Lecture 5, Sep 12, 2025

Solving $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$

- Consider the *autonomous* (i.e. no control input) LTI system, $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^n$; we will show that this is solved by $e^{\boldsymbol{A}t}\boldsymbol{x}_0$
- We will show that $\frac{\mathrm{d}}{\mathrm{d}t}e^{\boldsymbol{A}t} = e^{\boldsymbol{A}t}\boldsymbol{A}$
  - $$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}e^{\boldsymbol{A}t} &= \lim_{h\to 0}\frac{e^{\boldsymbol{A}(t+h)} - e^{\boldsymbol{A}t}}{h} \\ &= \lim_{h\to 0}\frac{e^{\boldsymbol{A}t}e^{\boldsymbol{A}h} - e^{\boldsymbol{A}t}}{h} && \text{Since } \boldsymbol{A}h \text{ and } \boldsymbol{A}t \text{ commute} \\ &= \lim_{h\to 0}\frac{e^{\boldsymbol{A}t}(e^{\boldsymbol{A}h} - \boldsymbol{I})}{h} \\ &= e^{\boldsymbol{A}t}\lim_{h\to 0}\frac{1}{h}\left(-\boldsymbol{I} + \sum_{k=0}^{\infty}\frac{(\boldsymbol{A}h)^k}{k!}\right) && \text{Matrix exponential definition} \\ &= e^{\boldsymbol{A}t}\lim_{h\to 0}\frac{1}{h}\left(-\boldsymbol{I} + \boldsymbol{I} + h\sum_{k=1}^{\infty}\frac{\boldsymbol{A}^k h^{k-1}}{k!}\right) && \text{Take out first term and factor } h \\ &= e^{\boldsymbol{A}t}\lim_{h\to 0}\sum_{k=1}^{\infty}\frac{\boldsymbol{A}^k \boldsymbol{h}^{k-1}}{k!} \\ &= e^{\boldsymbol{A}t}\lim_{h\to 0}\left(\frac{\boldsymbol{A}h^0}{1!} + \sum_{k=2}^{\infty}\frac{\boldsymbol{A}^k h^{k-1}}{k!}\right) && \text{Take out first term} \\ &= e^{\boldsymbol{A}t}\boldsymbol{A}\end{aligned}$$
  - Note due to commutativity, we could've also taken out $e^{\boldsymbol{A}t}$ on the right and get $\frac{\mathrm{d}}{\mathrm{d}t}e^{\boldsymbol{A}t} = \boldsymbol{A}e^{\boldsymbol{A}t}$

Theorem

The differential equation $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^n$ has the **unique solution**

$$\boldsymbol{x}(t) = e^{\boldsymbol{A}t}\boldsymbol{x}_0, t \geq 0$$

- To show existence:
  - $\dot{\boldsymbol{x}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}(e^{\boldsymbol{A}t}\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}t}(e^{\boldsymbol{A}t})\boldsymbol{x}_0 = \boldsymbol{A}e^{\boldsymbol{A}t}\boldsymbol{x}_0 = \boldsymbol{A}\boldsymbol{x}(t)$
  - $\boldsymbol{x}(0) = e^{\boldsymbol{0}}\boldsymbol{x}_0 = \boldsymbol{I}\boldsymbol{x}_0 = \boldsymbol{x}_0$
- To show uniqueness, let $\boldsymbol{y}(t)$ be any other solution to the differential equation; we want to show that $\boldsymbol{y} = \boldsymbol{x}$
  - $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{y}(t) = \boldsymbol{A}\boldsymbol{y}(t)$ and $\boldsymbol{y}(0) = \boldsymbol{x}_0$
  - We want to show $\boldsymbol{y}(t) = e^{\boldsymbol{A}t}x_0$, equivalently $e^{-\boldsymbol{A}t}\boldsymbol{y}(t) = e^{-\boldsymbol{A}t}e^{\boldsymbol{A}t}\boldsymbol{x}_0 = \boldsymbol{x}_0$
  - Notice that both sides are constants, so we can try taking the derivative and seeing what we get
  - $$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}e^{-\boldsymbol{A}t}\boldsymbol{y}(t) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}e^{-\boldsymbol{A}t}\right)\boldsymbol{y}(t) + e^{-\boldsymbol{A}t}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{y}(t) \\ &= -e^{-\boldsymbol{A}t}\boldsymbol{A}\boldsymbol{y}(t) + e^{-\boldsymbol{A}t}\boldsymbol{A}\boldsymbol{y}(t) \\ &= -e^{\boldsymbol{A}t}\left(-\boldsymbol{A}\boldsymbol{y}(t) + \boldsymbol{A}\boldsymbol{y}(t)\right) \\ &= \boldsymbol{0}\end{aligned}$$
  - Since $e^{-\boldsymbol{A}t}\boldsymbol{y}(t)$ has a zero derivative, it must be constant, so $e^{-\boldsymbol{A}t}\boldsymbol{y}(t) = e^{-0\boldsymbol{A}}\boldsymbol{y}(0) = \boldsymbol{y}(0)$, but $\boldsymbol{y}(0) = \boldsymbol{x}_0$ since $\boldsymbol{y}$ is a solution to the differential equation
  - Therefore we've shown $e^{-\boldsymbol{A}t}\boldsymbol{y}(t) = \boldsymbol{x}_0$ and so $\boldsymbol{y}(t) = e^{\boldsymbol{A}t}\boldsymbol{x}_0$, and thus $\boldsymbol{x}(t) = \boldsymbol{y}(t)$

