MATH 2040A Linear Algebra II 2015-16

Week 4: Diagonalizability, Matrix Limits (textbook §5.2 and 5.3)

Invariant subspaces, Cayley-Hamilton Theorem (§5.4)

Characterization of Diagonalizability

Thm B: Let $T: V \to V$ be a linear operator on V ($\dim V < +\infty$).

Suppose the characteristic polynomial of T splits with

distinct eigenvalues: $\lambda_1, \lambda_2, \ldots\ldots, \lambda_k$.

and algebraic multiplicity: $m_1, m_2, \ldots\ldots, m_k$.

Then (a) T diagonalizable $\iff$ $\boxed{\dim E_{\lambda_i} = m_i}$ for all $i = 1, \ldots, k$.

(b) If T is diagonalizable and β_i is an ordered basis for E_{λ_i} ($i=1,\ldots,k$), then $\boxed{\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k}$ is an eigenbasis of V for T.

Proof: We first prove (b), after establishing "$\Rightarrow$" part of (a).

(I): (a) "$\Rightarrow$" part: Assume T is diagonalizable, then $\exists$ eigenbasis β s.t.

$$[T]_\beta = \begin{pmatrix} \begin{array}{c}\scriptstyle m_1 \\ \boxed{\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}}\end{array} & & \\ & \begin{array}{c}\scriptstyle m_2 \\ \boxed{\begin{matrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_2\end{matrix}}\end{array} & \\ & & \begin{array}{c}\scriptstyle m_k \\ \boxed{\lambda_k}\end{array} \end{pmatrix} \qquad \text{Thus, } [T - \lambda_i I]_\beta = \begin{pmatrix} \begin{array}{c}\scriptstyle m_1 \\ \boxed{\lambda_1 - \lambda_i}\end{array} & & \\ & \begin{array}{c}\scriptstyle m_i \\ \boxed{O}\end{array} & \\ & & \begin{array}{c}\scriptstyle m_k \\ \boxed{\lambda_k - \lambda_i}\end{array} \end{pmatrix}$$

$$\Rightarrow \boxed{\dim E_{\lambda_i} = m_i}$$

(II): (b): Assume T is diagonalizable and β_i is a basis for E_{λ_i}.

By (a) "$\Rightarrow$", $\dim E_{\lambda_i} = m_i = \#\beta_i$. Therefore,

since char. poly. of T splits

$$\#\beta = \#\beta_1 + \#\beta_2 + \cdots + \#\beta_k = m_1 + m_2 + \cdots + m_k = \dim V$$

To show that β is an eigenbasis, it suffices to show that β is linearly independent.

[Recall that: If $0 \neq v_i \in E_{\lambda_i}$ belongs to distinct eigenspaces, then $\{v_1, \dots, v_k\}$ is linearly independent.] (*)

Let $\beta_i = \{v_{i1}, \dots, v_{in_i}\}$ and thus $\beta = \{v_{ij} : 1 \le i \le k, 1 \le j \le n_i\}$.

To show β is linearly indep., suppose $\exists\, a_{ij} \in \mathbb{F}$ s.t.

$$\sum_{i,j} a_{ij} v_{ij} = \vec{0}$$

my goal is to show all $a_{ij} = 0$

regrouping terms:

$$\underbrace{\Big(\sum_{j=1}^{n_1} a_{1j} v_{1j}\Big)}_{\in E_{\lambda_1}} + \underbrace{\Big(\sum_{j=1}^{n_2} a_{2j} v_{2j}\Big)}_{\in E_{\lambda_2}} + \cdots + \underbrace{\Big(\sum_{j=1}^{n_k} a_{kj} v_{kj}\Big)}_{\in E_{\lambda_k}} = \vec{0}$$

By (*), each term in the above expression vanishes:

$$\sum_{j=1}^{n_i} a_{ij} v_{ij} = \vec{0} \qquad \text{for } \underline{\text{each}}\ i = 1, \dots, k.$$

Since β_i is linealy indep, we have $a_{ij} = 0$ for $i = 1, \dots, k$, $j = 1, \dots, n_i$.

