Lectures on Linear Algebra

Lec 26 Eigenvalues & Eigenvectors

> Let T be a linear operator over V ($\dim V < \infty$).

Can we obtain a basis $\mathcal{B}$ of V s.t $[T]_{\mathcal{B}}$ is simple?
↳ diagonal?

> Does there exist a basis $\mathcal{B}$ of V s.t, $[T]_{\mathcal{B}} = \mathrm{diag}(\alpha_{11}, \alpha_{22}, \ldots, \alpha_{nn})$

$$= \begin{pmatrix} \alpha_{11} & & \\ & \alpha_{22} & 0 \\ & 0 & \ddots & \\ & & & \alpha_{nn} \end{pmatrix}$$

In this case T is diagonalizable.

If T is diagonalizable, $Tu^1 = \alpha_{11} u^1 + \alpha_{22} u^2 + \cdots$ where $\mathcal{B} = \{u^1, u^2, \ldots, u^n\}$

$\Rightarrow N(T) = \mathrm{span}\{u^i : \alpha_{ii} = 0\}$ & $R(T) = \mathrm{span}\{u^i : \alpha_{ii} \neq 0\}$

Defn: Eigenvalue

Let $T \in L(V)$, A number $\lambda \in \mathbb{F}$ is an eigenvalue of T if

$\exists$ $x \neq 0$, $x \in V$ s.t $Tx = \lambda x$. ($x \neq 0$)

Here any such x satisfying ↑ is called an eigenvector corresponding to λ.

> $T - \lambda I_V$ is not invertible $\Longleftrightarrow$ $A - \lambda I$ is not invertible. where $A = [T]_{\mathcal{B}}$.

$(A - \lambda I) x = 0$, $x \neq 0$

$\Rightarrow (A - \lambda I) x = 0$, $x \neq 0$

$\Leftrightarrow$ $\det(A - \lambda I) = 0$

$p(\lambda) = \det(A - \lambda I)$ is the characteristic poly of A.

(A monic pol. of deg n.)

If $B = S^{-1} A S$

$\det(B) = \det(A)$.

$\Rightarrow$ $\det(B - \lambda I) = \det(A - \lambda I)$ $\Rightarrow$ eigenvalues are same for any basis.

T is diagonalizable iff there is a basis of V each of whose vectors is an eigenvector for T.

Lec 27

$T \in L(V,V)$ is diagonalizable $\Rightarrow [T]_{\mathcal{B}}$ = diagonal matrix.

Let $\lambda_1; \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues with multiplicities $d_1, d_2, \ldots, d_k$.

$\Rightarrow \; p(\lambda) = \sum_{i=1}^{k} (\lambda-\lambda_i)^{d_i}$ → can only be done if T is diagonalizable.

$$\Rightarrow \quad [T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 I_1 & & 0 \\ & \lambda_2 I_2 & \\ & & \ddots & \\ 0 & & & \lambda_k I_k \end{bmatrix} \quad \text{where } I_i = I_{d_i \times d_i}$$

Also $d_1 + d_2 + \ldots + d_k = n$.

$\underbrace{N(T-\lambda_i I)}$ $= d_i$ since in $[T]_{\mathcal{B}}$, $\lambda_i I_i \to 0$ & gives the nullity.

The set of all eigenvectors of T corresponding to λ_i → Subspace called eigenspace.

⇒ The multiplicity of λ_i must be equal to the dim of it's eigenspace for T to be diagonalizable.

Lemma: Let $T \in L(V,V)$, f be a polynomial over $\mathbb{F}$.
If $Tx = \lambda x$ for $\lambda \in \mathbb{F}$, then $f(T)x = f(\lambda)x$.

Proof: $f(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_s t^s$

$$f(T) = a_0 I + a_1 T + \dots + a_s T^s$$

$$f(T)x = a_0 Ix + a_1 Tx + \dots + a_s T^s x$$

where $T^2 x = T(Tx) = T(\lambda x) = \lambda^2 x$.

By induction $T^r x = \lambda^r x$

$$\Rightarrow f(T)x = a_0 x + a_1 \lambda x + \dots + a_s \lambda^s x$$

$$= (a_0 + a_1 \lambda + \dots + a_s \lambda^s) x$$

$$= f(\lambda) x.$$

Lemma: Let $T \in L(V,V)$, have $\lambda_1 \dots \lambda_k$ as the distinct eigenvalues. Let W_i be the eigenspace due to λ_i. Let

$$W = W_1 + W_2 + \dots + W_k$$

Then $\dim(W) = \dim(W_1) + \dim(W_2) + \dots + \dim(W_k)$.

"This means that eigenvectors corresponding to distinct eigenvalues are linearly independant".

Proof:

Let $\mathcal{B}_1 = \{u^{11}, u^{12}, \dots u^{1l_1}\} \longrightarrow \lambda_1$

$\mathcal{B}_2 = \{u^{21}, u^{21}, \dots, u^{2l_2}\} \longrightarrow \lambda_2$

$\vdots$

$\mathcal{B}_K = \{u^{K1}, u^{K2}, \dots, u^{Kl_K}\} \longmapsto \lambda_K$

be basis for $W_1, W_2, \dots W_K$.

Claim: $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_K\}$ is basis for W.

Clearly this spans W.

Showing LI: Consider $\alpha_{11} u^{11} + \alpha_{12} u^{12} + \dots + \alpha_{1l_1} u^{1l_1}$

$\vdots$

$\alpha_{K1} u^{K1} + \dots + \alpha_{Kl_K} u^{Kl_K} = 0$

$\Rightarrow$ all scalars are 0.

For any polynomial $f(t)$,

$$0 = f(T)(\alpha_{11} u^{11} + \dots + \alpha_{1l_1} u^{1l_1} + \dots + \alpha_{Kl_K} u^{Kl_K})$$

$\Rightarrow \alpha_{11} f(T) u^{11} = \alpha_{11} f(\lambda_1) u^{11}$

$$\Rightarrow \quad 0 = \alpha_{11} f(\lambda_1) u^{11} + \dots + \alpha_{1l_1} f(\lambda_1) u^{1l_1}$$
$$+ \alpha_{21} f(\lambda_2) u^{21} + \dots + \alpha_{2l_2} f(\lambda_2) u^{2l_2}$$
$$+ \alpha_{K1} f(\lambda_K) u^{K1} + \dots + \alpha_{Kl_K} f(\lambda_K) u^{Kl_K}$$ for any f.

Choose $f(t) = (t-\lambda_2)(t-\lambda_3) \dots (t-\lambda_k)$.

$f(\lambda_1) \neq 0$, but $f(\lambda_i) = 0 \;\; \forall \; i \geq 2$

$\Rightarrow \quad 0 = f(\lambda_1) \left[\alpha_{11} u^{11} + \dots + \alpha_{1 l_1} u^{1 l_1} \right]$

$u^{11}, \dots, u^{1 l_1}$ form a basis $\Rightarrow$ $\alpha_{11}, \dots, \alpha_{1 l_1}$ are all 0.

Do the same to show other coeffs are also 0. //

<u>Thm:</u> $T \in L(V, V)$, have $\lambda_1, \lambda_2, \dots, \lambda_k$ as distinct eigenvalues. Let W_i be the eigenspace corresponding to λ_i. Then the following are equivalent.

a) T is diagonalizable.

b) $p(\lambda) = (\lambda-\lambda_1)^{d_1} (\lambda-\lambda_2)^{d_2} \dots (\lambda-\lambda_k)^{d_k}$

Where $d_1 + d_2 + \dots + d_k = n$.

c) $V = W_1 + W_2 + \dots + W_k$.

<u>Proof</u>: a) $\Rightarrow$ b) done.

b) $\Rightarrow$ c). Set $W = W_1 + \dots + W_k$.

Then $\dim W = \sum_{i=1}^{k} \dim W_i = \sum_{i=1}^{k} d_i = n$. So $W = V$.

c) $\Rightarrow$ a) follows from definition. Since each W_i is an eigenspace with an eigenbasis (of eigenvectors).

The minimal polynomial

The set of polynomials $f(t)$ such that $f(T) = 0$. is $\mathcal{A}$.

$\mathcal{A} = \mathbb{F}[t]$, $t \in \mathbb{F}$. $\mathcal{A}$ forms an Algebra. (i.e multiplication of vectors is defined). Hence we have an ideal. where this ideal is generated by a unique element (polynomial) called the minimal polynomial.

Lec 28

$f(t)$ such that $f(T) = 0$. → annihilating polynomial. Let $T \in L(V,V)$. $\dim(V) = n$.

$\dim(L(V,V)) = n^2$.

Consider the operators, $1_V, T, T^2, \cdots, T^{n^2}$ in $L(V,V)$.

$\exists$ scalars not all zero s.t $\alpha_0 I + \alpha_1 T + \cdots + \alpha_{n^2} T^{n^2} = 0$.

Define $f(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n^2} t^{n^2}$.

Then $f(T) = 0$.

$M_T = \{ f \in \mathbb{R}[t] : f(T) = 0 \} \neq \phi$

$\exists$ $m \in \mathbb{R}[t]$ s.t $M_T = (m)$ → generated by m.

m satisfies:

i) $m(T) = 0$

ii) m is monic.

iii) if $f \in \mathbb{R}[T]$ & $f(T) = 0$, then $\deg(m) \leq \deg(f)$.

Such an m is called <u>the</u> minimal polynomial for T.

<u>Lemma:</u> The zeros of m & p are the same.

ie $m(\lambda) = 0 \Leftrightarrow p(\lambda) = 0$.

<u>Proof:</u> Let λ be such that $Tx = \lambda x$ for some $x \neq 0$.

We show that $m(\lambda) = 0$.

$0 = m(T)x = m(\lambda)x$ from earlier Lemma.

So $m(\lambda) = 0$.

Conversely, $m(\lambda) = 0$

$\Rightarrow m(t) = (t-\lambda)q(t)$., where $q(T) \neq 0$. ($\because \deg(q) < \deg(m$

$\Rightarrow \exists\ x \neq 0$ s.t $q(T)x \neq 0$. Let $y = q(T)x \neq 0$.

$$\begin{aligned} 0 &= m(T)x. \\ &= (T-\lambda I)q(T)x \\ &= (T-\lambda I)y. \end{aligned}$$

ie. $Ty = \lambda y$, $y \neq 0$ ie. $p(\lambda) = 0$. ∎

If x_1, x_2, x_3 form eigenbasis for T or A. Then $P = [x_1\ x_2\ x_3]$

& $A = PDP^{-1}$ or $D = P^{-1}AP$ where $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_3 \end{bmatrix}$

Lec 29 **Cayley Hamilton Theorem.**

Before that,

If $T \in L(V,V)$. T is diagonalizable & $\lambda_1, \lambda_2, \dots, \lambda_k$ be the distinct eigenvalues of T. Then

$$m(t) = (t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_k)$$

We only need to show $m(T) = 0$.

