Structure Theorem of Finitely Generated Modules over PIDs, Applications to linear algebra

Jiang-Hua Lu

The University of Hong Kong

Monday March 17, 2025

- Structure Theorem of Finitely Generated Modules over PIDs: Applications to linear algebra (§2.5.2 of Lecture notes).

Consider an n-dim vector space V over a field K (eg: $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$), together w/ a $T \in \mathrm{End}_K(V)$. Regard V as a module of $K[x]$ where

$$\left(f = a_0 + a_1 x + \cdots + a_m x^m\right) \cdot v = a_0 v + a_1 T v + \cdots + a_m T^m v$$

Let $I = \{ f \in K[x] : f(T) = 0 \}$

$= \{ f = a_0 + a_1 x + \cdots + a_m x^m : a_0 + a_1 T + \cdots + a_m T^m = 0 \in \mathrm{End}_K(V) \}$

Caley-Hamilton $\Rightarrow I \neq \{0\}$

The unique monic generator of I, as an ideal of $K[x]$, is called the minimal polynomial of T over K.

Note that $I \subset \mathrm{ann}(v)$, $\forall\, v \in V$

Thus $V^{tor} = V$

Moreover, V, as a $K[x]$-module, is finitely generated.

Recall: Let K be a field, V an n-dimensional K-vector space, and $T \in \mathrm{End}_K(V)$. Then

- V is a $K[x]$-module via $x \cdot v = Tv$ for $v \in V$.
- If $\{v_1, \ldots, v_n\}$ is a basis of V, then have surjective morphism of $K[x]$-modules $\phi : K[x]^n \to V$,

$$\phi \begin{pmatrix} f_1 \\ f_2 \\ \cdots \\ f_n \end{pmatrix} = f_1 \cdot v_1 + \cdots + f_n \cdot v_n$$

$$= f_1(T)v_1 + f_2(T)v_2 + \cdots + f_n(T)v_n.$$

$\varphi|_{K^n}: K^n \to V, \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \lambda_1 v_1 + \cdots + \lambda_n v_n$

- Have $K[x]$-module isomorphism $V \cong K[x]^n / \ker(\phi)$.
- Need to better understand the sub-module $\ker(\phi)$ of $K[x]^n$.

Observation. Let $A = (a_{i,j}) \in M_{n,n}(K)$ be such that

$$T(v_1, \ldots, v_n) = (v_1, \ldots, v_n)A.$$

Then for each $j = 1, 2, \ldots, n$,

$$Tv_j = a_{1,j}v_1 + a_{2,j}v_2 + \cdots + a_{j,j}v_j + \cdots + a_{n,j}v_n,$$

which can be written as

$$-a_{1,j}v_1 - \cdots - a_{j-1,j}v_2 + (x - a_{j,j})v_j - a_{j+1,j}v_{j+1} - \cdots - a_{n,j} \cdot v_n = 0.$$

In other words, all columns of

$$xI_n - A = \begin{pmatrix} x - a_{1,1} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{2,1} & x - a_{2,2} & \cdots & -a_{2,n} \\ & \cdots & & \\ -a_{n,1} & -a_{n,2} & \cdots & x - a_{n,n} \end{pmatrix} \in M_{n,n}(K[x])$$

are in $\ker(\phi) \subset K[x]^n$

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & & a_{nn} \end{pmatrix}$$

$$\det(xI-A)=x^n+\lambda_1 x^{n-1}+\lambda_2 x^{n-2}+\cdots+\lambda_{n-1}x+\lambda_n$$

$$\lambda_1=\sum a_{ii}$$

$$\lambda_2=\sum_{i=1}^{n-1}\begin{vmatrix} a_{ii} & a_{i,i+1} \\ a_{i+1,i} & a_{i+1,i+1} \end{vmatrix}=\sum_{i<j}\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$$

$$\begin{vmatrix} \lambda-a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda-a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda-a_{33} \end{vmatrix}$$

Proposition. The sub-module $\ker(\phi)$ of $K[x]^n$ is generated by the columns of the characteristic matrix $xI_n - A$ of A.

so columns of $xI_n - A$ form a basis of $\ker\varphi$

Tutorial.

