A JOURNEY

IN

PURE MATHEMATICS

A JOURNEY

IN

PURE MATHEMATICS

MAT3006 & 3040 & 4002 Notebook

Prof. Daniel Wong

The Chinese University of Hongkong, Shenzhen

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is from the MAT3006,MAT3040,MAT4002 in spring semester, 2018-2019.

CUHK(SZ)

Notations and Conventions

$\mathbb{R}^n$	n-dimensional real space
$\mathbb{C}^n$	n-dimensional complex space
$\mathbb{R}^{m \times n}$	set of all $m \times n$ real-valued matrices
$\mathbb{C}^{m \times n}$	set of all $m \times n$ complex-valued matrices
x_i	ith entry of column vector $\boldsymbol{x}$
a_{ij}	(i,j)th entry of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	ith column of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i^{\mathrm{T}}$	ith row of matrix $\boldsymbol{A}$
$\mathbb{S}^n$	set of all $n \times n$ real symmetric matrices, i.e., $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and $a_{ij} = a_{ji}$ for all i, j
$\mathbb{H}^n$	set of all $n \times n$ complex Hermitian matrices, i.e., $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ and $\bar{a}_{ij} = a_{ji}$ for all i, j
$\boldsymbol{A}^{\mathrm{T}}$	transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ means $b_{ji} = a_{ij}$ for all i, j
$\boldsymbol{A}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$ means $b_{ji} = \bar{a}_{ij}$ for all i, j
$\mathrm{trace}(\boldsymbol{A})$	sum of diagonal entries of square matrix $\boldsymbol{A}$
$\mathbf{1}$	A vector with all 1 entries
$\mathbf{0}$	either a vector of all zeros, or a matrix of all zeros
$\boldsymbol{e}_i$	a unit vector with the nonzero element at the ith entry
$\mathcal{C}(\boldsymbol{A})$	the column space of $\boldsymbol{A}$
$\mathcal{R}(\boldsymbol{A})$	the row space of $\boldsymbol{A}$
$\mathcal{N}(\boldsymbol{A})$	the null space of $\boldsymbol{A}$
$\mathrm{Proj}_{\mathcal{M}}(\boldsymbol{A})$	the projection of $\boldsymbol{A}$ onto the set $\mathcal{M}$

Chapter 7

Week7

7.1. Monday for MAT3040

Reviewing. Define the characteristic polynomial for an linear operator T:

$$\mathcal{X}_T(x) = \det((T)_{\mathcal{A},\mathcal{A}} - x\boldsymbol{I})$$

We will use the notation "$I/\boldsymbol{I}$" in two different occasions:

1. I denotes the identity transformation from V to V with $I(\boldsymbol{v}) = \boldsymbol{v}, \forall \boldsymbol{v} \in V$
2. $\boldsymbol{I}$ denotes the identity matrix $(I)_{\mathcal{A},\mathcal{A}}$, defined based on any basis $\mathcal{A}$.

7.1.1. Minimal Polynomial

Definition 7.1 [Linear Operator Induced From Polynomial] Let $f(x) := a_m x^m + \cdots + a_0$ be a polynomial in $\mathbb{F}[x]$, and $T : V \to V$ be a linear operator. Then the mapping

$$f(T) = a_m T^m + \cdots + a_1 T + a_0 I : \quad V \to V,$$

is called a linear operator induced from the polynomial $f(x)$. ■

Definition 7.2 [Minimal Polynomial] Let $T : V \to V$ be a linear operator. The **minimal polynomial** $m_T(x)$ is a **nonzero monic polynomial** of least (minimal) degree such that

$$m_T(T) = \mathbf{0}_{V \to V}.$$

where $\mathbf{0}_{V \to V}$ denotes the zero vector in $\text{Hom}(V, V)$. ▪

▪ **Example 7.1** 1. Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then $\boldsymbol{A}$ defines a linear operator:

$$\begin{aligned} A: \quad & \mathbb{F}^2 \to \mathbb{F}^2 \\ \text{with} \quad & \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} \end{aligned}$$

Here $\mathcal{X}_A(x) = (x-1)^2$ and $\boldsymbol{A} - \boldsymbol{I} = \mathbf{0}$, which gives $m_A(x) = x - 1$.

2. Let $\boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, which implies

$$\mathcal{X}_B(x) = (x-1)^2,$$

The question is that can we get the minimal polynomial with degree 1?

The answer is no, since $\boldsymbol{B} - k\boldsymbol{I} = \begin{pmatrix} 1-k & 1 \\ 0 & 1-k \end{pmatrix} \neq \mathbf{0}$.

In fact, $m_B(x) = (x-1)^2$, since

$$(\boldsymbol{B} - \boldsymbol{I})^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

▪

Two questions naturally arises:

1. Does $m_T(x)$ exist? If exists, is it unique?
2. What's the relationship between $m_T(x)$ and $\mathcal{X}_T(x)$?

Regarding to the first question, the minimal polynomial $m_T(x)$ may not exist, if V has infinite dimension:

■ **Example 7.2** Consider $V=\mathbb{R}[x]$ and the mapping

$$\begin{aligned} T:\quad & V\to V \\ & p(x)\mapsto \int_0^x p(t)\,\mathrm{d}t \end{aligned}$$

In particular, $T(x^n)=\frac{1}{n+1}x^{n+1}$. Suppose $m_T(x)$ is with degree n, i.e.,

$$m_T(x)=x^n+\cdots+a_1x+a_0,$$

then

$$m_T(T)=T^n+\cdots+a_0I \text{ is a zero linear transformation}$$

It follows that

$$[m_T(T)](x)=\frac{1}{n!}x^n+a_{n-1}\frac{1}{(n-1)!}x^{n-1}+\cdots+a_1x+a_0=0_{\mathbb{F}},$$

which is a contradiction since the coefficients of x^k is nonzero on LHS for $k=1,\dots,n$, but zero on the RHS. ■

Proposition 7.1 The minimal polynomial $m_T(x)$ always exists for $\dim(V)=n<\infty$.