$\exists$ $\mathcal{B}$ basis of V s.t each $x \in \mathcal{B}$ is an eigenvector of T (for some eigenvalue of T).

Let $x \in \mathcal{B}$. $m(T) = (T-\lambda_1 I)(T-\lambda_2 I)\dots(T-\lambda_k I)x$.

$$\begin{aligned}(T-\lambda_j I)(T-\lambda_s I) &= T^2 - \lambda_s T - \lambda_j T + \lambda_j \lambda_s I \\ &= T^2 - \lambda_j T - \lambda_s T + \lambda_s \lambda_j I \\ &= (T-\lambda_s I)(T-\lambda_j I).\end{aligned}$$

Let $Tx = \lambda_i x$ for some i.

I.e. $(T-\lambda_i\ )x = 0$

So $m(T)x = (T-\lambda_1 I)(T-\lambda_2 I)\dots T(\lambda_{i-1}I)(T-\lambda_{i+1}I)\dots$
$\dots(T-\lambda_k I)\underbrace{(T-\lambda_i I)x}_{0}$

$= 0$

Note: $A = \begin{bmatrix} 1+t^2 & 2t-t^3 \\ 1 & 1-t^3 \end{bmatrix}$, $t \in \mathbb{R}$.

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} 1 + \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} t + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} t^2 + \begin{bmatrix} 0 & -1 \\ 0 & -1 \end{bmatrix} t^3$$

Cayley Hamilton Theorem

Let $T \in L(V,V)$ & V fdvs. Let p be the char. pol. for T. Then $p(T) = 0$. i.e. the minimal polynomial divides the characteristic polynomial.

Proof (For matrix case) Let $A = [T]_{\mathcal{B}}$.

We use the following $B \cdot adj(B) = \det(B) I_{n \times n}$.

Let $|A - \lambda I| = \alpha_0 \lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_{n-1}\lambda + \alpha_n = p(\lambda)$

We must show $\alpha_0 A^n + \dots + \alpha_{n-1} A + \alpha_n I = 0$.

We have,

$$(A - \lambda I)\, adj(A - \lambda I) = \det(A - \lambda I) \cdot I.$$

RHS : $\alpha_0 \lambda^n I + \alpha_1 \lambda^{n-1} I + \dots + \alpha_{n-1}\lambda I + \alpha_n I$ — (1)

LHS : Also $adj(A - \lambda I) = B_1 \lambda^{n-1} + B_2 \lambda^{n-2} + \dots + B_{n-1}\lambda + B_n$

where $B_1, B_2, \dots, B_n$ are matrices (whose entries depend on A).

LHS : $(A - \lambda I)(B_1 \lambda^{n-1} + B_2 \lambda^{n-2} + \dots + B_{n-1}\lambda + B_n)$

$$= (-B_1)\lambda^n + (AB_1 - B_2)\lambda^{n-1} + (AB_2 - B_3)\lambda^{n-2} + \dots$$

$$+ (AB_{n-1} - B_n)\lambda + AB_n \quad \text{— (2)}$$

$\Rightarrow \underbrace{-B_1 = \alpha_0 I}_{\times A^n}$; $\underbrace{AB_1 - B_2 = \alpha_1 I}_{\times A^{n-1}}$; $\dots$; $AB_{n-1} - B_n = \alpha_{n-1} I$; $\underbrace{AB_n = \alpha_n I}_{\times I}$.

$$\Rightarrow -A^n B_1 + A^n B_1 - A^{n-1} B_2 \cdots + A^2 B_{n-1} - AB_n + AB_n$$

$$= (\alpha_0 A^n + \alpha_1 A^{n-1} + \ldots + \alpha_n) I$$

LHS $= 0$

RHS $= p(A)$.

$\Rightarrow \quad \underline{p(A) = 0}$ ■

<u>Lec 30</u> <u>Invariant Subspaces</u>.

<u>Def</u>: $T \in L(V,V)$. A subspace W is called an invariant subspace if $T(W) \subseteq W$. i.e. $\forall x \in W$, $Tx \in W$.

<u>Eg</u>: $W = \{0\}$ & $W = V$ are T invariant.

Also $R(T)$ & $N(T)$ are T invariant.

Let $S \in L(V,V)$ s.t. $ST = TS$

Then $R(S)$ & $N(S)$ are invariant under T.

> Look at $T - \lambda I$ (for any eigenvalue λ.) commutes with T.

$\Rightarrow N(T - \lambda I)$ is invariant under T

$\Rightarrow$ <u>Eigenspaces are invariant under T</u>.

<u>Def</u>: Let W be an invariant subspace of T. Let $y \in V$ & $y \notin W$. The "T conductor of y into W" is defined by

$$S_T(y; W) = \{g \in F[t] : g(T)y \in W\}$$

Sends y into W

The unique monic generator of $S_T(y;W)$ is also called the "T conductor of y into W".

Let us denote the T conductor by g.

$$S_T(y;W) = (g)$$

Since $m \in S_T(y;W) \Rightarrow g/m$, g divides m.

$\Rightarrow$ In general g cannot be an annihilator.

<u>Def</u>: An operator $T \in L(V,V)$ is said to be triangulable if $\exists$ a basis $\mathcal{B}$ of V s.t $[T]_{\mathcal{B}}$ is triangular. $[T]_{\mathcal{B}} = \begin{pmatrix} a_{11} & & \neq 0 \\ & a_{22} & \\ 0 & & \ddots \\ & & & a_{nn} \end{pmatrix}$ (upper)

<u>Lemma</u>: Let $T \in L(V,V)$ and W be a proper invariant subspace of T. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T & let $m(t) = (t-\lambda_1)^{r_1}(t-\lambda_2)^{r_2} \ldots (t-\lambda_k)^{r_k}$

Then

i) $\exists\, x \notin W$ and

ii) $(T-\lambda I)x \in W$ for some eigenvalue λ.

<u>Proof</u>: $W \neq V \Rightarrow \exists\, y \in V$ s.t $y \notin W$. Let g be the T conductor of y into W. $g(T)y \in W \Rightarrow g$ is not a constant!

So $g(t) = (t-\lambda_i)h(t)$. Set $x = h(T)y \notin W$. Also

$$(T-\lambda_j I)x = (T-\lambda_j I)h(T)y = g(T)y \in W_{//}.$$

To summarize, $\exists\, x \notin W$ s.t the T conductor of x into W is a linear poly.

<u>Thm:</u> Let $T \in L(V,V)$ and $\lambda_1 \dots \lambda_k$ be the distinct eigenvalues of T. Then T is triangulable iff.

$$m(t) = (t-\lambda_1)^{s_1}(t-\lambda_2)^{s_2}\dots(t-\lambda_k)^{s_k}.$$

<u>Proof</u>: $\Rightarrow$

$\exists$ a basis $\mathcal{B}$ of V s.t $[T]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & \\ 0 & 0 & \ddots & \\ \vdots & & & a_{nn} \end{pmatrix}$

eigenvalues are the diagonal entries

$$\Rightarrow \quad p(t) = (t-a_{11})(t-a_{22})\dots(t-a_{nn})$$

$\Rightarrow$ $m(t)$ has the required form.

$\Leftarrow$ Suppose $m(t) = (t-\lambda_1)^{r_1}(t-\lambda_2)^{r_2}\dots(t-\lambda_k)^{r_k}$

$W_1 = \{0\}$ $\quad \exists\, x^1 \notin W_1 \Rightarrow x^1 \neq 0$ s.t $(T-\lambda_j I)x^1 \in W_1$

where λ_j is some eigenvalue.

$W_2 = \text{span}\{x^1\}$ $\exists\, x^2 \notin W_2$ (i.e. x^1, x^2 LI) s.t

$(T-\lambda_i I)x^2 \in W_2$

$$\Rightarrow \quad Tx^2 = \lambda_i x^2 + \underbrace{\alpha_{11} x^1}_{W_2}$$

$$\Rightarrow \quad Tx^2 = \alpha_{11}x^1 + \lambda_i x^2$$

$\vdots$

$$W_n = \text{span}\{x^1, x^2, \dots, x^{n-1}\} \quad \exists\, x^n \notin W_n \ \text{s.t}$$

$$(T - \lambda_s I)\, x^n \in W_n.$$

$$\Rightarrow T x^n = \lambda_s x^n + \overset{\in W_n}{u}$$

$$= \alpha_{n1} x^1 + \alpha_{n2} x^2 + \dots + \alpha_{nn-1} x^{n-1} + \lambda_s x^n.$$

$$\mathcal{B} = \{x^1, x^2, \dots, x^n\}$$

$$\Rightarrow [T]_{\mathcal{B}} = \begin{pmatrix} \lambda_j & \alpha_{11} & & \alpha_{n1} \\ 0 & \lambda_i & & \vdots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \vdots & & \\ & 0 & 0 & \lambda_s \end{pmatrix} \quad \blacksquare$$

(If all eigenvalues lie in the underlying field, it is triangulable).

> Over an algebraically closed field any operator is Triangulable!

> Algebraically closed $\Rightarrow$ any polynomial can be factored into irreducible polynomials that <u>are linear</u>.

Thm: $T \in L(V,V)$ is diagonalizable iff

$$m(t) = (t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_k), \text{ where}$$

$\lambda_1, \dots, \lambda_k$ are distinct eigenvalues of T.

Proof: $\Rightarrow$ done earlier.

$\Leftarrow$ Suppose $m(t)$ = product of distinct linear polynomials.

Set W to be the space generated by all the eigenvectors of T.

If $W=V$, then T is diagonalizable. Suppose that $W \neq V$.

$\exists\, x \notin W$ but $(T-\lambda I)x \in W$ for some eigenvalue λ.

Set $y = (T-\lambda I)x$.

Set $y = y^1 + y^2 + \dots + y^k$ where $Ty^i = \lambda_i y^i$, y^i comes from W_i which is the eigenspace of λ_i

(Recall: $Tx = \lambda x \Rightarrow g(T)x = g(\lambda)x. \ \forall g$)

For any $h(t)$,
$$h(T)y = h(T)y^1 + h(T)y^2 + \dots + h(T)y^k$$
$$= h(\lambda_1)y^1 + h(\lambda_2)y^2 + \dots + h(\lambda_k)y^k.$$

We have $0 = m(T)x$. $y = (T-\lambda_j I)x$

$m(t) = (t-\lambda_j)\,q(t). \Rightarrow q(\lambda_j) \neq 0$

Consider $h(t) = q(t) - q(\lambda_j)$ Then $h(\lambda_j) = 0 \Rightarrow h(t) = (t-\lambda_j)\,f(t)$

So $q(t) - q(\lambda_j) = (t-\lambda_j)\,f(t)$.

