corresponding to the smallest eigenvalue of C
&
$Q(x)$ is max $\Leftrightarrow$ x is a norm-1 eigenvector corresponding to the largest eigenvalue of C.

$Q(x) = x^T C x$, **C IS SYMMETRIC** $\Rightarrow$
$\max\{Q(x) : \|x\| = r\} = r^2 \max\{\lambda_1, \dots, \lambda_n\}$;
$\min\{Q(x) : \|x\| = r\} = r^2 \min\{\lambda_1, \dots, \lambda_n\}$
(T17.3)

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form, ie $Q(x) = x^T C x \ \forall x \in \mathbb{R}^n$,
where C is a symmetric matrix.
By T14.1 & T15.3, Q has n real eigenvalues, say
$\lambda_1, \dots, \lambda_n$.
Then necessarily
① $\max\{Q(x) : \|x\| = r\} = r^2 \max\{\lambda_1, \dots, \lambda_n\}$; and
② $\min\{Q(x) : \|x\| = r\} = r^2 \min\{\lambda_1, \dots, \lambda_n\}$.