<u>19/11/19</u>

Phrase to summarise the results of Prop 17.1: say the char poly / evalues / det / trace etc. of a (square) matrix are <u>invariant under similarity</u>.

However, the properties in 17.1 are not sufficient to determine a matrix up to similarity:

Eg. Let

$$A = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ & 1 & 0 & 0 \\ \hline & 0 & 1 & 1 \\ & & & 1 \end{array}\right),$$

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ & 1 & 1 & 0 \\ & 0 & 1 & 0 \\ & & & 1 \end{pmatrix}$$

Then A & B have the same

char poly $(x-1)^4$

evalues 1

det $\quad 1$

multiplicities $\quad a(1) = 4$

$\qquad\qquad\qquad g(1) = 2$

trace $\quad 4$

YET

A, B are <u>not similar</u>.

Why? Well,

$$A - I = \begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}$$

$$B - I = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

so $\quad (A - I)^2 = 0$

$\qquad (B - I)^2 \neq 0.$

<u>Aim</u> To find enough properties of a matrix A to determine A up to similarity.

This we shall do in the next chapter

18. Jordan Canonical Form

Defn Let $\lambda \in \mathbb{C}$, and define the $n \times n$ matrix

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

Such a matrix is a *Jordan block*.

Eg. $J_2(-2) = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}$

$J_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

$J_1(\lambda) = (\lambda)$.

Properties of Jordan blocks:

Prop 18.1 Let $J = J_n(\lambda)$.

1) Char. poly is $(x-\lambda)^n$.

2) Only evalue of J is λ.

Multiplicities:

$a(\lambda) = n$

$g(\lambda) = 1$

3) We have

$$J - \lambda I = J_n(0) = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 1 \\ & & & 0 \end{pmatrix}$$

and linear map $v \longrightarrow (J - \lambda I) v$ sends standard basis vectors

$$e_n \to e_{n-1} \to e_{n-2} \to \cdots \to e_1 \to 0.$$

4) $(J-\lambda I)^n = 0$, and for $1 \leqslant i \leqslant n-1$, $(J - \lambda I)^i$ sends

$$e_n \to e_{n-i},\ e_{n-1} \to e_{n-i-1}, \ldots$$

and

$$(J - \lambda I)^i = \begin{pmatrix} \overleftrightarrow{\,0 \ \cdots \ \overset{i+1}{}\ 1\,} & & \\ & \ddots & \ddots & 1 \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix}$$

Pf. (1) Clear (as J upper Δ).

2) The λ-eigenspace $E_\lambda(J)$ is the solution space of system

$$\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{pmatrix} v = 0$$

This space is $Sp(e_1)$, hence $g(\lambda) = \dim E_\lambda(J) = 1$.

3) Clear.

4) $(J-\lambda I)^n$ sends all $e_i \to 0$.

and $(J-\lambda I)^i$ sends $e_n \to e_{n-i}$, etc. //

<u>Block-diagonal matrices</u>

<u>Defn</u> If $A_1, \ldots, A_k$ are square matrices, where A_i is $n_i \times n_i$ define

$$A_1 \oplus A_2 \oplus \cdots \oplus A_k = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix}$$

a <u>block-diagonal</u> matrix

This is $n \times n$, where
$$n = \sum_1^k n_i.$$

Eg. If
$$A_1 = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}, \quad A_2 = (-1)$$
then
$$A_1 \oplus A_2 = \left(\begin{array}{cc|c} 2 & 1 & 0 \\ 3 & 0 & 0 \\ \hline 0 & 0 & -1 \end{array}\right).$$

<u>Prop 18.2</u> Let

$$A = A_1 \oplus \dots \oplus A_k$$

and let $p_i(x)$ be char. poly of A_i. Then

1) Char poly of A is $\prod_{i=1}^{k} p_i(x)$.

2) Evalues of A = $\bigcup_{1}^{k}$(evalues of A_i)

3) For each evalue λ, geometric mult

$$g_A(\lambda) = \sum_{i=1}^{k} g_{A_i}(\lambda)$$

where

$$g_A(\lambda) = \dim E_\lambda(A),$$

$$g_{A_i}(\lambda) = \dim E_\lambda(A_i).$$

4) For any poly. $q(x)$,

$$q(A) = q(A_1) \oplus \dots \oplus q(A_k).$$

Pf 1) Char poly of A is

$$|xI_n - A| = \det \begin{pmatrix} xI_{n_1} - A_1 & & \\ & \ddots & \\ & & xI_{n_k} - A_k \end{pmatrix}$$

$$= \prod_{i=1}^{k} \det(xI_{n_i} - A_i)$$

$$= \prod_{i=1}^{k} p_i(x).$$

2) Follows from (1).

