Cayley-Hamilton theorem

Theorem: For any matrix $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ with characteristic polynomial $p_{\boldsymbol{A}}(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_2 t^2 + a_1 t + a_0$ it holds that:

$$p_{\boldsymbol{A}}(\boldsymbol{A}) = (-1)^n \boldsymbol{A}^n + a_{n-1}\boldsymbol{A}^{n-1} + \cdots + a_2\boldsymbol{A}^2 + a_1\boldsymbol{A} + a_0\mathbf{I}_n = \mathbf{0}_n$$

Here $\mathbf{0}_n$ is the zero square matrix of order n.

Example:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 3 \\ 1 & -2 & 2 \end{pmatrix}, \qquad p_{\boldsymbol{A}}(t) = -t^3 + 2t^2 + t - 2$$

$$p_{\boldsymbol{A}}(\boldsymbol{A}) = -\begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 3 \\ 1 & -2 & 2 \end{pmatrix}^3 + 2\begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 3 \\ 1 & -2 & 2 \end{pmatrix}^2 + \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 3 \\ 1 & -2 & 2 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$-\begin{pmatrix} 13 & 2 & 12 \\ 9 & -1 & 9 \\ -5 & -2 & -4 \end{pmatrix} + 2\begin{pmatrix} 7 & 0 & 6 \\ 3 & 1 & 3 \\ -3 & 0 & -2 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 3 \\ 1 & -2 & 2 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Cayley-Hamilton theorem

Theorem: For any matrix $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ with characteristic polynomial $p_{\boldsymbol{A}}(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_2 t^2 + a_1 t + a_0$ it holds that:
$p_{\boldsymbol{A}}(\boldsymbol{A}) = (-1)^n \boldsymbol{A}^n + a_{n-1}\boldsymbol{A}^{n-1} + \cdots + a_2\boldsymbol{A}^2 + a_1\boldsymbol{A} + a_0\mathbf{I}_n = \mathbf{0}_n$

Proof: We use the fact that $\boldsymbol{M} \cdot \operatorname{adj}(\boldsymbol{M}) = \det(\boldsymbol{M})\mathbf{I}_n$ for $\boldsymbol{M} = \boldsymbol{A} - t\mathbf{I}_n$.

Entries of $\operatorname{adj}(\boldsymbol{A} - t\mathbf{I}_n)$ are determinants of its submatrices, i.e. polynomials in t of degree at most $n-1$. Hence we can write:
$\operatorname{adj}(\boldsymbol{A} - t\mathbf{I}_n) = t^{n-1}\boldsymbol{B}_{n-1} + \cdots + t\boldsymbol{B}_1 + \boldsymbol{B}_0$ for $\boldsymbol{B}_{n-1}, \ldots, \boldsymbol{B}_0 \in \mathbb{K}^{n\times n}$

Example:

$$\operatorname{adj}(\boldsymbol{A} - t\mathbf{I}_3) = \operatorname{adj}\begin{pmatrix} 1-t & 2 & 0 \\ 3 & -1-t & 3 \\ 1 & -2 & 2-t \end{pmatrix} = \begin{pmatrix} t^2 - t + 4 & 2t-4 & 6 \\ 3t-3 & t^2 - 3t + 2 & 3t-3 \\ t-5 & -2t+4 & t^2 - 7 \end{pmatrix}$$

$$= t^2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + t\begin{pmatrix} -1 & 2 & 0 \\ 3 & -3 & 3 \\ 1 & -2 & 0 \end{pmatrix} + \begin{pmatrix} 4 & -4 & 6 \\ -3 & 2 & -3 \\ -5 & 4 & -7 \end{pmatrix} = t^2\boldsymbol{B}_2 + t\boldsymbol{B}_1 + \boldsymbol{B}_0$$

Cayley-Hamilton theorem

Theorem: For any matrix $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ with characteristic polynomial $p_{\boldsymbol{A}}(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_2 t^2 + a_1 t + a_0$ it holds that:
$p_{\boldsymbol{A}}(\boldsymbol{A}) = (-1)^n \boldsymbol{A}^n + a_{n-1}\boldsymbol{A}^{n-1} + \cdots + a_2\boldsymbol{A}^2 + a_1\boldsymbol{A} + a_0\mathbf{I}_n = \mathbf{0}_n$

Proof: We use the fact that $\boldsymbol{M} \cdot \operatorname{adj}(\boldsymbol{M}) = \det(\boldsymbol{M})\mathbf{I}_n$ for $\boldsymbol{M} = \boldsymbol{A} - t\mathbf{I}_n$.

