Section 12.3 Jordan Canonical Form

Goal Suppose V is a finite dimensional vector space over a field F. Given linear transformation $T: V \longrightarrow V$, find a basis of V relative to which the matrix for T has a standard (canonical) simple form — as close to diagonal as possible.

Simplifying assumption F is algebraicly closed, i.e. every polynomial $f(x)$ in $F[x]$ factors into linear terms. (e.g. $F = \mathbb{C}$)

Then $f(x) = a(x-\lambda_1)^{\alpha_1}(x-\lambda_2)^{\alpha_2} \cdots (x-\lambda_k)^{\alpha_k}$

In particular, any prime polynomial in $F[x]$ is linear.

Recall

① Given $T: V \longrightarrow V$, space V is an $F[x]$-module with action

$$f(x).v = f(T)(v),$$

that is, $(a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n).v = (a_0 I + a_1 T + a_2 T^2 + \cdots + a_n T^n)(v)$
$= a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_n T^n(v)$

In particular, as F is a field, $F[x]$ is a PID, so V is an $F[x]$-module over PID $F[x]$.

② Finitely generated R-module M over PID R has form

$$M \cong R^r \oplus R/(p_1^{\alpha_1}) \oplus R/(p_2^{\alpha_2}) \oplus \cdots \oplus R/(p_m^{\alpha_m})$$

where the p_i's are primes in the ring R.

Now we will apply decomposition ② to $F[x]$-module V in ①

Observations

ⓐ $F[x]$ module is finitely generated (V is finite-dimensional)

ⓑ V is a torsion $F[x]$-module.

Reason Given $v \in V$, set $\{v, Tv, T^2v, \ldots T^n v\}$ is linearly dependent. Thus there is a linear combo

$a_0 v + a_1 Tv + a_2 T^2 v + \cdots + a_n T^n v = 0$

i.e. $(a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n).v = 0$ ← non-zero polynomial kills v

ⓒ Since the primes in $F[x]$ are linear binomals $(x-\lambda)$ and V is a torsion $F[x]$ module, decomposition ② is

$$V \cong F[x]/((x-\lambda_1)^{\alpha_1}) \oplus F[x]/((x-\lambda_2)^{\alpha_2}) \oplus \cdots \oplus F[x]/((x-\lambda_k)^{\alpha_k})$$

$$v \longleftrightarrow (\overline{f_1(x)}, \overline{f_2(x)}, \cdots \overline{f_k(x)})$$

$F[x]$ action
$f(x).v = f(T)v$
$x.v = Tv$

$F[x]$ action:
$f(x).(\overline{f_1(x)} \cdots \overline{f_k(x)}) = (\overline{f(x)f_1(x)}, \dots, \overline{f(x)f_k(x)})$
$x.(\overline{f_1(x)}, \cdots \overline{f_k(x)}) = (\overline{xf_1(x)}, \dots, \overline{xf_k(x)})$

Action of T

Each subspace $0 \oplus 0 \oplus \cdots \oplus R[x]/(x-\lambda_i)^{\alpha_i} \oplus 0 \oplus \cdots \oplus 0 \cong R[x]/(x-\lambda_i)^{\alpha_i}$ is T-stable (i.e. x-stable).

Also $R[x]/(x-\lambda_i)^{\alpha_i}$ contains linearly independent set $\{(x-\lambda_i)^{\alpha_i-1}, (x-\lambda_i)^{\alpha_i-2}, \dots (x-\alpha_i), 1\}$

[linearly independent because any linear combo is a polynomial of degree at most α_i-1, hence can't be in ideal $((x-\lambda_i)^{\alpha_i})$]

Put $\mathcal{B}_i = \{(0,0,\dots,(x-\lambda_i)^{\alpha_i-1},\dots 0), (0,0,\dots,(x-\lambda_i)^{\alpha_i-2},\dots 0), \dots (0,0\dots 1,\dots 0)\}$

Then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_k$ is basis for V because it contains $\dim(V)$ linearly independent vectors

