Lecture 3

Discussion: Last time we showed that if A has n-linearly independent ch. vectors then A is similar to a diagonal matrix, i.e. $\exists$ non-singular T s.t. $T^{-1}AT = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$.

In general, it's not necessarily possible to "diagonalize" A. Though it is always possible to "block diagonalize" by means of a similarity transformation.

Defn: A Jordan block J_λ has the form:

$$J_\lambda = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

This time we show that A is always similar to a block diagonal matrix $J_A = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_\eta \end{bmatrix}$, where $J_p = \begin{bmatrix} \lambda_p & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_p \end{bmatrix}$

and λ_p denotes a (possibly repeated) ch. value of A, $p = 1, \ldots, \eta$.

Defn: An order-p minor of the matrix A, corresponding to rows $i_1, \ldots, i_p$ and columns $j_1, \ldots, j_p$, is defined as

$$A\begin{pmatrix} i_1, \ldots, i_p \\ j_1, \ldots, j_p \end{pmatrix} \triangleq \det\left(\begin{bmatrix} [A]_{i_1 j_1} & \cdots & [A]_{i_1 j_m} \\ \vdots & & \\ [A]_{i_p j_1} & & [A]_{i_p j_m} \end{bmatrix}\right).$$

(HW)

Thm: The rank of A coincides with the largest r for which there exists a nonzero minor of order-r.

pf: The rank of A is defined as the number of lin. ind. cols. of A.

$$\Rightarrow \exists\ A\begin{pmatrix} i_1, \ldots, i_r \\ j_1, \ldots, j_r \end{pmatrix} \neq 0 \quad \text{and} \quad A\begin{pmatrix} i_1 \cdots i_{r+1} \\ j_1 \cdots j_{r+1} \end{pmatrix} = 0 \quad \forall \left.\begin{matrix} i_1, \ldots, i_{r+1} \in [1, \ldots, m] \\ j_1, \ldots, j_{r+1} \in [1, \ldots, n] \end{matrix}\right\} \text{(cond.)}$$

Conversely, if (cond.) is true then $\exists$ E, F (where E is row interchange matrix and F is column combining/interch matrix) s.t.

$$EAF = \begin{bmatrix} D & 0 \\ X & 0 \end{bmatrix},$$

$D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_r \end{bmatrix}$, and X is an $m-r \times r$ matrix. (pf: exercise)

It follows that $\text{rank}(A) = r$.

Thm: Binet-Cauchy formula

Let C denote an $m\times m$ matrix, A is $m\times n$, and B is $n\times m$.

Then, $$\det(C) = \sum_{1\le k_1 \le \cdots \le k_m \le n} A\begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_m \\ 1 & 2 & \cdots & m \end{pmatrix}$$

pf: $$\begin{vmatrix} c_{11} & \cdots & c_{1m} \\ c_{m1} & \cdots & c_{mm} \end{vmatrix} = \begin{vmatrix} \sum_{\alpha_1=1}^{n} a_{1\alpha_1} b_{\alpha_1 1} & \cdots & \sum_{\alpha_m=1}^{n} a_{1\alpha_m} b_{\alpha_m m} \\ \vdots & & \\ \sum_{\alpha_1=1}^{n} a_{m\alpha_1} b_{\alpha_1 1} & \cdots & \sum_{\alpha_m=1}^{n} a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix}$$

$$= \sum_{\alpha_1, \alpha_2 \cdots \alpha_m = 1}^{n} \begin{vmatrix} a_{1\alpha_1} b_{\alpha_1 1} & \cdots & a_{1\alpha_m} b_{\alpha_m m} \\ \vdots & & \\ a_{m\alpha_1} b_{\alpha_1 1} & & a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix}$$

$$= \sum_{\alpha_1, \cdots \alpha_m = 1}^{n} A\begin{pmatrix} 1 & 2 & \cdots & m \\ \alpha_1 & \alpha_2 & & \alpha_m \end{pmatrix} b_{\alpha_1 1} \cdots b_{\alpha_m m} = \begin{cases} 0, & m > n \\ (*), & m \le n \end{cases}$$