Consider

$$q(T)x - q(\lambda_j)x = f(T)(T-\lambda_j I)x = f(T)\,y \in W.$$

$\Rightarrow \quad q(T)x - q(\lambda_j)x \in W.$

If $z = q(T)x = 0$ then $q(\lambda_j) = 0$ $\divideontimes$

So $z \neq 0$

$$0 = m(T)x = (T-\lambda_j I)\,q(T)x = (T-\lambda_j I)z$$

$\Rightarrow \quad Tz = \lambda_j z, \quad z \neq 0.$, z is an eigenvector.

$\Rightarrow \quad z \in W \Rightarrow q(\lambda_j) = 0$ $\divideontimes$

$\Rightarrow \quad \underline{\underline{W = V}}$ $\blacksquare$

Lec 32 Independant Subspaces & Projection operators.

Def: Let $W_1, W_2, \ldots, W_k$ be subspaces of a vector space V.
Then $W_1, \ldots, W_k$ are called independant if

$$w^1 + w^2 + \ldots + w^k = 0 \quad \text{with } w^i \in W_i$$

holds only for $w^1 = w^2 = \ldots = w^k = 0$.

Eg: If $W_1, W_2 \ldots W_k$ be eigenspaces of distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then they are independant.

Note: Suppose $W_1, W_2, \ldots, W_k$ are independant.
and $W = W_1 + W_2 + \ldots + W_k$. Then any $x \in W$ has a unique representation $x = w^1 + w^2 + w^3 + \ldots + w^k$, $w^i \in W_i$

If $x = z^1 + z^2 + \ldots + z^k$; $z^k \in W_i$

$\Rightarrow (w^1 - z^1) + \ldots + (w^k - z^k) = 0.$

$\Rightarrow y^1 + y^2 + \ldots + y^k = 0$, $y^i \in W_i$

$\Rightarrow w^1 = z^1, \ldots w^k = z^k$!

Thm: Let $W_1, \ldots, W_k$ be subspaces of V. Then the following statements are equivalent.

a) $W_1, \ldots, W_k$ are independant

b) $\forall j,\ 2 \leq j \leq k$; $W_j \cap (W_1 + W_2 + \ldots + W_{j-1}) = \{0\}$.

c) If $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ are ordered basis for $W_1, W_2, \ldots, W_k$ then

$$\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k\} \text{ is a basis for } W = W_1 + W_2 + \ldots + W_k.$$

<u>Proof</u>: We only show that (a) $\Rightarrow$ (b).

Take $x \in W_j \cap (W_1 + W_2 + \ldots + W_{j-1})$

$$x = x^1 + x^2 + \ldots + x^{j-1}, \quad x^i \in W_i \quad 1 \le i \le j-1$$

$$\Rightarrow \quad x^1 + x^2 + \ldots + x^{j-1} + \underbrace{(-x)}_{\in W_j} = 0 \quad \text{\& use a)}$$

$$\Rightarrow \quad x = 0$$

<u>Eg</u>: $V = \mathbb{R}^{n \times n}$. Define $W_1 = \{A \in V : A = A^t\}$ symmetric
$W_2 = \{A \in V : A = -A^t\}$ skew symmetric.

$A \in V$: $A = A_1 + A_2$ where $A_1 = \frac{A + A^t}{2} \in W_1$, $A_2 = \frac{A - A^t}{2} \in W_2$

$W_1 \cap W_2 = \{0\}$.

<u>Def</u>: If $W_1, W_2, \ldots, W_k$ are independent subspaces, & if $W = W_1 + \ldots + W_k$ then we say that W is the <u>direct sum</u> of $W_1, W_2, \ldots, W_k$. This is denoted by $W = W_1 \oplus W_2 \oplus \ldots \oplus W_k$.

Def: $E \in L(V,V)$ is called a projection if

$E^2 = E$. (Idempotence).

Properties

1. If $x \in R(E) \iff Ex = x$.
2. $N(E) = R(I-E)$
3. $V = R(E) \oplus N(E)$.

Proof: 1) $x \in R(E) \Rightarrow x = Ey \Rightarrow Ex = Ey = x \Rightarrow Ex = x$.

$x = Ex \Rightarrow x \in R(E)$

2) $x \in N(E) \Rightarrow Ex = 0 \Rightarrow x - Ex = x$

$\Rightarrow (I-E)x = x$

$\Rightarrow x \in R(I-E)$.

$E^2 = E \iff (I-E)^2 = I-E$ } gives reverse implications.

3) $x \in V$: $x = Ex + x - Ex$

$$= \underbrace{Ex}_{R(E)} + \underbrace{(I-E)x}_{N(E)}$$

Take $u \in R(E) \cap N(E)$

$\Rightarrow u = Eu = 0$

$\Rightarrow V = R(E) \oplus N(E)$

Property 4 If E is a projection it is diagonalizable.

Let λ be an eigenvalue of E. Then $Ex = \lambda x$, $x \neq 0$

$$E^2 x = Ex = \lambda x = \lambda Ex = \lambda^2 x,$$

$$\Rightarrow \lambda = \lambda^2$$

$$\Rightarrow \lambda(\lambda - 1) = 0 \Rightarrow \lambda = 0, 1.$$

$$V = \underbrace{R(E)}_{} \oplus N(E)$$

$\mathcal{B}_1 = \{u^1, u^2, \ldots, u^r\}$

each vector is eigenvector with $\lambda = 1$

$$Eu^i = u^i, \quad 1 \le i \le r$$

$\mathcal{B}_2 = \{u^{r+1}, u^{r+2}, \ldots, u^n\}$

each is eigenvector with $\lambda = 0$.

$$Eu^i = 0 \quad r+1 \le i \le n.$$

Let $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Define E_i on V:

$$x \in V, \quad x = x^1 + x^2 + \cdots + x^{i-1} + x^i + x^{i+1} + \cdots + x^k$$

$$E_i(x) = x^i$$

$$E_i^2(x) = E_i(E_i(x)) = E_i(x^i) = x^i \quad \text{since } x^i = 0 + 0 + \cdots + x^i + 0 + \cdots + 0.$$
$$= E_i(x).$$

$$E_i E_j(x) = E_i(x_j) = 0 \quad (\text{assuming } i \neq j)$$

Lec 34 Direct sum decomposition & Projection Operators II.

Thm: Let V be fdvs. Suppose that

$V = W_1 \oplus W_2 \oplus \cdots \oplus W_K$ where $W_1, W_2, \ldots W_K$ are subspaces of V. Then.

$\exists$ k linear operators $E_1, E_2 \ldots E_k$ on V

Such that

i) $I = E_1 + E_2 + \cdots + E_k$

ii) $E_i^2 = E_i$

iii) $E_i E_j = 0$, if $i \neq j$

iv) $R(E_i) = W_i$.

Conversely, suppose that there exist k nonzero operators $E_1, E_2, \ldots E_k$ such that conditions (i)–(iii) hold. If $W_i = R(E_i)$, then

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

Proof: Assume $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Let $x \in V$, then $\exists$ unique $x^i \in W_i$ such that $x = x^1 + x^2 + \cdots + x^k$.

For each E_i, $E_i(x) = x^i$ (definition). it is unique & hence welldefined.

$$
\begin{aligned}
x &= x^1 + x^2 + \cdots + x^k \\
&= E_1(x) + E_2(x) + \cdots + E_k(x) \\
&= (E_1 + E_2 + \cdots + E_k)\, x = I x \quad \forall x \quad \text{i)} \checkmark
\end{aligned}
$$

$$E_i^2(x) = E_i(x_i) = E_i(0+0+\dots+0+\overset{i\text{th term}}{x_i}+0+\dots+0\} = x_i = E_i(x)$$

ii) ✓

$$E_i(E_j(x)) = E_i(x_j) = \overset{i^{th}\text{ term.}}{0} \quad \text{iii) ✓}$$

$E_i(x) = x_i \in W_i$ by definition

Also if $x \in R(E_i)$ then $x = E_i(x) = x^i \in W_i$

$\Rightarrow R(E_i) \subseteq W_i$

{Any idempotent operator acts as identity on its' range. (i.e. $T(T(x)) = T(x)$)}

If $x \in W_i$ then $x = 0+0+\dots+\overset{i\text{th term}}{x_i}+\dots+0$

$\Rightarrow E_i(x) = x \Rightarrow x \in R(E_i) \Rightarrow W_i \subseteq R(E_i)$

iv) ✓

Converse,

Recall, $I = E_1 + E_2 + \dots + E_k$.

Let $x \in V \Rightarrow x = Ix = \underset{W_1}{E_1 x} + \underset{W_2}{E_2 x} + \dots + \underset{W_k}{E_k x}$ } since $W_i = R(E_i)$

{Showing that $W_1 \dots W_k$ are unique is the same as showing that this representation is unique (by defn).}

Suppose that $x = y^1 + y^2 + \dots + y^k$ where $y^i \in W_i$

$$E_1 x = E_1(y^1 + y^2 + \dots + y^k) = E_1 y^1 + E_1 y^2 + \dots + E_1 y^k$$

$$= E_1 y^1 + \cancel{E_1 E_2 y^2}^{\,0} + \quad + \cancel{E_1 E_k y^k}^{\,0}$$

since $y^2 \in R(E_2)$

$$= E_1 y^1 = y^1$$

$\Rightarrow \quad E_1(x) = y^1$

Similarly $E_i(x) = y^i \quad \forall i$ so the representation is unique.

$\Rightarrow \quad V = W_1 \oplus W_2 \oplus \ldots \oplus W_k. \quad \blacksquare$

Thm: $T \in L(V,V)$, let $E_1, E_2, \ldots E_k$; $W_1, \ldots, W_k$ be as above. Then $TE_i = E_iT \iff T(W_i) \subseteq W_i$.

Proof: $\Rightarrow$ Suppose that $TE_i = E_iT$

Take $y \in T(W_i)$, let $y = Tx$ for some $x \in W_i$.