Proof. It's clear that $\{I,T,\dots,T^n,T^{n+1},\cdots,T^{n^2}\}\subseteq \mathrm{Hom}(V,V)$. Since $\dim(\mathrm{Hom}(V,V))=n^2$, we imply $\{I,T,\dots,T^n,T^{n+1},\cdots,T^{n^2}\}$ is linearly dependent, i.e., there exists a_i's that are not all zero such that

$$a_0I+a_1T+\cdots+a_{n^2}T^{n^2}=0$$

i.e., there is a polynomial $g(x)$ of degree less than n^2 such that $g(T)=0$.

The proof is complete. ■

Proposition 7.2 The minimal polynomial $m_T(x)$, if exists, then it exists uniquely.

Proof. Suppose f_1,f_2 are two distinct minimal polynomials with $\deg(f_1)=\deg(f_2)$. It follows that

- $\deg(f_1 - f_2) < \deg(f_1)$.
- $f_1 - f_2 \neq 0$
- $(f_1 - f_2)(T) = f_1(T) - f_2(T) = 0_{V\to V}$

By scaling $f_1 - f_2$, there is a monic polynomial g with lower degree satisfying $g(T) = 0$, which contradicts the definition for minimal polynomial. ■

Proposition 7.3 Suppose $f(x) \in \mathbb{F}[x]$ satisfying $f(T) = 0$, then

$$m_T(x) \mid f(x).$$

Proof. It's clear that $\deg(f) \geq \deg(m_T)$. The division algorithm gives

$$f(x) = q(x)m_T(x) + r(x).$$

Therefore, for any $\boldsymbol{v} \in V$

$$[r(T)](\boldsymbol{v}) = [f(T)](\boldsymbol{v}) - [q(T)m_T(T)](\boldsymbol{v}) = \mathbf{0}_V - q(T)\mathbf{0}_V = \mathbf{0}_V - \mathbf{0}_V = \mathbf{0}_V$$

Therefore, $r(T) = \mathbf{0}_{V\to V}$. By definition of minimal polynomial, we imply $r(x) \equiv 0$. ■

Proposition 7.4 If $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{F}^{n\times n}$ are similar to each other, then $m_A(x) = m_B(x)$.

Proof. Suppose that $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, then

$$m_A(x) = x^k + \cdots + a_1 x + a_0,\ m_B(x) = x^\ell + \cdots + b_0$$

It follows that

$$\begin{aligned}
m_A(\boldsymbol{B}) &= \boldsymbol{B}^k + \cdots + a_0 I \\
&= \boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P} + \cdots + a_0\boldsymbol{P}^{-1}\boldsymbol{P} \\
&= \boldsymbol{P}^{-1}(\boldsymbol{A}^k + \cdots + a_0\boldsymbol{I})\boldsymbol{P} \\
&= \boldsymbol{P}^{-1}(m_A(\boldsymbol{A}))\boldsymbol{P}
\end{aligned}$$

By proposition, $m_B(x) \mid m_A(x)$. Similarly, $m_A(x) \mid m_B(x)$. Since $m_A(x)$ and $m_B(x)$ are monic, we imply $m_A(x) = m_B(x)$. ■

Recall that we also have $\mathcal{X}_A(x) = \mathcal{X}_B(x)$.

We now focus on vanishing of a single vector $\boldsymbol{v} \in V$.

Proposition 7.5 Let $T : V \to V$ be a linear operator and $\boldsymbol{v} \in V$. Consider $m_{T,\boldsymbol{v}}(x)$, the unique monic polynomial of least possible degree such that $m_{T,\boldsymbol{v}}(T)(\boldsymbol{v}) = 0$. Then

$$\deg(m_{T,\boldsymbol{v}}(x)) \leq \dim(V).$$

Proof. Let $\dim(V) = n$, and since

$$\{\boldsymbol{v}, T\boldsymbol{v}, \ldots, T^n\boldsymbol{v}\} \subseteq V,$$

we imply there exists a_i's that are not all zero suc h that

$$a_n T^n \boldsymbol{v} + \cdots + a_0 I = 0$$

i.e., $\deg(m_{T,\boldsymbol{v}}(x)) \leq n$. ■

Proposition 7.6 Suppose that $m_{T,\boldsymbol{v}}(x) = f_1(x) f_2(x)$, where f_1, f_2 are both monic. Let $\boldsymbol{w} = f_1(T)\boldsymbol{v}$, then

$$m_{T,\boldsymbol{w}}(x) = f_2(x)$$

Proof. Note that

$$f_2(T)\boldsymbol{w} = f_2(T) f_1(T)\boldsymbol{v} = m_{T,\boldsymbol{v}}(T)\boldsymbol{v} = \mathbf{0}$$

Therefore, $m_{T,\boldsymbol{w}} | f_2$.

On the other hand,

$$\mathbf{0} = m_{T,\boldsymbol{w}}(T)(\boldsymbol{w}) = m_{T,\boldsymbol{w}}(T)(f_1(T)\boldsymbol{v})$$

Therefore, $\mathbf{0} = f_1(T)m_{T,\boldsymbol{w}}(T)\boldsymbol{v}$, which implies that

$$m_{T,\boldsymbol{v}}(x) \mid f_1(x)m_{T,\boldsymbol{w}}(x),$$

i.e.,

$$f_1 \cdot f_2 \mid f_1 \cdot m_{T,\boldsymbol{w}} \implies f_2 \mid m_{T,\boldsymbol{w}}.$$

The proof is complete. ■