$$\sum_{p\in\Pi(k_1\cdots k_m)} A\begin{pmatrix} 1 & 2 & \cdots & m \\ p(1) & p(2) & \cdots & p(m) \end{pmatrix} b_{p(1)1} \cdots b_{p(m)m} = \sum_{p\in\Pi(k_1\cdots k_m)} (-1)^{\mu(p)} A\begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} b_{p(1)1} \cdots b_{p(m)m}$$

$$= A\begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} \sum_{p\in\Pi(k_1\cdots k_m)} (-1)^{\mu(p)} b_{p(1)1} \cdots b_{p(m)m} = A\begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & \cdots & k_m \\ 1 & \cdots & m \end{pmatrix}$$

$$\Rightarrow (*) = \sum_{1\le k_1 \cdots \le k_m \le n} A\begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_m \\ 1 & 2 & & m \end{pmatrix}.$$

corollary $$C\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} = \sum_{1\le k_1 \le \cdots \le k_p \le n} A\begin{pmatrix} i_1 & \cdots & i_p \\ k_1 & & k_p \end{pmatrix} B\begin{pmatrix} k_1 & \cdots & k_p \\ j_1 & \cdots & j_p \end{pmatrix}, \quad p \le n$$

pf: HW

Defn A polynomial matrix is a matrix $A(s)$ whose elements are polynomials in s.

$$[A(s)]_{ij} = a_{ij}^{(0)} s^{\ell} + a_{ij}^{(1)} s^{\ell-1} + \cdots + a_{ij}^{(\ell)}, \qquad i = 1,2,\ldots,m \quad j = 1,2,\ldots,n,$$

where ℓ is the largest degree of the elements of $A(s)$.

Thus, $A(s) = A_0 s^{\ell} + A_1 s^{\ell-1} + \cdots + A_\ell$, where $A_0, \ldots, A_\ell$ are constant (degree-zero) matrices.

Defn The elementary operations on a polynomial matrix, A, are:

1. Multiplication of a row by $c \neq 0$: $A_1 = E^{(1)}A$, where

$$E^{(1)} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{bmatrix}, \quad \det(E^{(1)}) = c.$$

2. Addition of $p(s)$ times row k to row j: $A_2 = E^{(2)}A$, where

$$E^{(2)} = \begin{array}{c} \begin{matrix} k\downarrow & j\downarrow \end{matrix} \\ \begin{bmatrix} 1 & & & \\ \ddots & & & \\ p(s) & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \end{array}, \quad \det(E^{(2)}) = 1.$$

3. Interchange row k and row j: $A_3 = E^{(3)}A$

$$E^{(3)} = \begin{array}{c} \\ j \rightarrow \\ k \rightarrow \end{array} \begin{bmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{bmatrix}, \quad \det(E^{(3)}) = -1$$

4. multiplication of a column by $c \neq 0$: $A_4 = AF^{(1)}$

$$F^{(1)} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{bmatrix}, \quad \det(F^{(1)}) = c.$$

5. Addition of $p(s)$ times column k to column j: $A_5 = AF^{(2)}$

$$F^{(2)} = \begin{array}{c} \\ j \rightarrow \\ k \rightarrow \end{array} \begin{bmatrix} 1 & & \\ & 1 & \\ & p(s) & 1 \end{bmatrix}, \quad \det(F^{(2)}) = 1$$

6. Interchange column k and column j: $A_6 = AF^{(3)}$

$$F^{(3)} = \begin{bmatrix} 1 & 0 & 1 & \\ & 1 & 0 & 1 \end{bmatrix}, \quad \det(F^{(3)}) = -1.$$

<u>Defn</u>: The greatest common monic divisor (g.c.m.d.) of the order-p minors of a polynomial matrix $A(s)$ is denoted $\delta_A^{(p)}(s)$, where

$$\delta_A^{(p)}(s) = \text{g.c.m.d} \left\{ A\begin{pmatrix} i_1 \cdots i_p \\ j_1 \cdots j_p \end{pmatrix} : \begin{array}{l} 1 \le i_1 \le \cdots \le i_p \le m \\ \text{and } 1 \le j_1 \le \cdots \le j_p \le n \end{array} \right\}$$

<u>Thm</u>: $\delta_A^{(p-1)}(s)$ divides $\delta_A^{(p)}(s)$ without remainder, $p = 2, \dots, \min(m,n)$.