$\Rightarrow y = TE_ix = E_iTx \in R(E_i) = W_i$

$\Rightarrow y \in W_i$

$\Leftarrow$ Assume W_i invariant under $T \Rightarrow T(W_i) \subseteq W_i$

Let $x \in V$; $x = E_1x + E_2x + \ldots + E_kx$

$$\left.\begin{aligned} Tx &= TE_1x + TE_2x + \ldots + TE_kx \\ &= E_1y^1 + E_2y^2 + \ldots + E_ky^k \end{aligned}\right\}$$

Since $T(E_1x) \in W_1$
Let $T(E_1x) = y_1$ where $y_1 \in W_1$
& $y_1 = E_1y_1$

$$= \sum_{j=1}^{k} E_jy^j$$

So $E_i T x = \sum_{j=1}^{k} E_i E_j y^j = E_i y^i = T E_i x \quad \forall x.$

↳ from earlier expansion

$$\Rightarrow \quad E_i T = T E_i \quad \forall i$$

$\blacksquare$

<u>Thm</u>: Let V fdvs & $T \in L(V,V)$. Let $\lambda_1 \ldots \lambda_k$ be the distinct eigenvalues of T. <u>If T is diagonalizable</u> then $\exists$ k linear operators $E_1, E_2 \ldots E_k$ on V such that

i) $T = \lambda_1 E_1 + \lambda_2 E_2 + \ldots + \lambda_k E_k$.

ii) $I = E_1 + E_2 + \ldots + E_k$

iii) $E_i^2 = E_i \quad \forall i$

iv) $E_i E_j = 0$ if $i \neq j$

v) $R(E_i) =$ the eigenspace corr. to λ_i, $\forall i$

Conversely, suppose $\exists$ k non-zero linear operators $E_1, E_2, \ldots E_k$ and distinct numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$ s.t. 'i), ii) & iv) hold'. Then, $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T, T is diagonalizable & $R(E_i) =$ eigenspace of λ_i, $\forall i$. Also iii) is true.

<u>Proof</u>: $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$ where W_i is the eigenspace corresponding to λ_i (since T is diagonalizable)

For $x \in V$,

$$x = x^1 + x^2 + \ldots + x^k \ ; x^i \in W_i$$

Define E_i's as before. Then (ii) — (iv) hold.

Suppose (v) holds.

$x = Ix = (E_1 + E_2 + \cdots + E_k)x$

$Tx = (TE_1 + \cdots + TE_k)x = TE_1x + \ldots + TE_kx.$

$= \lambda_1 E_1 x + \lambda_2 E_2 x + \ldots + \lambda_k E_k x.$ (since $E_1 x \in$ eigenspace of λ_1 it is an eigenvector)

$= (\lambda_1 E_1 + \ldots + \lambda_k E_k)x \Rightarrow$ i) ✓

$x = x^1 + \ldots + x^i + \ldots + x^k$

$E_i(x) = x^i \Rightarrow E_i(x) \in W_i$ which is by defn. the eigenspace of λ_i.

$\Rightarrow R(E_i) =$ eigenspace of λ_i.

<u>Converse proof</u>

$I = E_1 + E_2 + \ldots + E_k$

$E_i = E_i E_1 + \ldots + E_i E_k$

$= E_i^2$ (since $E_i E_j = 0$) $\Rightarrow$ iii) ✓

We show that $R(E_i) \subseteq$ eigenspace of $\lambda_i = N(T - \lambda_i I)$

Let $x \in R(E_i) \Rightarrow x = E_i(x)$ (idempotence on range).

Consider $Tx - \lambda_i x$

$$Tx - \lambda_i x = \sum_{j=1}^{k} \Big(\underbrace{\lambda_j E_j}_{T} x - \lambda_i \underbrace{E_j}_{I} x \Big)$$

$$= \sum_{j=1}^{k} (\lambda_j - \lambda_i) E_j x$$

$$= \sum_{j=1}^{k} (\lambda_j - \lambda_i) E_j E_i x$$

$= 0$ since $\lambda_j = \lambda_i$ is the only term remaining.

So $R(E_i) \subseteq N(T - \lambda_i I)$

Since $E_i \neq 0$, $\exists\ x \neq 0,\ x \in R(E_i)$

$\Rightarrow \exists\ x \neq 0$ s.t. $Tx = \lambda_i x$ $\Rightarrow \lambda_1, \dots \lambda_k$ are eigenvalues of T.

Are they all the eigenvalues?

Let λ be such that, $Tx = \lambda x$, $x \neq 0$.

Then $0 = (T - \lambda I)x = \sum_{j=1}^{k} (\lambda_j E_j - \lambda E_j)\, x$

$$= \sum_{j=1}^{k} (\lambda_j - \lambda) E_j x$$

We have $0 = u^1 + u^2 + \dots + u^j + \dots + u^k$ where $u^j = (\lambda_j - \lambda) E_j x$.

$u^j \in R(E_j) \Rightarrow u^j$ is an eigenvector (from distinct λ_j) (since $R(E_i) \subseteq N(T - \lambda_i I)$).

$R(E_j)$ are independant for each j. $\Rightarrow$ each u^j must be 0, $\forall j$.

$$\Rightarrow (\lambda_j - \lambda) E_j x = 0 \quad \forall\ j$$

If $E_j x = 0\ \forall\ j$, then $x = 0$ ※ (since $x \neq 0$)
(since $I = \sum E_i$)

$\Rightarrow \quad E_j x \neq 0$ for some j.

So $\lambda = \lambda_j$ for that j.

$\Rightarrow \quad \lambda_1 \ldots \lambda_k$ exhaust all the eigenvalues!

By taking basis for $R(E_i)$, namely $\mathcal{B}_1 \ldots \mathcal{B}_k$ & by setting

$\mathcal{B} = \bigcup_{j=1}^{k} \mathcal{B}_j$, it follows that $[T]_{\mathcal{B}}$ is diagonal.

It remains to be shown that $N(T-\lambda_i I) \subseteq R(E_i)$

Let $u \in N(T-\lambda_i I)$.

$$\Rightarrow \quad 0 = (T-\lambda_i I)u$$

$$= \sum_{j=1}^{k} (\lambda_j - \lambda_i) E_j u$$

$\Rightarrow (\lambda_j - \lambda_i) E_j u = 0 \quad \forall j$ since E_j's are independant.

$j \neq i \Rightarrow E_j u = 0 \quad \Rightarrow u \in R(E_i)$ // $\blacksquare$ (Since $u = E_1 u + E_2 u + \ldots + E_k u$.
$\Rightarrow u = E_i(u)$ from $j \neq i \Rightarrow E_j u = 0$)

Exercise: $P_j(t) = \prod_{k \neq j} \left(\dfrac{t-\lambda_k}{\lambda_j - \lambda_k} \right) \quad \Rightarrow P_j(\lambda_i) = \delta_{ij}$

* We can show that $\boxed{E_j = P_j(T)}$ ie. projections are polys in T *

given by Lagrange interpolation.

Example.

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix}$$

Char poly.

$$p(\lambda) = \lambda^3 - 7\lambda^2 + 11\lambda - 5 = (\lambda-1)^2(\lambda-5)$$

A is diagonalizable since eigenspace of $\lambda = 1$ has dim 2.

$$E_1 = p_1(T) = \frac{T - 5I}{1-5} = -\frac{1}{4}(T - 5I)$$

$$= -\frac{1}{4}\begin{pmatrix} -3 & +2 & +1 \\ +1 & -2 & +1 \\ +1 & +2 & -3 \end{pmatrix}$$

$$E_2 = \frac{T - I}{4} = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix}$$

We get,

$E_1 + E_2 = I$, $A = E_1 + 5E_2$, $E_1E_2 = 0$, $E_1^2 = E_1$, it is also clear that $R(E_1)$ is the eigenspace of λ_1.

Note:

$$m(t) = (t-\lambda_1)^{r_1}(t-\lambda_2)^{r_2} \cdots (t-\lambda_k)^{r_k}$$

In this case what is the direct sum decomposition?

This gives the PDT.

Primary Decomposition Theorem.

V fdvs, $T \in L(V,V)$. Suppose that $m = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$ where r_i's are positive integers, p_i's are monic irreducible polynomials over $\mathbb{F}$. Then $\boxed{\text{i) } V = \bigoplus_{i=1}^{k} \underbrace{N(p_i(T)^{r_i})}_{W_i}}$. ii) Also W_i's are T invariant. iii) (If T_i is $T|_{W_i}$, $m_i(T_i) = p_i^{r_i}$)

↳ min poly of T_i.

(This is the most generalised version of the last few theorems. The name primary comes from the fact that p_i's are prime (irreducible).)

Proof:

Set $f_i = \dfrac{m}{p_i^{r_i}} = \prod_{j \neq i} p_j^{r_j}$, $1 \le i \le k$.

Consider $f_1, f_2, \dots, f_k$. These are relatively prime because f_j misses $p_j^{r_j}$. So $\exists$ polynomials $g_1, g_2 \cdots g_k$ such that

$$f_1 g_1 + f_2 g_2 + \dots + f_k g_k = 1$$

Let $h_i(t) = f_i(t)\, g_i(t)$. Then $\sum_{j=1}^{k} h_j = 1$

Set $E_j = h_j(T)$. $1 \le j \le k$.

$\Rightarrow E_1 + E_2 + \dots + E_k = I$.

$$h_i h_j = f_i g_i f_j g_j = \underbrace{f_i f_j}_{\text{divisible by } m} g_i g_j$$

$$\Rightarrow h_i h_j(T) = m(T) \times (\text{something}) = 0$$

$$\Rightarrow E_i E_j = 0 \quad ; \quad i \neq j$$

It follows as before that $E_i^2 = E_i$.

$\Rightarrow E_i$'s are projections & give rise to a Direct sum decomposition

We only need to show $R(E_i) = W_i = N(P_i(T)^{r_i})$

Let $x \in R(E_i) \Rightarrow x = E_i(x)$

Consider $P_i(T)^{r_i} x = P_i(T)^{r_i} E_i x = P_i(T)^{r_i} h_i(T) x$

$$= \underbrace{P_i(T)^{r_i} f_i(T)}_{m(T)} g_i(T) x$$

$$= 0$$

$\Rightarrow x \in N(P_i(T)^{r_i}) \Rightarrow R(E_i) \subseteq N(P_i(T)^{r_i})$

Conversely, suppose $x \in N(P_i(T)^{r_i})$

We show that $E_j x = 0 \ \forall j \neq i$. This gives $x = E_i x$ (same arg as before)

$P_i(T)^{r_i} x = 0$. Consider $E_j x = f_j g_j x$. If $j \neq i$ f_j is divisible by $P_i(T)^{r_i}$ $\Rightarrow f_j(T) g_j(T) x = 0 \Rightarrow E_j(x) = 0$.

$\Rightarrow x = E_i x \Rightarrow x \in R(E_i)$

It remains to show that W_i's are T invariant.

