Linear Algebra II

Minerva

2020-2nd

Contents

Chapter 1

Jordan form and rational form

1.1 What is a Jordan form?

1.1.1 Motivation

We assume that $\dim V \leq \infty, T : V \to V$ and $ch_T(x) = \prod_{i=1}^{k} (x - \lambda_i)^{m_i}$ splits over F. Previously, we have consider the case where $\dim E_{\lambda_i} = m_i$ for all i. In such a case, there exists a basis $\mathcal{B}$ for V such that $[T]_{\mathcal{B}}$ is diagonal. However, if $\dim E_{\lambda_i} < m_i$ for some i, can we find a "nice basis" such that $[T]_{\mathcal{B}}$ is "simple" and easy to do computation using it?

We give the result first.

1.1.2 Goal

If $ch_T(x)$ splits over F, then there exists a basis $\mathcal{B}$ such that $[T]_{\mathcal{B}}$ is of the form

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & & O \\ & A_2 & & \\ & & \ddots & \\ O & & & A_k \end{pmatrix}$$

where

$$A_i = \begin{pmatrix} \lambda_i & 1 & & & O \\ 0 & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ O & & & & \lambda_i \end{pmatrix} \tag{$*$}$$

Such a matrix $[T]_{\mathcal{B}}$ is called a **Jordan normal form** (or **Jordan canonical form**) of T. Each matrix of the form $(*)$ is called a **Jordan block** and $\mathcal{B}$ is called a **Jordan canonical basis**. Notice that a $[T]_{\mathcal{B}}$ is easy to do computation.

- Since $[T]_{\mathcal{B}}$ is block diagonal matrix, we have

$$[T]_{\mathcal{B}}^n = \begin{pmatrix} A_1^n & & & O \\ & A_2^n & & \\ & & \ddots & \\ O & & & A_k^n \end{pmatrix}$$

- $A_i^k = (\lambda_i I + N)^k$ where N is matrix with entry on counter diagonal is 1 and others are 0.
  - $\lambda_i I$ and N commute
  - $N^{m_i} = 0$

  So we have

$$A_i^k = (\lambda_i + N)^k = \sum_{j=0}^{k} \binom{k}{j} \lambda_i^{k-j} N^j = \sum_{0 \le j < m_i} \binom{k}{j} \lambda_i^{k-j} N^j$$
$$= \begin{pmatrix} z_{0k} & z_{1k} & z_{2k} & \cdots & \\ 0 & z_{0k} & \ddots & & \\ & & \ddots & & \vdots \\ & & & z_{0k} & z_{1k} \\ O & & & 0 & z_{0k} \end{pmatrix} \quad \left(\text{where } z_{jk} = \binom{k}{j} \lambda_i^{k-j} \right)$$

1.2 Review

In this section, we will review the knowledge what we had learned and we will use it when we prove Jordan normal form.

1.2.1 T-invariant subspace

Definition 1.2.1 (T-invariant subspace)**.** If $T : V \to V$ is a linear operator. A subspace of W of V is a T**-invariant subspace** if $T(W) \subseteq W$

Example 1.2.1. For any $f(x) \in F[x]$, $\ker f(T)$ is a T-invariant subspace.

Definition 1.2.2. Let $v \in V$. The subspace $Z(v;T) := \operatorname{span}\{T^k(v) : k \in \mathbb{N}_0\}$ is called the **cyclic T-invariant subspace** generated by v

Theorem 1.2.1. If $k = \dim Z(v;T) < \infty$, then

- $\{v, T(v), ..., T^{k-1}(v)\}$ is a basis for $Z(v;T)$
- If $a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$, then

$$ch_{T|_{Z(v;T)}} = x^k + a_{k-1} x^{k-1} + \cdots a_1 x + a_0$$

Theorem 1.2.2. Assume $\dim V \le \infty$. Let W be a T-invariant subspace of V. Then $ch_W | ch_T$

1.2.2 Direct sum

Definition 1.2.3. Let $W_1, ..., W_k$ be subspace of V. We say V is the **direct sum** of $W_1, ..., W_k$ if $V = W_1 + W_2 + \cdots W_k$ and $W_i \cap \sum_{j \ne i} W_j = \{0\}$
If V is direct sum of $W_1, ..., W_k$, we write $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$

Property 1.2.1. TFAE

- $V = W_1 \oplus \cdots \oplus W_k$
- $V = W_1 + \cdots + W_k$ and if $v_1 + \cdots + v_k = 0$ then $v_j = 0\ \forall j$.

- Each $v \in V$ can be written as $v = v_1 + \cdots + v_k$ for some $v_j \in W_j$ uniquely.
- If $\mathcal{B}_j$ is a basis for W_j, $j = 1, \ldots, k$, then $\mathcal{B} = \bigcup_{j=1}^{k} \mathcal{B}_j$ is a basis for V.
- $\exists$ basis $\mathcal{B}_j$ for W_j such that $\mathcal{B} = \bigcup_{j=1}^{k} \mathcal{B}_j$ is a basis for V.

Theorem 1.2.3. Assume that $\dim V < \infty$ and $V = W_1 \oplus \cdots \oplus W_k$. Then

$$ch_T(x) = \prod_{i=1}^{k} ch_{T|_W}(x)$$

Also, if $\mathcal{B}_i$ is a basis for W_i, then $\mathcal{B} = \bigcup\limits_{i=1}^{k} \mathcal{B}_i$ is a basis for V and

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T|_{W_1}]_{\mathcal{B}_1} & & \\ & \ddots & \\ & & [T|_{W_k}]_{\mathcal{B}_k} \end{pmatrix}$$

Remark 1.2.1. Thus, to prove that a Jordan form exists for T. We will prove that $\exists T$-invariant subspaces $W_1, W_2, ..., W_k$ such that $V = W_1 \oplus \cdots W_k$ and each W_i has a basis $\mathcal{B}_i$ such that

$$[T|_{W_i}]_{\mathcal{B}_i} = \begin{pmatrix} \lambda_i & 1 & & & O \\ 0 & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ O & & & & \lambda_i \end{pmatrix}$$

1.2.3 Polynomial rings

Theorem 1.2.4. If $f(x), g(x) \in F[x]$ and $g(x) \neq 0$, then there exists unique polynomial $q(x)$ and $r(x)$ such that

$$f(x) = q(x)g(x) + r(x)$$

and $r(x) = 0$ or $\deg r(x) < \deg g(x)$ (which means $F[x]$ is ED)

Definition 1.2.4. A nonempty set I of $F[x]$ is said to be an **ideal** if

- $f(x), g(x) \in I \implies f(x) - g(x) \in I$
- If $g(x) \in I$, then $f(x)g(x) \in I \ \forall f(x) \in F[x]$

Example 1.2.2. Let $T : V \to V$ be a linear operator.

- $$I = \{f(x) \in F[x] : f(T) = 0\}$$
  - If $f(x), g \in I$ i.e. $f(T) = g(T) = 0$, then $f(T) - g(T) = 0 \implies f - g \in I$
  - If $g(x) \in I$ i.e. if $g(T) = 0$, then $f(T) \cdot g(T) = 0 \implies fg \in I$

  Hence, I is an ideal in $F[x]$.

- Given $v \in V$, the set
$$I_T(v) = \{f(x) \in F[x] : f(T)v = 0\}$$
is an ideal

- Let W be a T-invariant subspace, then
$$I_{v,W} \text{ or } I_T(v, W) := \{f(x) \in F[x] : f(T)v \in W\}$$
is an ideal.

Theorem 1.2.5. If I is an ideal of $F[x]$, then $\exists$ a polynomial $g(x) \in F[x]$ such that
$$I = \{f(x)g(x) : f(x) \in F[x]\} = (g(x))$$
(Which means $F[x]$ is PID.)

Remark 1.2.2. Note that $g(x)$ has the smallest degree among all nonzero elements of I.

Definition 1.2.5 (principal ideal)**.** We say I is the **principal ideal** generated by $g(x)$ if $I = (g(x))$

Remark 1.2.3. $T : V \to V$. Recall that $I = \{f(x) \in F[x] | f(T) = 0\}$ is a ideal. Then the minimal polynomial $m_T(x)$ is defined to be the monic polynomial that generates I.

Definition 1.2.6. Let $f(x), g(x) \in F[x]$. If $h(x)$ is a polynomial such that $(h(x)) = (f(x)) + (g(x))$, then we say $g(x)$ is a **greatest common divisor**(GCD) of $f(x)$ and $g(x)$. If $(f(x)) + (g(x)) = (1)$, then we say f, g are relatively prime.

Definition 1.2.7. A nonconstant polynomial $f(x) \in F[x]$ is **irreducible** over if "$f(x) = g(x)h(x)$ for $g(x), h(x) \in F[x]$, then one of $g(x), h(x)$ is constant"

Compare $\mathbb{Z}$ and $F[x]$

	$\mathbb{Z}$	$F[x]$
ideal	$n\mathbb{Z}$	$(g(x))$
GCD	$n\mathbb{Z} + m\mathbb{Z} = \gcd(m, n)\mathbb{Z}$	$(f) + (g) = (\gcd(f, g))$
irreducible	prime number	irreducible polynomial
prime	$p \mid ab \rightsquigarrow p \mid a$ or $p \mid b$	f : irr, $f \mid gh \rightsquigarrow f \mid g$ or $f \mid h$
UFD	Fundamental theorem of arithmetic	$f(x) = ap_1(x)^{n_1} \cdots p_k(x)^{n_k}$ with unique decomposition

1.2.4 Kernel decomposition theorem

Theorem 1.2.6. (kernel decomposition theorem) $t : V \to V$. If $f(x)$ and $g(x)$ are relatively prime, then
$$\ker f(T)g(T) = \ker f(T) \oplus \ker g(T)$$

Corollary 1.2.1. Assume $\dim V < \infty$. Then $T : V \to V$ is diagonalizable $\iff$ $m_T(x)$ splits into a product of distinct linear factors over F.

1.3 Generalized eigenspace

1.3.1 Motivation and definition

We continue to prove Jordan form. Assume that $\dim V < \infty$. $T : V \to V$ and $ch_T(x)$ splits over F, say

$$ch_T(x) = \prod_{i=1}^{k} (x - \lambda_i)^{m_i}$$

where λ_i are distinct.
By Cayley-Hamilton theorem ($ch_T(T) = 0$).

$$V = \ker ch_T(T) = \ker \prod_{i=1}^{k} (T - \lambda_i I)^{m_i}$$

Then by the kernel decomposition theorem

$$V = \ker(T - \lambda_1)^{m_1} \oplus \cdots \oplus \ker(T - \lambda_k I)^{m_k} = \bigoplus_{i=1}^{k} \ker(T - \lambda_i I)^{m_i}$$

So we want to research the property of $\ker(T - \lambda_i I)^{m_i}$
Claim: If $v \in \ker(T - \lambda_i I)^{m_i}$, then

$$I = \{f(x) \in F[x], f(T)v = 0\} = ((x - \lambda_i)^p)$$

for some $p \leq m_i$

Proof: Assume $g(x)$ is a polynomial such that $I = (g(x))$. Now, $v \in \ker(T - \lambda_i I)^{m_i} \implies (T - \lambda_i I)^{m_i} v = 0 \implies (x - \lambda_i)^{m_i} \in I \implies g(x) \big| (x - \lambda)^{m_i} \implies g(x) = (x - \lambda_i)^p$ for some $p \leq m_i$ □

Definition 1.3.1. Let λ be an eigenvalue of $T : V \to V$, ($\dim V$ may be ∞). The set

$$K_\lambda = \{v \in V : (T - \lambda I)^p v = 0 \text{ for some } p \geq 1\}$$

is called the **generalized eigenspace** corresponding to λ. A nonzero element v in K_λ is called a **generalized eigenvector**.

