Lecture 12
Jordan canonical form

- Jordan canonical form
- generalized modes
- Cayley-Hamilton theorem

Jordan canonical form

what if A cannot be diagonalized?

any matrix $A \in \mathbf{R}^{n\times n}$ can be put in *Jordan canonical form* by a similarity transformation, $i.e.$

$$T^{-1}AT = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_q \end{bmatrix}$$

where

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathbf{C}^{n_i\times n_i}$$

is called a *Jordan block* of size n_i with eigenvalue λ_i (so $n = \sum_{i=1}^q n_i$)

- J is upper bidiagonal
- J diagonal is the special case of n Jordan blocks of size $n_i = 1$
- Jordan form is unique (up to permutations of the blocks)
- can have multiple blocks with same eigenvalue

note: JCF is a *conceptual tool*, never used in numerical computations!

$\mathcal{X}(s) = \det(sI - A) = (s-\lambda_1)^{n_1} \cdots (s-\lambda_q)^{n_q}$

hence distinct eigenvalues $\Rightarrow n_i = 1 \Rightarrow A$ diagonalizable

$\dim \mathcal{N}(\lambda I - A)$ is the number of Jordan blocks with eigenvalue λ

more generally,

$$\dim \mathcal{N}(\lambda I - A)^k = \sum_{\lambda_i = \lambda} \min\{k, n_i\}$$

so from $\dim \mathcal{N}(\lambda I - A)^k$ for $k = 1, 2, \ldots$ we can determine the sizes of the Jordan blocks associated with λ

- factor out T and T^{-1}, $\lambda I - A = T(\lambda I - J)T^{-1}$

- for, say, a block of size 3:

$$\lambda_i I - J_i = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad (\lambda_i I - J_i)^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (\lambda_i I - J_i)^3 = 0$$

- for other blocks (say, size 3, for $k \geq 2$)

$$(\lambda_i I - J_j)^k = \begin{bmatrix} (\lambda_i - \lambda_j)^k & -k(\lambda_i - \lambda_j)^{k-1} & (k(k-1)/2)(\lambda_i - \lambda_j)^{k-2} \\ 0 & (\lambda_j - \lambda_i)^k & -k(\lambda_j - \lambda_i)^{k-1} \\ 0 & 0 & (\lambda_j - \lambda_i)^k \end{bmatrix}$$

Generalized eigenvectors

suppose $T^{-1}AT = J = \mathbf{diag}(J_1, \ldots, J_q)$

express T as

$$T = [T_1 \ T_2 \ \cdots \ T_q]$$

where $T_i \in \mathbf{C}^{n \times n_i}$ are the columns of T associated with ith Jordan block J_i

we have $AT_i = T_i J_i$

let $T_i = [v_{i1} \ v_{i2} \ \cdots \ v_{in_i}]$

then we have:

$$Av_{i1} = \lambda_i v_{i1},$$

$i.e.$, the first column of each T_i is an eigenvector associated with e.v. λ_i

for $j = 2, \ldots, n_i$,

$$Av_{ij} = v_{i\ j-1} + \lambda_i v_{ij}$$

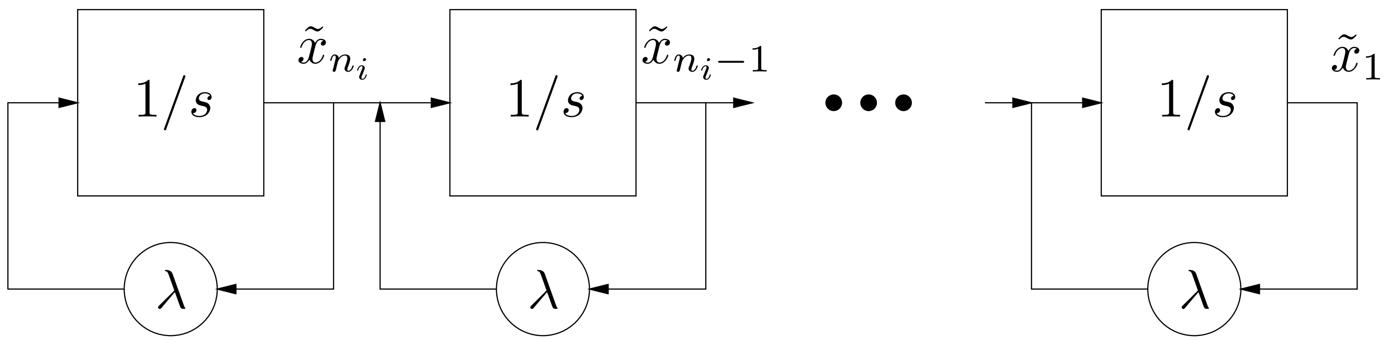
the vectors $v_{i1}, \ldots v_{in_i}$ are sometimes called *generalized eigenvectors*

Jordan form LDS

consider LDS $\dot{x} = Ax$

by change of coordinates $x = T\tilde{x}$, can put into form $\dot{\tilde{x}} = J\tilde{x}$

system is decomposed into independent 'Jordan block systems' $\dot{\tilde{x}}_i = J_i\tilde{x}_i$



Jordan blocks are sometimes called Jordan chains

(block diagram shows why)