3) Sheet 6 qu.

4) Well,

$$A^2 = \begin{pmatrix} A_1 & & O \\ & \ddots & \\ O & & A_k \end{pmatrix}^2$$

$$= \begin{pmatrix} A_1^2 & & \\ & \ddots & \\ & & A_k^2 \end{pmatrix}$$

and similarly for all powers of A

and all polys $q(A)$. //

<u>The Main Theorem!</u>

<u>Theorem 18.3</u> (JCF theorem)

Let A be an $n \times n$ matrix over $\mathbb{C}$. Then

(1) A is similar to a matrix of the form

$J_{n_1}(\lambda_1) \oplus \cdots\cdots \oplus J_{n_k}(\lambda_k)$ 

where $\lambda_i \in \mathbb{C}$ and $\sum_{1}^{k} n_i = n$.

(2) The block-diagonal matrix  is <u>unique</u>, apart from changing the order of the blocks.

It is called the <u>Jordan Canonical Form</u> (JCF) of the matrix A.

<u>Proof</u> Later.

Eg. Here are some JCF's

a) $$J_2(1) \oplus J_2(1) = \left(\begin{array}{cc|cc} 1 & 1 & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ 0 & 1 & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{0}} & 1 & 1 \\ & & 0 & 1 \end{array}\right)$$

$$J_3(1) \oplus J_1(1) = \left(\begin{array}{ccc|c} 1 & 1 & 0 & \\ & 1 & 1 & 0 \\ & & 1 & \\ \hline & 0 & & 1 \end{array}\right)$$

Uniqueness part (2) tells us these are <u>not</u> similar.

b) $J_2(-i) \oplus J_1(0) = \begin{pmatrix} -i & 1 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$J_1(0) \oplus J_2(-i) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -i & 1 \\ 0 & 0 & -i \end{pmatrix}$

These are similar (sheet 6, Q1).

c) Diagonal matrix

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ 0 & & \ddots & \\ & & & \lambda_n \end{pmatrix} = J_1(\lambda_1) \oplus J_1(\lambda_2) \oplus \dots \oplus J_1(\lambda_n)$$

21/11/19

Points about JCF Thm

1) In the JCF , the evalues $\lambda_1, \ldots, \lambda_k$ are <u>not</u> necessarily distinct.

2) JCF theorem is not true for $n \times n$ matrices over $\mathbb{R}$

eg. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ has no evalues in $\mathbb{R}$. However, there <u>is</u> a version of the JCF thm. for matrices over $\mathbb{R}$, in which each pair of complex evalues $\lambda, \bar{\lambda}$ gets replaced by a real 2×2 matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ with evalues $\lambda, \bar{\lambda}$.

(See next week's Project).

How to calculate the JCF of an arbitrary matrix over $\mathbb{C}$

Consider a JCF matrix J of form . For each evalue λ, collect all the Jordan blocks with evalue λ, and change the order of the blocks to write

$$J = \big(\overbrace{J_{n_1}(\lambda) \oplus \dots \oplus J_{n_a}(\lambda)}^{\lambda\text{-blocks}}\big)$$

$$\oplus \big(J_{m_1}(\mu) \oplus \dots \oplus J_{m_b}(\mu)\big)$$

$$\oplus \dots\dots$$

(where $\lambda, \mu, \dots$ are distinct).

Prop 18.4 For J as above:

1) $n_1 + \dots + n_a = a(\lambda)$, the alg. multiplicity of λ

2) a = no. of λ-blocks
$= g(\lambda)$, geom. mult. of λ.

Pf. 1) The power of $x-\lambda$ that appears in the char. poly of J is

$$(x-\lambda)^{\sum_1^a n_i}$$

Hence $a(\lambda) = \sum_1^a n_i$.

2) By 18.1, each λ-block $J_{n_i}(\lambda)$ has geometric mult 1.

Hence by Prop 18.2(3)

$$g(\lambda) = \text{no. of } \lambda\text{-blocks}$$
$$= a. \quad //.$$

Prop 18.4 already takes us some way towards computing JCF's.

Ex. 1) Find JCF of

$$A = \begin{pmatrix} -1 & 5 & 0 & 0 & 1 \\ & -1 & 0 & 0 & 0 \\ & & 1 & 0 & -1 \\ & 0 & & 1 & 0 \\ & & & & 1 \end{pmatrix}$$

<u>Ans</u> Char. poly $(x+1)^2(x-1)^3$

Evalues $-1, 1$ with alg. mults $2, 3$.