Example 1.3.1. $\begin{array}{cccc} T: & F[x] & \longrightarrow & F[x] \\ & f & \longmapsto & f' \end{array} \implies K_0 = F[x]$

Theorem 1.3.1. Let λ be an eigenvalue of $T : V \to V$ ($\dim V$ may be ∞). Then

(i) K_λ is a T-invariant subspace of V

(ii) For any $\mu \neq \lambda$, the restriction of $(T - \mu I)$ to K_λ is $1 - 1$

Remark 1.3.1. If V is finite dimensional, say λ has multiplicity m, then $K_\lambda = \ker(T - \lambda I)^m$ and K_λ is a T-invariant subspace since $\ker f(T)$ is T-invariant for any $f(x) \in F[x]$

Proof:

(i) We first prove that K_λ is a subspace.

Suppose that $v_1, v_2 \in K_\lambda$, say $(T-\lambda I)^{p_1}v_1 = (T-\lambda I)^{p_2} = 0$, let $p = \max(p_1, p_2)$, then $(T-\lambda I)^p(v_1 + cv_2) = 0$ and it clear $0 \in K_\lambda$

Hence, K_λ is a subspace and it is clear that K_λ is T-invariant subspace.

(ii) We need to show $\ker(T-\mu I) \cap K_\lambda = \{0\}$

Suppose thta $v \in \ker(T-\mu I) \cap K_\lambda$, say $(T-\lambda I)^p v = 0$

Since $(x-\mu)$ and $(x-\lambda)^p$ are relatively prime i.e. $(x-\mu) + ((x-\lambda)^p) = (1)$

$\implies \exists\, a(x), b(x) \in F[x]$ s.t. $1 = a(x)(x-\mu) + b(x)(x-\lambda)^p$

$\implies I = a(T)(T-\mu I) + b(T)(T-\lambda I)^p$

$\implies v = Iv = a(T)(T-\mu I)v + b(T)(T-\lambda I)^p v = 0$

Hence, $\ker(T-\mu I) \cap K_\lambda = \{0\}$

□

Theorem 1.3.2. Assume that $\dim V < \infty$ and $ch_T(x) = \prod_{i=1}^{k}(x-\lambda_i)^{m_i}$. Then $\dim K_i = m_i$

Proof: We have $V = \bigoplus_{i=1}^{k} K_{\lambda_i}$. Here each K_{λ_i} is a T-invariant subspace since $K_{\lambda_i} = \ker(T - \lambda_i I)^{m_i}$. Let $T_i : K_{\lambda_i} \to K_{\lambda_i}$ be the restriction of T to K_{λ_i}. By Theorem 1.2.3.

$$ch_T(x) = \prod_{i=1}^{k} ch_{T_i}(x)$$

Since λ_i is the only eigenvalue of T_i (by Theorem 1.3.1.(ii)). We have $ch_{T_i}(x) = (x-\lambda_i)^{n_i}$, where $n_i := \dim K_i$. Compare both side of

$$\prod_{i=1}^{k}(x-\lambda_i)^{m_i} = ch_T(x) = \prod_{i=1}^{k} ch_{T_i}(x) = \prod_{i=1}^{k}(x-\lambda_i)^{n_i}$$

then $m_i = n_i = \dim K_i$. □

Theorem 1.3.3. Assume $\dim V < \infty$. Let K_λ be the generalized eigenspace corresponding to an eigenvalue λ of T. Then $\exists v_1, ..., v_r \in K_\lambda$ such that

$$K_\lambda = Z(v_1; T) \oplus \cdots \oplus Z(v_r; T)$$

Moreover, let $s_i = \dim Z(v_i; T)$ and arrange the subscripts such that $s_1 \geq s_2 \geq \cdots \geq s_r$. Then the sequence $s_1, s_2, ..., s_r$ is unique.

We leave the proof in Appendix 2.1

Notation 1.3.1. If $I_T(v, W) = ((x-\lambda)^s)$ for some s , then denote $s(v, W)$ for this s.

Remark 1.3.2. Since $v_i \in K_\lambda$, we have $(T-\lambda I)^p(v_i) = 0 \implies I_T(v_i) = ((x-\lambda)^{s_i})$.

Proof: Since $\{v_i, T(v_i), ..., T^{s_i-1}(v_i)\}$ is a basis for $Z(v_i; T)$. Say $I_T(v_i) = ((x-\lambda)^s)$.

If $s \leq s_i$: Since $(T-\lambda I)^s(v_i) = 0 \implies v_i, T(v_i), ..., T^s(v_i)$ are linearly independent $(\to\leftarrow)$. Thus, $s \geq s_i$. On the other hand, there is a non-trivial relation

$$T^{s_i}(v_i) + a_{s_i-1}T^{s_i-1}(v_i) + \cdots a_1 T(v_i) + a_0 = 0$$

$\implies$ this polynomial in $I_T(v_i) \implies s \leq s_i$. Hence, $s = s_i$ □

Remark 1.3.3. Now we choose a basis $\mathcal{B}_i$ for $Z(v;T)$ to be

$$\mathcal{B} := \left\{(T-\lambda I)^{s_i-1}(v_i), ..., (T-\lambda I)v_i, v_i\right\}$$

We have

$$T\left((T-\lambda I)^j v_i\right) = (T-\lambda I)^{j+1} v_i + \lambda (T-\lambda I)^j v_i$$

$$\implies [T|_{Z(v_i;T)}] = \begin{pmatrix} \lambda & 1 & & & O \\ 0 & \lambda & 1 & & \\ & 0 & \lambda & & \\ & & & \ddots & 1 \\ O & & & & \lambda \end{pmatrix}$$

which is the form what we want.

1.3.2 Existence and uniqueness of Jordan normal form

Theorem 1.3.4 (Existence of Jordan normal form)**.**

Proof: If $ch_T(x) = \prod_{i=1}^{k} (x-\lambda)^{m_i}$ splits over F, then by Thm 1.3.2,

$$V = \bigoplus_{i=1}^{k} K_{\lambda_i} = \bigoplus_{i=1}^{k} \bigoplus_{j=1}^{r_i} Z(v_{ij};T)$$

for some $v_{ij} \in K_{\lambda_i}$. Choose a basis for V to be

$$\mathcal{B} = \bigsqcup_{i=1}^{k} \bigsqcup_{j=1}^{r_i} \mathcal{B}_{ij}$$

where $B_{ij} = \{T^\ell(v_{ij}) : 0 \le \ell \le s_{ij}-1\}$. Then

$$[T]_\mathcal{B} = \begin{pmatrix} J_1 & & & O \\ & J_2 & & \\ & & \ddots & \\ O & & & J_k \end{pmatrix}$$

where each J_m is of the form

$$J_i = \begin{pmatrix} \lambda_i & 1 & & & O \\ 0 & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ O & & & & \lambda_i \end{pmatrix}$$

Hence, $[T]_\mathcal{B}$ is the Jordan normal form of T. $\square$

Theorem 1.3.5 (Uniqueness of Jordan normal form)**.**

Proof: Let $\widetilde{T} := T - \lambda I$. Recall that

$$\mathcal{B} = \bigcup_{i=1}^{r_i} \underbrace{\left\{v_i, \widetilde{T}(v_i), ..., \widetilde{T}^{s_i-1}(v_i)\right\}}_{\text{basis for } Z(v_i;T)}$$

is a basis for K_λ.
Let's determine $\dim\ker \widetilde{T}^\ell$ for $\ell \geq 1$. Observe that for $1 \leq k \leq s_i$

$$\widetilde{T}^{s_i-k}(v_i) \in \ker \widetilde{T}^\ell \iff \widetilde{T}^\ell\left(\widetilde{T}^{s_i-k}(v_i)\right) = 0 \iff s_i - k + \ell \geq s_i \iff k \leq \ell$$

On other hand, the remaining vectors

$$B' := \bigcup_{i=1}^{r} \left\{v_i, ..., \widetilde{T}^{s_i-\ell-1}(v_i)\right\}$$

has linearly independent images in $\widetilde{T}^\ell$, since $\widetilde{T}^\ell(\mathcal{B}') \subseteq B'$
In summary

$$\text{basis for } Z(v_i;T) : \left\{ \underbrace{v_i, \widetilde{v}_i, ..., \widetilde{T}^{s_i-\ell-1}(v_i)}_{\text{form a basis for } \operatorname{Im}\widetilde{T}^\ell} \middle| \underbrace{\widetilde{T}^{s_i-\ell}(v_i), ..., \widetilde{T}^{s_i-1}(v_i)}_{\in\ker\widetilde{T}^\ell} \right\}$$

Hence,

$$\dim\ker\widetilde{T}^\ell = \sum_{i=1}^{r} \begin{cases} \ell & \text{if } s_i \geq \ell \\ s_i & \text{if } s_i < \ell \end{cases} = \sum_{i=1}^{r} \min(\ell, s_i)$$

$$\begin{aligned} \dim\ker\widetilde{T}^\ell - \dim\ker\widetilde{T}^{\ell-1} &= \sum_{i=1}^{r} (\min(\ell, s_i) - \min(\ell-1, s_i)) \\ &= \sum_{i=1}^{r} \begin{cases} s_i - s_i = 0 & \text{if } s_i \leq \ell - 1 \\ \ell - (\ell - 1) & \text{if } s_i \geq \ell \end{cases} = \#\{s_i \geq \ell\} \end{aligned}$$

$$\begin{aligned} \implies \#\{s_i = \ell\} &= \#\{s_i \geq \ell + 1\} - \#\{s_i \geq \ell\} \\ &= (\dim\ker\widetilde{T}^\ell - \dim\ker\widetilde{T}^{\ell-1}) - (\dim\ker\widetilde{T}^{\ell+1} - \dim\ker\widetilde{T}^\ell) \end{aligned}$$

RHS is a quantity intrinsic to T and is independent of choices of $v_1, v_2, ...$ $\implies$ The sequence $s_1, ...,$ is unique. $\square$

Remark 1.3.4. The proof of uniqueness can be visualized using the dot diagram.
Rule:

- r columns, each column represents on s_i
- The i-thm column has s_i dots representing vectors $\widetilde{T}^{s_i-1}(v_i), \widetilde{T}^{s_i-2}(v_i), ..., v_i$

Then first i rows forms a basis for $\ker\widetilde{T}^i$. The numbers of dots in the first ℓ rows $= \dim\ker\widetilde{T}^\ell$.
Thus, # on the ℓ-th row $= \dim\ker\widetilde{T}^\ell = \dim\ker\widetilde{T}^{\ell-1} = \#\{i : s_i = \ell\}$

Remark 1.3.5. # of Jordan block for $K_\lambda = \dim E_\lambda$

Recall that if λ is a eigenvalue with multiplity m, then by theorem 1.3.3, we can

Definition 1.3.2. A **partition** of a positive integer m is a non-increasing sequence of positive integers $s_1, ..., s_r$ such $s_1 + \cdots + s_r = m$
The # of partitions will be denoted by $p(m)$ called the **partition function**.