This is clear since if T_i is the operator induced on W_i by T, $p_i(T_i)^{r_i} = 0$, because by definition $p_i(T)^{r_i} = 0$ on the subspace W_i.

$\Rightarrow$ $p_i(T_i)^{r_i}$ is annihilating & is divided by $m_i(T_i)$.
↳ min poly of T_i

Conversely let $g(T_i) = 0$ be an annihilating polynomial.

$\Rightarrow$ $g(T) f_i(T) = 0$. Since $f_i(T) = 0$ when $\vec{x} \notin W_i$ & $g(T) = 0$ when $\vec{x} \in W_i$.

$\Rightarrow$ $g f_i$ is divisible by m. (min poly of T).

$m = p_i^{r_i} f_i$; $\Rightarrow$ $g f_i$ is divisible by $p_i^{r_i} f_i$

$\Rightarrow$ $p_i^{r_i}$ divides g

$\Rightarrow$ $p_i^{r_i}$ is the minimal polynomial of T_i. ∎

A particular case:

Suppose that each p_i is a linear poly. ie. $p_i(t) = (t - \lambda_i)$.

$\Rightarrow$ $W_i = N((T - \lambda_i)^{r_i})$. Set $D = \lambda_1 E_1 + \dots + \lambda_k E_k$.

where $E_j = h_j(T)$.

We then have: D is diagonalizable, $\lambda_1 \dots \lambda_k$ are the only distinct eigenvalues of D. $R(E_i) = W_i$

Define N by $N = T - D$. Then $T = D + N$.

$$N = T - D$$
$$= TI - D$$
$$= T(E_1 + E_2 + \dots + E_k) - (\lambda_1 E_1 + \dots + \lambda_k E_k)$$
$$= (T - \lambda_1 I)E_1 + \dots + (T - \lambda_k I)E_k.$$
$$= \sum_{j=1}^{k} (T - \lambda_j I) E_j$$

$N^2 = ?$;

Consider $(T - \lambda_1 I) E_1 . N = (T - \lambda_1 I) E_1 \sum_{j=1}^{k} (T - \lambda_j I) E_j$

$= (T - \lambda_1 I) E_1 . (T - \lambda_1 I) E_1$

$= (T - \lambda_1 I)^2 E_1$ since E_1 is a poly in T & commutes

$\Rightarrow N^2 = \sum_{j=1}^{k} (T - \lambda_j I)^2 E_j$

$\Rightarrow N^\ell = \sum_{j=1}^{k} (T - \lambda_j I)^\ell E_j \qquad \ell \in \mathbb{N}.$

If $\ell \geq r_i \quad 1 \leq i \leq k$, then

$N^\ell x = \sum_{j=1}^{k} (T - \lambda_j I)^\ell \underbrace{E_j x}_{\in R(E_j)} = 0$ (since $(T - \lambda_j I)^{\ell - r_j} \underbrace{(T - \lambda_j I)^{r_j} E_j x}_{0}$

(Recall $R(E_i) = N(p_i(T)^{r_i})$, $p_i(t) = (t - \lambda_i)$)

So $N^\ell = 0$ } Nilpotent operator!

$\Rightarrow$ $\boxed{T = \underbrace{D}_{\text{diagonalizable}} + \overbrace{N}^{\text{Nilpotent}}}$

and $\boxed{DN = ND}$ since they are both polys in T. } Jordan Decomp

<u>Theorem</u>: Let $T \in L(V)$. Then $\exists$ a diagonalizable operator D & a Nilpotent operator N. Such that

$$T = D + N \quad \& \quad ND = DN.$$

(Also D, N are unique).

[This is very useful in computing powers of A.

$A^r = (D+N)^r$ which can be expanded using binomial since D, N commute. And Nilpotency of N sends many terms to 0.

> $e^A = I + A + \frac{A^2}{2!} + \dots$ is used in solving systems of DEs.]

<u>Example</u>

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad ; \quad p(t) = (t-2)^2 (t+1) = m(t)$$

$\lambda_1 = 2, \quad \lambda_2 = -1$

$f_1(T) = T + I \quad ; \quad f_2(T) = (T - 2I)^2$

$g_1 = \frac{5I - T}{9} \qquad g_2 = \frac{I}{9}$. Then $f_1 g_1 + f_2 g_2 = I$.

$$E_1 = f_1 g_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad E_2 = f_2 g_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$\Rightarrow E_1 + E_2 = I.$

By def : $D = \lambda_1 E_1 + \lambda_2 E_2$

$= 2E_1 - E_2$

$$= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$ $\to$ already diagonal.

$$N = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$N^2 = 0$ //

Lec 36 Cyclic Subspaces and Annihilators.

Let V fdvs & $T \in L(V,V)$. For an $x \in V$, the <u>T-cyclic subspace generated by x</u> is defined by $Z(x;T)$.

$$Z(x;T) = \{g(T)x : g \in \mathbb{F}[t]\}.$$

($Z(x;T) \subseteq V$ is a subspace)

x is called a cyclic vector for T if x satisfies

$$Z(x;T) = V.$$

<u>Examples</u>: 1) $Z(0,T) = \{0\}$

2) $Z(x,I) = \text{span}\{x\}$, $x \neq 0$

3) $Z(x,T) = \text{span}\{x\} \iff x$ is an eigenvector for T.

4) Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x) = (0, x_1)$

$Z(e^1, T) = ?$

<u>Note</u>: $T^2 = 0$

$$Z(e^1, T) = \{g(T)e^1 : g \in \mathbb{F}[x]\}$$

Let $g(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_k t^k$

$$g(T) = \alpha_0 I + \alpha_1 T$$

$$g(T)e_1 = \alpha_0 e_1 + \alpha_1 T(e_1)$$

$$= \alpha_0 e_1 + \alpha_1 e_2$$

Def: The T-annihilator of a vector x is the subspace

$$S_T(x, \{0\}) := \{g \in \mathbb{F}[t] : g(T)x = 0\}$$

Let $M(x;T) = S_T(x, \{0\})$. It is an ideal.

Let p_x be the unique monic poly. generating $M(x;T)$.

Note that p_x divides m. → minimal poly. It is also called the T-annihilator of x

Thm: Let $\deg p_x = k$. Then $\dim(Z(x, U)) = k$.
↳ T

In fact $\mathcal{B} = \{x, Ux, \dots, U^{k-1}x\}$ is a basis for $Z(x, U)$.

Proof: Let $g \in \mathbb{F}[t]$. $\exists\, q, r \in \mathbb{F}[t]$ s.t.

$$g(t) = p_x(t)\,q(t) + r(t)$$ where either $r = 0$ or $\deg(r) < \deg p_x$.

$$Z(x; U) = \{g(U)x : g \in F[t]\}$$

$$g(U)x = p_x(U)\,q(U)x + r(U)x$$

$$= q(U)\underbrace{p_x(U)x}_{0} + r(U)x$$

$$= r(U)x.$$

$$r(t) = r_0 + r_1 t + \dots + \cancel{r_{k-1}} t^{k-1}$$ 1 because monic

$$\Rightarrow g(U)x = r(U)x = \left(r_0 I + r_1 U + \dots + r_{k-2}U^{k-2} + U^{k-1}\right)x$$

$$\Rightarrow \operatorname{span}\{x_1, Ux, U^2x, \dots, U^{k-1}x\} = Z(x, U)$$

Suppose that $x, Tx, \ldots, T^{k-1}x$ are dependant.

Then $\exists$ scalars $\delta_0, \delta_1, \ldots \delta_{k-1}$ not all zeros.

$$\delta_0 x + \delta_1 Tx + \cdots + \delta_{k-1} T^{k-1} x = 0.$$

Let i be the greatest subscript $(0 \leq i \leq k-1)$ such that $\delta_i \neq 0$.

$$0 = \delta_0 x + \delta_1 Tx + \cdots + \delta_i T^i x.$$

Let $h(t) = \delta_0 + \delta_1 t + \cdots + \delta_i t^i$, then $h(T)x = 0$.

$\deg h = i \leq k-1 < k$ $\divideontimes$ since $\deg p_x = k$ & is the ideal generator of all T annihilators of x.

> $V = Z(x^1; T) \oplus Z(x^2; T) \oplus \cdots \oplus Z(x^k; T)$. is basically the Cyclic Decomposition Thm.

Let U be an operator on W. Let $\deg(p_x) = k$. Then $\mathcal{B} = \{x, Ux, U^2x, \ldots, U^{k-1}x\}$ form a basis for W.

Set $v^i = U^{i-1}x$ for $1 \leq i \leq k$.

) $v^1 = x$; $v^2 = Ux = Uv^1$; $\ldots$; $v^k = U^{k-1}x = U^{k-1}v \underset{\text{by induction.}}{=} Uv^{k-1}$.

$\Rightarrow \mathcal{B} = \{v^1, v^2 \ldots v^k\}$. What is $[U]_{\mathcal{B}}$?

$Uv^1 = v^2$

$Uv^2 = v^3$

$\vdots$

$Uv^{k-1} = v^k$

$\deg p_x = k$; $p_x(t) = p_0 + p_1 t + \dots + p_{k-1} t^{k-1} + t^k$

$0 = p_x(U)x = p_0 x + p_1 U x + \dots + p_{k-1} U^{k-1} x + U^k x$

$= p_0 v^1 + p_1 v^2 + \dots + p_{k-1} v^k + U(v_k).$

$\Rightarrow U(v_k) = -p_0 v^1 - p_1 v^2 \dots - p_{k-1} v^k$

$$[U]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & & 0 & -p_0 \\ 1 & 0 & 0 & & 0 & -p_1 \\ 0 & 1 & 0 & & \vdots & \vdots \\ 0 & \vdots & 1 & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & 0 & \vdots \\ \vdots & \vdots & \vdots & & 1 & -p_{k-1} \end{pmatrix} \longrightarrow \text{companion matrix for } p_x.$$

$k-1^{th}$ column.

Is the converse true?

Does there exist a cyclic basis (generated by x) for $W = Z(x; U)$.

Yes. <u>Proof:</u>

Let $\mathcal{B}' = \{v^1, v^2 \dots v^k\}$ be a basis for $Z(x; U)$ such that $[U]_{\mathcal{B}'}$ = companion matrix for p_x. It can be shown that v^1 is a cyclic vector ie. $W = Z(v^1; U)$.

<u>Theorem</u>: If A is the companion matrix for p_x, then the minimal polynomial and the characteristic polynomial equal p_x.

<u>Proof</u>: <u>Hint</u>: $p_x(U) = 0$.

