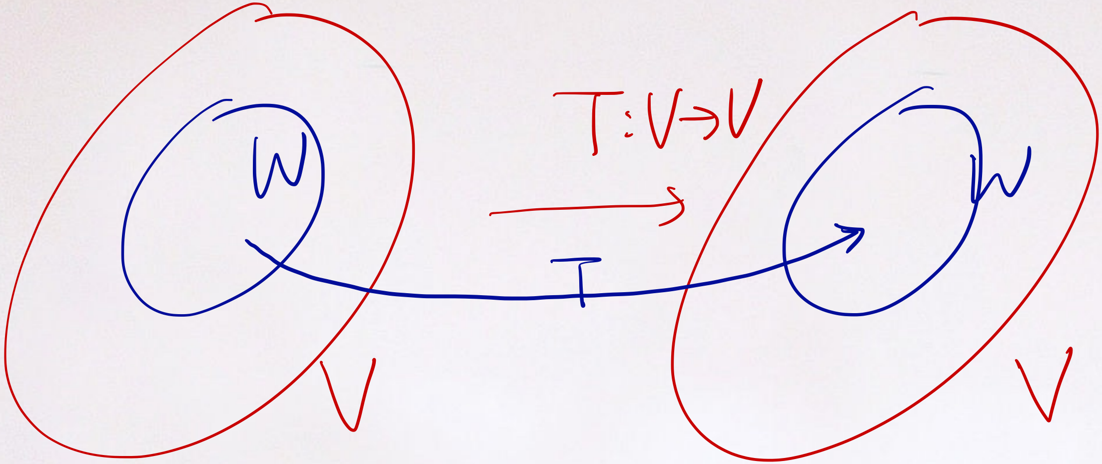
Lecture 14: Recall:

Definition: Let T be a linear operator on a vector space V.

A subspace $W \subset V$ is called T-invariant if $T(W) \subseteq W$.

That is, $T(\vec{w}) \in W$ for $\forall \vec{w} \in W$.



Def: $T|_W : W \to W$ defined by $T|_W(\vec{w}) = T(\vec{w})$

$f_{T|_W}(t)$ divides $f_T(t)$

Remark: Let $T: V \to V$ be a linear operator on a finite-dim vector space V, and let $W \subset V$ be a T-invariant subspace.

Then, the restriction of T to W, denote it by $T|_W : W \to W$, is well-defined and linear.

Proposition: $f_{T|_W}(t)$ divides $f_T(t)$.

Proof: Choose an ordered basis $\gamma = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ for W and extend it to an ordered basis $\beta = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k, \vec{v}_{k+1}, \dots, \vec{v}_n\}$ for V. Then:

$$[T]_\beta = \begin{pmatrix} | & | & \\ \underbrace{[T(\vec{v}_1)]_\beta}_{} \overset{\in W}{} & [T(\vec{v}_2)]_\beta & \cdots \\ | & | & \end{pmatrix} = \begin{array}{c} \overbrace{\qquad\quad}^{k} \qquad\qquad \\ \left(\begin{array}{cc} [T|_W]_\gamma & B \\ O & C \end{array}\right) \end{array} \begin{array}{l} \}\,k \\ \\ \end{array}$$

$$f_T(t) = \det\left[\begin{pmatrix} [T_W]_\gamma & B \\ O & C \end{pmatrix} - t\,I\right]$$

$$= \det\begin{pmatrix} [T_W]_\gamma - t\,I_k & B \\ O & C - t\,I_{n-k} \end{pmatrix}$$

$$= \det\left([T_W]_\gamma - t\,I_k\right)\,\underbrace{\det\left(C - t\,I_{n-k}\right)}_{g(t)}$$

$$= f_{T_W}(t)\; g(t)$$

$\therefore$ $f_{T_W}(t)$ divides $f_T(t)$

Theorem: Let $T: V \to V$ be a linear operator on a finite-dim vector space V and let $W \subset V$ be T-cyclic subspace of V generated by $\vec{v} \neq \vec{0} \in V$. $(W = \text{span}\{\vec{v}, T(\vec{v}), T^2(\vec{v}), \dots\})$

