

Lecture 5

Lecturer: Assoc. Prof. M. Mert Ankarali

5.1 Functions of a Matrix

In linear systems theory course, we are interested in matrix polynomials, specifically

- Matrix Exponential in CT Systems: e^{At}
- Matrix Power in DT Systems: A^k

which arise on the solution of state-space equations in their respective domains. Obviously A^k in DT systems is “easier” to analyze and understand compared to matrix exponential. Let’s first review the matrix exponential, e^{At} . Let $t \in \mathbb{R}$ and $A \in \mathbb{R}^{n \times n}$, then e^{At} defined as

$$e^{At} := I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \dots$$

$$e^{At} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

which converges for all $t \in \mathbb{R}$ and $A \in \mathbb{R}^{n \times n}$.

Now let’s review some properties

- **Claim:** $\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A$

Proof:

$$\begin{aligned} \frac{d}{dt}e^{At} &= \frac{d}{dt} \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \right) = \sum_{k=0}^{\infty} k \frac{t^{k-1}}{k!} A^k = \sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!} A^k \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} A^{n+1} = A \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n = \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} A^n \right) A \\ \frac{d}{dt}e^{At} &= Ae^{At} = e^{At}A \end{aligned}$$

- **Claim:** Let $A, B \in \mathbb{R}^{n \times n}$ and $AB = BA$, then

$$e^A e^B = e^B e^A = e^{(A+B)}$$

Proof:

$$e^A e^B = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k \right) \left(\sum_{j=0}^{\infty} \frac{1}{j!} B^j \right) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{A^k}{k!} \frac{B^j}{j!}$$

Let $n = k + j$ and $j = n - k$, then

$$e^A e^B = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \frac{A^k B^{n-k}}{k!(n-k)!} = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{A^k B^{n-k}}{n!} \frac{n!}{k!(n-k)!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n A^k B^{n-k} \binom{n}{k}$$

Note that if $AB \neq BA$ we have to stop at this point. However, since $AB = BA$, we can adopt binomial theorem

$$e^A e^B = \sum_{n=0}^{\infty} \frac{1}{n!} (A + B)^n = e^{A+B} = e^B e^A$$

- **Claim:** Let $t_1, t_2 \in \mathbb{R}$ then

$$e^{At_1} e^{At_2} = e^{At_2} e^{At_1} = e^{A(t_1+t_2)}$$

Proof: Let $A := At_1$ and $B := At_2$, obviously $(At_1)(At_2) = (At_2)(At_1)$ hence we can use the previous property, i.e.

$$e^{At_1} e^{At_2} = e^{(At_1)+(At_2)} = e^{A(t_1+t_2)} = e^{At_2} e^{At_1}$$

Now let $t_1 = t$ and $t_2 = -t$, then we have

$$e^{At} e^{-At} = e^{A(t-t)} = I \quad \rightarrow \quad (e^{At})^{-1} = e^{-At}$$

- **Claim:** Let $P \in \mathbb{R}^{n \times n}$ and $\det(P) \neq 0$, then

$$e^{(P^{-1}AP)t} = P^{-1} e^{At} P$$

Proof: Let's first show that $(P^{-1}AP)^k = P^{-1}A^kP$

$$\begin{aligned} (P^{-1}AP)^k &= (P^{-1}AP)(P^{-1}AP) \cdots (P^{-1}AP)(P^{-1}AP) \\ &= P^{-1}APP^{-1}APP^{-1} \cdots PP^{-1}APP^{-1}AP \\ &= P^{-1}AIAI \cdots IAIAP \\ &= P^{-1}A^kP \end{aligned}$$

Now let's expand

$$\begin{aligned} e^{(P^{-1}AP)t} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} (P^{-1}AP)^k \\ &= \sum_{k=0}^{\infty} \frac{t^k}{k!} P^{-1}A^kP \\ &= P^{-1} \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \right) P \\ &= P^{-1} e^{At} P \end{aligned}$$

5.1.1 Computation of e^{At} and A^k

5.1.2 Computation via Solution of State-Space Equations and Frequency Domain Expressions

An LTI CT state-space representation of an autonomous system has the form

$$\dot{x}(t) = Ax(t), \text{ where } x(t) \in \mathbb{R}^n$$

Let's test if $x(t) = e^{At}x_0$ is a solution of the homogeneous equation

$$\begin{aligned} x(0) &= e^{A0}x_0 = x_0 \\ \dot{x}(t) - Ax(t) &= (Ae^{At})x_0 - Ae^{At}x_0 = 0 \end{aligned}$$

