Lecture 4

Thm: A and B are similar iff $sI-A$ and $sI-B$ have the same invariant polynomials.

pf: $\Rightarrow$ Assume $A = T^{-1}BT$.

$$sI - A = sI - T^{-1}BT = T^{-1}(sI-B)T$$

$\Rightarrow$ $sI-A$ and $sI-B$ have same invariant polynomials

$\Leftarrow$ Assume $sI-A$ and $sI-B$ have same invariant polynomials

$\exists$ $E_1(s), E_2(s), F_1(s), F_2(s)$ with constant non-zero determinants s.t.

$$E_1(s)(sI-A)F_1(s) = E_2(s)(sI-B)F_2(s)$$

$$\Rightarrow (sI-B) = E_2^{-1}(s)\,E_1(s)(sI-A)F_1(s)\,F_2^{-1}(s)$$

$\Rightarrow$ $\exists$ nonsingular constant matrices P & Q

s.t. $(sI-B) = P(sI-A)Q$

$$\Rightarrow PQ = I \Rightarrow P = Q^{-1}$$

$$\Rightarrow B = PAQ = Q^{-1}AQ.$$

J_A is unique upto a perm. of its diag. elem.
J_A is the "Jordan normal form" of A.

Thm A is similar to a block diagonal matrix J_A whose diag. elements are Jordan blocks corresp. to the ch. values of A.

pf: Need to show $C(s) = sI - A$ and $sI - J_A$ have the same invariant polynomials.

Find $E(s)$ and $F(s)$ s.t. $E(s)C(s)F(s) = D_C(s)$

where $D_C(s) = \begin{bmatrix} h_1(s) & & \\ & \ddots & \\ & & h_n(s) \end{bmatrix}$, $\det(E(s)) \cdot \det(C(s)) \det(F(s)) = \det(D_C(s))$

$$\Rightarrow \quad c \cdot \det(C(s)) = \det(D_C(s)) \quad \Rightarrow \det(C(s)) = \prod_{i=1}^{n} h_i(s)$$

Let $\det(C(s)) = (s-\lambda_1)^{n_1} \cdots (s-\lambda_p)^{n_p}$, where $n_1 + n_2 + \cdots n_p = n$.

$h_{p-1}(s)$ divides $h_p(s)$, $p = 2, \ldots n \Rightarrow$

$$\left\{ \begin{array}{l} h_1(s) = (s-\lambda_1)^{a_1} \cdots (s-\lambda_p)^{a_p} \\ h_2(s) = (s-\lambda_1)^{b_1} \cdots (s-\lambda_p)^{b_p} \\ \vdots \\ h_n(s) = (s-\lambda_1)^{x_1} \cdots (s-\lambda_p)^{x_p} \end{array} \right\} \text{ where } \left\{ \begin{array}{l} 0 \le a_1 \le b_1 \le \cdots \le x_1 \le n_1 \\ 0 \le a_2 \le b_2 \le \cdots \le x_2 \le n_2 \\ \vdots \\ 0 \le a_p \le b_p \le \cdots \le x_p \le n_p \end{array} \right\}$$

and $a_1 + b_1 + \cdots + x_1 = n_1, \ldots, a_p + b_p + \cdots + x_p = n_p$

WLOG assume $h_1(s) = \cdots = h_k(s) = 1$.

Let $J_m = \begin{bmatrix} J_{m_1} & & \\ & \ddots & \\ & & J_{m_t} \end{bmatrix}$, where $J_{m_i} = \begin{bmatrix} \lambda_{m_i} & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_{m_i} \end{bmatrix}$ is an $l_{m_i} \times l_{m_i}$ Jordan block corresponding to the factor $(s-\lambda_{m_i})^{l_{m_i}}$ in the mth invariant polynomial $h_m(s) = (s-\lambda_{m_1})^{l_{m_1}} \cdots (s-\lambda_{m_t})^{l_{m_t}}$

$sI - J_m \overset{E}{=} \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & h_m(s) \end{bmatrix}$ (since $(s-\lambda_{m_i})^{l_{m_i}}$ and $(s-\lambda_{m_j})^{l_{m_j}}$ are rel. prime for $m_i \ne m_j$ pf: Hw)

Let $J_A = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{bmatrix}$. Then $sI - J_A \overset{E}{=} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & h_{k+1}(s) & \\ & & & & \ddots \\ & & & & & h_n(s) \end{bmatrix} = D_C(s)$

Hence, A is similar to J_A.