Cyclic Decomposition Theorem

> Can we write V fdvs as

$$V = Z(x^1, T) \oplus Z(x^2, T) \oplus \dots \oplus Z(x^k, T). \; ?$$

$V = W \oplus W'$ can always be done for some W & W'.

> Given W, such a W' exists (many do) & they are <u>complimentary</u> to W.

> Suppose that $T(W) \subseteq W$, does there exist W' s.t, $T(W') \subseteq W'$. In general No.

<u>Def</u>: V fdvs, $T \in L(V,V)$, W is a subspace of V & is called T admissible if.

i) W is T invariant

ii) If $f(T)y \in W$, then $\exists\, z \in W$ s.t $f(T)y = f(T)z$. for any polynomial f.

<u>Lemma</u>: Let $V = W \oplus W'$ with $T(W) \subseteq W$ & $T(W') \subseteq W'$. Then W is T-admissible.

<u>Proof</u>: Let $y \in V$; $y = y^1 + y^2 \Rightarrow y^1 \in W$ & $y^2 \in W'$.

For any $f \in \mathbb{F}[x]$, $f(T)y = \underbrace{f(T)y_1}_{\in W} + \underbrace{f(T)y_2}_{\in W'}$

If $f(T)y \in W \Rightarrow f(T)y_2 = 0 . \Rightarrow f(T)y = f(T)y_1$, $y^1 \in W$. //

> The converse is nontrivial & is a consequence of the CDT.

Statement of CDT

Let V fdvs and $T \in L(V,V)$. Let W_0 be a T-admissible proper subspace of V. Then $\exists$ non-zero vectors $x^1, x^2, \dots, x^r \in V$ such that

i) $V = W_0 \oplus Z(x^1;T) \oplus \dots \oplus Z(x^r;T)$

ii) $\exists$ T-annihilators p_k, $1 \le k \le r$ corresponding to x^k such that $p_k \mid p_{k-1}$ $\forall k$.

↳ divides

Further the integer r and p_k for which i) & ii) hold are unique.

Proof:

Step 1 We S.T, $\exists$ nonzero vectors $y^1, y^2, \dots, y^r \in V$ S.T

i) $V = W_0 + Z(y^1;T) + \dots + Z(y^r;T)$.

ii) If $W_k = W_0 + Z(y^1;T) + \dots + Z(y^k;T)$ for $1 \le k \le r$ then the conductor $p_k = s(y^k; W_{k-1})$ has maximum degree among all T-conductors in W_{k-1}.

$\deg p_k = \max_{x \in V} \deg s(x; W_{k-1})$.

notation of the ideal generator.

Recall:
$S_T(x;W) = \{g : g(T)x \in W\}$ – ideal generated by $s(x;W)$

Proof of Step 1

Let W be an invariant subspace of V.

ie. $T(W) \subseteq W$. If $W \neq V$. Then

$$\underset{\substack{| \\ \text{Since } W \neq V}}{0} < \max_{x \in V} s(x; W) \leq \dim V.$$

Let y be a vector in V for which this maximum is attained.

Consider a new subspace $W + Z(y, T)$

(Note, $y \notin W$, otherwise $\deg = 0$) Therefore $\dim(W + Z(y, T)) > \dim W$

since $y \notin W \Rightarrow Z(y, T)$ is of atleast 1 dim.

Applying W_0 to W.

$\exists\, y^1 \notin W_0$ s.t $\dim(W_1) > \dim(W_0)$ where $W_1 = W_0 + Z(y^1; T)$.

If $W_1 \neq V$, Then we construct y^2 s.t $W_2 = W_0 + Z(y^1; T) + Z(y^2; T)$

where $\dim(W_2) > \dim(W_1) > \dim(W_0)$.

Since this process must terminate we have $V = W_0 + Z(y_1; T) + Z(y_2; T) + \dots + Z(y_r; T)$.

Where $1 \leq k \leq r$ $W_k = W_0 + Z(y^1; T) + \dots + Z(y^k; T)$

$\Gamma s(y^1; W_0) = p_1$

Choice of y^k's ensure that p_k satisfies the maximum property.

Step 2 Let $y^1, y^2, \dots, y^k$ be nonzero vectors from step 1 satisfying (they cannot be zero since dim increases).

(i) & (ii)

For a fixed K $(1 \leq K \leq r)$ set $f = s(y^k; W_{K-1})$

Then $f(T)y^k \in W_{K-1}$. If $f(T)y^k = y^0 + \sum_{i=1}^{K-1} \underbrace{g_i}_{\text{Some poly.}}(T)y^i$

Then f divides each g_i, and $\exists\, z^0 \in W_0$ s.t $f(T)y^0 = f(T)z^0$.

Proof of Step 2

$\exists\, h_i, r_i$ s.t. $g_i = f h_i + r_i$,

where either $r_i = 0$ or $\deg r_i < \deg f$. $(i = 1, 2, \dots, K-1)$

We show that each $r_i = 0 \Rightarrow f$ divides g_i

Proof is by induction on K.

$K=1$: T- admissibility of W_0 is already assumed (so trivial)

Assume true for $K-1$. Prove for K.

Let $K>1$, $\exists$ polys h_i & r_i s.t $g_i(t) = f(t)h_i(t) + r_i(t)$

where $r_i(t) = 0$ or $\deg r_i < \deg f$.

Claim: $r_i = 0\ \forall i$.

Let $z = y - \underbrace{\sum_{i=1}^{K-1} h_i(T)y^i}_{\in W_{K-1}?}$

W_K's are a nested sequence of subspaces so $z - y \in W_{K-1}$

Then $S(z; W_{K-1}) = S(y; W_{K-1}) = f$.

$x, u \in W$ s.t $x - u \in W$

Then $S_T(x; W) = S_T(u; W)$.

$g \in S_T(x; W)$. Then $g(T)x \in W$; $g(T)x = w \in W$.

Consider $g(T)(u-x) = g(T)u - g(T)x$.

$\in W \quad = g(T)u - w \iff g(T)u \in W. \iff g \in S_T(u; W)$

Also $f(T)z = f(T)y - \sum_{i=1}^{k-1} f(T)h_i(T)y^i$

$$= y^0 + \sum_{i=1}^{k-1} g_i(T)y^i - \sum_{i=1}^{k-1} f(T)\, h_i(T)y^i$$

$$= y^0 + \sum_{i=1}^{k-1} (g_i - fh_i)(T)\, y^i$$

$\Rightarrow$ $f(T)z = y^0 + \sum_{i=1}^{k-1} r_i(T)\, y^i$

Suppose $r_i \neq 0$ for some i _ let j be the largest; s.t $r_i \neq 0$.

Then $r_j \neq 0$, $\deg r_j < \deg f$.

$$f(T)z = y^0 + \sum_{i=1}^{j} r_i(T)\, y^i$$

Set $p = s(z; W_{j-1})$.

Note $W_{j-1} \subseteq W_{k-1}$. Then the conductor $f = s(y; W_{k-1})$ divides p.

Then $p = fg$ for some g. Consider

$p(T)z = f(T)g(T)z = g(T)f(T)z = g(T)y^0 + \sum_{i=1}^{j} g(T)r_i(T)y^i$.

$p(T) \in W_{j-1}$

$$\Rightarrow \underbrace{p(T)z}_{\in W_{j-1}} = \underbrace{g(T)y^0}_{\in W_0 \subset W_{j-1}} + \underbrace{\sum_{i=1}^{j-1} g(T)r_i(T)y^i}_{\in W_{j-1}} + g(T)r_j(T)y^j$$

$\Rightarrow g(T)r_j(T)y^j \in W_{j-1}$

Compare gr_j with $s(y^j; W_{j-1})$

$\Rightarrow \deg(gr_j) \geq \underbrace{\deg(s(y^j; W_{j-1}))}_{= \deg p_j}$ → ideal gen!

But $\deg p_j \geq \deg \underbrace{s(z; W_{j-1})}_{\substack{\deg p \\ = \deg(fg)}}$ from its' prop.

Finally $\deg(gr_j) \geq \deg(gf)$

$\Rightarrow \deg(r_j) \geq \deg(f)$. ✱ since $\deg(r_j) < \deg(f)$ //

$\Rightarrow$ each r_i is $0 \Rightarrow f$ divides g_i $\forall i$

Since W_0 is T-admissible, $\exists\, z^0 \in W_0$ s.t. $g^0 = f(T)z^0$ //

// part 2 done //

Step 3 : $\exists$ nonzero vectors $x^1, x^2 \dots x^r$ and the corresponding T-annihilators $p_1 \dots p_k$ s.t

i) $V = W_0 \oplus Z(x^1; T) \oplus \dots \oplus Z(x^r; T)$.

ii) p_k divides p_{k-1}, $2 \le k \le r$.

Proof of step 3

Start with $y^1; y^2 \dots y^r$ from step①. Apply step 2 to $y = y^k$ & $f = p_k$. Then

$$p_k(T) y^k = y^0 + \sum_{i=1}^{k-1} \underbrace{p_k(T) h_i(T)}_{g_i(T)} y^i$$

Define x^k by $x^k = y^k - z^0 - \sum_{i=1}^{k-1} h_i(T) y^i$.

$\Rightarrow$ $x^k - y^k \in W_{k-1}$ so $s(x^k; W_{k-1}) = s(y^k; W_{k-1}) = p_k$.

Also $p_k(T) x^k = 0$ $\forall k$.

Thus $W_{k-1} \cap Z(x^k; T) = \{0\}$.

[$Z(x^k, T) = \{ g(T) x^k ; g \in \mathbb{F}[t] \}$ & $p_k = s(x^k; W_{k-1})$.

$\Rightarrow$ g is a multiple of p_k (since $g(T) x^k \in W_{k-1}$)

$\Rightarrow g(T) x^k = 0$] This guarantees independance.

Claim: p_k divides p_{k-1}.

Use $p_k x^k = 0$ $\forall k$.

So $p_K(T)x^k = 0 + p_1(T)x^1 + p_2(T)x^2 + \cdots + p_{K-1}(T)x^{K-1}$.

$$\Rightarrow \quad = 0 + \sum_{i=1}^{K-1} p_i(T)x^i$$

from step 2 we know $p_K \mid p_i \quad \forall i = 1, 2, \ldots, k-1$. //

(uniqueness is not done here).

Corollary:

Let W be T-admissible. Then $\exists\, W'$ s.t. W' is T invariant & $V = W \oplus W'$.

Proof: If $W = V$ nothing to show.

If $W \neq V$ apply CDT.

$$\Rightarrow V = W \oplus \underbrace{Z(x^1;T) \oplus \cdots \oplus Z(x^r;T)}_{W'}.$$

Matrix analog.