Now let's remember Laplace domain solution of the same equation

$$\begin{aligned} \mathcal{L}[\dot{x}(t)] &= \mathcal{L}[Ax(t)] \\ sX(s) - x(0) &= AX(s) \\ [sI - A]X(s) &= x(0) \\ X(s) &= [sI - A]^{-1}x_0 \end{aligned}$$

If we connect time and s-domain solutions we obtain

$$e^{At} = \mathcal{L}^{-1} [[sI - A]^{-1}]$$

Now let's focus on A^k . An LTI DT state-space representation of an autonomous system has the form

$$x[k+1] = Ax[k]$$

Unlike CT systems we can compute the response iteratively easily

$$\begin{aligned} x[1] &= Gx[0] \\ x[2] &= Gx[1] = G^2x[0] \\ x[3] &= Gx[2] = G^3x[0] \\ &\vdots \\ x[k] &= Gx[k-1] = G^kx[0] \end{aligned}$$

Now let's remember form of the response in Z-domain.

$$\begin{aligned} \mathcal{Z}[x[k+1]] &= \mathcal{Z}[Gx[k]] \\ zX(z) - zx[0] &= GX(z) \\ (zI - G)X(z) &= zX(z) \\ X(z) &= z(zI - G)^{-1}x[0] \end{aligned}$$

If we connect time and s-domain solutions we obtain

$$G^k = \mathcal{Z}^{-1}\{z(zI - G)^{-1}\}$$

5.1.3 Computation via Diagonalization

Theorem: $A \in \mathbb{C}^{n \times n}$ is diagonalizable, if and only if there exists a (nonsingular) similarity transformation, $V \in \mathbb{C}^{n \times n}$, such that $A = V^{-1}\Lambda V$ where Λ is a diagonal matrix,

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ & 0 & & \ddots \\ & & & & \lambda_n \end{bmatrix}$$

where $\lambda_i \in \mathbb{C}$'s are the eigenvalues of A , which are the roots of the characteristic equation $d(\lambda) = \det(\lambda I - A)$.

Now let's compute A^k and e^{At} for a diagonalizable matrix

$$\begin{aligned} A^k &= (V^{-1}\Lambda V)^k = V^{-1}\Lambda^k V = V^{-1} \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} V \\ e^{At} &= e^{V^{-1}\Lambda V t} = V^{-1}e^{\Lambda t}V = V^{-1} \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} V \end{aligned}$$

A sufficient but not necessary condition that A will have n distinct eigenvalues in such a case characteristic equation will have the following form

$$d(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n), \text{ where } \lambda_i \neq \lambda_j, \text{ if } i \neq j$$

In this case, we also have the following properties associated with A

- A has n linearly independent eigenvectors
- For each λ_i there exists an eigenvector, v_i , such that $Av_i = \lambda v_i$ and $\text{Span}\{v_1, \dots, v_n\} = \mathbb{C}^n$
- $\forall \lambda_i$, geometric multiplicity is equal to algebraic multiplicity and they are both equal to 1, i.e. $\text{GM}(\lambda_i) = \text{AM}(\lambda_i) = 1$
- Minimal polynomial is equal to the characteristic equation, $m(\lambda) = d(\lambda)$
- For each λ_i , $\mathcal{N}(\lambda_i I - A) = \text{Span}\{v_i\}$ and $\dim[\mathcal{N}(\lambda_i I - A)] = 1$

Let's remember some concepts from EE501, to better understand and generalize *diagonalizable* and *non-diagonalizable* square matrices.

Definition: Given a matrix $A \in \mathbb{C}^{n \times n}$, the characteristic polynomial $d(\lambda)$ is defined as

$$\begin{aligned} d(\lambda) &= (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_k)^{r_k} \\ k &: \# \text{ distinct eigenvalues where } k \leq n \\ r_i &= \text{AM}(\lambda_i) : \# \text{ algebraic multiplicity } \lambda_i, \text{ where } n = \sum_{i=1}^k r_i \end{aligned}$$

Theorem: Every $n \times n$ matrix satisfies its characteristic equation (Cayley-Hamilton Theorem)

$$d(\lambda) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_k)^{r_k} = \lambda^n + d_{n-1}\lambda^{n-1} + \cdots + d_0$$

$$d(A) = A^n + d_{n-1}A^{n-1} + \cdots + d_1A + d_0I = 0_{n \times n}$$

Remark: Any power of a matrix $A \in \mathbb{C}^{n \times n}$ can be written as a linear combination of $\mathcal{A}_n = \{I, A, A^2, \dots, A^{n-1}\}$.

Note that Cayley-Hamilton theorem does not guarantee that $\{I, A, A^2, \dots, A^{n-1}\}$ are linearly independent.