<u>(III): (a) "<=" part</u>: From the proof above, if β_i is a basis for E_{λ_i} then $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is linearly indep. If furthermore $\#\beta_i = m_i$ then $\#\beta = m_1 + m_2 + \cdots + m_k = \dim V$ and hence β is an eigenbasis. Therefore, T is diagonalizable.

$\square$

Using the notion of "direct sum", we can rephrase Thm. B above:

<u>Thm B rephrased</u>: $T : V \to V$ is diagonalizable if and only if

$$\boxed{V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}}$$

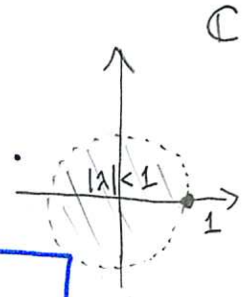
(see more details in textbook & tutorial)

Matrix Limits

Recall that if $A \in M_{n\times n}(\mathbb{C})$ is diagonalizable, then we can compute A^k easily by the formula

$$A^k = QD^kQ^{-1} = Q\begin{pmatrix} d_1^k & & & \\ & d_2^k & & \\ & & \ddots & \\ & & & d_n^k \end{pmatrix} Q^{-1}$$

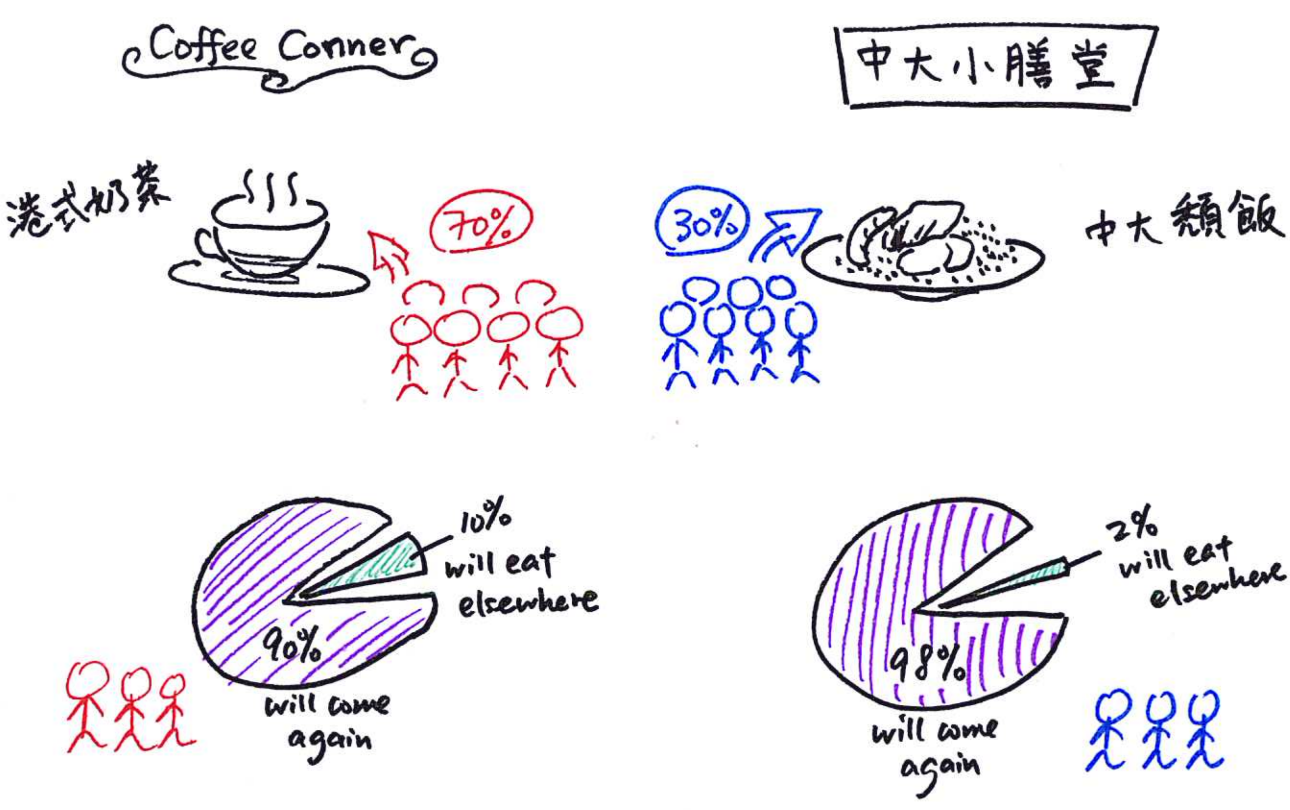
Therefore, $\lim_{k\to\infty} A^k$ exists $\iff$ $\lim_{k\to\infty} d_i^k$ exists for all i.