Any matrix $B \in \mathbb{F}^{n \times n}$ is similar to A, where.

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_r \end{pmatrix}$$

where A_i is the C_f of the polynomial p_i. Called the Rational Form.

Ex.

$$B = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix} \rightarrow B \text{ is diagonalizable.}$$

$p(t) = (t-1)(t-2)^2$

$m(t) = (t-1)(t-2)$

$$B \sim A = \begin{pmatrix} \boxed{\begin{matrix} 0 & -2 \\ 1 & 3 \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 2 \end{pmatrix}$$

companion matrix of the T-annihilator of eigenvector of $\lambda = 2$.

Start with $\lambda=2$, calc. the eigenvector, (say x_1). (Look at x_1, Tx_1 which forms a cyclic subspace?) Find T-annihilator of this subspace, its companion matrix gives this block.

Inner Product Spaces

- Let $V = \mathbb{C}^n$. An inner product on V is a function

$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ which satisfies:

i) $\langle x, x \rangle \geq 0 \quad \forall x \in V.$
$\langle x, x \rangle = 0 \iff x = 0.$ } (positive definiteness)

ii) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad \forall x, y, z \in V.$
iii) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle. \quad \forall \lambda \in \mathbb{C}$ & $\forall x, y \in V.$ } linearity wrt first argument

iv) $\langle y, x \rangle = \overline{\langle x, y \rangle} \quad \forall x, y \in V.$ } conjugate symmetry.

> A vector space V together with an inner product $\langle \cdot, \cdot \rangle$ is called an inner product space. $(V, \langle \cdot, \cdot \rangle)$.

Eg: i) $V = \mathbb{C}^{n \times n}$

If $A, B \in V$. Let $A^* = \bar{A}^t$

Define $\langle A, B \rangle = tr(AB^*) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \overline{b_{ij}}$ //

ii) $V = C([0,1])$ space of all complex valued cont. fns. on $[0,1]$.

$f, g \in V$, define $\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} \, dt$ //

Def: Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Let $x \in V$.

The "norm" of x denoted by $\|x\|$ is defined by

$$\|x\| = +\sqrt{\langle x, x \rangle}$$

So, $\|x\| \geq 0 \quad \forall x \in V$.

$$\|x\| = 0 \iff \|x\|^2 = 0 \iff \langle x, x \rangle = 0 \iff x = 0$$

Theorem: Let V IPS, then

(a) $\|\alpha x\| = |\alpha| \, \|x\| \quad \forall \alpha \in \mathbb{C}$ & $x \in V$.

(b) $|\langle x, y \rangle| \leq \|x\| \, \|y\| \quad \forall x, y \in V \to$ CS inequality

(c) $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V \to$ Triangle inequality.

Proof a) $\|\alpha x\|^2 = \langle \alpha x, \alpha x \rangle = \alpha \langle x, \alpha x \rangle = \alpha \bar{\alpha} \langle x, x \rangle$

$$= |\alpha|^2 \|x\|^2$$

$$\Rightarrow \|\alpha x\| = |\alpha| \, \|x\|$$

b) We have $\forall \lambda \in \mathbb{C}$

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle \quad \forall x, y \in V$$

$$= \langle x, x \rangle - \bar{\lambda} \langle x, y \rangle - \lambda \langle y, x \rangle + |\lambda|^2 \|y\|^2$$

Let $\alpha = \bar{\lambda} \langle x, y \rangle$

$$\Rightarrow \quad 0 \le \|x\|^2 + |\lambda|^2 \|y\|^2 - (\alpha + \bar{\alpha}).$$

$$= \|x\|^2 + |\lambda|^2 \|y\|^2 - 2\,\mathrm{Re}(\alpha) \quad \text{— (1)}$$

Let $\lambda = \dfrac{\langle x, y\rangle}{\|y\|^2}$, if $y \neq 0$ ($y = 0$ is trivial)

$$\Rightarrow (1) \Rightarrow \quad 0 \le \|x\|^2 + \frac{|\langle x, y\rangle|^2}{\|y\|^4}\|y\|^2 - 2\underbrace{\mathrm{Re}(\alpha)}_{\frac{|\langle x,y\rangle|^2}{\|y\|^2}}.$$

$$\left(\alpha = \bar{\lambda}\langle x, y\rangle = \frac{|\langle x, y\rangle|^2}{\|y\|^2}\right)$$

$$\Rightarrow \quad 0 \le \|x\|^2\|y\|^2 + |\langle x, y\rangle|^2 - 2|\langle x, y\rangle|^2$$

$$\Rightarrow \quad \|x\|^2\|y\|^2 \ge |\langle x, y\rangle|^2$$

$$\Rightarrow \quad \|x\|\,\|y\| \ge |\langle x, y\rangle|$$

c) $\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \|y\|^2 + \langle y, x\rangle + \langle x, y\rangle$

$$= \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}(\langle x, y\rangle)$$

For any $z \in \mathbb{C}$ $\quad |z|^2 = \alpha^2 + \beta^2 \Rightarrow \mathrm{Re}(z) \le |z|$

$$\Rightarrow \quad \|x+y\|^2 \le \|x\|^2 + \|y\|^2 + 2|\langle x, y\rangle|$$

$$\overset{CS}{\le} \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$

$$\Rightarrow \quad \|x+y\| \le \|x\| + \|y\|$$

Norms on Vector Spaces.

Let V be an IPS. Then

$$|\langle x, y\rangle| \leq \|x\| \, \|y\| \quad \forall x, y \in V.$$

In particular we have,

i) $$\left|\sum_{i=1}^{n} x_i \overline{y_i}\right| \leq \left(\sum_{j=1}^{n} |x_j|^2\right)^{1/2} \left(\sum_{j=1}^{n} |y_j|^2\right)^{1/2}$$

ii) $$|tr(AB^*)| \leq |tr(AA^*)|^{1/2} \, |tr(BB^*)|^{1/2}$$

iii) $$\left|\int_0^1 f(t)\overline{g(t)}\,dt\right| \leq \left[\int_0^1 |f(t)|^2 dt\right]^{1/2} \left[\int_0^1 |g(t)|^2 dt\right]^{1/2}$$

Defn A norm on a VS V, is a function $\|.\|: V \to \mathbb{R}$ Such that

i) $\|x\| \geq 0 \;\; \forall x \in V$ & $\|x\| = 0 \iff x = 0$

ii) $\|\lambda x\| = |\lambda| \, \|x\| \;\; \forall \lambda \in \mathbb{C} \;\; \forall x \in V.$

iii) $\|x+y\| \leq \|x\| + \|y\| \;\; \forall x, y \in V.$

A normed ~~linear~~ vector space is a pair $(V, \|.\|)$ where $\|.\|$ is a given norm on V.

$\Rightarrow$ Every IPS is a normed linear space w.r.t the induced norm (ie $\|x\| = \sqrt{\langle x, x\rangle}$.

> Consider $V = \mathbb{C}^n$ with the following norms:

i) $\|x\|_1 = \sum_{j=1}^{n} |x_j|$, $x \in \mathbb{C}^n$

ii) $\|x\|_\infty = \max \{ |x_i| : 1 \leq i \leq n \}$

} Not induced by any inner product.

iii) $\|x\|_2 = \sqrt{\sum_{j=1}^{n} |x_j|^2}$ (Euclidean norm)

Theorem: Parallelogram Law

Let V be an inner product space. Then

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \forall\ x, y \in V.$$ (if $\|\cdot\|$ is induced by the inner product).

Proof :- $\|x+y\|^2 + \|x-y\|^2 = \langle x+y, x+y \rangle + \langle x-y, x-y \rangle.$

$$= \|x\|^2 + \|y\|^2 + \langle y, x \rangle + \langle x, y \rangle + \|x\|^2 + \|y\|^2 - \langle y, x \rangle - \langle x, y \rangle$$

$$= 2\|x\|^2 + 2\|y\|^2$$

■

(We can show that 1 norm & ∞ norm are not induced by any inner product using the above Thm.)

On $C([0,1])$: $\|f\|_1 = \int_0^1 |f(t)|\,dt$

$\|f\|_\infty = \sup \{ |f(t)| : t \in [0,1] \}$

Orthogonal Bases and Orthonormal sets

Defn: Let V be an inner product space & $x, y \in V$. Then x is orthogonal to y if $\langle x, y \rangle = 0$.

- A subset $A \subset V$ is called orthogonal if distinct elements in A are mutually orthogonal. A is called orthonormal if A is orthogonal and each vector in A has norm 1. $\forall\, a, b \in A$ we have $\langle a, b \rangle = 0$ $a \neq b$ & $= 1$ if $a = b$.

Theorem: (Pythagoras)

If $x, y \in V$ & $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

(Trivial proof)

Let V be an IPS. $\mathcal{B} = \{u_1, u_2 \ldots u_n\}$ be a basis of V.

$\Rightarrow x \in V$: $x = \alpha_1 u_1 + \ldots + \alpha_n u_n$. $[x]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$

$x = (u_1, u_2 \ldots u_n) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ $\rightarrow$ Solve to find $(\alpha_1 \ldots \alpha_n)$. (RREF etc.)

In addition if $\mathcal{B}$ is an orthonormal basis, then

$x = \alpha_1 u_1 + \ldots + \alpha_n u_n$; $\langle x, u_i \rangle = \alpha_1 \langle u_1, u_i \rangle + \ldots + \alpha_n \langle u_n, u_i \rangle$

$$= \underline{\underline{\alpha_i}} \cancelto{1}{\langle u_i, u_i \rangle}$$

> Any orthogonal set is linearly independent. (but not conversely)

Gram Schmidt Orthonormalization Process.

Theorem: Let $\{u_1, u_2, \ldots\}$ be LI in an IPS V. Then we can construct an orthonormal set $\{v_1, v_2 \ldots\}$ such that

$$\text{span}(\{u_1, u_2 \ldots u_j\}) = \text{span}(\{v_1, v_2 \ldots v_j\}) \text{ for each } j.$$

Proof (By induction on j)

$j=1$: $u_1 \neq 0$ since LI. $\Rightarrow v_1 = \dfrac{u_1}{\|u_1\|} \Rightarrow \|v_1\| = 1$.

span u_1 = span v_1. ✓

Suppose that $\{v_1, v_2 \ldots v_n\}$ have been constructed such that $\{v_1 \ldots v_n\}$ is orthonormal & span$\{u_1 \ldots u_n\}$ = span$\{v_1 \ldots v_n\}$.

