JORDAN CANONICAL FORM

(SUPPLEMENTARY NOTES - FALL 2007)

<u>Motivation</u>: Let V be a finite dimensional vector space over a field $\mathbb{F}$. It is known that <u>not</u> every linear transformation

$$T: V \to V$$

is diagonizable. An example is $V = \mathbb{R}^2$, $T = L_A$ with $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. We saw that A has a single eigenvalue ($\lambda = 2$) and its corresponding eigenspace E_λ has dimension $= 1 \Rightarrow$ there is no basis of $\mathbb{R}^2$ consisting of eigenvectors for T. In other words, T is not diagonizable since

$$\text{algebraic multiplicity of } \lambda \neq \text{geometric multiplicity of } \lambda$$

One can think of this situation as having a "deficient" eigenspace.

<u>Idea</u>: We would like to be able to represent every linear transformation (even those with deficient eigenspaces) in a canonical fashion, specifically using so-called Jordan blocks of the form

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}_{p \times p} \quad \text{for some } p \geq 1 \quad \left[\text{The case } p = 1 \Rightarrow J = (\lambda)_{1\times 1}\right]$$

as the diagonal blocks of the matrix representation of T.

In short, we want to be able to find basis β such that

$$[T]_\beta = \begin{pmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & \\ & & \ddots & \\ & & & \boxed{J_s} \end{pmatrix}$$

where each J_r is a Jordan block of some size, $r=\overline{1,s}$.

Example of Jordan canonical forms (matrices)

$$A = \begin{pmatrix} 2 & 1 & & & & \\ & 2 & 1 & & & \\ & & 2 & & & \\ & & & 2 & & \\ & & & & 5 & 1 \\ & & & & & 5 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 1 & & & \\ & -1 & & & \\ & & 1 & 1 & \\ & & & 1 & \\ & & & & 4 \end{pmatrix}$$

etc.

<u>Key Ingredients</u>: <u>Generalized Eigenvectors</u> and <u>Generalized Eigenspaces</u>

<u>Def</u>: $T: V \to V$ linear. A vector $v \neq 0$ is called a generalized eigenvector for T is

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

<u>Note</u> 1 Usually, if v is a generalized eigenvector for T, we consider p to be the smallest integer with the property $(T-\lambda I)^p v = 0$, hence $(T-\lambda I)^{p-1} v \neq 0$.

Note ② If $p=1 \Rightarrow (T-\lambda I)v=0$ (and $v\neq 0$) $\Rightarrow$
$\Rightarrow$ v is an eigenvector for T. Hence, every eigenvector of T is also a generalized eigenvector for T.

Def: Let $\boxed{K_\lambda = \{v \in V : (T-\lambda I)^p v = 0 \text{ for some } p\}}$ be the generalised eigenspace of T corresponding to the eigenvalue λ. It therefore consists of all generalized eigenvectors of T and also 0 (the zero vector).

By the note #2 above, for each λ = eigenvalue,

$$\boxed{E_\lambda \subseteq K_\lambda}$$

Note ③ If $\lambda \in \mathbb{F}$ is a scalar such that K_λ contains nonzero vectors, then λ is automatically an eigenvalue for T.

Indeed, if $v \in K_\lambda$ and $v \neq 0$, then $(T-\lambda I)^p v = 0$ and $w = (T-\lambda I)^{p-1} v \neq 0$ for some $p \geq 1$.

Then $(T-\lambda I)w = (T-\lambda I)^p v = 0 \Rightarrow \left.\begin{matrix} Tw = \lambda w \\ w \neq 0 \end{matrix}\right] \Rightarrow \lambda$ = eigenvalue

with corresponding eigenvector $w = (T-\lambda I)^{p-1} v$.

Def: For $v \in K_\lambda$, $v \neq 0$, define the cycle generated by v to be the ordered set

$$\gamma_v = \{(T-\lambda I)^{p-1} v, (T-\lambda I)^{p-2} v, \dots, (T-\lambda I)v, v\}$$

Note that γ has p vectors (p may depend on v)

Proposition: For $v \in K_\lambda$, $v \neq 0$ with $(T-\lambda I)^p v = 0$ and $(T-\lambda I)^{p-1} v \neq 0$, the cycle generated by v is a linearly independent set in K_λ.

Proof: clearly $v \in K_\lambda \Rightarrow (T-\lambda I)v \in K_\lambda$ since

$$(T-\lambda I)^{p-1}(T-\lambda I)v = 0. \text{ etc } \Rightarrow \gamma_v \subset K_\lambda$$

To prove linear independence of γ_v simply take any linear combination that gives rise to 0

$$a_1 (T-\lambda I)^{p-1} v + \dots + a_p v = 0$$

and apply $(T-\lambda I)^{p-1}$ to both sides, to get $a_p = 0$, and iterate this procedure until all $a_i = 0$, $i = \overline{1,p}$.

