

Lecture 6, Sep 17, 2025

Computing the Matrix Exponential – Continued

Complex Eigenvalues

- Consider $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ -1 & -1 \end{bmatrix}$; the eigenvectors and eigenvalues are $\left(\begin{bmatrix} 2 \\ i-1 \end{bmatrix}, i\right)$ and $\left(\begin{bmatrix} 2 \\ -i-1 \end{bmatrix}, -i\right)$
 - Recall that eigenvector/eigenvalues always come in complex conjugate pairs, so for a 2x2 we can take the conjugate to find the other one
 - \mathbf{A} has two distinct eigenvalues, so it is diagonalizable
- Direct calculation: $e^{\mathbf{A}t} = \mathbf{P}e^{\mathbf{\Lambda}t}\mathbf{P}^{-1}$

$$= \dots$$

$$= \begin{bmatrix} \sin t + \cos t & 2 \sin t \\ -\sin t & \cos t - \sin t \end{bmatrix}$$
 - Note we applied Euler's formula
- Alternatively, we can define \mathbf{P} and $\mathbf{\Lambda}$ differently to avoid dealing with complex numbers
 - Denote $\mathbf{v} = \begin{bmatrix} 2 \\ i-1 \end{bmatrix}$ (the other eigenvector is $\bar{\mathbf{v}}$)
 - Let $\tilde{\mathbf{P}} = [\operatorname{Re}(v) \quad \operatorname{Im}(v)] = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$
 - Let $\tilde{\mathbf{\Lambda}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (the imaginary parts of λ_1, λ_2 are in the off-diagonal entries)
 - We can show that $e^{\begin{bmatrix} a & b \\ -b & a \end{bmatrix}t} = e^{at} \begin{bmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{bmatrix}$
 - * When we have complex eigenvalues, we see that this results in a rotation with a rate of decay
 - Now we can compute $\tilde{\mathbf{P}}e^{\tilde{\mathbf{\Lambda}}t}\tilde{\mathbf{P}}^{-1}$ and this leads to the same answer

Non-Diagonalizable Case – Jordan Forms

- Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ have less than n linearly independent eigenvectors, i.e. it is non-diagonalizable; in this case we cannot form an invertible \mathbf{P} with the eigenvectors
 - We will introduce the notion of *generalized eigenvectors* which allow us to form \mathbf{P} in this case
- Recall that the characteristic polynomial of \mathbf{A} is $x_A(s) = \det(s\mathbf{I} - \mathbf{A}) = \prod_{i=1}^{\sigma} (s - \lambda_i)^{m_i}$, where $\sum_{i=1}^{\sigma} m_i = n$
 - m_i is the *algebraic multiplicity* of the eigenvalue λ_i
- The *minimal polynomial* $\psi_A(s)$ is the polynomial of least degree such that $\psi_A(\mathbf{A}) = \mathbf{0}$
 - We can show that such a polynomial always exists
 - The minimal polynomial has the form $\psi_A(s) = \prod_{i=1}^{\sigma} (s - \lambda_i)^{l_i}$ where $l_i \leq m_i$
 - * Note that normally $s \in \mathbb{C}$; when we substitute \mathbf{A} into the polynomial we replace λ_i with $\lambda_i \mathbf{I}$
 - * This is the same form as the characteristic polynomial but we may not have to repeat each term as many times, i.e. some of the information in the characteristic polynomial is redundant
 - The l_i are known as the *geometric multiplicity* of λ_i
 - Each l_i is also the number of linearly independent eigenvectors corresponding to λ_i (i.e. $\dim(\mathcal{N}(\lambda_i \mathbf{I} - \mathbf{A}))$ where \mathcal{N} denotes null space)
- Using this, we can decompose \mathbb{C}^n into σ subspaces: $\mathbb{C}^n = \mathcal{N}(\lambda_1 \mathbf{I} - \mathbf{A})^{l_1} \oplus \dots \oplus \mathcal{N}(\lambda_{\sigma} \mathbf{I} - \mathbf{A})^{l_{\sigma}}$
 - σ denotes a direct sum, $\mathcal{V} \oplus \mathcal{W} = \{ \mathbf{v} + \mathbf{w} \mid \mathbf{v} \in \mathcal{V}, \mathbf{w} \in \mathcal{W} \}$, where it is required that $\mathcal{V} \cap \mathcal{W} = \{ \mathbf{0} \}$
- Example: Let $\mathbf{A} \in \mathbb{R}^{6 \times 6}$ have the characteristic polynomial $x_A(s) = (s - \lambda)^6$ (i.e. λ repeated 6 times) and $\psi_A(s) = (s - \lambda)^3$ (i.e. only 3 independent eigenvectors)
 - $\mathcal{N}(\lambda \mathbf{I} - \mathbf{A}) = \operatorname{span} \{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \}$ from the 3 independent eigenvectors