Entries of $\operatorname{adj}(\boldsymbol{A} - t\mathbf{I}_n)$ are determinants of its submatrices, i.e. polynomials in t of degree at most $n-1$. Hence we can write:
$\operatorname{adj}(\boldsymbol{A} - t\mathbf{I}_n) = t^{n-1}\boldsymbol{B}_{n-1} + \cdots + t\boldsymbol{B}_1 + \boldsymbol{B}_0$ for $\boldsymbol{B}_{n-1}, \ldots, \boldsymbol{B}_0 \in \mathbb{K}^{n\times n}$

Now we got: $(\boldsymbol{A} - t\mathbf{I}_n)(t^{n-1}\boldsymbol{B}_{n-1} + \cdots + t\boldsymbol{B}_1 + \boldsymbol{B}_0) = p_{\boldsymbol{A}}(t)\mathbf{I}_n =$
$= (-1)^n t^n\mathbf{I}_n + a_{n-1}t^{n-1}\mathbf{I}_n + \cdots + a_2 t^2\mathbf{I}_n + a_1 t\mathbf{I}_n + a_0\mathbf{I}_n$

Example:

$$\begin{pmatrix} 1-t & 2 & 0 \\ 3 & -1-t & 3 \\ 1 & -2 & 2-t \end{pmatrix} \left(t^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + t \begin{pmatrix} -1 & 2 & 0 \\ 3 & -3 & 3 \\ 1 & -2 & 0 \end{pmatrix} + \begin{pmatrix} 4 & -4 & 6 \\ -3 & 2 & -3 \\ -5 & 4 & -7 \end{pmatrix} \right)$$

$$= (-t^3 + 2t^2 + t - 2)\mathbf{I}_3 = -t^3\mathbf{I}_3 + 2t^2\mathbf{I}_3 + t\mathbf{I}_3 - 2\mathbf{I}_3$$

Cayley-Hamilton theorem

Theorem: For any matrix $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ with characteristic polynomial $p_{\boldsymbol{A}}(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_2 t^2 + a_1 t + a_0$ it holds that:
$p_{\boldsymbol{A}}(\boldsymbol{A}) = (-1)^n \boldsymbol{A}^n + a_{n-1}\boldsymbol{A}^{n-1} + \cdots + a_2\boldsymbol{A}^2 + a_1\boldsymbol{A} + a_0\mathbf{I}_n = \mathbf{0}_n$

Proof: We use the fact that $\boldsymbol{M} \cdot \mathrm{adj}(\boldsymbol{M}) = \det(\boldsymbol{M})\mathbf{I}_n$ for $\boldsymbol{M} = \boldsymbol{A} - t\mathbf{I}_n$.

Entries of $\mathrm{adj}(\boldsymbol{A} - t\mathbf{I}_n)$ are determinants of its submatrices, i.e. polynomials in t of degree at most $n-1$. Hence we can write:
$\mathrm{adj}(\boldsymbol{A} - t\mathbf{I}_n) = t^{n-1}\boldsymbol{B}_{n-1} + \cdots + t\boldsymbol{B}_1 + \boldsymbol{B}_0$ for $\boldsymbol{B}_{n-1}, \ldots, \boldsymbol{B}_0 \in \mathbb{K}^{n\times n}$

Now we got: $(\boldsymbol{A} - t\mathbf{I}_n)(t^{n-1}\boldsymbol{B}_{n-1} + \cdots + t\boldsymbol{B}_1 + \boldsymbol{B}_0) = p_{\boldsymbol{A}}(t)\mathbf{I}_n =$
$= (-1)^n t^n \mathbf{I}_n + a_{n-1}t^{n-1}\mathbf{I}_n + \cdots + a_2 t^2\mathbf{I}_n + a_1 t\mathbf{I}_n + a_0\mathbf{I}_n$

coefficients by t^n:	$-\boldsymbol{B}_{n-1}$	$=$	$(-1)^n\mathbf{I}_n$	$\cdot\boldsymbol{A}^n$ from the left
coefficients by t^i:	$\boldsymbol{A}\boldsymbol{B}_i - \boldsymbol{B}_{i-1}$	$=$	$a_i\mathbf{I}_n$	$\cdot\boldsymbol{A}^i$ from the left
coefficients by t^0:	$\boldsymbol{A}\boldsymbol{B}_0$	$=$	$a_0\mathbf{I}_n$	leave as is and $\sum$ all

The left side:
$-\boldsymbol{A}^n\boldsymbol{B}_{n-1} + \boldsymbol{A}^{n-1}(\boldsymbol{A}\boldsymbol{B}_{n-1} - \boldsymbol{B}_{n-2}) + \cdots + \boldsymbol{A}(\boldsymbol{A}\boldsymbol{B}_1 - \boldsymbol{B}_0) + \boldsymbol{A}\boldsymbol{B}_0 = \mathbf{0}_n$
The right side: $(-1)^n\boldsymbol{A}^n + a_{n-1}\boldsymbol{A}^{n-1} + \cdots + a_2\boldsymbol{A}^2 + a_1\boldsymbol{A} + a_0\mathbf{I}_n = p_{\boldsymbol{A}}(\boldsymbol{A})$