ECE240 Introduction to Linear Dynamical Systems
Lecture 6: Jordan Form

Dr. Yu Zhang
ECE Department, UC Santa Cruz

Fall 2025

Outline

Motivation: Simplifying Matrices

We often want to simplify a matrix $\mathbf{A}$ to a diagonal form:

$$\mathbf{A} = \mathbf{V}\,\mathbf{\Lambda}\,\mathbf{V}^{-1}, \quad \mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Then the matrix exponential is easy to compute:

$$e^{\mathbf{A}t} = \mathbf{V}\, e^{\mathbf{\Lambda}t}\, \mathbf{V}^{-1}, \qquad e^{\mathbf{\Lambda}t} = \operatorname{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}).$$

Each state evolves independently as $e^{\lambda_i t}$.

Diagonalization $\Rightarrow$ independent modes and simple dynamics.

When Diagonalization Fails

Consider $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, compute $(\mathbf{A} - \mathbf{I})\mathbf{v} = \mathbf{0}$:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathbf{0} \Rightarrow v_2 = 0,\; v_1 \text{ free.}$$

Thus, there is only one eigenvector

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The eigenspace is a line (1D), while a 2×2 diagonalizable matrix requires 2D eigenbasis $\Rightarrow$ matrix $\mathbf{A}$ is **not diagonalizable**.

Eigenvalue Multiplicity Inequality

- **Algebraic multiplicity (AM)** of an eigenvalue λ: the number of times λ appears as a root of the characteristic polynomial.
- **Geometric multiplicity (GM)** of λ: the dimension of the eigenspace $\mathcal{E}_\lambda$.

- For any eigenvalue λ, we always have:

$$\boxed{1 \leq \text{GM}(\lambda) \leq \text{AM}(\lambda)}.$$

Proof of Eigenvalue Multiplicity Inequality (1/2)

Setup: Let λ_i be an eigenvalue of $\mathbf{A} \in \mathbb{F}^{n \times n}$.

Lower bound: for the eigenvector $\mathbf{v} \neq \mathbf{0}$ associated with λ_i, we have $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{v} = \mathbf{0}$. Thus, $1 \leq \mathrm{GM}(\lambda_i)$.

Upper bound: Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_g\}$ be a basis of the eigenspace $\mathcal{E}_{\lambda_i}$, where $g = \mathrm{GM}(\lambda_i)$. Extend it to a full basis of $\mathbb{F}^n$ to form the matrix

$$\mathbf{V} := [\mathbf{v}_1, \ldots, \mathbf{v}_g, \mathbf{w}_1, \ldots, \mathbf{w}_{n-g}].$$

It can be seen that $\mathbf{A}$ is *similar* to a block upper-triangular matrix $\tilde{\mathbf{A}}$ (denoted as $\mathbf{A} \sim \tilde{\mathbf{A}}$), as follows

Proof of Eigenvalue Multiplicity Inequality (2/2)

$$
\begin{aligned}
\tilde{\mathbf{A}} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V} &= \mathbf{V}^{-1}\mathbf{A}[\,\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{w}_{n-g}\,] \\
&= [\,\mathbf{V}^{-1}\mathbf{A}\mathbf{v}_1 \mid \cdots \mid \mathbf{V}^{-1}\mathbf{A}\mathbf{v}_g \mid \cdots \mid \mathbf{V}^{-1}\mathbf{A}\mathbf{w}_{n-g}\,] \\
&= [\,\lambda_i \mathbf{e}_1 \mid \cdots \mid \lambda_i \mathbf{e}_g \mid \cdots \mid \mathbf{V}^{-1}\mathbf{A}\mathbf{w}_{n-g}\,] \\
&= \begin{bmatrix} \lambda_i \mathbf{I}_g & * \\ \mathbf{0} & \mathbf{B} \end{bmatrix}.
\end{aligned}
$$

Since similar matrices have the same characteristic polynomial, we get

$$
p_{\mathbf{A}}(\lambda) = p_{\tilde{\mathbf{A}}}(\lambda) = \det\left(\lambda \mathbf{I} - \tilde{\mathbf{A}}\right) = (\lambda - \lambda_i)^g \det(\lambda \mathbf{I} - \mathbf{B}).
$$

Thus, $(\lambda - \lambda_i)^g$ divides $p_{\mathbf{A}}(\lambda)$, implying

$$
\mathrm{AM}(\lambda_i) \geq g = \mathrm{GM}(\lambda_i).
$$

Diagonalizability Criterion

Theorem (Diagonalizability Criterion)