<u>pf</u>:
$$A\begin{pmatrix} i_1 \cdots & i_p \\ j_1 \cdots & j_p \end{pmatrix} = \sum_{k=1}^{p} [A]_{i_1 j_k} (-1)^{1+k} A\begin{pmatrix} i_2 & \cdots & & i_p \\ j_1 & \cdots j_{k-1} & j_{k+1} \cdots & j_p \end{pmatrix}$$

$\Rightarrow$ $\delta_A^{(p-1)}(s)$ divides any order-p minor of A

$\Rightarrow$ $\delta_A^{(p-1)}(s)$ divides $\delta_A^{(p)}(s)$.

<u>Defn</u>: The <u>invariant</u> <u>polynomials</u> of a polynomial matrix $A(s)$ are denoted $h_p(s)$, $p = 1, \dots, \min(m,n)$, where

$$h_1(s) = \delta_A^{(1)}(s) \quad \text{and} \quad h_p(s) = \frac{\delta_A^{(p)}(s)}{\delta_A^{(p-1)}(s)}, \quad p = 2, \dots, \min(m,n).$$

Defn: The polynomial matrices $A(s)$ and $B(s)$ are <u>equivalent</u> iff there exists $P(s)$ and $Q(s)$

s.t. $B(s) = P(s)\,A(s)\,Q(s)$, where $\det(P(s)) = c_1 \neq 0$ and $\det(Q(s)) = c_2 \neq 0$.

Thm: If $A(s)$ and $B(s)$ are equivalent, then they have the same invariant polynomials.

pf: Let $B(s) = P(s)\,A(s)\,Q(s)$, where $\det(P(s)) = c_1 \neq 0$ and $\det(Q(s)) = c_2 \neq 0$.

Define $C(s) = A(s)\,Q(s)$. Then $B(s) = P(s)\,C(s)$ and by Cauchy-Binet:

$$C\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} = \sum_{1 \le k_1 \le \cdots \le k_p \le n} A\begin{pmatrix} i_1 & \cdots & i_p \\ k_1 & \cdots & k_p \end{pmatrix} Q\begin{pmatrix} k_1 & \cdots & k_p \\ j_1 & \cdots & j_p \end{pmatrix}$$

$$B\begin{pmatrix} l_1 & \cdots & l_p \\ j_1 & \cdots & j_p \end{pmatrix} = \sum_{1 \le i_1 \le \cdots \le i_p \le m} P\begin{pmatrix} l_1 & \cdots & l_p \\ i_1 & \cdots & i_p \end{pmatrix} C\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix}$$

$$= \sum_{\substack{1 \le i_1 \le \cdots \le i_p \le m \\ 1 \le k_1 \le \cdots \le k_p \le n}} P\begin{pmatrix} l_1 & \cdots & l_p \\ i_1 & \cdots & i_p \end{pmatrix} A\begin{pmatrix} i_1 & \cdots & i_p \\ k_1 & \cdots & k_p \end{pmatrix} Q\begin{pmatrix} k_1 & \cdots & k_p \\ j_1 & \cdots & j_p \end{pmatrix}$$

$\Rightarrow$ $\delta_A^{(p)}(s)$ divides any order-p minor of B

$\Rightarrow$ $\delta_A^{(p)}(s)$ divides $\delta_B^{(p)}(s)$

But $A(s) = P^{-1}(s)\,B(s)\,Q^{-1}(s) \Rightarrow \delta_B^{(p)}(s)$ divides $\delta_A^{(p)}(s)$