It's clear that for a given eigenvalue λ with multiplicity m,

$$\{\text{possible Jordan-forms for } K_\lambda\} \xleftrightarrow{1-1} \{\text{partition of } m\}$$

$\implies$ # of possible Jordan forms for $K_\lambda = p(m)$

Property 1.3.1. Given $f(x) = \prod_{i=1}^{k} (x - \lambda_i)^{m_i}$. There are $\prod_{i=1}^{k} p(m_i)$ possible Jordan-forms with char. poly. $f(x)$.

Problem: How many possible Jordan Form are there for given $ch_T(x)$ and $m_T(x)$.
Ans: Say $ch_T(x) = \prod_{i=1}^{k} (x - \lambda_i)^{m_i}, m_T = \prod_{i=1}^{k} (x - \lambda_i)^{n_i}$. Then number of possible Jordan forms

$$\prod_{i=1}^{k} \# \text{ of partitions of } m_i \text{ with largest part is } m_i$$

1.4 Rational canonical forms

1.4.1 Motivation and Goal

Note that in order for Jordan forms to exists, a prerequisite is that $ch_T(x)$ splits over F. However, there are F and T whose $ch_T(x)$ does not split over F

eg. $\begin{array}{cccc} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (a,b) & \longmapsto & (b,-a) \end{array} \implies [T]_e = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and the char. poly is $x^2 + 1$, which has no root in $\mathbb{R}$.
Goal: Find a simple matrix representation

Definition 1.4.1. Let $T : V \to V$ (V possibly ∞-dimensional) and $p(x)$ be an irreducible polynomial over F. We let K_p denote the subspace

$$K_p = \{v \in V : p(T)^n v = 0 \text{ for some } n \geq 1\}$$

Note that if λ is an eigenvalue, then $K_{x-\lambda}$ is a simply the generalized eigenspace K_λ.

For now on, we assume $\dim V < \infty$
Let $ch_T(x) = \prod_{i=1}^{k} p_i(x)^{m_i}$ be the (unique) factorization of $ch_T(x)$ into a product of irreducible polynomial. We have

$$V = \ker ch_T(T) = \bigoplus_{i=1}^{k} \ker p_i(x)^{m_i} = \bigoplus_{i=1}^{k} K_{p_i^{m_i}} = \bigoplus_{i=1}^{k} K_{p_i}$$

Note that each K_{p_i} is a T-invariant subspace.

Theorem 1.4.1. $\dim K_{p_i} = m_i \deg p_i$

Proof: Let $T_i = T|_{K_{p_i}}$. Then $K_{p_i} = \ker p_i(T)^{m_i}$.

Claim: $ch_{T_i}(x) = p_i(x)^{n_i}$ for some n_i

pf. Assume that $ch_{T_i} = p_i(x)^{n_i} g(x)$ with $(g, p_i) = 1$. Then

$$K_{p_i} = \underbrace{\ker p_i(T)^{n_i}}_{:=U_1} \oplus \underbrace{\ker g(T)}_{:=U_2}$$

It's an easy exercise to show $m_{T_i} = \operatorname{lcm}(m_{T_{U_1}}, m_{T_{U_2}})$. Now, since $K_{p_i} = \ker p_i(T)^{m_i}$, we have $m_{T_i}(x)|p_i(x)^{m_i}$. Thus, the minimal polynomial of T_{U_2} is $p_i(x)^s$ for some $s \geq 0$. But $(g, p_i) = 1 \implies$ minimal polynomial of T_{U_2} is $1 \implies g(x) = 1$ i.e. $ch_{T_i}(x) = p_i(x)^{n_i}$ for some n_i □

By Claim, we have

$$\prod_{i=1}^{k} p_i(x)^{n_i} = ch_T(x) = \prod_{i=1}^{k} ch_{T_i}(x) = \prod_{i=1}^{k} p_i(x)^{m_i} \implies n_i = m_i$$

□

Remark 1.4.1. Another proof in Theorem 1.4.1 (Field extension)
By field theory $\exists F'/F$ such that $ch_T(x)$ splits over F'. Extend the scalar of V to F' and denoted the new vector space by $V \otimes_F F'$. Now assume $ch_{T_i}(x) = p_i(x)^{m_i} g(x)$ for $(p_j, g) = 1$. Let λ be a root of $g(x)$ in F', which is in fact an eigenvalue and there is an eigenvector $v \neq 0$ in $K_{p_i} \otimes_F F'$ corresponding to λ. However, $v \notin K_{p_i} \otimes F'(p_i(T)^n(v) \neq 0)$ $(\rightarrow\leftarrow)$

Property 1.4.1. $T : V \to V$ is linear with $\dim_F V = n < \infty$. Then $m_T(x)|ch_T(x)|m_T(x)^n$. In particular, the irreducible factors of $m_T(x)$ is the same as the irreducible factors of $ch_T(x)$.

Proof: $m_T(x)|ch_T(x)$ OK! We prove $ch_T(x)|m_T(x)^n$:

Suppose $\deg m_T = d$, by $x^k - y^k = (x-y)(x^{k-1} + x^{k-2}y + \cdots + y^{k-1})$ and xI_N commute with $[T]_\beta$, we have

$$m_T(xI_n) = m_T(xI_n) - m_T([T]_\beta) = (xI_n - [T]_\beta)P \text{ for some } P \in M_n(F[x])$$

Take determinant in both side,

$$m_T(x)^n = \det(m_T(xI_n)) = \det(xI_n - [T]_\beta)\det(P) = \mathrm{ch}_T(x)\det P \implies ch_T(x)|m_T(x)^n$$

□

1.4.2 Existence and uniqueness of Rational form

Theorem 1.4.2. $T : V \to V, \dim V < \infty$. $p(x)$ is an irreducible factor of $ch_T(x)$. Then $\exists v_1, ..., v_r \in K_p$ such that

$$K_p = Z(v_1; T) \oplus \cdots \oplus Z(v_r; T)$$

Moreover, let s_i be the smallest integer such that $p(T)^{s_i}(v_i) = 0$ and arrange the subscripts such that $s_1 \geq s_2 \geq \cdots \geq s_r$. Then the sequence $s_1, ..., s_r$ is unique.

Remark 1.4.2.

(1) For $v \in K_p$, let s be the smallest integer such that $p(T)^s(v) = 0$, Then $\dim Z(v; T) = s \deg p$ and $I_v = (p(x)^s)$.

pf. In general, the characteristic polynomial if $T_{Z(v;T)}$ is the same as the minimal polynomial of $T_{Z(v;T)}$. Here the assumption that s is the smallest integer such that $p(T)^s(v) = 0$ means the minimal poly. of $T_{Z(v;T)}$ is $p(x)^s$ i.e. $I_v = (p(x)^s)$

(2) More generally, let W be a T-invariant subspace of K_p and s be the smallest integer such that $p^s(T)(v) \in W$. Then $I_{v,W} = (p(x)^s)$.

Notation 1.4.1. For $v \in K_p$ and W is a T-invariant subspace of K_p. If $I_{v,W} = (p(x)^s)$, then define $s(v, W) = s$

Outline of proof Theorem 1.4.2

(i) Let $W_0 = \{0\}$

(ii) For $i \geq 1$, assume that W_{i-1} have been defined. Choose $u \in K_p$ such that

$$s(u, W_{i-1}) := \max_{v \in K_p} s(v, W_{i-1}) := s_i$$

Claim: $\exists w \in W_{i-1}$ such that $p(T)^{s_i}(w) = p(T)^{s_i}(u)$

(iii) Let $v_i = u - w$ and claim :

- $W_{i-1} \cap Z(v_i; T) = \{0\}$
- $I_{v_i} = (p(x)^{s_i})$

Let $W_i = W_{i-1} \oplus Z(v_i; T)$

(iv) Repeat (ii),(iii) until $W_i = K_p$

We leave the detail proof in Appendix 4.1.2
Recall that is $\dim Z(v; T) = k$, then

$$\mathcal{B} = \{v, T(v), ..., T^{k-1}(v)\} \text{ is a basis for } Z(v; T)$$

We have $T^k(v) = -a_{k-1}T^{k-1}(v) - \cdots - a_1 T(v) - a_0 v$ for some a_j, then

$$[T|_{Z(v,T)}]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & 0 & & & \vdots \\ & & & \ddots & & \\ & & & & 0 & -a_{k-2} \\ 0 & & & & 1 & -a_{k-1} \end{pmatrix} \tag{*}$$

Definition 1.4.2. For a polynomial $f(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0$, define the **companion matrix** of f is a matrix in $(*)$

Corollary 1.4.1. Assume $\dim V < \infty$, $T : V \to T$. Then exists a basis $\mathcal{B}$ for T such that $[T]_{\mathcal{B}}$ is a block matrix of the form

$$\begin{pmatrix} C_1 & & O \\ & \ddots & \\ O & & C_m \end{pmatrix}$$

where each C_j is the companion matrix of $p(x)^s$ for some irreducible factor of $ch_T(x)$ and some $s \geq 1$. Moreover, C_i is unique up to permutation.

Definition 1.4.3. A matrix representation of T of the form is called a rational canonical form and $\mathcal{B}$ is called a rational canonical basis. The factors $p_i(x)^{s_i}$ are called the **elementary divisors of** T

Remark 1.4.3. There is another definition of a rational canonical form, where

$$T = \begin{pmatrix} C_1' & & O \\ & \ddots & \\ O & & C_k' \end{pmatrix}$$

and C_i; is the companion matrix of some f_i such that $f_i | f_{i+1}$ $\forall i = 1, ..., k-1$. The polynomials $f_i(x)$ are $f_i(x)$ are called the invariant factors of T.

1.5 Real Jordan Form

Let $A \in M_n(\mathbb{R})$, $ch_A(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_\ell)^{m_\ell}(x^2+a_1x+b_1)^{n_1} \cdots (x^2+a_kx+b_k)^{n_k}$ with $\alpha_i \pm \beta_i\sqrt{-1}$ are roots of $x^2+a_ix+b_i = 0$. Since $ch_T(x)$ is not splits over $\mathbb{R}$. How can we find a good basis for A ?