A square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ *is diagonalizable if and only if, for every eigenvalue* λ_i *of* $\mathbf{A}$,

$$\mathrm{GM}(\lambda_i) = \mathrm{AM}(\lambda_i).$$

Equivalently, if there exists an eigenvalue λ *such that* $\mathrm{GM}(\lambda) < \mathrm{AM}(\lambda)$, *then* $\mathbf{A}$ *is not diagonalizable (*$\mathbf{A}$ *is defective).*

Jordan Form: The Next Best Thing

If **A** cannot be diagonalized, we can still find an invertible **V** such that

$$\mathbf{J} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \lambda\mathbf{I} + \mathbf{N}, \quad \mathbf{N} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \ \mathbf{N}^2 = \mathbf{0}.$$

- **J** is called **Jordan block**: upper bidiagonal, and only 1's just above the diagonal.
- $\mathbf{N} \in \mathbb{R}^{n\times n}$ is **nilpotent** if $\mathbf{N}^k = \mathbf{0}$ for some integer $k > 0$. The smallest such k is called the *index of nilpotency*.
- All eigenvalues of a nilpotent matrix are 0 (because $\mathbf{0} = \mathbf{N}^k\mathbf{v} = \lambda^k\mathbf{v}$).

Jordan form $\Rightarrow$ "diagonalization with small corrections."

$e^{\mathbf{A}t}$ via JCF (2×2 System)

If $\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$, then

$$e^{\mathbf{A}t} = \mathbf{V}\, e^{\mathbf{J}t}\, \mathbf{V}^{-1}.$$

Since $\mathbf{J} = \lambda\mathbf{I} + \mathbf{N}$ and $\mathbf{N}^2 = \mathbf{0}$, we have

$$e^{\mathbf{J}t} = e^{(\lambda\mathbf{I}+\mathbf{N})t} = e^{\lambda t} \times e^{\mathbf{N}t} = e^{\lambda t} \times (\mathbf{I} + \mathbf{N}t).$$

Hence,

$$e^{\mathbf{J}t} = e^{\lambda t}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \implies e^{\mathbf{A}t} = \mathbf{V}\, e^{\lambda t}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\mathbf{V}^{-1}.$$

- $\mathbf{N}$ creates the $te^{\lambda t}$ terms in $e^{\mathbf{A}t}$.
- Jordan form reveals how repeated eigenvalues cause coupling and polynomial growth $t^k e^{\lambda t}$.

Jordan Canonical Form Theorem

Theorem (JCF: Existence and Uniqueness)

Let $\mathbf{A} \in \mathbb{C}^{n\times n}$*. Then there exists an invertible matrix* $\mathbf{V}$ *such that* $\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$*, where* $\mathbf{J}$ *is a block diagonal matrix* $\mathbf{J} = \mathrm{diag}\big(\mathbf{J}_{r_1}(\lambda_1), \mathbf{J}_{r_2}(\lambda_2), \ldots, \mathbf{J}_{r_q}(\lambda_q)\big)$*, and each block*

$$\mathbf{J}_{r_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_i & 1 \\ 0 & \cdots & 0 & 0 & \lambda_i \end{bmatrix}_{r_i \times r_i}$$

is a Jordan block. The multiset of eigenvalues $\{\lambda_i\}$ *and the sizes of Jordan blocks associated with each eigenvalue are uniquely determined by* $\mathbf{A}$*, up to permutation of blocks.*

JCF Implications and Examples

- $\mathbf{A}$ is diagonalizable $\Leftrightarrow$ all $r_i = 1$.
- Each eigenvalue λ may have one or more Jordan blocks.
- $\text{GM}(\lambda) =$ number of blocks for λ.
- $\text{AM}(\lambda) =$ sum of block sizes for λ.

Example 1: Diagonalizable Case

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \mathbf{J}$$

- Eigenvalues: $\lambda_1 = 2$, $\lambda_2 = 3$
- For $\lambda = 2$: AM $= 2$, GM $= 2$
- Each Jordan block size is 1
- Two independent eigenvectors of $\lambda_1 = 2$, fully diagonalizable

Example 2: Single Jordan Chain

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \mathbf{J}$$

- Eigenvalue: $\lambda = 2 \implies \text{AM} = 3, \text{GM} = 1$
- One Jordan chain of length $r = 3$
- Chain relations: $(\mathbf{A} - 2\mathbf{I})\mathbf{v}_1 = 0, \quad (\mathbf{A} - 2\mathbf{I})\mathbf{v}_2 = \mathbf{v}_1, \quad (\mathbf{A} - 2\mathbf{I})\mathbf{v}_3 = \mathbf{v}_2$

Example 3: Two Jordan Chains

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

- Eigenvalue: $\lambda = 3 \implies \text{AM} = 3, \text{ GM} = 2$
- Two Jordan chains: $(\mathbf{A} - 3\mathbf{I})\mathbf{v}_1 = 0, \quad (\mathbf{A} - 3\mathbf{I})\mathbf{v}_2 = \mathbf{v}_1, \quad (\mathbf{A} - 3\mathbf{I})\mathbf{v}_3 = 0.$
- Chain lengths: $r_1 = 2,\ r_2 = 1 \Rightarrow \text{AM} = r_1 + r_2 = 3$