Resolvent, exponential of Jordan block

resolvent of $k \times k$ Jordan block with eigenvalue λ:

$$
\begin{aligned}
(sI - J_\lambda)^{-1} &= \begin{bmatrix} s-\lambda & -1 & & \\ & s-\lambda & \ddots & \\ & & \ddots & -1 \\ & & & s-\lambda \end{bmatrix}^{-1} \\
&= \begin{bmatrix} (s-\lambda)^{-1} & (s-\lambda)^{-2} & \cdots & (s-\lambda)^{-k} \\ & (s-\lambda)^{-1} & \cdots & (s-\lambda)^{-k+1} \\ & & \ddots & \vdots \\ & & & (s-\lambda)^{-1} \end{bmatrix} \\
&= (s-\lambda)^{-1} I + (s-\lambda)^{-2} F_1 + \cdots + (s-\lambda)^{-k} F_{k-1}
\end{aligned}
$$

where F_i is the matrix with ones on the ith upper diagonal by inverse Laplace transform, exponential is:

$$
\begin{aligned}
e^{tJ_\lambda} &= e^{t\lambda}\left(I + tF_1 + \cdots + (t^{k-1}/(k-1)!)F_{k-1}\right) \\
&= e^{t\lambda}\begin{bmatrix} 1 & t & \cdots & t^{k-1}/(k-1)! \\ & 1 & \cdots & t^{k-2}/(k-2)! \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}
\end{aligned}
$$

Jordan blocks yield:

- repeated poles in resolvent
- terms of form $t^p e^{t\lambda}$ in e^{tA}

Generalized modes

consider $\dot{x} = Ax$, with

$$x(0) = a_1 v_{i1} + \cdots + a_{n_i} v_{in_i} = T_i a$$

then $x(t) = Te^{Jt}\tilde{x}(0) = T_i e^{J_i t} a$

- trajectory stays in span of generalized eigenvectors
- coefficients have form $p(t)e^{\lambda t}$, where p is polynomial
- such solutions are called *generalized modes* of the system

with general $x(0)$ we can write

$$x(t) = e^{tA}x(0) = Te^{tJ}T^{-1}x(0) = \sum_{i=1}^{q} T_i e^{tJ_i}(S_i^T x(0))$$

where

$$T^{-1} = \begin{bmatrix} S_1^T \\ \vdots \\ S_q^T \end{bmatrix}$$

hence: all solutions of $\dot{x} = Ax$ are linear combinations of (generalized) modes

Cayley-Hamilton theorem

if $p(s) = a_0 + a_1 s + \cdots + a_k s^k$ is a polynomial and $A \in \mathbf{R}^{n\times n}$, we define

$$p(A) = a_0 I + a_1 A + \cdots + a_k A^k$$

Cayley-Hamilton theorem: for any $A \in \mathbf{R}^{n\times n}$ we have $\mathcal{X}(A) = 0$, where $\mathcal{X}(s) = \det(sI - A)$

example: with $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ we have $\mathcal{X}(s) = s^2 - 5s - 2$, so

$$\begin{aligned} \mathcal{X}(A) &= A^2 - 5A - 2I \\ &= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - 5 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 2I \\ &= 0 \end{aligned}$$

corollary: for every $p \in \mathbf{Z}_+$, we have

$$A^p \in \operatorname{span}\left\{ I,\ A,\ A^2, \ldots,\ A^{n-1} \right\}$$

(and if A is invertible, also for $p \in \mathbf{Z}$)

i.e., every power of A can be expressed as linear combination of $I,\ A, \ldots, A^{n-1}$

proof: divide $\mathcal{X}(s)$ into s^p to get $s^p = q(s)\mathcal{X}(s) + r(s)$

$r = \alpha_0 + \alpha_1 s + \cdots + \alpha_{n-1}s^{n-1}$ is remainder polynomial

then

$$A^p = q(A)\mathcal{X}(A) + r(A) = r(A) = \alpha_0 I + \alpha_1 A + \cdots + \alpha_{n-1}A^{n-1}$$

for $p = -1$: rewrite C-H theorem

$$\mathcal{X}(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_0 I = 0$$

as

$$I = A\left(-(a_1/a_0)I - (a_2/a_0)A - \cdots - (1/a_0)A^{n-1}\right)$$

(A is invertible $\Leftrightarrow a_0 \neq 0$) so

$$A^{-1} = -(a_1/a_0)I - (a_2/a_0)A - \cdots - (1/a_0)A^{n-1}$$

i.e., inverse is linear combination of A^k, $k = 0, \ldots, n-1$

Proof of C-H theorem

first assume A is diagonalizable: $T^{-1}AT = \Lambda$

$$\mathcal{X}(s) = (s - \lambda_1) \cdots (s - \lambda_n)$$

since
$$\mathcal{X}(A) = \mathcal{X}(T\Lambda T^{-1}) = T\mathcal{X}(\Lambda)T^{-1}$$
it suffices to show $\mathcal{X}(\Lambda) = 0$

$$\begin{aligned}
\mathcal{X}(\Lambda) &= (\Lambda - \lambda_1 I) \cdots (\Lambda - \lambda_n I) \\
&= \mathbf{diag}(0, \lambda_2 - \lambda_1, \ldots, \lambda_n - \lambda_1) \cdots \mathbf{diag}(\lambda_1 - \lambda_n, \ldots, \lambda_{n-1} - \lambda_n, 0) \\
&= 0
\end{aligned}$$

now let's do general case: $T^{-1}AT = J$

$$\mathcal{X}(s) = (s-\lambda_1)^{n_1} \cdots (s-\lambda_q)^{n_q}$$

suffices to show $\mathcal{X}(J_i) = 0$

$$\mathcal{X}(J_i) = (J_i - \lambda_1 I)^{n_1} \cdots \underbrace{\begin{bmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ & & & \ddots \end{bmatrix}^{n_i}}_{(J_i - \lambda_i I)^{n_i}} \cdots (J_i - \lambda_q I)^{n_q} = 0$$