Consider $w_{n+1} = u_{n+1} - \sum_{j=1}^{n} \langle u_{n+1}, v_j \rangle v_j$ ($w_{n+1} \neq 0$ trivial)

Set $v_{n+1} = \dfrac{w_{n+1}}{\|w_{n+1}\|}$. Then $\|v_{n+1}\| = 1$.

$$\langle v_\ell, w_{n+1} \rangle = \langle v_\ell, u_{n+1} \rangle - \sum_{j=1}^{n} \overline{\langle u_{n+1}, v_j \rangle} \langle v_\ell, v_j \rangle$$

$$= 0 \quad \forall \ell \text{ s.t. } 1 \leq \ell \leq n.$$

$\Rightarrow \{v_1 \ldots v_n, v_{n+1}\}$ is an orthonormal set.

$$\begin{aligned}\operatorname{span}(\{v_1 \dots v_n, v_{n+1}\}) &\subseteq \operatorname{span}(\{v_1 \dots v_n, w_{n+1}\}) \\ &\subseteq \operatorname{span}(\{v_1 \dots v_n, u_{n+1}\}) \\ &\subseteq \operatorname{span}(\{u_1 \dots u_n, u_{n+1}\})\end{aligned}$$

Opposite inclusion is similar. ∎

Canonical Forms

RCF

$\nearrow$ (p_k) annihilates x_k

Recall that the degree of the T-annihilator P_k of x_k is equal to the dimension of the T-cyclic subspace generated by x_k.

- $Z(x,T) = \text{span}\{x\}$ if x is an eigenvector of T.

> Let $B_i = \{x_i, Tx_i, \ldots, T^{k_i-1}x_i\}$ be the 'cyclic ordered basis' of $Z(x_i, T)$ from CDT.

T_i is the restriction of T to $Z(x_i;T)$. Due to T-invariance, in this basis T has matrix

$$\begin{bmatrix} [\quad] & \begin{matrix}0\\A_i\\0\end{matrix} & [\quad] \end{bmatrix}$$

> Given a matrix, find the char. poly. We know $C(A)=0$. Also the companion matrix of the char. poly is the original matrix.

matrix $\longrightarrow$ char poly $\longrightarrow$ companion matrix $\longrightarrow$ char poly.

RCF

> Factorize the minimal poly : $m(x) = p_1\ p_2\ p_3$.

> Take a vector & check if $p_1(x)$; $p_1 p_2(x)$; $p_1 p_3(x)$ ... So on are 0.

> This gives the P_i for this x_i . If it is of deg 3 then $Z(x_i, T)$ is spanned by x_i, Tx_i & $T^2 x_i$.

> Take a vector not in $Z(x_i, T)$. If 3 dimensions are left & 3 eigenvectors exists check that they dont lie in $Z(x_i, T)$ & each would give a 1-D Z that completes the decomp. Use the basis $(x_i, Tx_i, T^2x_i, V_1, V_2\ V_3)$.

> You can always find a vector whose T-ann is the $m(x)$ (since otherwise $m(x)$ is not the ideal gen). Therefore there is always a $Z(x, T)$ with dim = deg of m.

> Finding p_i : Compute x_i, Tx_i, $T^2 x_i$ ... until you get a LD seq. & factor it to check it divides $m(x)$.

JCF

- If N is a nilpotent matrix. $m(x) = x^k$ for some k.

  Therefore $C_f = \begin{bmatrix} 0 & & & 0 \\ 1 & 0 & & 0 \\ & 1 & \ddots & 0 \\ & & & 0 \end{bmatrix}$.

- Given A & $m(x) = (x-c)^k$ if we choose $N = A - cI$ we have $m'(x) = (x')^k$ for N. Decompose this N into the RCF to get $\tilde{N} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & \ddots & \\ & & & 0 \end{bmatrix}$. The basis that does this gives $B^{-1}NB = \tilde{N} \Rightarrow B^{-1}AB = \tilde{N} + B^{-1}cIB = \tilde{N} + cI$

  $\Rightarrow \tilde{A} = \begin{bmatrix} c & & & \\ 1 & c & & \\ & 1 & \ddots & \\ & & & c \end{bmatrix}$.

- If $m(x) = \sum_{i}^{r} (x - c_i)^{k_i}$ then use PDT to get A_i which is the restriction of A to each null $(x-c_i)^{k_i} = W_i$. from there get N_i & a basis (using RCF) s.t $\tilde{A}_i$ is in JCF. Then JCF of A is $\begin{bmatrix} \tilde{A}_1 & & & \\ & \tilde{A}_2 & & \\ & & \ddots & \\ & & & \tilde{A}_r \end{bmatrix}$.

Eg ① (XI)

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ -1 & -1 & 2 & 0 & -1 & 0 \\ 1 & 1 & -1 & 1 & 1 & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 & 1 \end{bmatrix}$$

$m(x) = (x-1)^3$
$c(x) = (x-1)^6$.

JCF

> $N = A - I$.

This vector has $m_{T,v} = m_T$

We start by finding a vector in $Null(A-I)^3$ that is not in $Null(A-I)^2$. (finding RCF of $A-I$)

Choose $V_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = e_6$. Note $V_1 \notin Null(A-I)^2 \Rightarrow (x-1)^3$ is the T-annihilator of v_1. $\Rightarrow p_1 = (x-1)^3$

$\Rightarrow$ Cyclic Decomp Thm says $Z(v_1, T)$ has dim 3 & basis v_1, Nv_1, N^2v_1

Let $v_2 = Nv_1 = e_1$; $v_3 = N^2 v_1 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$.

> Next we look for vectors whose T-annihilator is of degree 2. Since $p_2 \mid p_1$ the choice is $p_2 = (x-1)^2$. We want to find a vector whose T-annihilator is p_2 and it is not in $Z(v_1; T)$. Therefore it should also not be in $Null(A-I)$ (ie. an eigenvector). Such a vector satisfies

$RREF([v_1 \ v_2 \ v_3 | v_4])$ has no solution & $(A-I)v_4 \neq 0$.

⑤

> $v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$ Satisfies this condition & $v_5 = N v_4 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$.

> The remaining T-ann. needed should satisfy $p_3 | p_2 \Rightarrow p_3 = x-1$.

$\Rightarrow$ we look for an eigenvector not in the span of $(v_1, v_2, v_3, v_4, v_5)$

$\Rightarrow$ $RREF(\, v_1\ v_2\ v_3\ v_4\ v_5 \mid v_6\,)$ has no soln.

Both $v_6 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ & $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}$ Satisfy this condition. Which is fine since ⊕ is not equivalent to orthogonality of the subspaces.

the bases need not be orthogonal.

$\Rightarrow$ $B = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \end{bmatrix}$ & $B^{-1}AB = \begin{bmatrix} 1 & 1 & & & & 0 \\ & 1 & 1 & & & \\ & & 1 & 0 & & \\ & & & 1 & 1 & \\ & & & & 1 & 0 \\ 0 & & & & & 1 \end{bmatrix}$ //

Eg② X2

$$B = \begin{bmatrix} -1 & -3 & 3 & 0 & -2 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & -3 & 0 \\ 0 & 0 & 1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

$c(x) = (x+1)^3 (x-2)^3$

$m(x) = (x+1)^2 (x-2)^2$ //

Start with PDT.

$\Rightarrow$ $V = W_1 \oplus W_2$ where $W_1 = \text{null}(B+I)^2$ & $W_2 = \text{null}(B-2I)^2$.

Let $N_1 = B+I$ & $N_2 = B-2I$.

N_1 is nilpotent & $p_1 = (x+1)^2$ (This p_1 is w.r.t $B|_{W_1}$).

$\Rightarrow$ There is a T-cyclic subspace generated by some w_1 of dim $= 2$.

We look for w_1 in $Null(B+I)^2$ S.T $(B+I)w_1 \neq 0$.

$w_1 = e_6$ is such a vector. $\Rightarrow$ $w_2 = N_1 w_1 = e_1$.

Next we have p_2 / p_1 & p_2 is of order 1, so $p_2 = (x+1)$.

We want w_3 S.T $(B+I)w_3 = 0$ & $w_3 \notin Z(w_1; N_1)$ $\Rightarrow$ $RREF([w_1 w_2 | w_3])$ has no solution. Look for w_3 in $Null(B+I)$.

$w_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$ is such a vector.

Moving to W_2 & repeating the same process. i.e. $w_4 \in W_2$ S.T

$(B+2I)^2 w_4 = 0$ & $(B+2I)w_4 \neq 0$ $\Rightarrow$ $w_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ (This is a choice, there are others).

$\Rightarrow$ $w_5 = (B+2I)w_4 = e_4$.

Choose $w_6 \in Null(B+2I)$ & $\notin Z(w_4; N_2)$ (again use $RREF(w_4 w_5 | w_6)$)

$\Rightarrow$ $w_6 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. $\Rightarrow$ $C = [w_1 w_2 w_3 w_4 w_5 w_6]$ &

$$C^{-1}BC = \begin{bmatrix} -1 & & & & & \\ 1 & -1 & & & & \\ & 0 & -1 & & & \\ & & 0 & 2 & & \\ & & & 1 & 2 & \\ & & & & 0 & 2 \end{bmatrix} \quad //$$

Notes

> PDT is a decomposition into independant subspaces that are T-invariant. $V = \oplus W_i$

> By choosing an ordered basis in these W_i's, the transformation has a matrix in block diagonal form. Because a vector written as $v = v_1 + v_2 \dots$ where $v_i \in W_i$ will have $T(v_i) \in W_i$ so $T(v_i)$ can be written in the basis for W_i itself. All other coordinates go to 0.

> CDT decomposes a subspace into T invariant, independant subspaces that are the T-cyclic subspaces of some vectors v_i that have a T-annihilator p_i. These p_i's are such that p_i / p_{i-1} & $\deg(p_i) = \dim(Z(v_i;T))$.

> In a matrix representation CDT gives the RCF, since CDT is also a decomp into T-inv, independant subspaces it must give a block diagonal form. Furthermore, since each subspace is generated by a vector with a T-annihilator, the restriction of T to this subspace must satisfy the T-annihilator's companion matrix. Therefore the blocks are the companion matrices of the T-annihilators.

> PDT + CDT gives JCF. Note that the companion matrix is $\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & & 0 \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{bmatrix}$ for a Nilpotent matrix & any matrix $A - c_i I$ is nilpotent when restricted to the subspace from PDT since $(A - c_i I)^{r_i} v\big|_{W_i} = 0$ (Subspaces in PDT are nullspaces)
For each $A - c_i I$ under restriction to W_i we find it's RCF which is a companion matrix & since $A - c_i I$ is nilpotent, A is in JCF. This is true for each W_i so the entirety of A is in JCF.