Summary Table

Eigenvalue	AM	GM	Jordan Block Sizes
$\lambda = 2$ (Ex. 1)	2	2	$1 + 1$
$\lambda = 2$ (Ex. 2)	3	1	3
$\lambda = 3$ (Ex. 3)	3	2	$2 + 1$

$$\boxed{\text{AM} = \sum_{i=1}^{\text{GM}} r_i, \quad \text{GM} = \text{number of of Jordan blocks}}$$

Jordan Chains

Definition

For an eigenvalue λ_0 of $\mathbf{A}$, a **Jordan chain of length** r is a sequence of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ satisfying

$$\mathbf{A}\mathbf{v}_i = \begin{cases} \lambda_0 \mathbf{v}_i, & i = 1, \\ \lambda_0 \mathbf{v}_i + \mathbf{v}_{i-1}, & i > 1, \end{cases}$$

where $\mathbf{v}_1$ is the eigenvector and $\{\mathbf{v}_2, \ldots, \mathbf{v}_r\}$ are the generalized eigenvectors.

Action of $\mathbf{A}$ *on this chain:*

$$(\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{v}_1 = \mathbf{0}, \quad (\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{v}_{k+1} = \mathbf{v}_k, \; k = 1, \ldots, r-1.$$

Computing Generalized Eigenvectors

Consider $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 2 \\ 0 & -1 & 1 \end{bmatrix}$.

1. Clearly, we have eigenvalue $\lambda = 1$ and eigenvector
$$\mathbf{v}_1 = (-2, 0, 1)^\mathsf{T}.$$
2. Solve $(\mathbf{A} - \mathbf{I})\mathbf{v}_2 = \mathbf{v}_1$ to get
$$\mathbf{v}_2 = (0, -1, 0)^\mathsf{T}.$$
3. Solve $(\mathbf{A} - \mathbf{I})\mathbf{v}_3 = \mathbf{v}_2$ to get
$$\mathbf{v}_3 = (-1, 0, 0)^\mathsf{T}.$$

Relationship among AM, GM, and Chain Length

- For each eigenvalue λ:

$$\sum_{i=1}^{\mathrm{GM}(\lambda)} r_i = \mathrm{AM}(\lambda)$$

- There are $\mathrm{GM}(\lambda)$ independent Jordan chains.
- Each chain length r_i equals the size of each Jordan block.
- The longest chain corresponds to the highest order of generalized eigenvectors.

Nullity Ladder and Block Sizes

Theorem

Let $\mathbf{A}$ *have Jordan blocks for eigenvalue* λ *of sizes* $r_1, r_2, \ldots, r_m$*. Then for all* $k \geq 1$*,*

$$\dim \mathcal{N}\big((\lambda \mathbf{I} - \mathbf{A})^k\big) = \sum_{i=1}^{m} \min(k, r_i).$$

Intuition: At step k, each Jordan block of length $\geq k$ contributes exactly one new independent generalized eigenvector. The nullity “climb” one step along each chain until reaching its maximum length.

Nullity Ladder: Concrete Example

Consider eigenvalue λ with three Jordan blocks of sizes $r_1 = 4, r_2 = 2, r_3 = 1$.

For each k, the nullity is

$$d_k = \dim \mathcal{N}\big((\lambda \mathbf{I} - \mathbf{A})^k\big) = \min(k, 4) + \min(k, 2) + \min(k, 1).$$

k	$\min(k, 4)$	$\min(k, 2)$	$\min(k, 1)$	d_k	$d_k - d_{k-1}$	$\#$ blocks of size $\geq k$
1	1	1	1	3	–	3
2	2	2	1	5	2	2
3	3	2	1	6	1	1
4	4	2	1	7	1	1
5	4	2	1	7	0	0

These increments are exactly what determines the block sizes.

Nullity Ladder Theorem: Succinct Proof

Let the Jordan blocks for eigenvalue λ have sizes $r_1, r_2, \ldots, r_m$.

Each block is $\mathbf{J}_{r_i}(\lambda) = \lambda \mathbf{I} + \mathbf{N}_i, \quad \mathbf{N}_i^{r_i} = \mathbf{0}$.

Claim:

$$\dim \mathcal{N}(\mathbf{N}_i^k) = \min(k, r_i).$$

Reason: $\mathbf{N}_i$ shifts coordinates upward; $\mathbf{N}_i^k$ kills exactly the last $\min(k, r_i)$ basis vectors in the Jordan chain.