Let $k = \dim(W)$. Then:

(a) $\{\vec{v}, T(\vec{v}), T^2(\vec{v}), \dots, T^{k-1}(\vec{v})\}$ is a basis for W

(b) If $a_0\vec{v} + a_1 T(\vec{v}) + a_2 T^2(\vec{v}) + \dots + a_{k-1}T^{k-1}(\vec{v}) + T^k(\vec{v}) = \vec{0}$, then the characteristic polynomial of $T|_W$ is:

$$f_{T_W}(t) = (-1)^k (a_0 + a_1 t + a_2 t^2 + \dots + a_{k-1} t^{k-1} + t^k)$$

Proof: (a) Since $\vec{v} \neq \vec{0}$, then $\{\vec{v}\}$ is linearly independent.

Let j be the largest +ve integer s.t.

$$\beta = \{\vec{v}, T(\vec{v}), \dots, T^{j-1}(\vec{v})\} \text{ is linearly independent.}$$

Such j exists because V is finite-dim.

Let $Z = \mathrm{span}(\beta)$. $\therefore$ $Z \subset W$

Then, $\underbrace{\beta}_{\text{L.I.}} \cup T^j(\vec{v})$ is linearly dependent. $\therefore$ $T^j(\vec{v}) \in \overset{Z}{\overset{\parallel}{\mathrm{span}(\beta)}}$

$\therefore$ $T^j(\vec{v}) \in Z$.

Now, let $\vec{w} \in Z$. Then $\exists\, b_0, b_1, \dots, b_{j-1} \in F$ s.t.

$$\vec{w} = b_0 \vec{v} + b_1 T(\vec{v}) + \dots + b_{j-1} T^{j-1}(\vec{v})$$

$$T(\vec{w}) = b_0 \underset{\ni Z}{T(\vec{v})} + b_1 \overset{\in Z}{T^2(\vec{v})} + \dots + b_{j-2} \overset{\in Z}{T^{j-1}(\vec{v})} + b_{j-1} \overset{\in Z}{T^j(\vec{v})} \in Z$$

$\therefore$ If $\vec{w} \in Z$, then $T(\vec{w}) \in Z$.

$\therefore$ Z is T-invariant subspace containing $\vec{v}$.

$\therefore$ $\underbrace{W}_{\substack{\shortparallel \\ \text{T-cyclic subspace} \\ \text{containing } \vec{v}}} \subset Z$. ($\because$ W is smallest T-invariant subspace containing $\vec{v}$)

$$\therefore \quad W = Z = \operatorname{span}(\overbrace{\beta}^{\text{L.I.}})$$

$\therefore$ β is a basis of W and $j = k$.

(b) By (a), $\beta = \{\vec{v}, T(\vec{v}), \dots, T^{k-1}(\vec{v})\}$ is an ordered basis for W.

Let $a_0, \dots, a_{k-1} \in F$ s.t.

$$a_0\vec{v} + a_1 T(\vec{v}) + \dots + a_{k-1} T^{k-1}(\vec{v}) + T^k(\vec{v}) = \vec{0}$$

$$\Rightarrow T^k(\vec{v}) = -a_0\vec{v} - a_1 T(\vec{v}) - \dots - a_{k-1} T^{k-1}(\vec{v}).$$

Then:
$$[T]_\beta = \begin{pmatrix} | & | & & | \\ [T(\vec{v})]_\beta & [T(T(\vec{v}))]_\beta & \dots & [T^k(\vec{v})]_\beta \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & 0 & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}$$

$$f_{T_W}(t) = \det\left(\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & \ddots & 0 & \cdots\ 0 & -a_1 \\ & \ddots & \ddots & & \vdots \\ & & & 1\ \ 0 & -a_{k-1} \end{pmatrix} - t I_k\right)$$

$$= (-1)^k \left(a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k\right) \quad (HW)$$

Theorem: (Cayley-Hamilton) Let T be a linear operator on a finite-dim. vector space V and let $f(t) = f_T(t)$ be a char poly of T. Then: $f(T) =$ zero transformation.