As each subspace is T-stable (i.e. x-stable) the matrix for T (i.e. x) has block form

$$[T]^{\mathcal{B}}_{\mathcal{B}} = \begin{array}{cl} \begin{bmatrix} J_1 & 0 & & 0 \\ 0 & J_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{bmatrix} & \begin{array}{l} \}\,\mathcal{B}_1 \\ \}\,\mathcal{B}_2 \\ \\ \}\,\mathcal{B}_k \end{array} \\ \quad\underbrace{}_{\mathcal{B}_1}\ \underbrace{}_{\mathcal{B}_2}\ \cdots\ \underbrace{}_{\mathcal{B}_k} & \end{array}$$

Subscripts dropped for clarity.
$\lambda_1 = \lambda$
$\alpha_1 = \alpha$

__Calculation of J_1__

Look at effect of T on basis $\mathcal{B}_1' = \{(x-\lambda)^{\alpha-1}, (x-\lambda)^{\alpha-2}, \dots, 1\}$ of $R[x]/((x-\lambda)^{\alpha})$

$$\begin{aligned} T(x-\lambda)^{\alpha-1} &= x(x-\lambda)^{\alpha-1} = (\lambda+(x-\lambda))(x-\lambda)^{\alpha-1} = \lambda(x-\lambda)^{\alpha-1} + 1(x-\lambda)^{\alpha} \\ T(x-\lambda)^{\alpha-2} &= x(x-\lambda)^{\alpha-2} = \cdots\cdots = \lambda(x-\lambda)^{\alpha-2} + 1\cdot(x-\lambda)^{\alpha-1} \\ &\vdots \\ T(x-\lambda)^{1} &= x(x-\lambda) = \cdots = \lambda(x-\lambda) + 1\cdot(x-\lambda)^2 \\ T(1) &= x\cdot 1 = \underbrace{\lambda\cdot 1 + 1\cdot(x-\lambda)}_{\text{Expressed in terms of } \mathcal{B}_1} \end{aligned}$$

Therefore $J_1 = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$

Conclusion

Relative to $\mathcal{B}$, T has matrix $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & & 0 \\ 0 & 0 & J_3 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_k \end{bmatrix}$

where $J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \lambda_i & 1 & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}$. This is called the Jordan Canonical Form of the transformation T

- Note: $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal $\iff$ each J_i is one-by-one
- Every transformation T (hence very matrix) has a Jordan Canonical form.
- J.F.C is the matrix for T relative to a change of basis
  Conclusion: Any matrix is similar to one in JCF.
- Each similarity class of matrices contains one in JCF. In this sense the JFC matrices "name" the similarity classes of matrices.

Cayley-Hamilton Theorem

Def The chacteristic polynomial of a $n \times n$ matrix A is

$$f(x) = \det(xI - A) = (x-\lambda_1)^{\alpha_1}(x-\lambda_2)^{\alpha_2}\cdots(x-\lambda_k)^{\alpha_k}$$

Note similar matrices A, B have same char. poly $(A = PBP^{-1})$

$$\det(xI - A) = \det(xI - PBP^{-1}) = \det(P(xI-B)P^{-1}) = \det(P)\det(xI-B)\det(P^{-1})$$
$$= \det(xI - B)$$

Cayley-Hamilton Theorem

Given a square matrix A with characteristic polynomial $f(x)$, then $f(A) = 0$.

Proof Let $J = PAP^{-1}$ be JCF of A. Then $f(x) = \det(xI - A)$ $= \det(xI - J) = (x-\lambda_1)^{\alpha_1}(x-\lambda_2)^{\alpha_2}\cdots(x-\lambda_k)^{\alpha_k}$. By decomposition

© $f(A)(v) = f(x).(\overline{f_1(x)} \cdots \overline{f_k(x)}) = (\bar{0}, \bar{0}, \bar{0} \cdots, \bar{0})$ for all $v \in V$.

Therefore $f(A) = 0$.