Find geom. mults:

$$g(-1) = 1$$
$$g(1) = 2$$

So the JCF has 1 -1-block

2 1-blocks

Hence the JCF must be

$$J_2(-1) \oplus J_2(1) \oplus J_1(1).$$

2) Find JCF of

$$A = \begin{pmatrix} 2 & 1 & -1 & 0 & 0 \\ & 2 & 0 & 1 & 0 \\ & & 2 & 0 & 0 \\ & 0 & & 2 & 1 \\ & & & & 2 \end{pmatrix}$$

<u>Ans</u> Char. poly is $(x-2)^5$.

To compute $g(2)$:

$$A - 2I = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ & 0 & 0 & 1 & 0 \\ & & 0 & 0 & 0 \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

This has rank 3, so

$$g(2) = 2.$$

So JCF has 2 Jordan blocks.

Hence JCF of A is

either $J_3(2) \oplus J_2(2) = J_1$

or $J_4(2) \oplus J_1(2). = J_2$

<u>Which?</u>

Observe that

$$J_1 - 2I = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & & \\ & 0 & 1 & & \\ & & 0 & & \\ \hline & & & 0 & 1 \\ & & & & 0 \end{array}\right)$$

and

$$J_2 - 2I = \left(\begin{array}{cccc|c} 0 & 1 & 0 & 0 & \\ & 0 & 1 & 0 & \\ & & 0 & 1 & \\ & & & 0 & \\ \hline & & & & 0 \end{array}\right)$$

So

$$(J_1 - 2I)^2 = \left(\begin{array}{ccc|cc} 0 & 0 & 1 & & \\ & 0 & 0 & & \\ & & 0 & & \\ \hline & & & 0 & 0 \\ & & & & 0 \end{array}\right)$$

$$(J_2 - 2I)^2 = \left(\begin{array}{cccc|c} 0 & 0 & 1 & 0 & \\ & 0 & 0 & 1 & \\ & & 0 & 0 & \\ & & & 0 & \\ \hline & & & & 0 \end{array}\right)$$

Hence

$$\operatorname{rank}(J_1-2I)^2 = 1$$

$$\operatorname{rank}(J_2-2I)^2 = 2.$$

Now compute

$$\operatorname{rank}(A-2I)^2 = 2.$$

Therefore JCF of A is

$$J_2 = J_4(2) \oplus J_1(2)$$

Uniqueness of JCF

We prove the uniqueness part (2) of JCF Thm 18.3.

Theorem 18.5 Suppose A is $n \times n$ matrix over $\mathbb{C}$, and that A is similar to a JCF

$$J = J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_k}(\lambda_k).$$

Then J is <u>unique</u>, apart from changing the order of the Jordan blocks.

Note Changing the order of the blocks *does* give a similar matrix, since

$$A_1 \oplus A_2 \overset{\text{similar}}{\sim} A_2 \oplus A_1$$

(Sheet 7, Q1).

Pf of 18.5

(A) First consider the case where A has just *one* evalue λ, so has char poly $(x-\lambda)^n$.

So A is similar to JCF J with Jordan blocks $J_{n_i}(\lambda)$. For each i with $1 \leq i \leq n$, let a_i be the number of Jordan blocks $J_i(\lambda)$ in J. So

$$J = J_1(\lambda)^{a_1} \oplus J_2(\lambda)^{a_2} \oplus \dots \oplus J_r(\lambda)^{a_r}$$

where all $a_i \geq 0$.

Define, for $i \geq 1$,

$$m_i = \operatorname{rank}(A - \lambda I)^i,$$

so ~~att~~ also $m_i = \operatorname{rank}(J - \lambda I)^i$.

<u>Claim</u> Given the values $m_1, m_2, \ldots\ldots$, we can compute the values $a_1, a_2, \ldots, a_r$. (hence can compute the JCF J).

Recall notation: A similar to

$$J = J_1(\lambda)^{a_1} \oplus \ldots \oplus J_r(\lambda)^{a_r}$$

and for $i \geq 1$,

$$m_i = \operatorname{rank}(A-\lambda I)^i$$
$$= \operatorname{rank}(J-\lambda I)^i.$$

<u>Claim</u> The values $m_1, m_2, \ldots$ determine the values $a_1, \ldots, a_r$.