$\Rightarrow$ $\delta_B^{(p)}(s) = \alpha\,\delta_A^{(p)}(s)$

$\Rightarrow$ $\delta_B^{(p)}(s) = \delta_A^{(p)}(s)$ (since δ_A and δ_B are monic)

$\Rightarrow$ A and B have same inv. polynomials. $\left(\text{since } h_p(s) = \dfrac{\delta^{(p)}(s)}{\delta^{(p-1)}(s)}\right)$

Thm: Equivalent diagonal form of a polynomial matrix.
Let $A(s)$ denote an $m\times n$ poly. matrix. $A(s)$ is equivalent to the diagonal matrix
$$D_A(s) = \begin{bmatrix} h_1(s) & & & \\ & \ddots & & \\ & & h_r(s) & \\ & & & 0 \end{bmatrix},$$
where $h_1(s),\dots,h_r(s)$ denote the invariant polynomials of $A(s)$ and r is the rank of $A(s)$.

pf: Let $A_1(s) = E_1(s)A(s)F_1(s)$, where $E_1(s)$ and $F_1(s)$ denote elem. op. matrices for which $\deg([A_1(s)]_{11})$ is minimal.

We have:
(by long division)
$$[A_1(s)]_{1j} = m_{1j}(s)[A_1(s)]_{11} + r_{1j}(s),\quad \deg(r_{1j}(s)) < \deg([A_1(s)]_{11}),\quad j=2,\dots,n$$
$$[A_1(s)]_{k1} = m_{k1}(s)[A_1(s)]_{11} + r_{k1}(s),\quad \deg(r_{k1}(s)) < \deg([A_1(s)]_{11}),\quad k=2,\dots,m$$
If any $r_{1j}(s)\neq 0$, $j=2,\dots,n$ or $r_{k1}(s)\neq 0$, $k=2,\dots,m$, then choose one with least degree, e.g. $r_{k1}(s)$. Subtract $m_{k1}(s)$ times row 1 from row k and move the $(k,1)$th element to the $(1,1)$th position. Denote the resulting matrix $A_2(s) = E_2(s)A_1(s)F_2(s)$. Then $\deg([A_2(s)]_{11}) < \deg([A_1(s)]_{11})$. Repeat above until there are no non-zero remainders (upon dividing by the $(1,1)$th element) in first row and column.
Thus, by elem. op. matrices, have transformed A to:

$\begin{bmatrix} x_1 & 0 \cdots 0 \\ 0 & \\ \vdots & M_1 \\ 0 & \end{bmatrix}$. If $M_1\neq 0$, check for non-zero remainders (upon dividing by the $(1,1)$th element). Move remainder with least degree into $(1,1)$ position and repeat above steps until A is transformed to:

$\begin{bmatrix} x_2 & 0 \cdots 0 \\ 0 & \\ \vdots & M_2 \\ 0 & \end{bmatrix}$, where $\deg(x_2(s)) < \deg(x_1(s))$. Since $\deg(x_2(s))$ is finite, we must arrive at the matrix:

$\begin{bmatrix} x_j & 0 \cdots 0 \\ 0 & \\ \vdots & M_j \\ 0 & \end{bmatrix}$, where $x_j(s)$ divides every element of M_j.

Repeat above on matrix M_j, etc. until A is transformed to $\begin{bmatrix} \tilde h_1(s) & & & 0 \\ & \ddots & & \vdots \\ & & \tilde h_r(s) & 0 \\ 0 & \cdots & 0 & M \end{bmatrix}$.