Thm: Let $g(s) = s^m + g_1 s^{m-1} + \cdots + g_m$ denote an mth order monic polynomial for which $g(A) = 0$.

Then: $g(s)\,Adj(sI-A) = a(s)\left[s^{m-1} I + s^{m-2}(A + g_1 I) + \cdots + (A^{m-1} + g_1 A^{m-2} + \cdots + g_{m-1} I) \right]$

pf: $\lambda^j - \mu^j = (\lambda-\mu)(\lambda^{j-1} + \lambda^{j-2}\mu + \lambda^{j-3}\mu^2 + \cdots + \mu^{j-1})$

(this factorization assumes λ and μ commute) (pf: exercise)

$$\Rightarrow \quad g(\lambda) - g(\mu) = (\lambda-\mu)\{ (\lambda^{m-1} + \lambda^{m-2}\mu + \cdots + \mu^{m-1}) + g_1(\lambda^{m-2} + \lambda^{m-3}\mu + \cdots + \mu^{m-2}) + \cdots + g_{m-1} \}$$

Let $\lambda = sI$ and $\mu = A$. Then, since sI commutes with A:

$$g(s)I = g(sI) - g(A) = (sI-A)\{ s^{m-1}I + s^{m-2}(A + g_1 I) + s^{m-3}(A^2 + g_1 A + g_2 I) + \cdots + (A^{m-1} + g_1 A^{m-2} + \cdots + g_{m-1} I) \}$$

$Adj(sI-A)(sI-A) = a(s)I$

$$\Rightarrow \quad Adj(sI-A)\,g(s) = a(s)\left[s^{m-1}I + s^{m-2}(A + g_1 I) + \cdots + (A^{m-1} + g_1 A^{m-2} + \cdots + g_{m-1} I) \right]$$

Corollary: $Adj(sI-A) = s^{n-1} I + s^{n-2}(A + a_1 I) + \cdots + (A^{n-1} + a_1 A^{n-2} + \cdots + a_{n-1} I)$

pf: HW.

Cayley-Hamilton (alt. pf)

If $a(s) = \det(sI - A)$, then $a(A) = 0$.

pf: $(sI-A)\operatorname{Adj}(sI-A) = a(s)\,I = a(sI)$

$$B(s) \triangleq \operatorname{Adj}(sI-A) = s^{n-1} I + s^{n-2}(A + a_1 I) + \cdots + (A^{n-1} + a_1 A^{n-2} + \cdots + a_{n-1} I)$$

$$= (sI)^{n-1} + (sI)^{n-2}(A + a_1 I) + \cdots + (A^{n-1} + a_1 A^{n-2} + \cdots + a_{n-1} I)$$

$$\Rightarrow B(A) = A^{n-1} + A^{n-2}(A + a_1 I) + \cdots + (A^{n-1} + a_1 A^{n-2} + \cdots + a_{n-1} I)$$

$$\Rightarrow a(A) = (A - A)\,B(A) = 0.$$

Defn: The minimal polynomial, $\Psi(s)$, of an $n\times n$ matrix A is the monic polynomial of least degree for which $\Psi(A)=0$.

Thm: If $g(A)=0$ for the polynomial $g(s)$, then $\Psi(s)$ divides $g(s)$ without remainder.

pf: Let $g(s) = m(s)\Psi(s) + r(s)$, where $\deg(r(s)) < \deg(\Psi(s))$.

Then $0 = g(A) = m(A)\Psi(A) + r(A) = r(A) \Rightarrow r(A)=0$.

But $\deg(r(s)) < \deg(\Psi(s))$ and since $\Psi(s)$ is the minimal poly. $r(s)=0$.

Thm: Let $\delta_c^{(n-1)}(s)$ denote the greatest common monic divisor (g.c.m.d) of $\mathrm{Adj}(C(s))$, where $C(s)=sI-A$. The minimal polynomial of A is given by $\Psi(s) = \dfrac{a(s)}{\delta_c^{(n-1)}(s)}$.

pf: Let $\mathrm{Adj}(C(s)) = \delta_c^{(n-1)}(s)\,\Gamma(s)$, $\Gamma(s)$ is "reduced adjoint matrix", elements of $\Gamma(s)$ are rel. prime.