Lemma-Definitioon

$d_1 \mid d_2 \mid \cdots \mid d_n$

1. The invariant factors $d_1, \ldots, d_n \in K[x]$ of

$$xI_n - A \in M_{n,n}(K[x])$$

are independent of the choices of the basis of V, and they are called the invariant factors of T.

$\Rightarrow V \simeq K[x]/d_1K[x] \oplus \cdots \oplus K[x]/d_nK[x]$

2. Let $d_{k+1}, \ldots, d_n \in K[x]$ be the invariant factors of T that are not constant polynomials. Then, V as a $K[x]$-module in which $x \in K[x]$ acts as T, is isomorphic to

$$R/(d_{k+1}) \oplus R/(d_{k+2}) \oplus \cdots \oplus R/(d_n).$$

$R = K[x]$

Define

$$I = \operatorname{ann}(V) = \{f \in K[x] : f(T) = 0\}$$

The minimal polynomial of T is defined to be monic polynomial $f \in K[x]$ that generates the ideal I.

The

Proposition. If $d_1, d_2, \ldots, d_n$ are the invariant factors of T, then

1. the minimal polynomial of T is $d_n \in K[x]$;
2. the characteristic polynomial of T is $\det(xI_n - A) = d_1 d_2 \cdots d_n$.

$\deg d_n = g_n$

$\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \notin \ker\varphi \iff d_1(T)v_1 + \cdots + d_n(T)v_n \neq 0$

$T = \begin{pmatrix} \square & & & \\ & \square & & \\ & & \square & \\ & & & \square \end{pmatrix}$ $g_n \times g_n$

$$A \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -a_0 \\ 1 & 0 & 0 & 0 & 0 & -a_1 \\ 0 & 1 & 0 & 0 & 0 & -a_2 \\ 0 & 0 & 1 & 0 & 0 & -a_3 \\ 0 & 0 & 0 & 1 & 0 & -a_4 \\ 0 & 0 & 0 & 0 & 1 & -a_5 \end{pmatrix}$$

$$ch = \det(xI - A) = a_0 + a_1 x + \cdots + a_5 x^5 + x^6$$

$$= \text{minimal poly}$$

Rational forms of matrices. Consider the $K[x]$-module

$$W = K[x]/(f = a_0 + a_1 x + \cdots + a_n x^n).$$

- As a K-vector space have $\dim_K W = n$
- Have basis $\{v_0 = \bar{1},\ v_1 = \bar{x},\ \ldots,\ v_{n-1} = \overline{x^{n-1}}\}$ over K;
- Have $T(v_0, v_1, \ldots, v_{n-2}, v_{n-1}) = (v_0, v_1, \ldots, v_{n-2}, v_{n-1})R(a)$, where

$$R(a) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ & & \cdots & \cdots & \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

$$A = \begin{pmatrix} 2 & 1 & 1 & -1 \\ -1 & \pi & 0 & 0 \\ 1 & 2 & 3 & -1 \\ 4 & -1 & 0 & 2 \end{pmatrix} \qquad xI - A = \begin{pmatrix} x-2 & -1 & -1 & 1 \\ 1 & x-\pi & 0 & 0 \\ -1 & -2 & x-3 & 1 \\ -4 & 1 & 0 & x-2 \end{pmatrix} \in M_{4,4}(\mathbb{R}[x])$$

$\rightsquigarrow$ Find $d_1 = d_2 = 1$ $\quad d_3 \mid d_4$

$d_3 = ? \qquad d_4 = ?$

$m_3 = \gcd$ of all 3×3 minors

$m_4 = \det$

Definition. For $a = (a_0, a_1, \ldots, a_{n-1}) \in K^n$, the $n \times n$ matrix

$$R(a) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ & & \cdots & \cdots & \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \tag{1}$$

is called the rational matrix defined by a or the companion matrix of a. A matrix of the form $R(a)$ for some a is said to be in rational form.