$M=0$ since elem. op. do not change rank.
Finally normalize the leading coefficients of $\tilde h_i(s)$, $i=1,\dots,r$, to obtain

$D_A(s) = \begin{bmatrix} h_1(s) & & & \\ & \ddots & & \\ & & h_r(s) & \\ & & & 0 \end{bmatrix} \stackrel{E}{=} A(s)$. It follows that $h_1(s),\dots,h_r(s)$ are the inv. polynomials of A (since equiv matrices have the same inv. polynomials).
(pf: exercise)

<u>Thm</u>: $A(s)$ is equivalent to $B(s)$ <u>iff</u> $A(s)$ and $B(s)$ have the same invariant polynomials.

<u>pf</u>: $\Rightarrow$] If $A(s)$ and $B(s)$ have the same invariant polynomials

then $D_A(s) = D_B(s) \Rightarrow \exists$ non-singular matrices $E_1(s), E_2(s), F_1(s), F_2(s)$

with constant, non-zero determinants s.t.

$$E_1(s)\,A(s)\,F_1(s) = E_2(s)\,B(s)\,F_2(s)$$

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

i.e. $A(s)$ is equivalent to $B(s)$.

$\Leftarrow$] proved already.

<u>Thm</u>: If $\det(A(s)) = c \neq 0$, then $A(s) = \prod_{i=1}^{l} E_i(s)$.

<u>pf</u>: $D_A(s) = \begin{bmatrix} h_1(s) & & \\ & \ddots & \\ & & h_n(s) \end{bmatrix} = E(s)\,A(s)\,F(s)$

$\det(D_A) = \det(E(s))\det(A(s))\det(F(s)) = c_1$ (since $\det(E(s)) = c_2 \neq 0$, $\det(F(s)) = c_3 \neq 0$)

$\Rightarrow \prod h_i(s) = \prod \beta_i$, β_i is non-zero constant, $i = 1, \dots, n$

$\Rightarrow I = \begin{bmatrix} \beta_1^{-1} & & \\ & \ddots & \\ & & \beta_n^{-1} \end{bmatrix} E(s)\,A(s)\,F(s)$

$\Rightarrow A(s) = \prod E_i(s)$. where $E_i(s)$ are elementary op. matrices.

<u>Thm</u>: If $sI-A$ and $sI-B$ are equivalent, then there exist non-singular constant matrices P and Q s.t. $sI-B = P(sI-A)Q$.

<u>pf</u>: By assumption $sI-B = P(s)(sI-A)Q(s)$,
where $P(s)$ and $Q(s)$ have constant non-zero determinants.

Let $R(s) = P^{-1}(s)$. Then $R(s)(sI-B) = (sI-A)Q(s)$.

$R(s)$ and $Q(s)$ are polynomial matrices of finite degree (finite products of elem. op. matrices).

$$\Rightarrow \left.\begin{aligned} R(s) &= (sI-A)M_1(s) + R, \quad R \text{ constant} \\ Q(s) &= M_2(s)(sI-B) + Q, \quad Q \text{ constant} \end{aligned}\right\} \text{pf: exercise}$$

$$\Rightarrow [(sI-A)M_1(s) + R](sI-B) = (sI-A)[M_2(s)(sI-B) + Q]$$

$$\Rightarrow \underbrace{(sI-A)(M_1(s)-M_2(s))(sI-B)}_{\text{degree} \geq 2} = \underbrace{(sI-A)Q - R(sI-B)}_{\text{degree} \leq 1}.$$

$$\Rightarrow M_1(s) = M_2(s) \Rightarrow R(sI-B) = (sI-A)Q$$

Need to show R is nonsingular:

$$P(s) = (sI-B)M_3(s) + P, \quad P \text{ constant}$$

$$\Rightarrow \underset{\text{deg.}=0}{I} = R(s)P(s) = R(s)(sI-B)M_3(s) + R(s)P$$

$$= (sI-A)Q(s)M_3(s) + (sI-A)M_1(s)P + RP$$

$$= \underbrace{(sI-A)(Q(s)M_3(s) + M_1(s)P)}_{\text{deg.} \geq 1} + RP$$

$$\Rightarrow I = RP$$

$$\Rightarrow R^{-1} = P$$

$$\Rightarrow sI-B = P(sI-A)Q$$

□