$a(s)I = (sI-A)\,\mathrm{Adj}(C(s)) = (sI-A)\,\delta_c^{(n-1)}(s)\,\Gamma(s)$

$\Rightarrow a(s) = \delta_c^{(n-1)}(s)\cdot g(s)$ (where $g(s)$ is monic)

$\Rightarrow g(sI) = g(s)I = (sI-A)\Gamma(s)$

$\Rightarrow g(A)=0. \Rightarrow \Psi(s)$ divides $g(s)$ without remainder

Need to show $g(s)$ divides $\Psi(s)$ without remainder, and thus $\Psi(s) = g(s) = \dfrac{a(s)}{\delta_c^{(n-1)}(s)}$.

$\mathrm{Adj}(sI-A)\Psi(s) = a(s)\,p(sI,A)$

$\Rightarrow \delta_c^{(n-1)}(s)\,\Gamma(s)\Psi(s) = \delta_c^{(n-1)}(s)\,g(s)\cdot p(sI,A)$

$\Rightarrow \Gamma(s)\Psi(s) = g(s)\cdot p(sI,A)$.

$\Rightarrow g(s)$ divides $\Psi(s)$ w/o rem. (since the elements of $\Gamma(s)$ are rel. prime)

Corollary: Let $h_1(s), \ldots, h_n(s)$ denote the invariant polynomials of $C(s) = sI - A$. Then $h_n(s) = \psi(s)$.

pf:
$$h_n(s) = \frac{\delta_C^{(n)}(s)}{\delta_C^{(n-1)}(s)} = \frac{a(s)}{\delta_C^{(n-1)}(s)} = \frac{\psi(s)\,\delta_C^{(n-1)}(s)}{\delta_C^{(n-1)}(s)}$$
$$= \psi(s).$$

We showed that $A = T J_A T^{-1}$, where

$J_A = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{bmatrix}$ is a block-diagonal matrix

whose diagonal elements $J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_k \end{bmatrix}$ are $l_k \times l_k$ Jordan blocks corresponding to (possibly repeated) ch. values of A and $n = l_1 + \cdots + l_p$.

$\Big\}$ l_k is the exponent of the factor $(s-\lambda_k)$ in one of the invariant polynomials

The structure of J_A can be used to determine the transformation matrix T.

Let $\underline{t}_1, \ldots, \underline{t}_{l_k}$ denote the l_k columns of T that multiply the kth Jordan block in the expression TJ_A.

Then $AT = TJ_A \Rightarrow$

$$\left.\begin{aligned} At_1 &= \lambda_k t_1 &&\Rightarrow (A-\lambda_k I)t_1 = 0 \\ At_2 &= t_1 + \lambda_k t_2 &&\Rightarrow (A-\lambda_k I)t_2 = t_1 \\ &\ \ \vdots && \qquad\qquad \vdots \\ At_{l_k} &= t_{l_k-1} + \lambda_k t_{l_k} &&\Rightarrow (A-\lambda_k I)t_{l_k} = t_{l_k-1} \end{aligned}\right\} \begin{array}{l} t_1, \ldots, t_{l_k} \\ \text{form} \\ \text{a} \\ \text{"\underline{Jordan} } \\ \text{\underline{chain}"} \end{array}$$

Hence, if $T = [t_1 \cdots t_n]$ satisfies the above equations (including every Jordan block) and $\det(T) \neq 0$, then $A = TJ_AT^{-1}$.

We next show how to find a Jordan chain.

Let $\psi(s) = (s-\lambda_1)^{\chi_1} \cdots (s-\lambda_p)^{\chi_p}$ denote the minimal polynomial.

<u>Thm</u>: Let $G_j(s) = \frac{1}{(j-1)!} \frac{d^{j-1}}{ds^{j-1}} \Gamma(s)$.