FACT: For $\lambda \in \mathbb{C}$, $\lim_{k\to\infty} \lambda^k$ exists $\iff$ $\lambda = 1$ or $|\lambda| < 1$

Thm: If $A \in M_{n\times n}(\mathbb{C})$ is diagonalizable and that for each eigenvalue $\lambda \in \mathbb{C}$, either $\lambda = 1$ or $|\lambda| < 1$, then $\lim_{k\to\infty} A^k$ exists.

Example - a stochastic process



Q: What will happen in the long run?

Let $P = \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix}$ be the initial distribution of customers.

After 1 day, the proportion of people going to

coffee conner: $0.9 \times 0.7 + 0.02 \times 0.3 = 0.636$

中大小膳堂: $0.1 \times 0.7 + 0.98 \times 0.3 = 0.364$

in matrix form:

$$\underbrace{\begin{pmatrix} 0.9 & 0.02 \\ 0.1 & 0.98 \end{pmatrix}}_{\substack{= A \\ \text{"transition matrix"}}} \underbrace{\begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix}}_{= P} = \begin{pmatrix} 0.636 \\ 0.364 \end{pmatrix}$$

Therefore, AP = proportion of customers after 1 day

Similarly. $A^2P = A(AP)$ = proportion of customers after 2 days

$\vdots$

A^kP = proportion of customers after k days.

$$\boxed{\text{Q: What is } \lim_{k\to\infty} A^kP\ ?}$$

An easy computation shows that A is diagonalizable and that

$$A = QDQ^{-1} \quad \text{where} \quad Q = \begin{pmatrix} 1/6 & -1/6 \\ 5/6 & 1/6 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 0.88 \end{pmatrix}$$

therefore, $\lim_{k\to\infty} A^k = Q \cdot \lim_{k\to\infty} D^k \cdot Q^{-1} = Q \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} = \begin{pmatrix} 1/6 & 1/6 \\ 5/6 & 5/6 \end{pmatrix}$

$$\Rightarrow \lim_{k\to\infty} A^kP = \begin{pmatrix} 1/6 & 1/6 \\ 5/6 & 5/6 \end{pmatrix} \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix} = \begin{pmatrix} 1/6 \\ 5/6 \end{pmatrix}$$

As a result, "eventually" 1/6 of the people will go to Coffee Conner and 5/6 of the people will go to 中大小膳堂, <u>independent of what the initial proportion P</u>! (can you explain why?)

Another Application - Linear ODE system

Consider a linear system of ODE (ordinary differential equations)

$x = x(t)$
$y = y(t)$

(#) $\begin{cases} x' = x + y \\ y' = 3x - y \end{cases}$ i.e. $\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

$$\boxed{\vec{x}' = A\vec{x}}$$

Idea: If there were only one (scalar) differential equation:

$\boxed{x' = a x} \Rightarrow$ general solution: $\boxed{x(t) = Ce^{at}}$, $c \in \mathbb{R}$

For a system of n ODE's:

$\boxed{\vec{x}' = A\vec{x}}$ $A \in M_{n\times n}(\mathbb{R})$ $\overset{?}{\Longrightarrow}$ $\boxed{\vec{x}(t) = c\, e^{At}}$

Q: How to define e^{At} for a matrix $A \in M_{n\times n}(\mathbb{R})$?

Recall: $e^{at} := 1 + at + \frac{1}{2}a^2t^2 + \cdots + \frac{1}{k!}a^k t^k + \cdots\cdots$

Just define: $\boxed{e^{At} := I + At + \frac{1}{2}A^2t^2 + \cdots + \frac{1}{k!}A^k t^k + \cdots\cdots}$

If A is diagonal, i.e.

$$A = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}, \text{ then } e^{At} = \begin{pmatrix} e^{d_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{d_n t} \end{pmatrix}.$$ (verify this!)