Pf. Now

$$m_1 = \operatorname{rank}(J-\lambda I).$$

Also

$$J-\lambda I = \begin{pmatrix} \underbrace{0 \ \cdots \ 0}_{a_1} & & \\ & \underbrace{\boxed{\begin{smallmatrix}0&1\\0&0\end{smallmatrix}} \cdots \boxed{\begin{smallmatrix}0&1\\0&0\end{smallmatrix}}}_{a_2} & \\ & & \ddots \end{pmatrix}$$

This has rank

$$a_2 + 2a_3 + 3a_4 + \ldots + (r-1)a_r$$

This gives equation

(1) $\underline{m_1 = a_2 + 2a_3 + \ldots + (r-1)a_r}$

Next consider

$$m_2 = \operatorname{rank}(J-\lambda I)^2.$$

By 18.1,

$$(J-\lambda I)^2 = \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & \boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}} & & & \\ & & & & \ddots & & \\ & & & & & \boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}} & \\ & & & & & & \boxed{\begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}} \\ & & & & & & & \ddots \end{pmatrix}$$

This has rank

$$a_3 + 2a_4 + \dots + (r-2)a_r.$$

Hence

(2) $m_2 = a_3 + 2a_4 + \dots + (r-2)a_r.$

Continue taking further powers of $(J-\lambda I)$. Get $r-1$ eqns:

(1) $m_1 = a_2 + 2a_3 + \dots + (r-1)a_r$

(2) $m_2 = a_3 + 2a_4 + \dots + (r-2)a_r$

$\vdots$

($r-2$) $m_{r-2} = a_{r-1} + 2a_r$

($r-1$) $m_{r-1} = a_r$

Note $m_r = \operatorname{rank}(J-\lambda I)^r = 0$ as $J_r(\lambda)$ is the largest block of J.

Given the values of $m_1, \ldots, m_{r-1}$, we can solve for $a_1, \ldots, a_r$.

This proves the Claim.

Hence we've proved the uniqueness theorem 18.5 in the case where A has just <u>one</u> evalue λ.

(B) <u>General case of Thm 18.5</u>

We are given that the matrix A is similar to a JCF J.

Let λ be ~~to~~ an evalue of A, and let J_λ be the block-diagonal matrix consisting of all the λ-blocks in J. So (changing order of blocks),

$$J = J_\lambda \oplus L$$

where λ is <u>not</u> an evalue of L. Let L be $\ell \times \ell$. Then $L - \lambda I_\ell$ has rank ℓ (it is invertible), so

$$\underline{\operatorname{rank}(L - \lambda I_\ell)^i = \ell \quad \forall i \geq 1.}$$

Again define

$$m_i = \operatorname{rank}(A - \lambda I)^i \quad (i \geq 1).$$

Then

$$m_i = \operatorname{rank}(J - \lambda I)^i$$
$$= \operatorname{rank}(J_\lambda - \lambda I)^i + \operatorname{rank}(L - \lambda I)^i$$

so

$$m_i = \operatorname{rank}(J_\lambda - \lambda I)^i + \ell \quad \forall i \geq 1.$$

Hence the values of $m_1, m_2, \ldots$ determine ~~\$~~ all the values of

$$\operatorname{rank}(J_\lambda - \lambda I)^i \quad (i \geq 1).$$

As J_λ has one the single evalue λ, by part (A) of the proof, these ranks determine the number of λ-blocks of each size in J_λ.

We can repeat this process for all the other evalues of A, and hence determine the JCF J.

This completes the proof of the uniqueness theorem 18.5 //.

One final example on computing JCFs:

Eg. Suppose A is an ~~8~~ delightful 8×8 matrix with properties

- char poly is $(x-1)^8$.
- ranks of powers $(A-I)^i$ $(i=1,2,3)$ are:

i	1	2	3
$\operatorname{rank}(A-I)^i$	6	4	2

What are the possible JCF's similar to A?

<u>Ans</u> Well,

1) $\text{rank}(A-I)=6 \Rightarrow g(1)=2$

$\Rightarrow$ JCF has 2 blocks

$\Rightarrow$ JCF is $J_7 \oplus J_1$, $J_6 \oplus J_2$, $J_5 \oplus J_3$ or $J_4 \oplus J_4$.

2) $\text{rank}(A-I)^2 = 4$

$\Rightarrow J_6 \oplus J_2, J_5 \oplus J_3, J_4 \oplus J_4$

3) $\text{rank}(A-I)^3 = 2$

$\Rightarrow J_5 \oplus J_3, J_4 \oplus J_4$

So there are two JCFs possible

To determine the JCF of A, need to calculate

$\text{rank}(A-I)^4$.

If this is 1, JCF is $J_5 \oplus J_3$

If 0, JCF is $J_4 \oplus J_4$.

Next aim: proof of JCF Thm 18.3, part (1).

First:

19. Minimal polynomial

Defn Let $T: V \to V$ linear map, where V is over F. We say a poly. $m(x) \in F[x]$ is a minimal polynomial for T if

1) $m(T) = 0$
2) Leading coeff of $m(x)$ is 1, i.e. $m(x)$ is a monic poly.
3) Degree $\deg(m(x))$ is as small as possible.

Note ~~S~~ Since by Cayley-Ham, $\exists$ nonzero polys. $p(x)$ s.t. $p(T) = 0$, minimal poly. exists.