Theorem 1.5.1. Let $A \in M_n(\mathbb{R})$, $ch_A(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_\ell)^{m_\ell}(x^2+a_1x+b_1)^{n_1} \cdots (x^2+a_kx+b_k)^{n_k}$ with $\alpha_i \pm \beta_i\sqrt{-1}$ are roots of $x^2+a_ix+b_i = 0$. Then $\exists$ invertible $P \in M_n(\mathbb{R})$ s.t.

$$P^{-1}AP = \begin{pmatrix} I_{\lambda_1} & & & & & \\ & \ddots & & & & \\ & & I_{\lambda_\ell} & & & \\ & & & J_{\mu_1} & & \\ & & & & \ddots & \\ & & & & & J_{\mu_k} \end{pmatrix}$$

where $I_{\lambda_i} = J_1(\lambda_i) \oplus \cdots \oplus J_r(\lambda_i)$ with $J_k(\lambda_i)$ is Jordan blocks corresponding λ_i and $I_{\mu_j} = J_1(\mu_j) \oplus \cdots \oplus J_r(\mu_j)$ with $J_k(\mu_j)$ is form

$$J_k(\mu_j) = \begin{pmatrix} \alpha_j & \beta_j & 1 & 0 & & & & \\ -\beta_j & \alpha_j & 0 & 1 & & & & \\ & & \alpha_j & \beta_j & 1 & 0 & & \\ & & -\beta_j & \alpha_j & 0 & 1 & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & \ddots & \ddots & 1 & 0 \\ & & & & & \ddots & 0 & 1 \\ & & & & & & \alpha_j & \beta_j \\ & & & & & & -\beta_j & \alpha_j \end{pmatrix}$$

is called **real Jordan block**.

Before proving the theorem, we see some property on vector space over $\mathbb{C}$.

Property 1.5.1.

- For a vector $v = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$, we write $\overline{v} = \begin{pmatrix} \overline{z}_1 \\ \overline{z}_2 \\ \vdots \\ \overline{z}_n \end{pmatrix} \in \mathbb{C}^n$
- For subspace $W \subseteq \mathbb{C}^n$, we write $\overline{W} = \{\overline{w} : w \in W\}$

- If $W = \text{span}_{\mathbb{C}}\{w_1, ..., w_k\} \implies \overline{W} = \text{span}_{\mathbb{C}}\{\overline{w}_1, ..., \overline{w}_k\}$
- $v_1, ..., v_k \in \mathbb{C}^n$ are linearly independent, then $\overline{v}_1, ..., \overline{v}_k$ are linearly independent.
- $A \in M_{n\times n}(\mathbb{R}),\ \lambda \in \mathbb{C} \setminus \mathbb{R}$, then

$$W = \ker(A - \lambda I)^r \implies \overline{W} = \ker(A - \overline{\lambda} I)^r$$

  and $W \cap \overline{W} = \{0\}$

Back to proof of Theorem 1.5.1

Proof: For real part I_{λ_i} : OK! Now, we deal with I_{μ_j} part!
Consider A in $M_{n\times n}(\mathbb{C})$, then we focus on the eigenvalues $\lambda = \alpha + \beta\sqrt{-1}$ and $\overline{\lambda} = \alpha - \beta\sqrt{-1}$
By Jordan form theory, we can find $v_1, ..., v_r \in K_\lambda \subseteq \mathbb{C}^n$ s.t. $K_\lambda = \bigoplus_{i=1}^{r} Z(v_i; A)$
Looking one Jordan block, we have basis $\{(a - \lambda I)^{s-1}v, ..., (A - \lambda I)v, v\}$ in $\mathbb{C}^n$ and write $(A - \lambda I)^{i-1}v = w_i = x_i + y_i + \sqrt{-1}$. Since $Aw_i = \lambda w_i + w_{i+1}$ (assume $w_{s+1} = 0$) and compare real part and imaginary part, then

$$\begin{cases} Ax_i = \alpha x_i - \beta y_i + x_{i+1} \\ Ay_i = \beta x_i + \alpha y_i + y_{i+1} \end{cases} \text{for } i = 1, ..., s-1 \text{ and } \begin{cases} Ax_s = \alpha x_s - \beta y_s \\ Ay_i = \beta x_s + \alpha y_s \end{cases}$$

Since $\{w_s, \overline{w}_s, w_{s-1}, \overline{w}_{s-1}, ..., w_1, \overline{w}_1\}$ are linearly independent, then

$$\left\{ x_i = \frac{w_i + \overline{w}_i}{2}, y_i = \frac{w_i + \overline{w}_i}{2i} \middle| i = 1, ..., s \right\}$$

are linearly independent. Finally, we use basis $\{x_s, y_s, x_{s-1}, y_{s-1}, ..., x_1, y_1\}$ to write down matrix representation :

$$\begin{pmatrix} \alpha_j & \beta_j & 1 & 0 & & & & & \\ -\beta_j & \alpha_j & 0 & 1 & & & & & \\ & & \alpha_j & \beta_j & 1 & 0 & & & \\ & & -\beta_j & \alpha_j & 0 & 1 & & & \\ & & & & \ddots & \ddots & \ddots & & \\ & & & & & \ddots & \ddots & 1 & 0 \\ & & & & & & \ddots & 0 & 1 \\ & & & & & & & \alpha_j & \beta_j \\ & & & & & & & -\beta_j & \alpha_j \end{pmatrix}$$

□

Chapter 2

Matrix exponential

2.1 Definition

Definition 2.1.1. Let $A \in M_{n\times n}(\mathbb{C})$. Then the **exponential** of A denoted by e^A or $\exp A$ is defined to

$$e^A = I_n + \sum_{k=1}^{\infty} \frac{A^k}{k!}$$

Example 2.1.1.

- $A = \begin{pmatrix} \lambda_1 & & O \\ & \ddots & \\ O & & \lambda_k \end{pmatrix}$, then $e^A = \begin{pmatrix} e^{\lambda_1} & & O \\ & \ddots & \\ O & & e^{\lambda_k} \end{pmatrix}$

- Nilpotent matrix is easy to calculate, since it only finite sum.

The example show that when A is diagonal or nilpotent, it is easy to computes e^A. Thus the theory of Jordan forms will be very useful in computing e^A. More specifically, let Q be an invertible matrix such that $J = Q^{-1}AQ$ is the Jordan form of A.

Example 2.1.2. If $A = QJQ^{-1}$, then $e^A = Qe^JQ^{-1}$. Note

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix} \implies e^{Jz} = \begin{pmatrix} e^{J_1 z} & & \\ & \ddots & \\ & & e^{J_k z} \end{pmatrix}$$

Now, Write $J_i = \lambda_i I + N_i$ and notice that : if $AB = BA$, then $e^{A+B} = e^A e^B$, so

$$e^{J_i z} = e^{\lambda_i z I} e^{N_i z} = e^{\lambda_i z} \cdot e^{N_i z}$$

If N_i is $n \times n$ matrix, then $N_i^n = O$ and

$$e^{N_i z} = \sum_{k=0}^{\infty} \frac{1}{k!}(N_i z)^k = \sum_{k=0}^{n-1} \frac{1}{k!}(N_i z)^k = \begin{pmatrix} 1 & \frac{z}{1!} & \frac{z^2}{2!} & \cdots & \cdots & \frac{z^{n-1}}{(n-1)!} \\ 0 & 1 & \frac{z}{1!} & \cdots & \cdots & \frac{z^{n-2}}{(n-2)!} \\ 0 & 0 & 1 & \ddots & \cdots & \frac{z^{n-3}}{(n-3)!} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \ddots & \frac{z}{1!} \\ 0 & 0 & 0 & \cdots & \cdots & 1 \end{pmatrix}$$

If we consider the real Jordan form of A, so it is sufficiently calculate the exponent of real Jordan block J. If

$$J = \begin{pmatrix} D & I_2 & & & \\ & D & I_2 & & \\ & & D & \ddots & \\ & & & \ddots & I_2 \\ & & & & D \end{pmatrix} \in M_{2d\times 2d}(\mathbb{R}) \text{ where } D = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

Similarly to calculate exponent of Jordan form, we calculate $\exp(D)$ first.

$$D = \alpha I_2 + \beta \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{:=K} \implies \exp(D) = \exp(\alpha I_2)\exp(\beta K) = e^{\alpha}\exp(\beta K)$$

$$K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ K^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},\ K^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\ K^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

(Note: K have similar structure with $\sqrt{-1}$). So

$$\exp(\beta K) = \sum_{s=0}^{\infty} \frac{1}{s!}(\beta K)^s = \sum_{s=0}^{\infty} \frac{(-1)^s \beta^{2s+1}}{(2s+1)!} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \sum_{s=0}^{\infty} \frac{(-1)^s \beta^{2s}}{(2s)!} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix}$$

$$\implies e^D = e^{\alpha} \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix}.\ \text{Similarly } e^{Jz} = e^{\alpha z} \begin{pmatrix} \cos(\beta z) & \sin(\beta z) \\ -\sin(\beta z) & \cos(\beta z) \end{pmatrix}$$

Hence,

$$e^{Mz} = \operatorname{diag}\{e^{Jz}, ..., e^{Jz}\} \begin{pmatrix} I_2 & \frac{z}{1!}I_2 & \frac{z^2}{2!}I_2 & \cdots & \frac{z^{d-1}}{(d-1)!}I_2 \\ & I_2 & \frac{z}{1!}I_2 & \cdots & \frac{z^{d-2}}{(d-2)!}I_2 \\ & & I_2 & \cdots & \frac{z^{d-3}}{(d-3)!}I_2 \\ & & & \ddots & \vdots \\ & & & & I_2 \end{pmatrix}$$

2.2 System of linear differential equation with constant coefficient

Theorem 2.2.1. Let $A \in M_{n\times n}(\mathbb{C})$. Then the unique solution of

$$y'(z) = \begin{pmatrix} y_1'(z) \\ \vdots \\ y_n'(z) \end{pmatrix} = A \begin{pmatrix} y_1(z) \\ \vdots \\ y_n(z) \end{pmatrix} =: Ay(z)$$

with the initial condition $y(0) = y_0$ is $y(z) = e^{Az}y_0$

Before prove this theorem, we need define a norm on matrix space to describe the limits of matrix.

Definition 2.2.1 (max norm). For $A \in M_n(\mathbb{C})$, define the **max norm** of A

$$\|A\| = \max_{1\le i,j\le n} |a_{ij}|$$

Property 2.2.1. For all $A, B \in M_n(\mathbb{C})$,

$$\|A \cdot B\| \leq n \cdot \|A\| \cdot \|B\|$$

Now, we can prove Theorem 2.2.1

Proof: We prove the case in $\mathbb{R}$ and the case in $\mathbb{C}$ is similar.