(Char poly "kills" the linear operator T)

Remark:

$$f(t) = a_0 1 + a_1 t + a_2 t^2 + \dots + a_n t^n$$

$$f(T) = a_0 I + a_1 T + a_2 T^2 + \dots + a_n T^n$$

Proof: We want to show $f(T)(\vec{v}) = \vec{0}$ for all $\vec{v} \in V$.

$f(T)(\vec{0}) = \vec{0}$ ($\because$ $f(T)$ is linear)

So, suppose $\vec{v} \neq \vec{0}$. Let W = T-cyclic subspace generated by $\vec{v}$.

Let $k = \dim(W)$

By Thm we have shown last time:

$\exists\, a_0, a_1, \dots, a_{k-1} \in F$ such that:

$$\begin{cases} \bullet\ a_0\vec{v} + a_1 T(\vec{v}) + \dots + a_{k-1}T^{k-1}(\vec{v}) + T^k(\vec{v}) = \vec{0} \\ \bullet\ g(t) \overset{def}{=} f_{T|_W} = (-1)^k(a_0 + a_1 t + \dots + a_{k-1}t^{k-1} + t^k) \end{cases}$$

$$\begin{cases} \bullet \; a_0\vec{v} + a_1 T(\vec{v}) + \dots + a_{k-1} T^{k-1}(\vec{v}) + T^k(\vec{v}) = \vec{0} \\ \bullet \; g(t) \overset{\text{def}}{=} f_{T|_W} = (-1)^k (a_0 + a_1 t + \dots + a_{k-1} t^{k-1} + t^k) \end{cases}$$

$$\Downarrow$$

$$g(T)(\vec{v}) = \vec{0}$$

Now, $g(t) \mid f(t)$ implies $\rightsquigarrow$ $\exists\, q(t)$ s.t. $f(t) = q(t)\,g(t)$

$$f(T) = q(T)\,g(T)$$

$$q(T) \circ g(T)$$

$$\therefore \; f(T)(\vec{v}) = q(T) \circ g(T)(\vec{v}) = \vec{0}$$

$$\vec{0}$$

Corollary: Let $A \in M_{n\times n}(F)$ and $f(t)$ be its char. poly. Then: $f(A) = O$, the zero matrix.

Inner product and norm

Assume $F = \mathbb{R}$ or $\mathbb{C}$.

Definition: Let V be a vector space over F. An inner product on V is a map $\langle \cdot , \cdot \rangle : V \times V \to F$ s.t. $\forall \vec{x}, \vec{y}, \vec{z} \in V$ and $c \in F$, it satisfies:

(a) $\langle \vec{x} + \vec{z}, \vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{z}, \vec{y} \rangle$

(b) $\langle c\vec{x}, \vec{y} \rangle = c \langle \vec{x}, \vec{y} \rangle$

(c) $\overline{\langle \vec{x}, \vec{y} \rangle} = \langle \vec{y}, \vec{x} \rangle$

(d) $\underbrace{\langle \vec{x}, \vec{x} \rangle}_{\in \mathbb{R}} > 0$ if $\vec{x} \neq \vec{0}$

Remark:

- (a), (b) say that the inner product is linear in its argument.
- If $F = \mathbb{R}$, $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$

Example: For $\vec{x} = (a_1, a_2, \ldots, a_n)$, $\vec{y} = (b_1, b_2, \ldots, b_n) \in F^n$

($F = \mathbb{R}, \mathbb{C}$)

We have: standard inner product

$$\langle \vec{x}, \vec{y} \rangle :\stackrel{\text{def}}{=} \sum_{i=1}^{n} a_i \overline{b_i}$$

- If $\langle \cdot, \cdot \rangle$ is an inner product on V, and $r > 0$, then: $\langle \vec{x}, \vec{y} \rangle' :\stackrel{\text{def}}{=} r \langle \vec{x}, \vec{y} \rangle$ is another inner product on V.

✓

- Let $V = C([0,1])$ be vector space of real-valued continuous functions on $[0,1]$. Then: for $f, g \in V$,

$$\langle f, g \rangle \overset{\text{def}}{=} \int_0^1 f(t)\,g(t)\,dt$$ defines an inner product on V.