Then $\exists\, m_k \in [1, \cdots, n]$ s.t. $\underline{t}_1^{(k)}, \ldots, \underline{t}_{\chi_k}^{(k)}$ form a <u>lin. ind</u> <u>Jordan chain</u>

where $[\underline{t}_j^{(k)}]_{l1} = [G_j(\lambda_k)]_{l m_k}$, $l = 1, \cdots, n$, $j = 1, \cdots, \chi_k$

pf:
$$\psi(s) = (s-\lambda_1)^{\chi_1} \cdots (s-\lambda_p)^{\chi_p}$$
$$\Rightarrow \psi^{(g)}(\lambda_k) = 0, \quad g = 0, 1, \ldots \chi_k - 1$$
$$\psi(s) I = (sI - A)\Gamma(s)$$
$$\Rightarrow \psi'(s) I = \Gamma(s) + (sI-A)\Gamma'(s)$$
$$\psi''(s) I = 2\Gamma'(s) + (sI-A)\Gamma''(s)$$
$$\vdots$$
$$\psi^{(g)}(s) I = g\Gamma^{(g-1)}(s) + (sI-A)\Gamma^{(g)}(s)$$

$$\left.\begin{aligned} \Rightarrow (A-\lambda_k I)\Gamma(\lambda_k) &= 0 \\ (A-\lambda_k I)\Gamma'(\lambda_k) &= \Gamma(\lambda_k) \\ &\vdots \\ (A-\lambda_k I)\frac{\Gamma^{(\chi_k-1)}(\lambda_k)}{(\chi_k-1)!} &= \frac{\Gamma^{(\chi_k-2)}(\lambda_k)}{(\chi_k-2)!} \end{aligned}\right\} \begin{aligned} &(A-\lambda_k I) G_1(\lambda_k) = 0. \\ &(A-\lambda_k I) G_2(\lambda_k) = G_1(\lambda_k) \\ &\vdots \\ &(A-\lambda_k I) G_{\chi_k}(\lambda_k) = G_{\chi_k-1}(\lambda_k) \end{aligned}$$

$\Gamma(\lambda_k) \neq 0 \Rightarrow \exists\, m_k \in [1, \cdots, n]$

$$\text{s.t.} \left.\begin{aligned} (A-\lambda_k)\underline{t}_1^{(k)} &= 0 \\ (A-\lambda_k)\underline{t}_2^{(k)} &= \underline{t}_1^{(k)} \\ &\vdots \\ (A-\lambda_k)\underline{t}_{\chi_k}^{(k)} &= \underline{t}_{\chi_k-1} \end{aligned}\right\} t_1^{(k)}, \ldots, t_{\chi_k}^{(k)} \text{ form a Jordan chain}$$

$\underline{t}_1^{(k)}, \ldots, \underline{t}_{\chi_k}^{(k)}$ are lin. ind. for suppose $\sum_{j=1}^{\chi_k} \alpha_j t_j^{(k)} = 0$.

then $(A-\lambda_k)^{\chi_k-1} \sum_{j=1}^{\chi_k} \alpha_j t_j^{(k)} = \alpha_{\chi_k} \underline{t}_1^{(k)} = 0 \Rightarrow \alpha_{\chi_k} = 0$.

$(A-\lambda_k)^{\chi_k-2} \sum_{j=1}^{\chi_k} \alpha_j t_j^{(k)} = \alpha_{\chi_k-1}\underline{t}_1^{(k)} + \alpha_{\chi_k}\underline{t}_2^{(k)} = 0 \Rightarrow \alpha_{\chi_k-1} = 0$. etc.

Thm: Let $t_g^{(k)}$, $g=1,\dots,l_k$, $k=1,\dots,K$, denote K Jordan chains, where $(A-\lambda_k I)\,t_1^{(k)}=0$ and $(A-\lambda_k I)\,t_g^{(k)} = t_{g-1}^{(k)}$, $g=2,\dots,l_k$, $k=1,\dots,K$, and $\{t_1^{(k)}\}$ are lin. ind. $k=1,\dots,K$.
Then $\{t_g^{(k)} : g=1,\dots,l_k,\ k=1,\dots,K\}$ is a lin. ind. set.

pf: Suppose $\displaystyle\sum_{k=1}^{K}\sum_{g=1}^{l_k} \alpha_g^{(k)}\, t_g^{(k)} = 0.$

Let $\mathcal{U}(\lambda_j)$ denote the index set of all the Jordan blocks corresponding to the ch. value λ_j.