Since the Jordan form is block diagonal,

$$\dim \mathcal{N}\big((\lambda \mathbf{I} - \mathbf{A})^k\big) = \sum_{i=1}^{m} \min(k, r_i).$$

This completes the proof.

Generalized Eigenvector

Definition

A nonzero vector $\mathbf{v}$ is a **generalized eigenvector** of $\mathbf{A}$ associated with eigenvalue λ if

$$(\mathbf{A} - \lambda\mathbf{I})^k\mathbf{v} = \mathbf{0} \quad \text{for some integer } k \geq 1.$$

- If $k = 1$, $\mathbf{v}$ is an ordinary eigenvector
- If $k > 1$, $\mathbf{v}$ is a higher-order generalized eigenvector.
- Accordingly, $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})^k$ is the eigenspace and generalized eigenspaces.

Jordan Chain vis-a-vis Generalized Eigenvectors

Jordan chain $\Rightarrow$ generalized eigenvector, but not vice versa.

- Every vector in a Jordan chain satisfies $(\mathbf{A} - \lambda\mathbf{I})^k \mathbf{v} = \mathbf{0}$ for some $k \in \mathbb{Z}_+$.
- However, not every generalized vector automatically forms a Jordan chain: the chain requires the linking relations $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v}_{i+1} = \mathbf{v}_i$.

An example

Consider $\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Clearly, canonical vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are all generalized eigenvectors (since $(\mathbf{A} - 2\mathbf{I})^2 = \mathbf{0}$). But only $\{\mathbf{e}_1, \mathbf{e}_2\}$ form a Jordan chain while $\mathbf{e}_3$ belongs to a different Jordan block $\rightarrow$ not part of the same chain.

Generalized Eigenspace Decomposition Theorem

Theorem

Let $\mathbf{A}$ *be an* $n \times n$ *matrix over* $\mathbb{C}$ *(or* $\mathbb{R}$*). For each eigenvalue* λ *of* $\mathbf{A}$ *with algebraic multiplicity* k*, the generalized eigenspace*

$$\mathcal{G}_\lambda = \mathcal{N}\left((\mathbf{A} - \lambda \mathbf{I})^k\right)$$

has dimension k*.*

Corollary

The space $\mathbb{C}^n$ *decomposes as a direct sum of generalized eigenspaces:*

$$\mathbb{C}^n = \bigoplus_\lambda \mathcal{G}_\lambda.$$

Consequently, $\mathbb{C}^n$ *has a basis consisting entirely of generalized eigenvectors of* $\mathbf{A}$*.*

Matrix Exponential via JCF

If $\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$ then $e^{t\mathbf{A}} = \mathbf{V}\, e^{t\mathbf{J}}\, \mathbf{V}^{-1}$ with

$$
\begin{aligned}
e^{t\mathbf{J}_i} &= e^{\lambda_i t}\left(\mathbf{I} + t\mathbf{N} + \frac{t^2}{2!}\mathbf{N}^2 + \cdots + \frac{t^{r_i-1}}{(r_i-1)!}\mathbf{N}^{r_i-1}\right) \\
&= e^{\lambda_i t}\begin{bmatrix} 1 & t & \cdots & \dfrac{t^{r_i-1}}{(r_i-1)!} \\ & 1 & \cdots & \dfrac{t^{r_i-2}}{(r_i-2)!} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}_{r_i\times r_i}
\end{aligned}
$$

where $\mathbf{N}$ is the nilpotent superdiagonal matrix ($\mathbf{N}^{r_i} = \mathbf{0}$).

Implication: each mode has the form $p(t)e^{\lambda t}$ with $\deg p(t) \le r_i - 1$.

Generalized Modes (1/2)

Consider $\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathbf{J} = \mathrm{diag}(\mathbf{J}_1, \dots, \mathbf{J}_q)$, where $\mathbf{V} = [\, \mathbf{V}_1 \ \mathbf{V}_2 \ \cdots \ \mathbf{V}_q \,]$ with $\mathbf{V}_i = [\, \mathbf{v}_{i1} \ \mathbf{v}_{i2} \ \cdots \ \mathbf{v}_{ir_i} \,] \in \mathbb{C}^{n \times r_i}$ containing the columns of $\mathbf{V}$ associated with the ith Jordan block $\mathbf{J}_i$.

Suppose $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ with $\mathbf{x}(0) = a_1\mathbf{v}_{i1} + \cdots + a_{r_i}\mathbf{v}_{ir_i} = \mathbf{V}_i\, \mathbf{a}$.
Then the solution is

$$\mathbf{x}(t) = \mathbf{V}\, e^{t\mathbf{J}}\tilde{\mathbf{x}}(0) = \mathbf{V}_i\, e^{t\mathbf{J}_i}\mathbf{a}.$$

- Trajectory stays in the span of generalized eigenvectors.
- Coefficients have the form $p(t)e^{\lambda_i t}$, where $p(t)$ is a polynomial.
- Such solutions are called *generalized modes* of the system.