- **Existence:** $y(z) = e^{Az} \cdot y_0$ is a solution

  Claim: $\frac{d}{dz}(e^{Az}) = Ae^{Az}$

  pf. For $h \in \mathbb{R}$, we have

$$\frac{e^{A(z+h)} - e^{Az}}{h} - Ae^{Az} = e^{Az}\left(\frac{e^{Ah} - I_n - Ah}{h}\right)$$

  Observe for all $x \in \mathbb{R}$, we have

$$|e^x - 1 - x| = \left|\sum_{k=2}^{\infty} \frac{x^k}{k!}\right| \leq \sum_{k=2}^{\infty} \frac{|x|^k}{k!} \leq |x| \sum_{k=1}^{\infty} \frac{|x|^k}{k!} = |x|\left(e^{|x|} - 1\right)$$

  With similar ideal, in $M_n(\mathbb{R})$, we have

$$\begin{aligned} \|e^B - I_n - B\| &= \left\|\sum_{k=2}^{\infty} \frac{B^k}{k!}\right\| \leq \frac{1}{n} \sum_{k=2}^{\infty} \frac{(n\|B\|)^k}{k!} \qquad \text{(By Property 2.2.1)} \\ &= \frac{1}{n}\left(e^{n\|B\|} - 1 - n\|B\|\right) \leq \left(e^{n\|B\|} - 1\right)\|B\| \end{aligned}$$

  Apply $B = Ah$, then

$$\left\|\frac{e^{Ah} - I_n - Ah}{h}\right\| \leq \frac{1}{h}\left(e^{n\|Ah\|} - 1\right)\|Ah\| = \left(e^{nh\|A\|} - 1\right)\|A\| \longrightarrow 0 \text{ as } h \longrightarrow 0$$

$$\Longrightarrow \lim_{h \to 0}\left(\frac{e^{A(z+h)} - e^{Az}}{h}h - Ae^{Az}\right) = 0$$

- **Uniqueness of solution** :

  Suppose $x(z), y(z)$ are two solutions. Consider $u(z) = x(z) - y(z)$. Then $u(z)$ satisfies $u'(z) = Au(z)$ and $u(0) = 0$.

$$\begin{aligned} \Longrightarrow u(t) &= A \cdot \underbrace{\int_0^t u(s_1)\ ds_2}_{\text{entrywise integral}} = A \int_0^t A \int_0^{s_1} u(s_2)\ ds_2 ds_1 = \cdots \\ &= A^k \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} u(s_k) ds_k ds_{k-1} \cdots ds_1 \end{aligned}$$

  We will prove that $u(z) = 0$ on any closed interval $[a, b] \subset \mathbb{R}$. Suppose $t \in [a, b]$ and let $M = \max_{t \in [a,b]} \|u(t)\|$, where $\|u(t)\| = \max_{1 \leq i \leq n} |u_i(t)|$ (Since $u_i(x)$ are continuous on $[a, b]$, thus M is exists). Thus, for any $t \in [a, b]$, we have

$$\begin{aligned} \|u(t)\| &= \left\|A^k \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} u(s_k) ds_k ds_{k-1} \cdots ds_1\right\| \\ &\leq n \cdot \|A^k\| \cdot \left\|\int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} u(s_k) ds_k ds_{k-1} \cdots ds_1\right\| \\ &\leq (n\|A\|)^k M \left|\int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} 1 \cdot ds_k ds_{k-1} \cdots ds_1\right| \leq (n\|A\|)^k M \frac{|b-a|^k}{k!} \longrightarrow 0 \end{aligned}$$

as $k \longrightarrow \infty$. Thus $\|u(t)\| = 0$ for all $t \in [a,b] \implies u(t) = 0$ for all $t \in [a,b] \implies u(t) = 0$ on whole $\mathbb{R} \implies x(t) = y(t)\ \forall t \in \mathbb{R}$.

$\square$

Example 2.2.1. Solve $y''(z) + 2y'(z) + y(z) = 0$.

Let $y_1(z) = y(z), y_2(z) = y'(z)$, then

$$\begin{pmatrix} y_1'(z) \\ y_2'(z) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}}_{:=A} \begin{pmatrix} y_1(z) \\ y_2(z) \end{pmatrix} \implies A = -I + \underbrace{\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}}_{\text{nilpotent}}$$

$$e^{Az} = e^{-Iz} e^{\left(\begin{smallmatrix} 1 & 1 \\ -1 & -1 \end{smallmatrix}\right) z} = e^{z}\left(I + \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} z \right) = \begin{pmatrix} (1+z)e^{z} & ze^{z} \\ -ze^{z} & (1-z)e^{z} \end{pmatrix}$$

Thus, solution are $y(z) = c_1(1+z)e^{-z} + c_2 z e^{-z} = c_1' e^{-z} + c_2' z e^{-z}$

If $y_1(z), y_2(z)$ are solutions of a linear differential equation, then $y_1(z) + cy_2(z)$ is also the solution. Thus, we usually write solutions of the DE as a linear combination of functions form a basis. For example, in example above, $c_1 e^{-z} + c_2 z e^{-z}$ is the general solution of $y''(z) + 2y'(z) + y(z) = 0$. If some additional conditions are given, e.g. $y(0) = a_1, y'(0) = a_2$. Then c_1, c_2 will be determined by additional condition.

Example 2.2.2. Solve $y''(z) + 9y(z) = 0$

Proof: We have $\begin{pmatrix} y(z) \\ y'(z) \end{pmatrix}' = \underbrace{\begin{pmatrix} 0 & 1 \\ -9 & 0 \end{pmatrix}}_{:=A} \begin{pmatrix} y(z) \\ y'(z) \end{pmatrix}$. The eigenvalues of A are $\pm 3i$. and $\begin{pmatrix} 1 \\ 3i \end{pmatrix}, \begin{pmatrix} 1 \\ -3i \end{pmatrix}$ are eigenvalue corresponding to $3i, -3i$, respectively. Thus, setting $Q = \begin{pmatrix} 1 & 1 \\ 3i & -3i \end{pmatrix}$, we have $Q^{-1}AQ = \begin{pmatrix} 3i & 0 \\ 0 & -3i \end{pmatrix} =: J$

$$\implies e^{Az} = Qe^{Jz}Q^{-1} = Q \begin{pmatrix} e^{3iz} & 0 \\ 0 & e^{-3iz} \end{pmatrix} Q^{-1} = \begin{pmatrix} \frac{e^{3iz}+e^{-3iz}}{2} & \frac{e^{3iz}-e^{-3iz}}{6i} \\ \frac{3i(e^{3iz}-e^{-3iz})}{2} & \frac{e^{3iz}+e^{-3iz}}{2} \end{pmatrix} = \begin{pmatrix} \cos 3z & \frac{1}{3}\sin 3z \\ -3\sin 3z & \cos 3z \end{pmatrix}$$

If the initial conditions are given as $y(0) = a_1, y'(0) = a_2$, then the solution is

$$a_1 \cos 3z + \frac{a_2}{3} \sin 3z$$

$\square$

2.3 Matrix limits

Theorem 2.3.1. Let $A \in M_{n\times n}(\mathbb{C})$. Then $\lim_{k\to\infty} A^k$ exists if and only if

- All eigenvalue of A are in
$$\{z \in \mathbb{C} : |z| < 1\} \cup \{1\}$$
- If 1 is an eigenvalue, then $\dim E_1 =$ multiplicity of 1.

Proof: The proof is left as an exercise to the reader. $\square$

Question: Suppose that each year 90% of city population stay in the city and 10% move to suburbs. 80% of suburbs population stay in suburbs and 20% move to city. Assume the number of all population will not change, will the populations of city and suburbs stabilize, oscillate or ?

Solution: Let a_n, b_n be the populations of city, suburbs, respectively in year n. We have

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}$$

Since $\begin{pmatrix} 1 & 1 \end{pmatrix} A = \begin{pmatrix} 1 & 1 \end{pmatrix} \implies 1$ is an eigenvalue and $\operatorname{tr} A = 1.7 \implies 0.7$ is another eigenvalue. Let v_1, v_2 be an eigenvector corresponding to $1, 0.7$, respectively. Let $Q = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$. Then $Q^{-1}AQ^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 0.7 \end{pmatrix} \implies \lim\limits_{k\to\infty} Q^{-1}A^kQ = Q^{-1} \lim\limits_{k\to\infty} \begin{pmatrix} 1^n & 0 \\ 0 & 0.7^n \end{pmatrix} Q = \begin{pmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{pmatrix}$ (Notice that $\begin{pmatrix} 2/3 & 1/3 \end{pmatrix}$ is an eigenvalue corresponding to 1)

$$\implies \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 2(a_0+b_0)/3 \\ (a_0+b_0)/3 \end{pmatrix}$$

Chapter 3

Inner products

Through out the chapter, we assume that $F = \mathbb{R}$ or $\mathbb{C}$.

3.1 Definition

Definition 3.1.1. Let V be a vector space over F. An **inner product** $\langle \cdot, \cdot \rangle : V \times V \to F$ is a function such that $\forall\ x, y, z \in V, \forall\ c \in F$

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle cx, y \rangle = c \langle x, y \rangle$
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- $\langle x, x \rangle > 0$ if $x \neq 0$
  (Note that the condition $\langle x, x \rangle > 0$ implicitly say that $\langle x, x \rangle \in \mathbb{R}$)

Theorem 3.1.1. Let V be an inner product space. Then $\forall\ x, y, z \in V$, $\forall\ c \in F$

- $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$
- $\langle x, cy \rangle = \bar{c} \langle x, y \rangle$
- $\langle x, 0 \rangle = \langle 0, x \rangle = 0$
- $\langle x, x \rangle = 0 \iff x = 0$
- If $\langle x, y \rangle = \langle x, z \rangle$ holds for all $x \in V$ then $y = z$.

Remark 3.1.1. A function $h : V \times V \to F$ is said to be

- **linear** in the first argument if

$$h(x + y, z) = h(x, z) + h(y, z),\ \ h(cx, y) = ch(x, y)$$

- **semilinear** in the second argument if

$$h(x, y + z) = h(x, y) + h(x, z),\ \ h(x, cy) = \bar{c}h(x, y)$$

- **sesquilinear** or **Hermitian** if it is linear in the first argument and semilinear in the second argument.

- **positive definite** if $h(x,x) > 0 \ \forall \ x \neq 0$ and $h(0,0) = 0$
- **semi-positive definite** if $h(x,x) \geq 0 \ \forall x \in V$
- **nondegenarate** if $\langle x, y\rangle = \langle x, z\rangle \ \forall \ x \in V \iff y = z$

Thus, an inner product can also be defined as a positive definite Hermition form.

Example 3.1.1. Define $\langle \cdot, \cdot \rangle$ on F^n by

$$\langle (a_1, ..., a_n), (b_1, ..., b_n) \rangle = \sum_{i=1}^{n} a_i \overline{b_i} \text{ or say } \langle x, y \rangle = x^T \overline{y}$$

which is called the **standard inner product** on F^n.