Definition: For an $A \in \mathbb{R}^{n \times n}$, the minimal polynomial $m(\lambda)$ is the monic polynomial with the smallest degree such that $m(A) = 0_{n \times n}$

Ex 5.1

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \rightarrow d_1(\lambda) = (\lambda + 1)^2, \text{ Let } m_1(\lambda) = (\lambda + 1) \rightarrow m_1(A) = A + I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0_{n \times n} \checkmark$$

$$A_2 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \rightarrow d_2(\lambda) = (\lambda + 1)^2, \text{ Let } m_2(\lambda) = (\lambda + 1) \rightarrow m_1(A) = A + I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq 0_{n \times n} \times$$

$$\text{Let } m_2(\lambda) = (\lambda + 1)^2 \rightarrow m_1(A) = (A + I)^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0_{n \times n} \checkmark$$

Theorem: Given $A \in \mathbb{R}^{n \times n}$, let $m(\lambda)$ be its minimal polynomial

1. $m(\lambda)$ is unique
2. $m(\lambda)$ divides $d(\lambda)$ without any reminder. $\exists q(\lambda)$ such that $d(\lambda) = q(\lambda)m(\lambda)$
3. Each root of $d(\lambda)$ is a root of $m(\lambda)$, then

$$m(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k} = \lambda^l + c_{l-1}\lambda^{l-1} + \cdots + c_0$$

where $1 \leq m_i \leq r_i, k \leq l \leq n, m_1 + \cdots + m_k = l$

$$q(\lambda) = (\lambda - \lambda_1)^{r_1 - m_1} \cdots (\lambda - \lambda_k)^{r_k - m_k}$$

Proof: of (3)

Since $m(\lambda)$ is the minimal polynomial, we know that $m(A) = 0_{n \times n}$. Let (λ_i, v_i) is an eigenvalue, eigenvector pair of A such that $Av_i = \lambda_i v_i$, then

$$m(A) = 0_{n \times n} \rightarrow m(A)v_i = 0_{n \times n}$$

$$(A^l + c_{l-1}A^{l-1} + \cdots + c_1A + c_0I)v_i = 0_{n \times n}$$

$$(\lambda_i^l v_i + c_{l-1}\lambda_i^{l-1}v_i + \cdots + c_1\lambda_i v_i + c_0v_i) = 0_{n \times n}$$

$$\lambda_i^l + c_{l-1}\lambda_i^{l-1} + \cdots + c_1\lambda_i + c_0 = 0$$

$$m(\lambda_i) = 0 \checkmark$$

Corollary: Let l be the order of the minimal polynomial, then the elements of $\mathcal{A}_l = \{I, A, A^2, \dots, A^{l-1}\}$ are linearly independent and higher order A^i 's can be written as a linear combination of $\{I, A, A^2, \dots, A^{l-1}\}$

Theorem: $\mathbb{C}^n = \mathcal{N}((\lambda_1 - A)^{m_1}) \oplus \cdots \oplus \mathcal{N}((\lambda_k I - A)^{m_k})$ where $\text{Dim}[\mathcal{N}((\lambda_i I - A)^{m_i})] = r_i, \forall i$

Now let's state the necessary sufficient condition(s) such that a matrix can be diagonalizable

Theorem: $A \in \mathbb{C}^{n \times n}$ is diagonalizable, if and only if

- Minimal polynomial has no repeated root. i.e.

$$m(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_n)^{m_k} \text{ where } m_i = 1 \ \forall i$$

$$m(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$$

- Geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity, i.e.

$$\forall i \text{ GM}(\lambda_i) = \dim(\mathcal{N}(\lambda_i - A)) = r_i = \text{AM}(\lambda_i)$$

- There exist n linearly independent eigenvectors of A

$$\exists \mathcal{V} = \{v_1, \dots, v_n\}, \text{ where } Av_i = \lambda_j \text{ \& Span}(\mathcal{V}) = \mathbb{C}^n$$

Ex 5.2

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \rightarrow m(\lambda) = (\lambda + 1), \text{ GM}(-1) = \dim(\mathcal{N}((-1) - A)) = 2 = \text{AM}(-1), \mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \checkmark$$

$$A_2 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \rightarrow m(\lambda) = (\lambda + 1)^2, \text{ GM}(-1) = \dim(\mathcal{N}((-1) - A)) = 1 \leq \text{AM}(-1), \mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \times$$

5.1.4 Computation via Jordan Form

When a matrix A is not diagonalizable, i.e. does not satisfy necessary and sufficient conditions listed above, there exists a similarity transformation, $A = G^{-1}JG$, such that J is Jordan canonical form (as close as possible to a diagonal form). Note that each $A \in \mathbb{R}^{n \times n}$ is similar to only one such J (except for a reordering of the blocks):

$$J = \begin{bmatrix} J_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & J_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & & \\ \mathbf{0} & & & \ddots & J_{N_J} \end{bmatrix} \text{ where } J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots \\ \vdots & & \ddots & & \\ & & & \ddots & 1 \\ 0 & & & 0 & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}$$