Computing the Matrix Exponential

Matrix Exponential by Laplace Transform

- Consider $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^n$, taking the Laplace transform:
  - $$\begin{aligned} &\mathcal{L}\{\dot{\boldsymbol{x}}\} = \mathcal{L}\{\boldsymbol{A}\boldsymbol{x}\} \\ \implies &s\boldsymbol{X}(s) - \boldsymbol{x}(0) = \boldsymbol{A}\boldsymbol{X}(s) \\ \implies &s\boldsymbol{X}(s) - \boldsymbol{A}\boldsymbol{X}(s) = \boldsymbol{x}(0) \\ \implies &(s\boldsymbol{I} - \boldsymbol{A})\boldsymbol{X}(s) = \boldsymbol{x}_0 \\ \implies &\boldsymbol{X}(s) = (s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{x}_0 \\ \implies &\boldsymbol{x}(t) = \mathcal{L}^{-1}\left\{(s\boldsymbol{I} - \boldsymbol{A})^{-1}\right\}\boldsymbol{x}_0 \end{aligned}$$
- Because we know that the unique solution is $\boldsymbol{x}(t) = e^{\boldsymbol{A}t}\boldsymbol{x}_0$, $\boldsymbol{e}^{\boldsymbol{A}t}\boldsymbol{x}_0 = \mathcal{L}^{-1}\left\{(s\boldsymbol{I} - \boldsymbol{A})^{-1}\right\}\boldsymbol{x}_0$
  - Since this holds for all $\boldsymbol{x}_0$, it must be that $\boldsymbol{e}^{\boldsymbol{A}t} = \mathcal{L}^{-1}\left\{(s\boldsymbol{I} - \boldsymbol{A})^{-1}\right\}$
  - Formally, to justify this, consider the case where $\boldsymbol{x}_0 = \boldsymbol{e}_i$, i.e. all zeros except 1 in the ith row; substituting this into the equation we get that the ith column of the LHS must be equal to the ith column of the RHS, so do this for all n columns

Matrix Exponential by Modal Decomposition (Eigenvectors & Eigenvalues)

- Recall that $\lambda \in \mathbb{C}$ is an eigenvalue of $\boldsymbol{A}$ if and only if $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = \det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$; i.e. λ are the roots of the characteristic polynomial of $\boldsymbol{A}$
  - To find eigenvectors corresponding to each λ, we find a basis for $\mathcal{N}(\lambda\boldsymbol{I} - \boldsymbol{A})$ where $\mathcal{N}$ denotes the null space

Definition

If there exists a nonsingular matrix $\boldsymbol{P} \in \mathbb{C}^{n\times n}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is diagonal, then $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is *diagonalizable.*

Theorem

$\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is diagonalizable if and only if it has n linearly independent eigenvectors.

- Suppose $\boldsymbol{A}$ has n linearly independent eigenvectors, and let $\boldsymbol{P} = \begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{bmatrix}$
  - $$\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix}\boldsymbol{A}\boldsymbol{v}_1 & \cdots & \boldsymbol{A}\boldsymbol{v}_n\end{bmatrix} = \begin{bmatrix}\lambda_1\boldsymbol{v}_1 & \cdots & \lambda_n\boldsymbol{v}_n\end{bmatrix} = \boldsymbol{P}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} = \boldsymbol{P}\boldsymbol{\Lambda}$$
  - Since $\boldsymbol{P}$ has all linearly independent columns it is invertible, therefore $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$
- Note $\boldsymbol{A}$ is diagonalizable if it has n distinct eigenvalues (but diagonalizability does not always imply distinct eigenvalues); $\boldsymbol{A}$ is also diagonalizable if it is symmetric (the *spectral theorem*)

Theorem

If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is diagonalizable, then $\boldsymbol{e}^{\boldsymbol{A}} = \boldsymbol{P}e^{\boldsymbol{\Lambda}}\boldsymbol{P}^{-1}$, where

$$e^{\boldsymbol{\Lambda}} = \begin{bmatrix}e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n}\end{bmatrix}$$

and λ_i are eigenvalues of $\boldsymbol{A}$.

- We can show by induction that $\boldsymbol{A}^n = (\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1})^n = \boldsymbol{P}\boldsymbol{\Lambda}^n\boldsymbol{P}^{-1}$, then we can prove the above by substituting this into the definition of the matrix exponential, and noting that taking a diagonal matrix to a power is equivalent to taking each of the components to that power