Generalized Modes (2/2)

For a general initial condition $\mathbf{x}(0)$:

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}(0) = \mathbf{V}e^{t\mathbf{J}}\mathbf{V}^{-1}\mathbf{x}(0) = \sum_{i=1}^{q} \mathbf{V}_i e^{t\mathbf{J}_i}\left(\mathbf{S}_i^\top \mathbf{x}(0)\right),$$

where

$$\mathbf{V}^{-1} = \begin{bmatrix} \mathbf{S}_1^\top \\ \vdots \\ \mathbf{S}_q^\top \end{bmatrix}.$$

Hence, all solutions of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ are linear combinations of those generalized modes.

Powers of a Jordan Block

For a single Jordan block $\mathbf{J} = \lambda\mathbf{I} + \mathbf{N}$ with $\mathbf{N}^s = \mathbf{0}$,

$$\mathbf{J}^k = \sum_{j=0}^{s-1} \binom{k}{j} \lambda^{k-j} \mathbf{N}^j = \begin{bmatrix} \lambda^k & \binom{k}{1}\lambda^{k-1} & \cdots & \binom{k}{s-1}\lambda^{k-s+1} \\ & \lambda^k & \ddots & \vdots \\ & & \ddots & \binom{k}{1}\lambda^{k-1} \\ 0 & & & \lambda^k \end{bmatrix}.$$

Useful for discrete-time stability and closed-form recurrences.

For any matrix norm $\|\cdot\|$, one can find $C > 0$ (depending on $\mathbf{N}$ and s) such that

$$\|\mathbf{J}^k(\lambda)\| \ \le \ C \sum_{j=0}^{s-1} \binom{k}{j} |\lambda|^{k-j} \ \le \ C' k^{s-1} |\lambda|^k$$

for some constant $C' > 0$ (use that $\binom{k}{j}$ grows polynomially in k, roughly like k^j).

Thus, if every Jordan block $\mathbf{J}(\lambda_i)$ has eigenvalue $|\lambda_i| < 1$, we get $\mathbf{J}^k(\lambda_i) \ \to \ \mathbf{0}$ as $k \ \to \ \infty$.

Stability Criteria

- We say the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ is *stable* if $e^{t\mathbf{A}} \to \mathbf{0}$ as $t \to \infty$.
  - The state $\mathbf{x}(t)$ converges to $\mathbf{0}$ as $t \to \infty$, regardless of the initial condition $\mathbf{x}(0)$.
  - All trajectories of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ converge to $\mathbf{0}$ as $t \to \infty$.
- The system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ is stable if and only if all eigenvalues of $\mathbf{A}$ have negative real parts:

$$\boxed{\mathrm{Re}(\lambda_i) < 0, \quad i = 1, \ldots, n.}$$

- The discrete–time system $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k)$ is stable (i.e., $\mathbf{A}^k \to \mathbf{0}$ as $k \to \infty$) if and only if all eigenvalues of $\mathbf{A}$ satisfy

$$\boxed{|\lambda_i| < 1, \quad i = 1, \ldots, n.}$$

- Jordan blocks yield terms of form $t^p e^{\lambda t}$ (or $k^p \lambda^k$ for a discrete-time system).

Cayley–Hamilton Theorem

Theorem (Matrix satisfies its own characteristic equation)

For any $\mathbf{A} \in \mathbb{F}^{n \times n}$ *with characteristic polynomial*

$$p_{\mathbf{A}}(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0,$$

we have

$$p_{\mathbf{A}}(\mathbf{A}) = \mathbf{A}^n + c_{n-1}\mathbf{A}^{n-1} + \cdots + c_0\mathbf{I} = \mathbf{0}.$$

C-H Theorem: Concrete Example

As a concrete example, let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Its characteristic polynomial is

$$p(\lambda) = \det(\lambda \mathbf{I}_2 - \mathbf{A}) = \det \begin{pmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 4 \end{pmatrix} = \lambda^2 - 5\lambda - 2.$$

We can verify directly:

$$\mathbf{A}^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad 5\mathbf{A} = \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix}, \quad 2\mathbf{I}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

$$\mathbf{A}^2 - 5\mathbf{A} - 2\mathbf{I}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Proof (via Jordan Form) of C-H Theorem

Idea.