Example 3.1.2. Let $V = \mathbb{C}^2$. For $x = (a_1, a_2), y(b_1, b_2)$ define $\langle x, y\rangle$ by

$$\langle x, y \rangle := x \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \overline{y}^t$$

Check $\langle \cdot, \cdot \rangle$ is an inner product :

- linear conditions : Ok!
- $\langle y, x\rangle = \overline{\langle x, y\rangle}$:
$$\overline{\langle x, y\rangle} = \overline{x} A y^t = \left(\overline{x} A y^t\right)^t = y A \overline{x}^t = \langle y, x \rangle$$
- $\langle x, x\rangle \geq 0$

  Observe that
$$A = \begin{pmatrix} 1 & 1/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3/2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1/2 & 1 \end{pmatrix}$$
  For $0 \neq x \in V$, let $(a, b) = x \begin{pmatrix} 1 & 1/2 \\ 0 & 1 \end{pmatrix} \neq 0$, then
$$\langle x, x \rangle = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 3/2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \overline{a} \\ \overline{b} \end{pmatrix} = \frac{3}{2}|a|^2 + 2|b|^2 > 0$$

Example 3.1.3. $V = \{\text{complex value continuous function on } \mathbb{R}\}$. Define

$$\langle x, y \rangle = \int_0^1 f(t)\overline{g(t)} \, dt$$

is an inner product.

Example 3.1.4. Let $A \in M_{n \times n}(F)$. Then the **conjugate transpose** or **adjoint** of A is the matrix $A^* := \overline{A}^t$ i.e. $(A^*)_{ij} = \overline{A_{ji}}$

Remark 3.1.2. The standard inner product on $\mathbb{C}^n$ is often written xy^*

Example 3.1.5. $V = M_{n \times n}(F)$. Define $\langle A, B \rangle := \operatorname{tr}(AB^*)$
Check $\langle \cdot, \cdot \rangle$ is an inner product :

- linear conditions : OK!

- $\overline{\langle A, B\rangle} = \overline{\mathrm{tr}(AB^*)} = \mathrm{tr}(\overline{A}B^t) = \mathrm{tr}((\overline{A}B^t)^t) = \mathrm{tr}(A^*B) = \mathrm{tr}(BA^*) = \langle B, A\rangle$
- $\langle A, A\rangle = \mathrm{tr}(AA^*) = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}(A^*)_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\overline{A}_{ij} > 0$

This inner product is called the **Frobenius inner product** of $M_{n\times n}(F)$

Remark 3.1.3. $(AB)^* = B^*A^*$

Definition 3.1.2. Let V be an inner product space. For $v \in V$, the **norm** or the **length** of x is defined to be

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

Theorem 3.1.2. $\forall x \in V$, $c \in F$

- $\|cx\| = \|c\|\|x\|$
- $\|x\| \geq 0$ and $\|x\| = 0 \iff x = 0$
- **Cauchy-Schwarz inequality** : $|\langle x, y\rangle| \leq \|x\|\|y\|$
- **triangle inequality** : $\|x + y\| \leq \|x\| + \|y\|$

Proof:

- $\|cx\|^2 = \langle cx, cx\rangle = c\langle x, cx\rangle = c\overline{c}\langle x, x\rangle = \|c\|^2\|x\|^2$
- obvious
- If $y = 0$, then inequality holds trivially. Now, we assume $y \neq 0$. We have $\|x - cy\| \geq 0 \ \forall c \in F$ i.e.

$$\begin{aligned} 0 &\leq \langle x - cy, x - cy\rangle = \langle x, x\rangle - \langle cy, x\rangle - \langle x, cy\rangle + \langle cy, cy\rangle \\ &= \|x\|^2 - c\overline{\langle x, y\rangle} - \overline{c}\langle x, y\rangle + \|c\|^2\|y\|^2 \end{aligned}$$

  Now, we choose $c = \dfrac{\langle x, y\rangle}{\langle y, y\rangle}$, then we obtain

$$0 \leq \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2} - \frac{|\langle x, y\rangle|^2}{\|y\|^2} + \frac{|\langle x, y\rangle|^2}{\|y\|^4}\|y^2\| \implies |\|x, y\||^2 \leq \|x\|^2\|y\|^2$$

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

□

Example 3.1.6. In the case of standard inner product on $\mathbb{C}^n$. Cauchy-Schwarz inequality

$$\left|\sum_{j=1}^{n} a_j\overline{b_j}\right| \leq \left(\sum_{j=1}^{n} |a_j|^2\right)^{1/2} \left(\sum_{j=1}^{n} |b_j|^2\right)^{1/2}$$

Recall that in $\mathbb{R}^2$ or $\mathbb{R}^3$, we have

$$\frac{\langle v_1, v_2\rangle}{\|v_1\|\|v_2\|} = \cos\theta,\ 0 \le \theta \le \pi$$

In particular, $\langle v_1, v_2\rangle = 0 \iff \theta = \pi/2$

Definition 3.1.3. Let V be an inner product space

- Two vector $x, y \in V$ are **orthogonal** or **perpendicular** if $\langle x, y\rangle = 0$
- A subset S of V is **orthogonal** if any 2 vectors in S are orthogonal.
- A vector space x in V is a unit vector if $\|x\| = 1$
- A subset S of V is **orthonormal** if S is orthogonal and every vector in S is a unit vector.

(Note that under our definition, an orthogonal subset may contain 0, but an orthonormal subset cannot have 0)

Example 3.1.7. $V = \{\text{continuous complex-valued functions on } \mathbb{R}\}$. Define

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\ dt$$

Then the set $\{f_n(t) := e^{int} : n \in \mathbb{Z}\}$ is orthonormal. Since

$$\langle f_n, f_m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)t}\ dt = \begin{cases} 1 & \text{, if } n = m \\ \left.\dfrac{e^{i(n-m)t}}{2\pi i(n-m)}\right|_0^{2\pi} = 0 & \text{, if } n \neq m \end{cases}$$

Note that the symbol δ_{ij} defined by $\delta_{ij} = \begin{cases} 1 & \text{, if } i = j \\ 0 & \text{, if } i \neq j \end{cases}$ is called the **kronecker symbol**.

Chapter 4

Appendix

4.1 Appendix of Chapter 1

4.1.1 Proof of Theorem 1.3.3

First we introduce a notation for a T-invariant subspace W of K_λ and $v \in K_\lambda$. Observe that $(T-\lambda I)^p(v) = 0$ for some p and hence $I_T(v, W) = ((x-\lambda)^s)$ for some s. We let $s(v, W)$ denote this s.

Outline of proof of existence:

Set $W_0 = \{0\}$. Let v_1 be a vector in K_λ such that

$$s(v_1, W_0) = \max_{v \in K_\lambda} s(v, W_0) =: s_1$$

Let $W_1 = Z(v_1; T)$. Note that $\dim Z(v_1; T) = s_1$, since $I_T(v_1, W_0)$ is simply $I_T(v_1)$ which by assumption is $((x-\lambda)^{s_1})$. We have seen earlier that the integer s, such that $I_T(v) = ((x-\lambda)^s)$ is preceisely the dimension of $Z(v_1, T)$. To find v_2 a natural ideal is to choose v_2 such that

$$s(v_2; W_1) = \max_{v \in K_\lambda} s(v, W_1) =: s_2$$

However, there is one problem here. That is $Z(v_1; T) \cap Z(v_2; T)$ may not be $\{0\}$ i.e. $Z(v_1; T) + Z(u; T)$ may not be a direct sum of $Z(v_1; T)$ and $Z(u; T)$. In order for v_1, v_2 to satisfy $Z(v_1; T) \cap Z(u; T) = \{0\}$, we need to modify it. We claim that there exists w in W_1 such that $(T-\lambda)^{s_2}(u) = (T-\lambda)^{s_2}(w)$ and replace v_2 by $v_2 - w$. We claim that :

- $Z(v_1; T) \cap Z(v_2; T) = \{0\}$
- $\dim Z(v_2, T) = s_2$

In general for $i \geq 3$, choose u such that

$$s(u, W_{i-1}) = \max_{v \in K_\lambda} s(v; W_{i-1}) =: s_i$$

We claim that there exists w in W_{i-1} such that $(T-\lambda)^{s_i}(u) = (T-\lambda)^{s_i}(w)$ and let $v_i = u - w$. We claim that :

- $W_{i-1} \cap Z(v_i; T) = \{0\}$
- $\dim Z(v_i, T) = s_i$

Let $W_i = W_{i-1} \oplus Z(w_i; T)$ and continuous until $W_i = K_\lambda$.

Outline of outline of proof of existence:

(1) Let $W_0 = \{0\}$

(2) Choose $u \in K_\lambda$ s.t.

$$s(u, W_{i-1}) = \max_{v \in K_\lambda} s(v, W_{i-1}) =: s_i$$

and prove that $\exists w \in W_{i-1}$ s.t. $(T-\lambda)^{s_i}(u) = (T-\lambda)^{s_i}(w)$

(3) Let $v_i = u - w$, prove that

- $W_{i-1} \cap Z(v_i; T) = \{0\}$
- $\dim Z(v_i; T) = s_i$

Let $W_i = W_{i-1} \oplus Z(v_i; T)$

(4) Repeat (2),(3) until $W_i = K_\lambda$ (Since $\dim W_i > \dim W_{i-1}$)

Detailed proof: Let $W_0 = \{0\}$, we induction on i :
$i = 1 : W_0 = \{0\}$, so $w = 0$. We indeed have $0 = (T-\lambda)^{s_1}(0) = (T-\lambda)^{s_1}(v_1)$.
Now, assume that $(2), (3)$ holds for $W_1, ..., W_{i-1}$.

Lemma: Let u be a vector in K_λ such that

$$s(u, W_{i-1}) = \max_{v \in K_\lambda} s(v, W_{i-1}) =: s_i$$

Then $\exists w \in W_{i-1}$ such that $(T-\lambda)^{s_i}(w) = (T-\lambda)^{s_i}(u)$

Proof: For convenience, let $\widetilde{T} = T - \lambda I$. For $j \leq i-1$, $\mathcal{B}_j := \{v_j, \widetilde{T}(v_j), ..., \widetilde{T}^{s_j - 1}(v_j)\}$ is a basis for $Z(v_j; T)$. Hence, $\mathcal{B} := \bigsqcup_{j=1}^{i-1} \mathcal{B}_j$ is a basis for $W_{j-1} = \bigoplus_{j=1}^{i-1} Z(v_j; T)$. Thus

$$\widetilde{T}^{s_i}(u) = \sum_{j=1}^{i-1} \sum_{k=0}^{s_j - 1} a_{jk} \widetilde{T}^k(v_j) \tag{*}$$

for some unique a_{jk}. Note that since $W_0 \subset W_1 \subset \cdots \implies s(v, W_1) \geq s(v, W_2) \geq \cdots$ for any $v \in K_\lambda$, thus $s_1 \geq s_2 \geq \cdots \geq s_{i-1} \geq s_i$. Our goal is to show that $a_{jk} = 0$ for any (j, k) with $k \leq s_i - 1$. Then

$$\widetilde{T}^{s_i}(u) = \sum_{j=1}^{i-1} \sum_{k=0}^{s_j-1} a_{jk} \widetilde{T}^k(v_j) = \sum_{j=1}^{i-1} \sum_{k=s_i}^{s_j-1} a_{jk} \widetilde{T}^k(v_j) = \widetilde{T}^{s_i} \left(\sum_{j=1}^{i-1} \sum_{k=s_i}^{s_j-1} a_{jk} . \widetilde{T}^{k-s_i}(v_j) \right)$$

which complete our lemma.
For $m \leq i-1$, we may apply $\widetilde{T}^{s_m - s_i}$ to $(*)$, we have

$$\widetilde{T}^{s_m}(u) = \sum_{j=1}^{i-1} \sum_{k=0}^{s_j-1} a_{jk} \widetilde{T}^{s_m - s_i + k}(v_j)$$