Let $\mathcal{U}^{(1)}(\lambda_j)$ denote the index set of all the Jordan blocks for λ_j of the largest order, $\mathcal{U}^{(2)}(\lambda_j)$ second largest order, etc.

$$\prod_{m\notin\mathcal{U}(\lambda_j)} (A-\lambda_m I)^{l_m}\cdot\left(\sum_{k=1}^{K}\sum_{g=1}^{l_k}\alpha_g^{(k)}\,t_g^{(k)}\right)=0 \Rightarrow$$

$$\sum_{k\in\mathcal{U}(\lambda_j)}\sum_{g=1}^{l_k}\alpha_g^{(k)}\prod_{m\notin\mathcal{U}(\lambda_j)}(A-\lambda_m I)^{l_m}\,t_g^{(k)} = 0 \Rightarrow$$

$$\sum_{k\in\mathcal{U}^{(1)}(\lambda_j)}\alpha_{l_k}^{(k)}\prod_{m\notin\mathcal{U}(\lambda_j)}(A-\lambda_m I)^{l_m}\,t_1^{(k)} = \left(\prod_{m\notin\mathcal{U}(\lambda_j)}(\lambda_j-\lambda_m)^{l_m}\right)\sum_{k\in\mathcal{U}^{(1)}(\lambda_j)}\alpha_{l_k}^{(k)}\,t_1^{(k)} = 0.$$

$$\Rightarrow \sum_{k\in\mathcal{U}^{(1)}(\lambda_j)}\alpha_{l_k}^{(k)}\,t_1^{(k)} = 0 \Rightarrow \alpha_{l_k}^{(k)}=0 \quad \forall k\in\mathcal{U}^{(1)}(\lambda_j) \text{ (since } t_1^{(k)} \text{ are l.i.)}$$

Continuing this argument: $\alpha_g^{(k)}=0$, $g=1,\dots,l_k-1$ $\forall k\in\mathcal{U}^{(1)}(\lambda_j)$

$\alpha_g^{(k)}=0$, $g=1,\dots,l_k$ $\forall k\in\mathcal{U}^{(2)}(\lambda_j)$, etc.

i.e. $\alpha_g^{(k)}=0$, $g=1,\dots,l_k$, $\forall k\in\mathcal{U}(\lambda_j)$

and: $\alpha_g^{(k)}=0$, $g=1,\dots,l_k$, $k=1,\dots,K$.

Example

#4 10/10

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix} \qquad sI - A = \begin{bmatrix} s-1 & & & \\ -1 & s-1 & & \\ & 0 & s-1 & \\ & & -2 & s-1 \end{bmatrix}$$

$$sI-A \overset{E}{=} \begin{bmatrix} -1 & s-1 & 0 & 0 \\ s-1 & 0 & 0 & 0 \\ 0 & 0 & s-1 & 0 \\ 0 & 0 & -2 & s-1 \end{bmatrix} \overset{E}{=} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & (s-1)^2 & 0 & 0 \\ 0 & 0 & s-1 & 0 \\ 0 & 0 & & s-1 \end{bmatrix}$$

$$\overset{E}{=} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & (s-1)^2 & 0 \\ 0 & s-1 & 0 & 0 \\ 0 & -2 & 0 & s-1 \end{bmatrix} \overset{E}{=} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -2 & 0 & s-1 \\ 0 & s-1 & 0 & 0 \\ 0 & 0 & (s-1)^2 & 0 \end{bmatrix}$$

$$\overset{E}{=} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2}(s-1)^2 \\ 0 & 0 & (s-1)^2 & 0 \end{bmatrix} \overset{E}{=} \begin{bmatrix} 1 & & & \\ & 1 & & 0 \\ & & (s-1)^2 & \\ 0 & & & (s-1)^2 \end{bmatrix} \Rightarrow J_A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\psi(s) = (s-1)^2 = s^2 - 2s + 1$$

$$\frac{\psi(\lambda) - \psi(\mu)}{\lambda - \mu} = \frac{\lambda^2 - 2\lambda + 1 - (\mu^2 - 2\mu + 1)}{\lambda - \mu} = (\lambda + \mu) - 2$$

$$\Rightarrow \Gamma(s) = (sI + A - 2I) = \begin{bmatrix} s-1 & 0 & 0 & 0 \\ 1 & s-1 & 0 & 0 \\ 0 & 0 & s-1 & 0 \\ 0 & 0 & 2 & s-1 \end{bmatrix}$$

$$\Gamma'(s) = I$$

$$\Gamma(1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix} \qquad \Gamma'(1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \Rightarrow AT = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix}$$

$$TJ_A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix}$$

$$|T| = -1 \cdot \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 2 & 0 \end{vmatrix} = -1 \begin{vmatrix} 0 & 1 \\ 2 & 0 \end{vmatrix} = 2$$