1. Every $\mathbf{A}$ is similar to its Jordan form: $\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$.
2. Similar matrices share the same characteristic polynomial and satisfy
$$p_{\mathbf{A}}(\mathbf{A}) = \mathbf{0} \iff p_{\mathbf{A}}(\mathbf{J}) = \mathbf{0}.$$
3. Each Jordan block $\mathbf{J}_k(\lambda) = \lambda\mathbf{I} + \mathbf{N}$, where $\mathbf{N}$ is nilpotent: $\mathbf{N}^k = \mathbf{0}$.
4. Since $(x-\lambda)^k$ divides $p_{\mathbf{A}}(x)$, we can write $p_{\mathbf{A}}(x) = (x-\lambda)^k q(x)$, hence
$$p_{\mathbf{A}}(\mathbf{J}_k(\lambda)) = (\mathbf{J}_k - \lambda\mathbf{I})^k q(\mathbf{J}_k) = \mathbf{N}^k q(\mathbf{J}_k) = \mathbf{0}.$$
5. Therefore $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{P}\, p_{\mathbf{A}}(\mathbf{J})\, \mathbf{P}^{-1} = \mathbf{0}$.

$\square$

Alternative Proof (via Adjugate Matrix)

Key Idea.

For any matrix $\mathbf{M}$:

$$\mathbf{M}\,\mathrm{adj}(\mathbf{M}) = \det(\mathbf{M})\,\mathbf{I}.$$

Let $\mathbf{M} = \lambda\mathbf{I} - \mathbf{A}$. Then

$$(\lambda\mathbf{I} - \mathbf{A})\,\mathrm{adj}(\lambda\mathbf{I} - \mathbf{A}) = p_{\mathbf{A}}(\lambda)\mathbf{I}. \qquad (\star)$$

Because $\mathrm{adj}(\lambda\mathbf{I} - \mathbf{A})$ is a polynomial in λ, we can safely substitute $\lambda \mapsto \mathbf{A}$ in $(\star)$:

$$(\mathbf{A} - \mathbf{A})\,\mathrm{adj}(\mathbf{A} - \mathbf{A}) = p_{\mathbf{A}}(\mathbf{A})\,\mathbf{I}.$$

Hence $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$. $\square$

Corollary of C–H Theorem

Corollary

For every $p \in \mathbb{Z}_+$, we have

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

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

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

Implication: The C–H theorem reduces all higher powers of $\mathbf{A}$ to a finite basis $\{\mathbf{I}, \mathbf{A}, \mathbf{A}^2, \ldots, \mathbf{A}^{n-1}\}$. This enables compact expressions for $e^{t\mathbf{A}}$, simplifies analysis, and underlies controller and observer design in linear dynamical systems.

Proof of the Corollary

By Cayley–Hamilton, we have

$$\mathbf{A}^n = -c_{n-1}\mathbf{A}^{n-1} - \cdots - c_1\mathbf{A} - c_0\mathbf{I} \tag{$\star$}$$

- Multiply both sides of ($\star$) by $\mathbf{A}^{k-n}$ gives:

  $$\boxed{\mathbf{A}^k = -c_{n-1}\mathbf{A}^{k-1} - c_{n-2}\mathbf{A}^{k-2} - \cdots - c_1\mathbf{A}^{k-n+1} - c_0\mathbf{A}^{k-n}, \quad k > n.}$$

  Thus, every $\mathbf{A}^k$ with $k > n$ can be written as a linear combination of $\{\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}\}$, where the linear combination coefficients can be found recursively from the characteristic polynomial coefficients.

- Multiply both sides of ($\star$) by $\mathbf{A}^{-1}$ and rearrange, we get

  $$\boxed{\mathbf{A}^{-1} = -\frac{1}{c_0}\left(\mathbf{A}^{n-1} + c_{n-1}\mathbf{A}^{n-2} + \cdots + c_2\mathbf{A} + c_1\mathbf{I}\right).}$$

  Note that $c_0 \neq 0$ because the characteristic polynomial has no zero root.

Using C–H to Simplify Powers of a Matrix

As an example, for $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, the theorem gives $\mathbf{A}^2 = 5\mathbf{A} + 2\mathbf{I}_2$.

To compute $\mathbf{A}^4$, observe:

$$\mathbf{A}^3 = (5\mathbf{A} + 2\mathbf{I}_2)\mathbf{A} = 5\mathbf{A}^2 + 2\mathbf{A} = 5(5\mathbf{A} + 2\mathbf{I}_2) + 2\mathbf{A} = 27\mathbf{A} + 10\mathbf{I}_2,$$

$$\mathbf{A}^4 = \mathbf{A}^3\mathbf{A} = 27\mathbf{A}^2 + 10\mathbf{A} = 27(5\mathbf{A} + 2\mathbf{I}_2) + 10\mathbf{A} = 145\mathbf{A} + 54\mathbf{I}_2.$$

Likewise,

$$\mathbf{A}^{-1} = \tfrac{1}{2}(\mathbf{A} - 5\mathbf{I}_2),$$

$$\mathbf{A}^{-2} = \mathbf{A}^{-1}\mathbf{A}^{-1} = \tfrac{1}{4}(\mathbf{A}^2 - 10\mathbf{A} + 25\mathbf{I}_2) = \tfrac{1}{4}(-5\mathbf{A} + 27\mathbf{I}_2).$$

Key insight: any power $\mathbf{A}^k$ can be written as a polynomial in $\mathbf{A} \in \mathbb{R}^{n\times n}$ of degree at most $n-1$. This is a direct application of the C–H theorem.