- Let $\mathbf{v}_1, \mathbf{v}_2$ be linearly independent solutions to $\begin{cases} (\lambda\mathbf{I} - \mathbf{A})\mathbf{v}_1 = -\mathbf{e}_1 \\ (\lambda\mathbf{I} - \mathbf{A})\mathbf{v}_2 = -\mathbf{e}_2 \end{cases}$ and \mathbf{w}_1 be a solution to $(\lambda\mathbf{I} - \mathbf{A})\mathbf{w}_1 = -\mathbf{v}_1$ and all of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1$ are linearly independent
 - * Multiplying by $(\lambda\mathbf{I} - \mathbf{A})$ on both sides, the right hand side goes to zero, so we see that $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{N}((\lambda\mathbf{I} - \mathbf{A})^2)$ and $\mathbf{w}_1 \in \mathcal{N}((\lambda\mathbf{I} - \mathbf{A})^3)$
 - * In general, a *generalized eigenvector* is a vector such that $(\lambda\mathbf{I} - \mathbf{A})^n \mathbf{v} = \mathbf{0}$ where $n \in \mathbb{N}$
 - * For each of the eigenvectors, we can create an entire chain of these generalized eigenvectors
- Now notice that if we rearrange the expressions we get $\mathbf{A}\mathbf{v}_1 = \lambda\mathbf{v}_1 + \mathbf{e}_1$, and $\mathbf{A}\mathbf{w}_1 = \lambda\mathbf{w}_1 + \mathbf{v}_1$ (and so on for longer chains)
- Let $\mathbf{P} = [\mathbf{e}_1 \ \mathbf{v}_1 \ \mathbf{w}_1 \ \mathbf{e}_2 \ \mathbf{v}_2 \ \mathbf{e}_3]$
 - * We group together the chain related to \mathbf{e}_1 , then the chain of \mathbf{e}_2 and so on
- $\mathbf{AP} = [\lambda\mathbf{e}_1 \ \lambda\mathbf{v}_1 + \mathbf{e}_1 \ \lambda\mathbf{w}_1 + \mathbf{v}_1 \ \lambda\mathbf{e}_2 \ \lambda\mathbf{v}_2 + \mathbf{e}_2 \ \lambda\mathbf{e}_3] = \mathbf{P} \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{bmatrix} = \mathbf{PJ}$
- The resulting matrix \mathbf{J} is the *Jordan form*, which is block diagonal; each of the blocks is known as a *Jordan block*
 - * The number of Jordan blocks is always equal to the number of linearly independent eigenvectors
- Note there are multiple other forms that are possible, depending on how we choose to do the generalized eigenvectors: we can have 3 Jordan blocks of size 2, 1 block of size 4 and 2 blocks of size 1; as long as we only have 3 blocks, any combination is possible
 - * Also, in general we have multiple distinct eigenvalues, so we have a block for each distinct eigenvalue, and then within each block we have a block for each linearly independent eigenvector for that eigenvalue (and its chain)
- In this way we can generalize the notion of diagonalizability to matrices without a full set of independent eigenvectors

Definition

The *Jordan form* of $\mathbf{A} \in \mathbb{R}^{n \times n}$ has the form

$$\mathbf{P}^{-1} \mathbf{AP} = \mathbf{J} = \begin{bmatrix} \mathbf{J}_{\lambda_1} & & \\ & \ddots & \\ & & \mathbf{J}_{\lambda_k} \end{bmatrix}$$

where k is the number of distinct eigenvalues, and each \mathbf{J}_{λ_i} has form

$$\mathbf{J}_{\lambda_i} = \begin{bmatrix} \mathbf{J}_{\lambda_i}^1 & & \\ & \ddots & \\ & & \mathbf{J}_{\lambda_i}^{l_i} \end{bmatrix} \in \mathbb{C}^{m_i \times m_i}$$

where m_i is the *algebraic multiplicity* of λ_i (number of times it appears as a root in the characteristic equation) and l_i is its *geometric multiplicity* (number of linearly independent eigenvectors for λ_i); each $\mathbf{J}_{\lambda_i}^j$ takes the form

$$\mathbf{J}_{\lambda_i}^j = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

with λ_i repeated along the diagonal, and 1s above each λ .

- With this we can now write $e^{\mathbf{A}t} = \mathbf{P}e^{\mathbf{J}t}\mathbf{P}^{-1}$ (following the same proof as the case of normal diagonalization for \mathbf{A})

- The block diagonal form means $e^{\mathbf{J}t} = \begin{bmatrix} e^{\mathbf{J}_{\lambda_1}t} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & e^{\mathbf{J}_{\lambda_k}t} \end{bmatrix}$

- Similarly $e^{\mathbf{J}_{\lambda_i}t} = \begin{bmatrix} e^{\mathbf{J}_{\lambda_i}^1} & & & \\ & \ddots & & \\ & & e^{\mathbf{J}_{\lambda_i}^{l_i}} & \end{bmatrix}$

- $\mathbf{J}_{\lambda_i}^j t = (\lambda_i \mathbf{I} + \mathbf{N}) t$ where \mathbf{N} is a matrix with 1s above the diagonal

- Therefore $e^{\mathbf{J}_{\lambda_i}^j t} = e^{\lambda_i t} e^{\mathbf{N}t}$

- * \mathbf{N} is a nilpotent matrix, so we can show that eventually the higher order terms in the infinite series expansion for $e^{\mathbf{N}t}$ all go to zero

- We can show that $e^{\mathbf{N}t}$ has form $\begin{bmatrix} 1 & t & \cdots & \frac{t^{p-1}}{(p-1)!} \\ \ddots & \ddots & & \vdots \\ & 1 & t & \\ & & & 1 \end{bmatrix} \in \mathbb{C}^{p \times p}$

- * Here p is the size of the block $\mathbf{J}_{\lambda_i}^j$ (i.e. the size of the chain of generalized eigenvectors corresponding to λ_i), which is not unique