Note that LHS$\in W_{m-1}$ according to the definition of $S_m := \max_{v \in K_\lambda} s(v, W_{m-1})$. Thus

$$\sum_{j=m}^{i-1} \underbrace{\sum_{k=0}^{s_j-1} a_{jk} \widetilde{T}^{s_m - s_i + k}(v_j)}_{\in Z(v_j; T)} = 0 \implies \sum_{k=0}^{s_j-1} a_{jk} \widetilde{T}^{s_m - s_i + k}(v_j) = 0 \ \forall j$$

In particular when $j = m$, if $k \geq s_i \rightsquigarrow s_m - s_i + k \geq s_m$ and thus $\widetilde{T}^{s_m - s_i + k}(v_m) = 0$. Since $\mathcal{B}_j$ is linearly independent over F, $a_{mk} = 0 \ \forall k \leq s_i - 1$ $\square$
Let $w \in W_{i-1}$ be a vector such that $(T-\lambda)^{s_i}(w) = (T-\lambda)^{s_i}(u)$ as in the previous lemma. Let $v_i = u - w$. Then

- $W_{i-1} \cap Z(v_i;T) = \{0\}$ ($\implies W_{i-1} + Z(v_i;T)$ is a direct sum of $Z(v_1;T), ..., Z(v_i;T)$) :

  pf. Recall that if W is a T-invariant and $v - v' \in W$, then $I_{v,W} = I_{v',W}$.

  Thus, $I_{v_i,W_{i-1}} = I_{u,W_{i-1}} = ((x-\lambda)^{s_i})$ i.e. $s(v_i, W_{i-1}) = s(u, W_{i-1}) = s_i$. Suppose that $v \in W_{i-1} \cap Z(v_i;T)$. Say $v = a_0 v_i + a_1 T(v_i) + \cdots + a_n T^n(v_i)$. Let $f(x) := a_0 + a_1 x + \cdots + a_n x^n$. Since $v \in W_{i-1}$ and $v = f(T)(v_i) \implies f(x) \in I_{v_i,W_{i-1}} = ((x-\lambda)^{s_i}) \implies f(x) = g(x)(x-\lambda)^{s_i}$. Then $v = f(T)(v_i) = g(T)(T-\lambda I)^{s_i}(v_i) = g(T)(T-\lambda I)^{s_i}(u-w) = 0$ □

- $\dim Z(v_i;T) = s_i$:

  pf. Recall that for $v \in K_\lambda$, $\dim Z(v_i;T) =$ the smallest integer s such that $(T-\lambda)^s(v) = 0$.

  Here we have found that $(T-\lambda)^{s_i}(v_i) = 0 \implies s \leq s_i$

  On the other hand. Since $I_{v_i,W_{i-1}} = ((x-\lambda)^{s_i}) \implies s \geq s_i$ and thus $s = s_i$

By induction, (2),(3) holds for all i. Since $\dim W_{i-1} < \dim W_i \leq \dim K_\lambda$. We can repeat (2),(3) until $W_i = K_\lambda$. □

4.1.2 Proof of Theorem 1.4.2

Claim in (ii): Let $u \in K_p$ be such that

$$s(u, W_{i-1}) = \max_{v \in K_p} s(v, W_{i-1}) = s_i$$

Then $\exists w \in W_{i-1}$ such that $p(T)^{s_i}(w) = p(T)^{s_i}(u) = 0$

Proof: Since $s(u, W_{i-1}) = s_i$, we have $p(T)^{s_i}(u) \in W_{i-1}$

$$\implies p(T)^{s_i}(u) = \sum_{j=1}^{i-1} f_j(T)(v_j) \tag{*}$$

for some polynomials $f_j(x)$. Observe that $s_1 \geq s_2 \geq \cdots \geq s_i$. For $m \leq i-1$, we have $s_m \geq s_i$ and we can apply $p(T)^{s_m - s_i}$ to $(*)$ and obtain

$$p(T)^{s_m}(u) = \sum_{j=1}^{i-1} p(T)^{s_m - s_i} f_j(T)(v_j) \tag{**}$$

Recall that s_m is defined to be $\max\limits_{v \in K_p} s(v, W_{m-1})$. Thus, the LHS of $(**)$ belongs to W_{m-1} i.e. LHS of (**) equal to $\sum\limits_{j=1}^{m-1} g_j(T)(v_j)$ for some $g_j(x)$

$$\implies \sum_{j=1}^{m-1} g_j(T)(v_j) = \sum_{j=1}^{i-1} p(T)^{s_m - s_i} f_j(T)(v_j) \implies \underbrace{p(T)^{s_m} f_m(T)(v_m)}_{\in Z(v_m;T)} + \sum_{j \neq m} \underbrace{h_j(T)(v_j)}_{\in Z(v_j;T)} = 0$$

for some $h_j(x) \in F[x]$. By Property 1.2.1, we have $p(T)^{s_m} f_m(T)(v_m) = 0$
By induction hypothesis of (iii), $I_{v_m} = (p(x)^{s_m}) \implies p(x)^{s_m} \big| p(x)^{s_m - s_i} f_m(x) \implies p(x)^{s_i} \big| f_m(x)$, say $f_m(x) = p(x)^{s_i} \widetilde{f}_m(x)$ for some $\widetilde{f}_m$. Recall $(*)$,

$$p(T)^{s_i}(u) = \sum_{j=1}^{i-1} f_j(T)(v_j) = \sum_{j=1}^{i-1} p(T)^{s_i} \widetilde{f}_j(T)(v_j) = p(T)^{s_i} \underbrace{\left(\sum_{j=1}^{i-1} \widetilde{f}_j(T)(v_j) \right)}_{:= w \in W_{j-1}}$$

Hence, we find $w \in W_{i-1}$ s.t. $p(T)^{s_i}(u) = p(T)^{s_i}(w)$. □

Claim in (iii): Let u, w be as in (ii). Let $v_i := u - w$. Then

- $W_{i-1} \cap Z(v_i; T) = \{0\}$
- $I_{v_i} = (p(x)^{s_i})$

Proof:

- Assume $v \in W_{i-1} \cap Z(v_i; T)$. We have $v = f(T)(v_i)$ for some $f(x) \in F[x]$. This vector is in $W_{i-1} \implies f(x) \in I_{v,W_{i-1}}$. Recall that in Remark 1.4.2, we have seen that $I_{v,W_{i-1}} = (p(x)^{s_i})$. So we have $p(x)^{s_i} \big| f(x)$, say $f(x) = p(x)^{s_i} \widetilde{f}(x) \implies v = f(T)(v_i) = \widetilde{f}(T) p(T)^{s_i}(v_i) = 0$
- Since $v_i \in K_p := \{v \in V : p(T)^n v = 0 \text{ for some } n\} \implies I_{v_i} = (p(x)^s)$. Now recall that $s(v_i, W_{i-1}) = s_i \implies p(T)^{s_i} \in W_{i-1}$ and s_i is the smallest integer with the property $\implies s \geq s_i$

  On the other hand, $p(T)^{s_i}(v_i) = p(T)^{s_i}(u - w) = 0 \implies s_i \geq s \implies s = s_i$

  (Note that $p(x)$ is irreducible is necessarily)

$\square$

Chapter 5

Homework and bonus

We only include interesting or useful problems. Absolutely, we will not put any calculation questions. Some notation will be defined in homework and will not be defined again in class.

5.1

Problem 5.1.1. Let $T : V \to V$ be a linear operator on V. Check the following sets are ideals of $F[x]$:

(1) the set $\{f(x) \in F[x] : f(T) = 0\}$.

(2) the set $\{f(x) \in F[x] : f(T)(v) = 0\}$, where $v \in V$ is a fixed given vector.

(3) the set $I_{v,W} := \{f(x) \in F[x] : f(T)(v) \in W\}$, where W is a T-invariant subspace of V and $v \in V$ is a given vector.

(4) In part (3), if W is only a subspace of V but not T-invariant, does the statement still hold? Prove it or disprove it by giving a counterexample.

Remark 5.1.1. If V is finite-dimensional, then we know that the first set

$$\{f(x) \in F[x] : f(T) = 0\} = (m_T(x))$$

is a principal ideal generated by the minimal polynomial of T.

Problem 5.1.2. Prove also that if W is a T-invariant subspace of V and $v_1 - v_2 \in W$, then $I_{v_1,W} = I_{v_2,W}$.

Problem 5.1.3. Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V and let v be a non-zero vector in V. The set

$$\{f(x) \in F[x] : f(T)(v) = 0\} = (g(x))$$

is a principal ideal generated by a monic polynomial $g(x) \in F[x]$.

(1) If U is the T-cyclic subspace generated by v, show that $g(x)$ is the minimal polynomial of $T|_U$, and $\dim(U)$ equals the degree of $g(x)$.

(2) Show that the degree of $g(x)$ is 1 if and only if v is an eigenvector of T.

Problem 5.1.4. Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V, let W_1 be a T-invariant subspace of V, and let v be a non-zero vector in V. The set

$$I_{v,W_1} = \{f(x) \in F[x] : f(T)(v) \in W_1\} = (g_1(x))$$

is a principal ideal generated by a monic polynomial $g_1(x) \in F[x]$.

(1) Show that $g_1(x)$ divides the minimal and the characteristic polynomials of T.

(2) Let W_2 be a T-invariant subspace of V such that $W_2 \subseteq W_1$ and let $g_2(x)$ be a monic polynomial such that the set

$$I_{v,W_2} = \{f(x) \in F[x] : f(T)(v) \in W_2\} = (g_2(x))$$

is a principal ideal generated by $g_2(x)$. Show that $g_1(x)$ divides $g_2(x)$.

5.2

Problem 5.2.1. Let T be a diagonalizable linear operator on a finite-dimensional vector space V. Prove that V is a T-cycle subspace if and only if each of the eigenspaces of T is one-dimensional.

Problem 5.2.2. Let T be a linear operator on a finite-dimensional vector space V.

(1) Let λ be an eigenvalue of T. Prove that if $\operatorname{rank}((T-\lambda I)^m) = \operatorname{rank}\left((T-\lambda I)^{m+1}\right)$ for some positive integer m, then $K_\lambda = \ker((T-\lambda I)^m)$.

(2) (Second Test for Diagonalizability.) Suppose that the characteristic polynomial $\mathrm{ch}_T(x)$ splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be all the distinct eigenvalues of T. Show that T is diagonalizable if and only if $\operatorname{rank}(T-\lambda_i I) = \operatorname{rank}\left((T-\lambda_i I)^2\right)$ for $1 \le i \le k$.

Bonus 1. Let $T : V \to V$ be a linear transformation on a finite-dimensional vector space V over F, where F is an infinite field. Show that V is T-cycle if and only if V has only finitely many T-invariant subspace.