Matrix Functions via C–H Theorem

Given an analytic function

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

and the characteristic polynomial $p(x)$ of degree n of an $n \times n$ matrix $\mathbf{A}$, long division gives

$$f(x) = q(x)p(x) + r(x),$$

where $q(x)$ is some quotient polynomial and the remainder polynomial

$$r(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}.$$

By the C–H theorem $p(\mathbf{A}) = \mathbf{0}$, so $f(\mathbf{A}) = r(\mathbf{A})$.

Matrix Functions via C–H Theorem

Thus, $f(\mathbf{A})$ is always expressible as a matrix polynomial

$$f(\mathbf{A}) = c_0\mathbf{I} + c_1\mathbf{A} + \cdots + c_{n-1}\mathbf{A}^{n-1}.$$

To determine c_k, evaluating at the n eigenvalues λ_i of $\mathbf{A}$ gives

$$f(\lambda_i) = r(\lambda_i) = c_0 + c_1\lambda_i + \cdots + c_{n-1}\lambda_i^{n-1}, \quad i = 1, \ldots, n.$$

This yields a linear system for $c_0, \ldots, c_{n-1}$.

Matrix Functions with Repeated Eigenvalues

When eigenvalues repeat (e.g. $\lambda_i = \lambda_j$), the linear equations for c_k are no longer unique.

For an eigenvalue λ with multiplicity m, the first $m-1$ derivatives of f and r must match:

$$\left.\frac{d^k f(x)}{dx^k}\right|_{x=\lambda} = \left.\frac{d^k r(x)}{dx^k}\right|_{x=\lambda}, \qquad k = 1, \ldots, m-1.$$

These derivative conditions, together with $f(\lambda) = r(\lambda)$, provide the full set of n equations for determining $c_0, \ldots, c_{n-1}$.

Interpretation: finding $r(x)$ is an interpolation problem at points $(\lambda_i, f(\lambda_i))$, solvable using Lagrange or Newton interpolation, leading to Sylvester's formula.

Example: Polynomial Form of $e^{\mathbf{A}t}$

Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \qquad p(x) = (x-1)(x-3) = x^2 - 4x + 3.$$

Let $r(x) = c_0 + c_1 x$. At the eigenvalues $\lambda = 1, 3$:

$$e^t = c_0 + c_1, \qquad e^{3t} = c_0 + 3c_1.$$

Solving:

$$c_0 = \tfrac{1}{2}(3e^t - e^{3t}), \qquad c_1 = \tfrac{1}{2}(e^{3t} - e^t).$$

Thus

$$e^{\mathbf{A}t} = c_0 \mathbf{I}_2 + c_1 \mathbf{A} = \begin{pmatrix} e^t & e^{3t} - e^t \\ 0 & e^{3t} \end{pmatrix}.$$

Example: $e^{\mathbf{A}t}$ for a Skew-Symmetric Matrix

Let

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad p(x) = x^2 + 1, \quad \lambda = \pm \mathrm{j}.$$

Evaluate $r(x) = c_0 + c_1 x$ at $\lambda = \mathrm{j}, -\mathrm{j}$:

$$e^{\mathrm{j}t} = c_0 + ic_1, \quad e^{-\mathrm{j}t} = c_0 - ic_1 \implies c_0 = \cos t, \quad c_1 = \sin t.$$

Thus

$$e^{\mathbf{A}t} = (\cos t)\mathbf{I}_2 + (\sin t)\mathbf{A} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix},$$

which is a rotation matrix: rotated clockwise with angle t.

Example: Polynomial Form of $\sin(\mathbf{A}t)$

Using the same $\mathbf{A}$ and $r(x) = c_0 + c_1 x$, evaluate:

$$\sin t = c_0 + c_1, \qquad \sin 3t = c_0 + 3c_1.$$

Solving:

$$c_0 = \tfrac{3\sin t - \sin 3t}{2}, \qquad c_1 = \tfrac{\sin 3t - \sin t}{2}.$$

Thus

$$\sin(\mathbf{A}t) = c_0 \mathbf{I}_2 + c_1 \mathbf{A} = \begin{pmatrix} \sin t & \sin 3t - \sin t \\ 0 & \sin 3t \end{pmatrix}.$$

Computing the Matrix Function $\sin(\mathbf{A}t)$

Compute $\sin(\mathbf{A}t)$ for a square matrix $\mathbf{A}$.