Fact. The characteristic polynomial of $R(a)$ is

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Let V be a finite dim. K-vector space and $T \in \mathrm{End}_K(V)$.

Theorem

There exists a direct sum decomposition $V = V_1 + \cdots + V_k$ such that for each $1 \leq j \leq k$, $T(V_j) \subset V_j$ and there exists a basis of V_j with respect to which T is in rational form.

Theorem

Every square matrix with entries in a field K is similar to a block diagonal one for which each block is in rational form.

Jordan canonical form of matrices. Consider the $K[x]$-module

$$V = K[x]/((x-a)^n).$$

- Have $\dim V = n$ with K-basis

$$\{v_1 = \overline{(x-a)^{n-1}},\ \cdots,\ v_{n-1} = \overline{(x-a)},\ v_n = \overline{1}\};$$

- Have $T(v_1, \ldots, v_n) = (v_1, \ldots, v_n)J_n(a)$, where

$T =$ mult. by x

$$J_n(a) = \begin{pmatrix} a & 1 & 0 & \cdots & 0 & 0 \\ 0 & a & 1 & \cdots & 0 & 0 \\ 0 & 0 & a & \cdots & 0 & 0 \\ & & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & \cdots & 0 & a \end{pmatrix}.$$

$x \cdot \overline{(x-a)^{n-1}} = \overline{x(x-a)^{n-1}} = \overline{(x-a)(x-a)^{n-1} + a(x-a)^{n-1}}$

$= a\,\overline{(x-a)^{n-1}}$

For $a \in K$, the matrix

$$J_n(a) = \begin{pmatrix} a & 1 & 0 & \cdots & 0 & 0 \\ 0 & a & 1 & \cdots & 0 & 0 \\ 0 & 0 & a & \cdots & 0 & 0 \\ & & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & \cdots & 0 & a \end{pmatrix}$$

is called a Jordan block of size n.

Theorem

Let V be a finite dimensional vector space over an algebraically closed field K and $T \in \mathrm{End}_K(V)$. Then there exists a direct sum decomposition

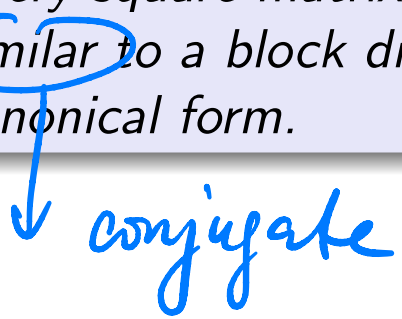
$$V = V_1 + \cdots + V_k$$

such that for each $1 \leq j \leq k$, $T(V_j) \subset V_j$ and there exists a basis of v_j with respect to which T is in Jordan canonical form.

Equivalently,

Theorem

Every square matrix with entries in an algebraically closed field K is similar to a block diagonal one for which each block is in Jordan canonical form.



$$V \simeq K[x]^n / \ker\varphi \simeq K[x]/\langle d_1\rangle \oplus \cdots \oplus K[x]/\langle d_n\rangle$$

$$d_1 = (x-a_1)^{k_1} \cdots (x-a_m)^{k_m}$$

$$\Rightarrow K[x]/\langle d_1\rangle \simeq K[x]/\langle (x-a_1)^{k_1}\rangle \oplus \cdots \oplus K[x]/\langle (x-a_m)^{k_m}\rangle$$

Example. The matrix $A = \begin{pmatrix} -1 & 0 & 0 & 0 & 3 \\ 1 & 2 & 0 & -4 & 0 \\ 3 & 1 & 2 & -4 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ -2 & 0 & 0 & 0 & 4 \end{pmatrix} \in M_{5,5}(\mathbb{Q})$, and $T_A : \mathbb{R}^5 \to \mathbb{R}^5$ defined by $T_A(x) = Ax$, has invariant factors $d_1 = d_2 = d_3 = 1$, $d_4 = x - 1$, and $d_5 = (x-2)^3(x-1)$. Its elementary divisors are then

$$x - 1, \qquad (x-1), \qquad (x-2)^3.$$

It thus has Jordan form