If A is NOT diagonal but diagonalizable, then $\exists\, Q$ invertible s.t.

$$A = Q \underset{\substack{\uparrow \\ \text{diagonal}}}{D} Q^{-1} \Rightarrow e^{At} = Q e^{Dt} Q^{-1}$$ (since $A^k = QD^kQ^{-1}$)

Let us look at the example (#) again now.

Let $A = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}$. Char. Poly. $= (1-\lambda)(-1-\lambda) - 3 = \lambda^2 - 4 = (\lambda+2)(\lambda-2)$

eigenvalues	eigenspaces	eigenbasis
$\lambda_1 = -2$	$E_{\lambda_1} = \text{span}\left\{\begin{pmatrix} -1 \\ 3 \end{pmatrix}\right\}$	$\beta = \left\{\begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$, $Q = \begin{pmatrix} -1 & 1 \\ 3 & 1 \end{pmatrix}$
$\lambda_2 = 2$	$E_{\lambda_2} = \text{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$	$A = Q \begin{pmatrix} -2 & \\ & 2 \end{pmatrix} Q^{-1}$

Therefore,

$$e^{At} = Q \begin{pmatrix} e^{-2t} & 0 \\ 0 & e^{2t} \end{pmatrix} Q^{-1} = \overbrace{\begin{pmatrix} -1 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} e^{-2t} & 0 \\ 0 & e^{2t} \end{pmatrix}}^{\begin{pmatrix} e^{-2t} & e^{2t} \\ 3e^{-2t} & e^{2t} \end{pmatrix}} \begin{pmatrix} -1/4 & 1/4 \\ 3/4 & 1/4 \end{pmatrix}$$

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

The general solution to (#) is:

$$\vec{X}(t) = e^{At}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \iff \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 \begin{pmatrix} -\frac{1}{4}e^{-2t} + \frac{3}{4}e^{2t} \\ -\frac{3}{4}e^{-2t} + \frac{3}{4}e^{2t} \end{pmatrix} + c_2 \begin{pmatrix} \frac{1}{4}e^{-2t} + \frac{1}{4}e^{2t} \\ \frac{3}{4}e^{-2t} + \frac{1}{4}e^{2t} \end{pmatrix}$$

Q: Why does it work?

$$\begin{pmatrix} x \\ y \end{pmatrix}' = A \begin{pmatrix} x \\ y \end{pmatrix} = \boxed{Q \begin{pmatrix} -2 & \\ & 2 \end{pmatrix} Q^{-1}} \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \underbrace{\boxed{Q^{-1}\begin{pmatrix} x \\ y \end{pmatrix}'}}_{=\begin{pmatrix} u \\ v \end{pmatrix}'} = \begin{pmatrix} -2 & \\ & 2 \end{pmatrix} \underbrace{\boxed{Q^{-1}\begin{pmatrix} x \\ y \end{pmatrix}}}_{=\begin{pmatrix} u \\ v \end{pmatrix}}$$

If we do the change of variables $\begin{pmatrix} u \\ v \end{pmatrix} = Q^{-1}\begin{pmatrix} x \\ y \end{pmatrix}$, then we have

$$\begin{cases} u' = -2u \\ v' = 2v \end{cases} \quad \text{decoupled !!} \quad \Rightarrow \begin{cases} u = c_1 e^{-2t} \\ v = c_2 e^{2t} \end{cases} \quad c_1, c_2 \in \mathbb{R}.$$

Then back to $\begin{pmatrix} x \\ y \end{pmatrix}$-variable, we have

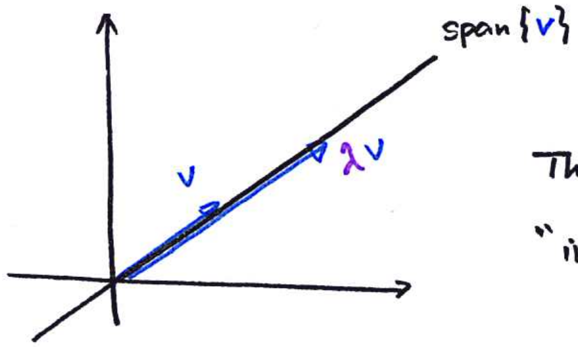
$$\begin{pmatrix} x \\ y \end{pmatrix} = Q \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -c_1 e^{-2t} + c_2 e^{2t} \\ 3c_1 e^{-2t} + c_2 e^{2t} \end{pmatrix}$$

(verify this gives the same answer as above !!)