- **1. Eigen-decomposition (if diagonalizable)**

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}, \qquad \sin(\mathbf{A}t) = \mathbf{V}\,\sin(\boldsymbol{\Lambda}t)\,\mathbf{V}^{-1}.$$

  Diagonal: $\sin(\boldsymbol{\Lambda}t) = \mathrm{diag}(\sin(\lambda_i t))$.

- **2. Jordan form (general theory)**

$$\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}, \qquad \sin(\mathbf{A}t) = \mathbf{V}\,\sin(\mathbf{J}t)\,\mathbf{V}^{-1}.$$

  Each Jordan block uses derivatives of $\sin(\lambda t)$ times powers of $t\mathbf{N}$.

- **3. Exponential identity (always valid)**

$$\sin(\mathbf{A}t) = \frac{e^{j\mathbf{A}t} - e^{-j\mathbf{A}t}}{2j}.$$

  Compute two matrix exponentials and combine.

Caveat: Jordan Form is Ill-Conditioned (1/2)

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \quad \text{Jordan block for } \lambda = 1.$$

- $\mathbf{A}$ is defective (only one eigenvector). Its Jordan form is exact but **numerically fragile**.

Now perturb by a tiny ε:

$$\mathbf{A}_\varepsilon = \begin{bmatrix} 1 & 1 \\ 0 & 1+\varepsilon \end{bmatrix}.$$

- Eigenvalues: 1 and $1+\varepsilon$ (now distinct) $\Rightarrow$ diagonalizable.
- Eigenvectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix}.$$

Caveat: Jordan Form is Ill-Conditioned (2/2)

- Eigenvector matrix:

$$\mathbf{P}_\varepsilon = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}, \qquad \mathbf{A}_\varepsilon = \mathbf{P}_\varepsilon \begin{bmatrix} 1 & 0 \\ 0 & 1+\varepsilon \end{bmatrix} \mathbf{P}_\varepsilon^{-1}.$$

- Condition number:

$$\kappa(\mathbf{P}_\varepsilon) \sim \frac{1}{|\varepsilon|} \ \to \ \infty \text{ as } \varepsilon \to 0.$$

Interpretation:

- Tiny perturbations in **A** lead to huge changes in its eigenbasis.
- The Jordan structure is **structurally unstable**, which is never used numerically.
- Instead, use the **Schur form** (unitary, well-conditioned).

JCF Summary

- Diagonalization fails when AM $>$ GM.
- Jordan form provides a canonical structure for every square matrix.
- Diagonalizable case: all Jordan blocks are size 1.
- Defective case: at least one Jordan block of size ≥ 2.

- JCF characterizes nondiagonalizable structure and yields closed forms for $e^{t\mathbf{A}}$ and $\mathbf{A}^k$.
- Stability depends on eigenvalues; Jordan blocks add polynomial factors.

Matrix Factorizations: Summary and Applications (1/2)

- **QR / RRQR:** Core for least-squares, rank estimation, and feature/column selection.

$$\mathbf{A} = \mathbf{QR} \quad \text{or} \quad \mathbf{AP} = \mathbf{QR} \text{ (rank-revealing)}$$

- **LU / Cholesky:** Essential for solving linear systems.

$$\mathbf{A} = \mathbf{LU}, \quad \mathbf{A} = \mathbf{LL}^{\mathsf{T}} \text{ (for SPD matrices)}$$

  Fast and numerically efficient in simulation and optimization.

- **SVD:** Fundamental for energy, norm, and subspace analysis.

$$\mathbf{A} = \mathbf{U\Sigma V}^{\mathsf{H}}, \quad \mathbf{U}, \mathbf{V} \text{ unitary, } \mathbf{\Sigma} \text{ diagonal with singular values.}$$

  Applications: PCA, low-rank approximation, noise filtering.

Matrix Factorizations: Summary and Applications (2/2)

- **Eigendecomposition (for diagonalizable A):**

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}, \quad \boldsymbol{\Lambda} \text{ diagonal (eigenvalues)}$$

  Used in modal analysis, system stability, and diagonalization.

- **Jordan Form:** Theoretical and structural analysis; reveals eigenvalue multiplicity, generalized eigenvectors, and qualitative dynamics. Avoided in numerical computation due to ill-conditioning.

$$\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$$

- **Schur Form:** Numerically stable eigen-analysis; always exists for any square matrix.

$$\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^{\mathsf{H}}, \quad \mathbf{Q} \text{ unitary (orthonormal basis)}, \ \mathbf{T} \text{ upper triangular}$$