$(F = \mathbb{R}, \mathbb{C})$

- Let $V = M_{n \times n}(F)$. For $A, B \in V$, we define:

sum of diagonal entries.

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

where B^* is the conjugate transpose of B defined by:

$$B^* = \overline{B}^{\,T}$$

For $A, B, C \in V$ and $\lambda \in F$, we check:

(a) $\langle A+B, C\rangle = \mathrm{tr}(C^*(A+B)) = \mathrm{tr}(C^*A + C^*B)$

$= \mathrm{tr}(C^*A) + \mathrm{tr}(C^*B)$

$= \langle A, C\rangle + \langle B, C\rangle$

(b) $\langle \lambda A, B\rangle = \mathrm{tr}(B^*(\lambda A)) = \mathrm{tr}(\lambda(B^*A))$

$= \lambda\, \mathrm{tr}(B^*A)$

$= \lambda \langle A, B\rangle$

(c) $\overline{\langle A, B\rangle} = \overline{\mathrm{tr}(B^*A)} = \mathrm{tr}(\overline{\underbrace{B^*}_{=\overline{B^T}}A}) = \mathrm{tr}(B^T\overline{A})$

$= \mathrm{tr}((B^T\overline{A})^T) = \mathrm{tr}(\overline{A}^T\underbrace{(B^T)^T}_{=B})$

$\mathrm{Tr}(C) = \mathrm{Tr}(C^T)$

$= \mathrm{tr}(A^*B) = \langle B, A\rangle.$

(d) $\langle A, A \rangle = \operatorname{tr}(A^* A) = \sum_{i=1}^{n} (A^* A)_{ii}$

$$= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \underbrace{(A^*)}_{=\bar{A}^T}{}_{ik} \, A_{ki} \right)$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{n} \underbrace{\bar{A}_{ki} A_{ki}}_{= |A_{ki}|^2}$$

$$\langle A, A \rangle = \sum_{i=1}^{n} \sum_{k=1}^{n} |A_{ki}|^2 \geq 0$$

and $\langle A, A \rangle = 0$ iff $A_{ki} = 0 \ \forall \, k, i$ (i.e. $A = 0$)

Definition: A vector space V equipped with an inner product is called an inner product space.

If $F = \mathbb{C}$, we call V a complex inner product space.

If $F = \mathbb{R}$, we call V a real inner product space.

✓

Proposition: Let V be an inner product space. Then, $\forall \vec{x}, \vec{y}, \vec{z} \in V$ and $\forall c \in F$, we have:

(a) $\langle \vec{x}, \vec{y} + \vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$

(b) $\langle \vec{x}, c\vec{y} \rangle = \bar{c} \langle \vec{x}, \vec{y} \rangle$

(c) $\langle \vec{x}, \vec{0} \rangle = \langle \vec{0}, \vec{x} \rangle = 0$

(d) $\langle \vec{x}, \vec{x} \rangle = 0$ iff $\vec{x} = \vec{0}$

(e) If $\langle \vec{x}, \vec{y} \rangle = \langle \vec{x}, \vec{z} \rangle$ for $\forall \vec{x} \in V$, then $\vec{y} = \vec{z}$.

✓

Proof: (a) $\langle \vec{x}, \vec{y}+\vec{z} \rangle = \overline{\langle \vec{y}+\vec{z}, \vec{x} \rangle}$

$$= \overline{\langle \vec{y}, \vec{x} \rangle + \langle \vec{z}, \vec{x} \rangle}$$

$$= \overline{\langle \vec{y}, \vec{x} \rangle} + \overline{\langle \vec{z}, \vec{x} \rangle} = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$$

(b) $\langle \vec{x}, c\vec{y} \rangle = \overline{\langle c\vec{y}, \vec{x} \rangle} = \overline{c \langle \vec{y}, \vec{x} \rangle} = \bar{c}\, \overline{\langle \vec{y}, \vec{x} \rangle}$

$$= \bar{c}\, \langle \vec{x}, \vec{y} \rangle$$

(c) $\langle \vec{x}, \vec{0} \rangle = \langle \vec{x}, \vec{0}+\vec{0} \rangle = \langle \vec{x}, \vec{0} \rangle + \langle \vec{x}, \vec{0} \rangle$

So, $\langle \vec{x}, \vec{0} \rangle = 0$. Similarly, $\langle \vec{0}, \vec{x} \rangle = 0$

(d) If $\vec{x} = \vec{0}$, then $\langle \vec{x}, \vec{x} \rangle = 0$ by (c)

If $\vec{x} \neq \vec{0}$, then $\langle \vec{x}, \vec{x} \rangle > 0$ by definition.

