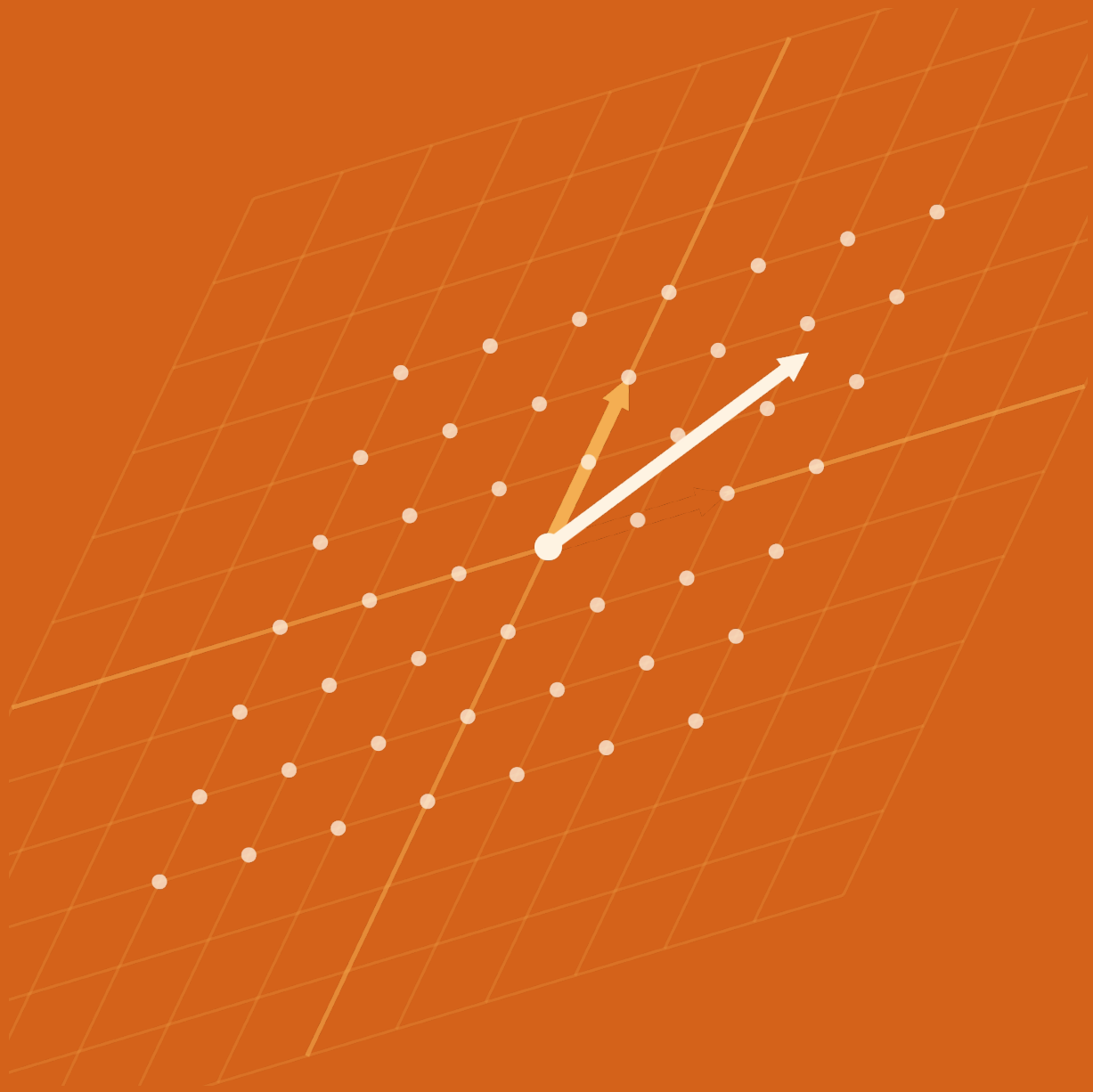
LINEAR ALGEBRA FOR AIML AND 6G

Cayley–Hamilton Theorem

Lecture 25

Prof. Aditya K. Jagannatham
IIT Kanpur



Outline

1. **Cayley–Hamilton theorem**
   Every matrix satisfies its characteristic equation
2. **Proof**
   Proof for a diagonalizable matrix
3. **Examples**
   Inverse and matrix polynomials from the theorem

01

Cayley–Hamilton theorem

Every matrix satisfies its characteristic equation



Cayley–Hamilton theorem

- The **Cayley–Hamilton (CH)** theorem can be stated as follows
- Consider the **characteristic polynomial**

$$p(\lambda) = |\mathbf{A} - \lambda \mathbf{I}| = a_0 + a_1\lambda + a_2\lambda^2 + \ldots + a_n\lambda^n$$

Cayley–Hamilton theorem

- Let λ_i denote the **eigenvalues**. We know

$$p(\lambda_i) = a_0 + a_1\lambda_i + a_2\lambda_i^2 + \ldots + a_n\lambda_i^n = 0$$

Cayley–Hamilton theorem

- The **Cayley–Hamilton theorem** states that

$$p(\mathbf{A}) = \mathbf{0}$$

$$\Rightarrow a_0\mathbf{I} + a_1\mathbf{A} + a_2\mathbf{A}^2 + \ldots + a_n\mathbf{A}^n = \mathbf{0}$$

02

Proof

Proof for a diagonalizable matrix



CH theorem – proof

- The **Cayley–Hamilton theorem** can be proved as follows
- For simplicity, consider a **diagonalizable matrix**

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$$

- It is easy to see that

$$\begin{aligned}\mathbf{A}^n &= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} \times \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} \times \ldots \times \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} \\ &= \mathbf{U}\mathbf{\Lambda}^n\mathbf{U}^{-1}\end{aligned}$$