Note: with different c_1, c_2 which are arbitrary!

Invariant Subspaces

If $v \in V$ is an eigenvector of $T: V \to V$ then $Tv = \lambda v$



<u>Note</u>: $T(\text{span}\{v\}) = \text{span}\{v\}$. $\underline{\lambda \neq 0}$.

The line span{v} is "preserved" or "invariant" under T.

<u>Defn</u>: Let $T: V \to V$ be linear. A <u>subspace</u> $W \subseteq V$ is a T-invariant subspace if $T(W) \subseteq W$.

<u>Note</u>: We may have $T(W) \neq W$!!

<u>Examples of T-invariant subspace</u>:

- $\{0\}$, V trivial subspaces
- $R(T)$, $N(T)$ (verify this!)
- E_λ eigenspaces.

  <u>Pf</u>: If $v \in E_\lambda$, then $Tv = \lambda v$ and $T(Tv) = T(\lambda v) = \lambda(Tv)$

  therefore $Tv \in E_\lambda$.

Q: Given $v \in V$, can we find a T-invariant subspace which contains v?

smallest

$$W = \text{span}\{v, Tv, T^2v, \dots\dots\}$$

<u>T-invariant</u>

(check this!)

T-cyclic subspace generated by v

Example: $T: P_3(\mathbb{R}) \to P_3(\mathbb{R})$, $T(f) = f''$.

The T-cyclic subspace generated by $x^3 \in P_3(\mathbb{R})$ is:

$$W = \operatorname{span}\{x^3, \underbrace{T(x^3)}_{6x}, \underbrace{T^2(x^3)}_{0}, \underbrace{\cdots\cdots}_{0}\} \quad \leftarrow \text{terminates!}$$

$$= \operatorname{span}\{x^3, 6x\} = \{c_1 x^3 + c_2 x \mid c_1, c_2 \in \mathbb{R}\}$$

Example: $T: \mathbb{R}^2 \to \mathbb{R}^2$ rotation by $2\pi/\sqrt{2}$.

For any $v \in \mathbb{R}^2$, $v \neq \vec{0}$, the T-cyclic subspace it generates is

$$W = \operatorname{span}\{\underbrace{v, Tv, T^2 v, \ldots\ldots}\} = \mathbb{R}^2$$

the vectors do not "terminate"
but the span remains unchanged
after finitely many terms!

- We care about T-invariant subspaces because we can <u>restrict</u> T to this subspace to get a "new" linear operator.

Given

$$\begin{array}{ccccl} T: & V & \longrightarrow & V & \text{linear operator on } V \\ & \cup\!\!\shortmid & & \cup\!\!\shortmid & \\ & W & \longrightarrow & W & T\text{-invariant subspace.} \end{array}$$

then $\boxed{T_W : W \longrightarrow W}$ linear operator on W

restriction of T to W

Lemma: char. poly. of T_W divides char. poly. of T.

Proof: Let $\gamma = \{v_1, \ldots, v_k\}$ be a basis for W and extend it to a basis $\beta = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ for V.

$$[T]_\beta = \begin{pmatrix} [T_W]_\gamma & * \\ O & * \end{pmatrix} \Rightarrow \det([T_W]_\gamma - \lambda I_k) \mid \det([T]_\beta - \lambda I_n)$$

$\because W$ is T-invariant.

$\square$

Properties of det(A) : A Review

(1) $\boxed{\det(A^t) = \det(A)}$ and $\boxed{\det(AB) = \det(A)\cdot\det(B)}$.

(2) $\det(A) \neq 0 \iff A$ is invertible.

Moreover, in this case, $\boxed{\det(A^{-1}) = \dfrac{1}{\det(A)}}$.