(e) If $\langle \vec{x}, \vec{y} \rangle = \langle \vec{x}, \vec{z} \rangle$ for all $\vec{x} \in V$.

then $\langle \vec{x}, \vec{y} - \vec{z} \rangle = 0 \quad \forall \vec{x} \in V$.

In particular, we can choose $\vec{x} = \vec{y} - \vec{z}$.

Then: $\langle \vec{y} - \vec{z}, \vec{y} - \vec{z} \rangle = 0 \Rightarrow \vec{y} - \vec{z} = \vec{0}$ (by (d))

$\Rightarrow \vec{y} = \vec{z}$.

<u>Remark</u>: (a) + (b) together say that the inner product is <u>conjugate linear</u> in the second argument.

Definition: Let V be an inner product space. For $\vec{x} \in V$, we can define the <u>length</u> or <u>norm</u> of $\vec{x}$ by:

$$\|\vec{x}\| \overset{\text{def}}{:=} \sqrt{\langle \vec{x}, \vec{x} \rangle}$$

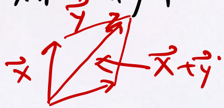
Proposition: Let V be an inner product space over F. Then, $\forall \vec{x}, \vec{y} \in V$ and $\forall c \in F$, we have:

(a) $\|c\vec{x}\| = |c| \cdot \|\vec{x}\|$

(b) $\|\vec{x}\| \geq 0$, and $\|\vec{x}\| = 0$ iff $\vec{x} = \vec{0}$.

(c) $|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \, \|\vec{y}\|$ ((Cauchy-Schwarz Inequality)

(d) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ (Triangle inequality)



Proof: (a) $\|c\vec{x}\| = \sqrt{\langle c\vec{x}, c\vec{x}\rangle} = \sqrt{c\,\underset{|c|^2}{\underset{\shortparallel}{\bar{c}}}\langle \vec{x}, \vec{x}\rangle}$

$= |c|\sqrt{\langle \vec{x}, \vec{x}\rangle} = |c|\,\|\vec{x}\|.$

(b) $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x}\rangle} \geq 0$ (by definition)

$\|\vec{x}\| = 0 \Leftrightarrow \langle \vec{x}, \vec{x}\rangle = 0$ iff $\vec{x} = \vec{0}$

(c) and (d) are shown in the tutorial.

Orthogonality

Definition: Let V be an inner product space. We say $\vec{x}, \vec{y} \in V$ are orthogonal (or perpendicular) if $\langle \vec{x}, \vec{y} \rangle = 0$.

A subset $S \subset V$ is called orthogonal if any two distinct vectors in S are orthogonal.

A unit vector in V is a vector $\vec{x} \in V$ with $\|\vec{x}\| = 1$.

A subset $S \subset V$ is called orthonormal if S is orthogonal and all vectors in S are unit vectors.

✓

e.g. Let H be the space of continuous complex-valued functions on $[0, 2\pi]$. We have inner product defined by:

$$\langle f, g \rangle \overset{\text{def}}{=} \frac{1}{2\pi} \int_0^{2\pi} f(t)\, \overline{g(t)}\, dt \quad \text{for } f, g \in H$$

For any $n \in \mathbb{Z}$ ← integer, let

$$f_n(t) = e^{int} \overset{\text{def}}{:=} \cos nt + i \sin nt \quad \text{for } t \in [0, 2\pi]$$

($i = \sqrt{-1}$)

and consider $S = \{f_n : n \in \mathbb{Z}\} \subset H$

Then, S is orthonormal.