Observation: For a fixed generalised eigenspace K_λ, it may happen that no single cycle spans the entire K_λ.

Example:

$$[T] = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}_{5\times 5}$$

has only one eigenvalue and $\dim K_\lambda = 5$.

But there are two cycles

$$\{v_1, v_2, v_3\} = \{(T-\lambda I)^2 v_3, (T-\lambda I)v_3, v_3\}$$

and

$$\{v_4, v_5\} = \{(T-\lambda I)v_5, v_5\}$$

where $\beta = \{v_1, v_2, v_3, v_4, v_5\}$.

<u>Theorem</u> $T: V \to V$ linear with characteristic polynomial $f_T(t)$ split over $\mathbb{F}$. Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T and $m_1, \dots, m_k$ be their algebraic multiplicities, respectively.

Hence $f_T(t) = (-1)^n (t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k}$, $n = \dim V$

Then (1) $\dim K_{\lambda_i} = m_i$, $i = \overline{1,k}$

(2) $V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \dots \oplus K_{\lambda_k}$

<u>Sketch of proof</u>:

We first prove $V = K_{\lambda_1} + \dots + K_{\lambda_k}$, that is, for all $v \in V$, there exist $v_i \in K_{\lambda_i}$ ($i = \overline{1,k}$) such that

$$v = v_1 + \dots + v_k$$

(see proof by induction on page 486).

Next, we show $\dim K_{\lambda_i} \le m_i$ (see proof page 486 top)

Finally, since

$$n = \dim V = \dim(K_{\lambda_1} + \dots + K_{\lambda_k}) \le \dim K_{\lambda_1} + \dots + \dim K_{\lambda_k}$$

$$\le m_1 + \dots + m_k = n$$

it follows that $\dim K_{\lambda_i} = m_i$, $i = \overline{1,k}$ and

$$\dim(K_{\lambda_1} + \dots + K_{\lambda_k}) = \dim K_{\lambda_1} + \dots + \dim K_{\lambda_k}$$

which implies $K_{\lambda_i} \cap \sum_{j \neq i} K_{\lambda_j} = \{0\}$ $(i = \overline{1,k})$, hence

$V = K_{\lambda_1} \oplus \dots \oplus K_{\lambda_k}$. $\square$

<u>Corollary</u>: If $\beta_1, \dots, \beta_k$ are basis for $K_{\lambda_1}, \dots, K_{\lambda_k}$ resp, then
$$\beta = \beta_1 \cup \dots \cup \beta_k \text{ is a basis for } V.$$

! <u>Next step</u>: Identify bases β_i for K_{λ_i}, $i = \overline{1,k}$ such that $T|_{K_{\lambda_i}}$ is represented using Jordan blocks on the diagonal:

$$[T|_{K_{\lambda_i}}]_{\beta_i} = \begin{pmatrix} \boxed{\begin{matrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{matrix}} & & & \\ & \boxed{\begin{matrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{matrix}} & & \\ & & \ddots & \\ & & & \boxed{\begin{matrix} \lambda_i & 1 & \\ & \ddots & \ddots \\ & & \lambda_i \end{matrix}} \end{pmatrix}$$

Obviously, such a basis β_i must consist of disjoint cycles generated by generalized eigenvectors in K_{λ_i}.

<u>Theorem</u>: For any λ_i = eigenvalue for T, there exists a basis β_i for K_{λ_i} of the form
$$\beta_i = \gamma_1 \cup \dots \cup \gamma_{n_i}, \text{ where}$$
γ_j are cycles of generalized eigenvectors.

Proof: (by mathematical induction, see pages 490-491). $\square$

Note that along the way, one also proves that the leading vectors in each cycle, which is an eigenvector for T corresponding to eigenvalue λ_i, form a linearly independent set.

More specifically if

$$\gamma_1 = \{(T-\lambda_i I)^{p_1-1} v_1, \dots, (T-\lambda_i I)v_1, v_1\}$$

$$\vdots$$

$$\gamma_{n_i} = \{(T-\lambda_i I)^{p_{n_i}-1} v_{n_i}, \dots, (T-\lambda_i I)v_{n_i}, v_{n_i}\}$$

are the disjoint cycles that form a basis for K_{λ_i}, then $w_1 = (T-\lambda_i I)^{p_1-1} v_1, \dots, w_{n_i} = (T-\lambda_i I)^{p_{n_i}-1} v_{n_i}$ are eigenvectors that form a basis for E_{λ_i}

$$E_{\lambda_i} = \text{span}\{w_1, \dots, w_{n_i}\}$$

Hence $n_i = \dim E_{\lambda_i}$ (= geom. multiplicity of λ_i).