5.3

Problem 5.3.1. Let A be an $n \times n$ matrix whose characteristic polynomial splits. Prove that A and A^t have the same Jordan canonical form, and conclude that A and A^t are similar.

5.4

Problem 5.4.1. Let T be a linear operator on a finite-dimensional vector space V whose characteristic polynomial splits, and let J be the Jordan canonical form of T. Let D be the diagonal matrix whose diagonal entries are the diagonal entries of J, and let $M = J - D$.

(1) Show that M is nilpotent, i.e., $M^k = O$ for some integer k.

(2) Show that $MD = DM$.

(3) If J is given by

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \in M_m(\mathbb{R}).$$

Show that $(J - \lambda I_m)^m = 0$ and that if $r \geq m$, then

$$J^r = \begin{pmatrix} \lambda^r & r\lambda^{r-1} & \frac{r(r-1)}{2!}\lambda^{r-2} & \cdots & \frac{r(r-1)\cdots(r-m+2)}{(m-1)!}\lambda^{r-m+1} \\ 0 & \lambda^r & r\lambda^{r-1} & \cdots & \frac{r(r-1)\cdots(r-m+3)}{(m-2)!}\lambda^{r-m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda^r \end{pmatrix}.$$

Moreover, if $f(x) \in \mathbb{R}[x]$, show that

$$f(J) = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f'(\lambda) & \frac{1}{2!}f''(\lambda) & \cdots & \frac{1}{(m-1)!}f^{(m-1)}(\lambda) \\ 0 & f(\lambda) & \frac{1}{1!}f'(\lambda) & \cdots & \frac{1}{(m-2)!}f^{(m-2)}(\lambda) \\ 0 & 0 & f(\lambda) & \cdots & \frac{1}{(m-3)!}f^{(m-3)}(\lambda) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f(\lambda) \end{pmatrix}.$$

Problem 5.4.2. Let T be a linear operator on a finite-dimensional vector space, and suppose that $\phi(t)$ is a monic factor of the characteristic polynomial of T. Suppose that x and y are two vectors such that the following ideals are the same and generated by $\phi(t)$:

$$I_T(x) := \{f(t) \in F[t] : f(T)(x) = 0\} = (\phi(t)) = \{f(t) \in F[t] : f(T)(y) = 0\} =: I_T(y).$$

Prove that $x \in Z(y;T)$ if and only if $Z(x;T) = Z(y;T)$, where $Z(v;T)$ is the T-cyclic subspace of V generated by the vector v.

Problem 5.4.3. Let T be a linear operator on a finite-dimensional vector space V. Suppose that $v_1, v_2 \in V$ are two vectors such that the following ideals are generated by $\phi_1(x)$ and $\phi_2(x)$.

$$I_T(v_1) := \{f(x) \in F[x] : f(T)(v_1) = 0\} = (\phi_1(x))$$
$$I_T(v_2) := \{f(x) \in F[x] : f(T)(v_2) = 0\} = (\phi_2(x)).$$

Suppose further that $\phi_1(x)$ and $\phi_2(x)$ are coprime monic polynomials.

(1) Show that $Z(v_1;T) + Z(v_2;T)$ is a direct sum.

(2) Show that there is a vector $v_3 \in V$ such that the following ideal is generated by $\phi_1(x)\cdot\phi_2(x)$:

$$I_T(v_3) := \{f(x) \in F[x] : f(T)(v_3) = 0\} = (\phi_1(x) \cdot \phi_2(x)).$$

(3) Show that there exists $v_3 \in V$ such that

$$Z(v_1;T) \oplus Z(v_2;T) = Z(v_3;T).$$

Remark 5.4.1. This exercise allows us to combine two T-cyclic subspaces into a single T-cyclic subspace provided that the minimal polynomials of this two subspaces are coprime.

Bonus 2. Prove the **Jordan-Chevalley decomposition theorem**: Let $A \in M_n(F)$ be a matrix whose characteristic polynomial splits. Then, there exist two unique matrices $S, N \in M_n(F)$ satisfying the conditions: $A = S+N$, S is diagonalizable, N is nilpotent, and $SN = NS$.

5.5

Problem 5.5.1. Let T be a linear operator on a finite-dimensional vector space V over F. Show that V is itself a T-cyclic subspace if and only if $\mathrm{ch}_T(x) = m_T(x)$.

Problem 5.5.2. For any $A \in M_{n\times n}(\mathbb{C})$, we write $A = (a_{ij})_{1\le i,j\le n}$ and define

$$||A|| = \max\{|a_{ij}| : 1 \le i, j \le n\}.$$

(1) Prove that for any $A, B \in M_{n\times n}(\mathbb{C})$, $||AB|| \le n||A|| \cdot ||B||$.

(2) Prove that e^A exists for every $A \in M_{n\times n}(\mathbb{C})$.

Problem 5.5.3. Suppose that J is a single Jordan block:

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \in M_m(\mathbb{C}).$$

Show that

$$e^{Jz} = e^{\lambda z} \begin{pmatrix} 1 & \frac{z}{1!} & \frac{z^2}{2!} & \cdots & \cdots & \frac{z^{m-1}}{(m-1)!} \\ 0 & 1 & \frac{z}{1!} & \cdots & \cdots & \frac{z^{m-2}}{(m-2)!} \\ 0 & 0 & 1 & \ddots & \cdots & \frac{z^{m-3}}{(m-3)!} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \ddots & \frac{z}{1!} \\ 0 & 0 & 0 & \cdots & \cdots & 1 \end{pmatrix},$$

where z is a variable.

Problem 5.5.4. Let A be a square matrix with complex entries. Write

$$S = \{z \in \mathbb{C} \mid |z| < 1 \text{ or } z = 1\}$$

to be a subset of $\mathbb{C}$. Show that $\lim_{m\to\infty} A^m$ exists if and only if both of the following conditions hold.

(1) Every eigenvalue of A is contained in S.

(2) If 1 is an eigenvalue of A, then the dimension of the eigenspace corresponding to 1 equals the multiplicity of 1 as an eigenvalue of A.

Bonus 3. Let $S, T : V \to V$ be two linear operators on a finite-dimensional vector space V. Suppose that any linear operator U on V with $US = SU$ has the property $UT = TU$. Show that T is a polynomial of S, that is, there is an $f(x) \in F[x]$ such that $T = f(S)$.

Chapter 6

Advance Problem

Problems proposed by Ping-Hsun Chuang (TA).

6.1

Problem 6.1.1. Suppose that $A \in M_n(F)$ and the characteristic polynomial $ch_A(x)$ splits in F. Let $\lambda_1, ..., \lambda_m$ be all distinct eigenvalues of A. Show that

$$n(m-1) \leq \sum_{j=1}^{m} \operatorname{rank}(A - \lambda_j I)$$

Problem 6.1.2. Show that the eigenvalues of the tridiagonal matrix

$$A = \begin{pmatrix} a_1 & -b_1 & 0 & \cdots & 0 & 0 & 0 \\ -c_1 & a_2 & -b_2 & \cdots & 0 & 0 & 0 \\ 0 & -c_2 & a_3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-2} & -b_{n-2} & 0 \\ 0 & 0 & 0 & \cdots & -c_{n-2} & a_{n-1} & -b_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & -c_{n-1} & a_n \end{pmatrix} \in M_n(\mathbb{R}),$$

where $b_i c_i > 0$ for all i, are all real and of multiplicity one.

Problem 6.1.3. Let $A \in M_n(\mathbb{R})$. Show that $\operatorname{rank} A = \operatorname{rank} A^2$ if and only if $\lim_{\lambda \to 0} (A + \lambda I)^{-1} A$ exists.

6.2

Problem 6.2.1. Suppose that $A \in M_n(F)$ and the characteristic polynomial $ch_A(x)$ splits in F. Let $\lambda_1, ..., \lambda_m$ be all distinct eigenvalues of A. Also, let $a_1(\lambda_i), ..., a_{r_i}(\lambda_i)$ be all size of Jordan blocks corresponding to the eigenvalue λ_i (counting multiplicity) for each i. Show that the subspace

$$\{X \in M_n(F) | XA = AX\} \subseteq M_n(F)$$

has dimension

$$\sum_{i=1}^{m} \sum_{1 \leq i, k \leq r_i} \min\{a_j(\lambda_i), a_k(\lambda_i)\}$$

Problem 6.2.2. Show that a matrix A can be represented as the product of two involutions (A matrix B is an involution if $B^2 = I$) if and only if the matrices A and A^{-1} are similar.

Problem 6.2.3. Suppose that $T : V \to V$ is a linear operator on a vector space V and that $\{v_1, ..., v_n\}$ is a basis for V that us a single Jordan chain (in other words, a cycle of generalized eigenvectors) for T. Determine a Jordan canonical basis for T^2.

6.3

Problem 6.3.1. Let A be an $n \times n$ diagonalizable matrix with all eigenvalues being $\lambda_1, ..., \lambda_n$. Show that there are right eigenvectors (column vector) $x_1, ..., x_n$ and left eigenvectors (row vectors) $y_1, ..., y_n$ such that

$$A = \sum_{i=1}^{n} \lambda_i x_i y_i$$

Problem 6.3.2. Let $A, B \in M_{n\times n}(\mathbb{C})$. Show that the following are equivalent :

(a) A and B are similar.

(b) Either

$$\begin{pmatrix} 0 & A \\ \overline{A} & 0 \end{pmatrix}, \begin{pmatrix} 0 & B \\ \overline{B} & 0 \end{pmatrix}$$

are similar or

$$\begin{pmatrix} 0 & A \\ -\overline{A} & 0 \end{pmatrix}, \begin{pmatrix} 0 & B \\ -\overline{B} & 0 \end{pmatrix}$$

are similar.

Problem 6.3.3. Let

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & 0 & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & -a_{n-3} \\ 0 & 0 & \cdots & 0 & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 0 & 1 & -a_{n-1} \end{pmatrix}$$

Show that there exists an invertible symmetric matrix S such that $A = SA^TS^{-1}$.

6.4

Problem 6.4.1. Let $A = I + xy^T$, where x and y are nonzero n real column vectors. Show that $\det(A) = 1 + x^Ty$. Also, determine the Jordan canonical form of the matrix A.

Problem 6.4.2. Let T be a linear transformation on a finite-dimensional vector space over F. Show that V is T-cyclic if and only if any linear operator S on V commuting with T is a polynomial in T.

Problem 6.4.3. Let $A \in M_{n\times n}(\mathbb{C})$. Define

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$
$$\sin(A) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} A^{2k+1}$$
$$\cos(A) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} A^{2k}$$
$$\log(I_n + A) = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} A^k$$

(a) For any $A \in M_{n\times n}(\mathbb{C})$, show that

$$\sin^2(A) + \cos^2(A) + I_n$$

(b) For any $A \in M_{n\times n}(\mathbb{C})$ with all eigenvalue $|\lambda_i| < 1$, show that

$$\exp(\log(I_n + A)) = I_n + A$$