(3) If B is obtained by switching two rows (or columns) of A, then $\boxed{\det(B) = -\det(A)}$

(<u>Cor</u>: If A has two identical rows (or columns), then $\det(A) = 0$)

(4) If B is obtained by multiplying a row (or column) of A by a scalar c, then $\boxed{\det(B) = c\det(A)}$.

(<u>Cor</u>: $\det(cA) = c^n\det(A)$ where $A \in M_{n\times n}(\mathbb{F})$.)

(5) If B is obtained by adding a multiple of a row (or column) to another of A, then $\boxed{\det(B) = \det(A)}$.

(6) $\boxed{\det(A) = \text{product of its diagonal entries}}$ if A is upper (or lower) triangular.

[<u>Ex</u>: If $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ is a "block matrix", then where $M \in M_{n\times n}(\mathbb{F})$, $A \in M_{k\times k}(\mathbb{F})$.

$$\boxed{\det(M) = \det(A)\cdot\det(B)}$$]

<u>Caution!</u>

$$\det\begin{pmatrix} A & C \\ D & B \end{pmatrix} \neq \det(A)\det(B) - \det(C)\det(D)$$

(<u>Ex</u>: find an example.)

Example: $T : P_3(\mathbb{R}) \to P_3(\mathbb{R})$, $T(f) = f''$.

Recall $W = \text{span}\{x^3, x\}$ T-invariant (generated by x^3)

For $T_W : W \to W$ with basis $\gamma = \{x^3, x\}$

$$[T_W]_\gamma = \begin{pmatrix} 0 & 0 \\ 6 & 0 \end{pmatrix} \quad \text{since} \begin{cases} T_W(x^3) = 6x \in W \\ T_W(x) = 0 \in W \end{cases}$$

Extend γ to a basis $\beta = \{x^3, x, x^3+x^2, 1\}$ for $P_3(\mathbb{R})$

$$[T]_\beta = \left(\begin{array}{cc|cc} \multicolumn{2}{c|}{\overbrace{\begin{matrix} 0 & 0 \\ 6 & 0 \end{matrix}}^{[T_W]_\gamma}} & 0 & 0 \\ & & 6 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right) \quad \text{since} \begin{cases} T(x^3) = 6x \in W \\ T(x) = 0 \in W \\ T(x^3+x^2) = 6x + 2\cdot 1 \\ T(1) = 0 \end{cases}$$

Therefore, the characteristic polynomial of T is

$$\det([T]_\beta - \lambda I) = \det\begin{pmatrix} -\lambda & 0 & 0 & 0 \\ 6 & -\lambda & 6 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 2 & -\lambda \end{pmatrix} = \underbrace{\det\begin{pmatrix} -\lambda & 0 \\ 6 & -\lambda \end{pmatrix}}_{\det([T_W]_\gamma - \lambda I)} \cdot \det\begin{pmatrix} -\lambda & 0 \\ 2 & -\lambda \end{pmatrix}$$

"block matrix"!!

Hence,

$$\det([T_W]_\gamma - \lambda I) = \lambda^2 \mid \lambda^4 = \det([T]_\beta - \lambda I).$$

$\square$

Cayley-Hamilton Theorems

Theorem (Cayley-Hamilton)

(matrix form) Let $f(\lambda)$ be the characteristic polynomial of $A \in M_{n\times n}(\mathbb{F})$.

Then, $\boxed{f(A) = O}$, ie. A "satisfies" the char. equation.

(operator form) Let $f(\lambda)$ be the char. poly. of $T : V \to V$ ($\dim V < +\infty$).

Then, $\boxed{f(T) = O}$.

zero transformation

Example: Let $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$.

Char. Poly. $= f(\lambda) = \det(A - \lambda I) = \det\begin{pmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 + 4$

i.e. $f(\lambda) = \lambda^2 - 2\lambda + 5$.

Direct calculation:

$$A^2 = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 4 \\ -4 & -3 \end{pmatrix}$$

Hence,

$$f(A) = A^2 - 2A + 5I = \begin{pmatrix} -3 & 4 \\ -4 & -3 \end{pmatrix} - 2\cdot\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = O \qquad \underline{\text{zero matrix!}}$$

Application: We can make use of this to find A^{-1}. (if A is invertible)

$$\boxed{A^2 - 2A + 5I = O} \xrightarrow[\text{by } A^{-1}]{\text{multiply}} \boxed{A - 2I + 5A^{-1} = O}$$

rearrange $\Rightarrow$ $A^{-1} = -\frac{1}{5}A + \frac{2}{5}I = -\frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} + \frac{2}{5}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix}$

$\square$

To prove Cayley-Hamilton Theorem, we need the following:

Lemma: Let $T: V \to V$ be a linear operator ($\dim V < +\infty$).