Ex 5.3

$$A_1 = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{array}{l} m(\lambda) = (\lambda + 1) \\ \text{GM}(-1) = \dim(\mathcal{N}((-1) - A)) = 4 \\ \mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \end{array}$$

$$A_2 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{array}{l} m(\lambda) = (\lambda + 1)^2 \\ \text{GM}(-1) = \dim(\mathcal{N}((-1) - A)) = 3 \\ \mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \end{array}$$

$$A_3 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{array}{l} m(\lambda) = (\lambda + 1)^2 \\ \text{GM}(-1) = \dim(\mathcal{N}((-1) - A)) = 2 \\ \mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\} \end{array}$$

$$A_4 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{array}{l} m(\lambda) = (\lambda + 1)^3 \\ \text{GM}(-1) = \dim(\mathcal{N}((-1) - A)) = 2 \\ \mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \end{array}$$

$$A_5 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{array}{l} m(\lambda) = (\lambda + 1)^4 \\ \text{GM}(-1) = \dim(\mathcal{N}((-1) - A)) = 1 \\ \mathcal{V} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\} \end{array}$$

Proposition: Jordan transformation of a matrix $A \in \mathbb{C}^{n \times n}$ with $d(\lambda) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_n)^{r_k}$ and $m(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_n)^{m_k}$, for each λ_i

- m_i : Size of the largest Jordan block for λ_i
- $\text{GM}(\lambda) = \dim(\mathcal{N}(\lambda_i - A))$: # Jordan blocks for λ_i
- $\dim(\mathcal{N}(\lambda_i - A)^k) - \dim(\mathcal{N}(\lambda_i - A)^{k-1})$: # Jordan blocks for λ_i with size $\geq k$

5.1.4.1 Construction of G for $A = G^{-1}JG$

If A is a diagonalizable matrix transformation matrices, V & V^{-1} , are simply composed of right and left eigenvectors

$$V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}, \text{ where } Av_i = \lambda_i v_i$$

$$V^{-1} = \begin{bmatrix} \bar{v}_1^T \\ \vdots \\ \bar{v}_n^T \end{bmatrix}, \text{ where } \bar{v}_i^T A = \lambda_i \bar{v}_i^T$$

$$V^{-1}V = I, \bar{v}_i^T v_i = 1, \bar{v}_i^T v_j = 0 \text{ for } i \neq j$$

When the matrix, $A \in \mathbb{C}^{n \times n}$, is not diagonalizable, i.e. does not have n linearly independent eigenvectors to construct the transformation matrix, G , we need to add (special) linearly independent vectors to complete the transformation matrix, such that $A = G^{-1}JG$. These vectors are generated from the eigenvectors and are called *generalized eigenvectors* of A . Suppose (λ, g_1) is an eigenvalue–eigenvector pair of A associated with Jordan block of size k . $k - 1$ generalized eigenvectors, $\{g_2, \dots, g_k\}$, are constructed as follows

$$\begin{aligned} Ag_1 &= \lambda g_1 \rightarrow (A - \lambda I)g_1 = 0 \\ Ag_2 &= \lambda g_2 + g_1 \rightarrow (A - \lambda I)g_2 = g_1, \text{ note } (A - \lambda I)^2 g_2 = 0 \text{ \& } (A - \lambda I)g_2 \neq 0 \\ Ag_3 &= \lambda g_3 + g_2 \rightarrow (A - \lambda I)g_3 = g_2, \text{ note } (A - \lambda I)^3 g_3 = 0 \text{ \& } (A - \lambda I)^2 g_3 \neq 0 \\ &\vdots \\ Ag_k &= \lambda g_k + g_{k-1} \rightarrow (A - \lambda I)g_k = g_{k-1}, \text{ note } (A - \lambda I)^k g_k = 0 \text{ \& } (A - \lambda I)^{k-1} g_k \neq 0 \end{aligned}$$

The string $\begin{bmatrix} g_1 & g_2 & \cdots & g_k \end{bmatrix}$ is called a Jordan chain. We can also re-write the relation between the generalized eigenvectors in matrix form using the Jordan chain

$$A \begin{bmatrix} g_1 & g_2 & \cdots & g_k \end{bmatrix} = \begin{bmatrix} g_1 & g_2 & \cdots & g_k \end{bmatrix} \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots \\ \vdots & & \ddots & & \\ & & & \ddots & 1 \\ 0 & & & 0 & \lambda \end{bmatrix}$$