Conclusion ① The number of disjoint cycles (Jordan blocks) that form a basis for K_λ equals the geometric multiplicity of the eigenvalue λ.

When is T diagonizable?

As a consequence, the situation when $E_{\lambda_i} = K_{\lambda_i}$ is characterized by the fact that there are m_i Jordan blocks corresponding to eigenvalue λ_i hence all blocks are $1\times 1 \Leftrightarrow$ there are m_i linearly independent eigenvectors corresponding to λ_i.

Therefore $E_{\lambda_i} = K_{\lambda_i}$, $i=\overline{1,k} \Leftrightarrow$ T is diagonizable.

Example: Let $A = \begin{bmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{bmatrix}$

One computes $f_A(t) = -(t-1)(t-2)^2$, hence $\lambda_1 = 1$, $\lambda_2 = 2$

$m_1 = 1$, $m_2 = 2$

$\dim E_{\lambda_1} = 1$ (since $1 \le \dim E_{\lambda_1} \le m_1 = 1$) $\Rightarrow$ there is one (1×1) single (1×1) Jordan block for λ_1

$$\dim E_{\lambda_2} = \dim N(A-2I) = 3 - \operatorname{rank}(A-2I)$$
$$= 3 - \operatorname{rank}\begin{bmatrix} -5 & 3 & -2 \\ -7 & 4 & -3 \\ 1 & -1 & 0 \end{bmatrix} = 3-2 = 1$$

$\Rightarrow$ one eigenvector generated $E_{\lambda_2} \Rightarrow$

$\Rightarrow$ one single (2×2) Jordan block for λ_2

$$\Rightarrow \quad J_A = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 2 & 1 \\ 0 & 0 & 2 \end{array}\right)$$

General guidelines for finding a Jordan basis for T

1. Find all (distinct) eigenvalues of T
2. For each eigenvalue, determine the algebraic and geometric multiplicities.
3. For each generalized eigenspace K_λ, determine a Jordan canonical basis for $T|_{K_\lambda}$ as the union of distinct cycles of generalized eigenvectors.

See solution to sample final exam (#5b) for one method of finding such basis of generalized eigenvectors.

See also examples 2, 3, 4 pages 500-508.

Minimal polynomial

Let $T: V \to V$ linear, V = finite dimensional vector space, and $f_T(t)$ its characteristic polynomial. We know (Cayley-Hamilton) that f_T annihilates T, that is

$$f_T(T) = 0.$$

<u>Def:</u> The minimal polynomial of T is the <u>monic</u> polynomial of least positive degree $p(t)$ such that

$$p(T) = 0.$$

We denote the minimal polynomial of T by $p_T(t)$.

Theorem: The minimal polynomial $p_T(t)$ of T exists and is unique. Moreover, $p(t)$ divides $f_T(t)$:

$$f_T(t) = p_T(t)\, q(t) \quad \text{for some } q = q(t).$$

Proof: See theorem 7.12, page 516.

Theorem: Let λ be an eigenvalue for T. Then $p_T(\lambda) = 0$, where p is the minimal polynomial of T.

Proof: See Theorem 7.14.

Corollary If $f_T(t)$ splits, $f_T(t) = (-1)^n (t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k}$

then $p_T(t) = (t-\lambda_1)^{d_1} \cdots (t-\lambda_k)^{d_k}$ with some $1 \leq d_i \leq m_i$

$i = \overline{1,k}$.

Example: If A has $f_A(t) = -(t-1)(t-2)^2$

then $p_A(t)$ is either $(t-1)(t-2)$ or $(t-1)(t-2)^2$.

Question: Does $(A - I)(A - 2I) = 0$?

If yes, then $p_A(t) = (t-1)(t-2)$.

Otherwise, $p_A(t) = (t-1)(t-2)^2 = -f_T(t)$.

Theorem $(t-\lambda)^d$ divides $p_T(t) \overset{d = \max}{\Longleftrightarrow}$ there exists a Jordan block of size $d \times d$ corresp. to λ.

(See exercise 13 page 523).

Consequence: The largest Jordan block corresp. to an eigenvalue λ equals the multiplicity of λ in the minimal polynomial of T.

Consequence: T is diagonalizable if and only if

$$P_T(t) = (t-\lambda_1) \cdots (t-\lambda_k)$$

(that is $d_i = 1$, $i = \overline{1,k}$).