CH theorem – proof

- It therefore follows that

$$
\begin{aligned}
& p\left(\mathbf{A}\right) \\
&= a_0\mathbf{U}\mathbf{U}^{-1} + a_1\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} + a_2\mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^{-1} + \ldots + a_n\mathbf{U}\mathbf{\Lambda}^n\mathbf{U}^{-1} \\
&= \mathbf{U}\left(a_0\mathbf{\Lambda}^0 + a_1\mathbf{\Lambda} + \ldots + a_n\mathbf{\Lambda}^n\right)\mathbf{U}^{-1}
\end{aligned}
$$

CH theorem – proof

- Consider now

$$a_0\mathbf{\Lambda}^0 + a_1\mathbf{\Lambda} + \ldots + a_n\mathbf{\Lambda}^n$$

- Note that the above matrix is **diagonal** with the ith diagonal element as

$$a_0\lambda_i^0 + a_1\lambda_i + a_2\lambda_i^2 + \ldots + a_n\lambda_i^n = 0$$

CH theorem – proof

- Therefore

$$a_0\mathbf{\Lambda}^0 + a_1\mathbf{\Lambda} + \ldots + a_n\mathbf{\Lambda}^n = \mathbf{0}$$

$$\Rightarrow \mathbf{U}\left(a_0\mathbf{\Lambda}^0 + a_1\mathbf{\Lambda} + \ldots + a_n\mathbf{\Lambda}^n\right)\mathbf{U}^{-1} = \mathbf{0}$$

$$\Rightarrow p\left(\mathbf{A}\right) = \mathbf{0}$$

03

Examples

Inverse and matrix polynomials from the theorem



CH – example

- Consider the matrix below and verify the **CH theorem**

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

CH – example

- The **characteristic equation** can be found as

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} 2-\lambda & 3 \\ 4 & 1-\lambda \end{vmatrix}$$

$$= (2-\lambda)(1-\lambda) - 12$$

$$\Rightarrow p(\lambda) = \lambda^2 - 3\lambda - 10$$

CH – example

- Now check

$$p(\mathbf{A}) = \mathbf{A}^2 - 3\mathbf{A} - 10\mathbf{I}$$

$$= \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} - 3 \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} - 10 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 16 & 9 \\ 12 & 13 \end{bmatrix} - \begin{bmatrix} 6 & 9 \\ 12 & 3 \end{bmatrix} - \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = \mathbf{0}$$

CH – example

- Compute $\mathbf{A}^{-1}$ using the **CH theorem**

$$\mathbf{A}^2 - 3\mathbf{A} - 10\mathbf{I} = \mathbf{0}$$

$$\Rightarrow \mathbf{A}^2 - 3\mathbf{A} = 10\mathbf{I}$$

$$\Rightarrow \mathbf{A}^{-1}\left(\mathbf{A}^2 - 3\mathbf{A}\right) = 10\mathbf{A}^{-1}$$

$$\Rightarrow \mathbf{A} - 3\mathbf{I} = 10\mathbf{A}^{-1}$$

$$\Rightarrow \mathbf{A}^{-1} = \frac{1}{10}\left(\mathbf{A} - 3\mathbf{I}\right)$$

CH – example

- Continuing

$$\mathbf{A}^{-1} = \frac{1}{10}(\mathbf{A} - 3\mathbf{I})$$

$$= \frac{1}{10}\left(\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} - 3\mathbf{I}\right)$$

$$= \frac{1}{10}\begin{bmatrix} -1 & 3 \\ 4 & -2 \end{bmatrix}$$

CH – example

- Evaluate

$$f(\mathbf{A}) = \mathbf{A}^5 - 3\mathbf{A}^4 - 9\mathbf{A}^3 - 5\mathbf{A}^2 - 2\mathbf{A} + 23\mathbf{I}$$

CH – example

- Consider the polynomial

$$f(\lambda) = \lambda^5 - 3\lambda^4 - 9\lambda^3 - 5\lambda^2 - 2\lambda + 23$$

- Divide this by $p(\lambda) = \lambda^2 - 3\lambda - 10$ to obtain

$$\begin{aligned} f(\lambda) &= \left(\lambda^2 - 3\lambda - 10\right)\left(\lambda^3 + \lambda - 2\right) + 2\lambda + 3 \\ &= p(\lambda)\left(\lambda^3 + \lambda - 2\right) + 2\lambda + 3 \end{aligned}$$

CH – example

- Therefore, one can write $f(\mathbf{A})$ as

$$f(\mathbf{A}) = \underbrace{\left(\mathbf{A}^2 - 3\mathbf{A} - 10\mathbf{I}\right)}_{p(\mathbf{A})=\mathbf{0}} \left(\mathbf{A}^3 + \mathbf{A} - 2\mathbf{I}\right) + 2\mathbf{A} + 3\mathbf{I}$$

$$= 2\mathbf{A} + 3\mathbf{I}$$

$$= 2\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} + 3\mathbf{I} = \begin{bmatrix} 7 & 6 \\ 8 & 5 \end{bmatrix}$$

Summary

- The characteristic polynomial is $p(\lambda) = |\mathbf{A} - \lambda \mathbf{I}|$
- **Cayley–Hamilton**: every matrix satisfies $p(\mathbf{A}) = \mathbf{0}$
- For diagonalizable $\mathbf{A}$, $p(\mathbf{A}) = \mathbf{U}\text{diag}(p(\lambda_i))\mathbf{U}^{-1} = \mathbf{0}$
- The theorem gives $\mathbf{A}^{-1}$ as a polynomial in $\mathbf{A}$
- Any polynomial $f(\mathbf{A})$ reduces to its remainder after dividing $f(\lambda)$ by $p(\lambda)$



Thank You!

Questions?