$W = \text{span}\{v, Tv, T^2v, \ldots\}$ T-cyclic subspace gen. by $v \neq \vec{0}$

Suppose $k = \dim W$. Then,

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

(b) If $\boxed{a_0 v + a_1 Tv + \cdots + a_{k-1}T^{k-1}v + T^k v = 0}$,

then char. poly. of $T_W = f(\lambda) = \underline{(-1)^k(a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k)}$.

Example: $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -b+c \\ a+c \\ 3c \end{pmatrix}$.

W = T-cyclic subspace generated by $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

$T(e_1) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$; $T^2(e_1) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = -e_1$

Thus, $\dim W = 2$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} = \{ e_1, Te_1 \}$ basis.

Since $T^2 e_1 = -e_1 \Rightarrow e_1 + T^2 e_1 = 0$. By Lemma (b),

char. poly. of $T_W = 1 + \lambda^2$. $\left[\underline{\text{Check}}: [T_W]_\beta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right]$.

Proof of Lemma:

(a) Let j be the largest integer st.

$\beta = \{ v, Tv, \dots, T^{j-1} v \}$ is <u>linearly independent</u> $\left(\substack{\text{Note:} \\ j \geq 1} \right)$

$\boxed{\text{Claim: } j = k = \dim W.}$ $\Rightarrow$ (a)

Let $Z = \operatorname{span} \beta$. We will show that $Z = W$

Clearly, $Z \subseteq W$. To prove $W \subseteq Z$, it suffices to prove that Z is T-invariant (since W is the <u>smallest</u> T-invariant subspace containing v).

Pick any $w \in Z$, $\exists\ a_0, \dots, a_{j-1} \in \mathbb{F}$ st.

$$w = a_0 v + a_1 Tv + \cdots + a_{j-1} T^{j-1} v$$

$$\Rightarrow \quad Tw = \underbrace{a_0 Tv + a_1 T^2 v + \cdots}_{\in Z} + \underbrace{a_{j-1} T^j v}_{\in Z \text{ by the choice of } j}$$

So $Tw \in Z$. We are done!

(b) $T^k v \in Z = \text{span}\,\beta$

$$\Rightarrow \quad T^k v = -a_0 v - a_1 T v - a_2 T^2 v - \cdots - a_{k-1} T^{k-1} v$$

for some $a_i \in \mathbb{F}$.

Therefore, in the basis $\beta = \{v, Tv, T^2 v, \ldots, T^{k-1} v\}$ for W

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}$$

and the char. poly. is given by

$$f(\lambda) = \det([T_W]_\beta - \lambda I) = \det \begin{pmatrix} -\lambda & 0 & \cdots & 0 & -a_0 \\ 1 & -\lambda & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & -\lambda & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} - \lambda \end{pmatrix}$$

$$\boxed{\text{Ex: By induction on } k} = (-1)^k (a_0 + a_1 \lambda + \cdots + a_{k-1} \lambda^{k-1} + \lambda^k)$$

□

Proof of Cayley-Hamilton Theorem: (operator form)

Need to show $f(T)(v) = \vec{0}$ for all $v \in V$.

WLOG, assume $v \neq \vec{0}$, and let $W = T$-cyclic subspace gen. by v.

with $\dim W = k$.

Denote $f_W(\lambda)$ as the char. poly. of T_W.

Previous lemma (b) $\Rightarrow$ $\boxed{f_W(T)(v) = \vec{0}}$ (why?)

An earlier lemma $\Rightarrow$ $f_W(\lambda) \mid f(\lambda)$

$$\Rightarrow f(T)(v) = \vec{0}.$$

We have proved the theorem since v is arbitrary.

□

[Caution: $f_W(T) \neq O$, it only gives $\vec{0}$ when acting on v.]