Cayley-Hamilton Theorem

An n☓n matrix A satisfies its own characteristic equation

➔Every **square** matrix satisfies its own characteristic equation.

Example: Verify the Cayley-Hamilton theorem for the given matrix A

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

We get the characteristic equation

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = \lambda^2 - (a+d)\lambda + (ad - cb) = 0$$

and find

$$\begin{aligned} p(A) &= A^2 - (a+d)A + (ad-bc)I \\ &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 - (a+d)\begin{bmatrix} a & b \\ c & d \end{bmatrix} + (ad-bc)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{bmatrix} - \begin{bmatrix} a^2+ad & ab+bd \\ ac+cd & ad+d^2 \end{bmatrix} + \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} \\ &= \begin{bmatrix} a^2+bc-a^2-ad+ad-bc & ab+bd-ab-bd \\ ac+cd-ac-cd & bc+d^2-ad-d^2+ad-bc \end{bmatrix} \\ &= 0 \end{aligned}$$

Example:

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

characteristic equation ➔ $\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = \lambda^2 - 3\lambda + 1 = 0$

By Cayley-Hamilton theorem ➔ $A^2 - 3A + I = 0$

➔ $A^2 = 3A - I = 3\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$

Or directly $A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4+1 & 2+1 \\ 2+1 & 1+1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$

Example:

$A=\begin{pmatrix}-1 & 3\\ 2 & 4\end{pmatrix}$; $m=3$ (page 403, Problem 3)

characteristic equation ➔ $\det(A-\lambda I)=\begin{vmatrix}-1-\lambda & 3\\ 2 & 4-\lambda\end{vmatrix}=\lambda^2-3\lambda-10=0$, $\lambda=-2,\ 5$

By Cayley-Hamilton theorem ➔ $A^2-3A-10I=0$

Moreover, $A^m=c_0I+c_1A$ and $\lambda^m=c_0+c_1\lambda$, for the same pair of constants $c_0,\ c_1$

For $\lambda_1=-2$ ➔ $(-2)^m=c_0+c_1(-2)$

$\lambda_2=5$ ➔ $(5)^m=c_0+c_1(5)$

➔ $c_1=\dfrac{5^m-(-2)^m}{7}$, $c_0=\dfrac{2(5^m)+5(-2)^m}{7}$

$A^m=c_0I+c_1A$

➔ $=\dfrac{2(5)^m+5(-2)^m}{7}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\dfrac{(5)^m-(-2)^m}{7}\begin{bmatrix}-1 & 3\\ 2 & 4\end{bmatrix}$

$=\begin{bmatrix}\dfrac{5^m+6(-2)^m}{7} & \dfrac{3(5)^m-3(-2)^m}{7}\\ \dfrac{2(5)^m-2(-2)^m}{7} & \dfrac{6(5)^m+(-2)^m}{7}\end{bmatrix}$

➔ $A^3=\begin{bmatrix}11 & 57\\ 38 & 106\end{bmatrix}$

Example:

Show that the given matrix has an eigenvalue λ_1 of multiplicity two. As a consequence, the equations $\lambda^m=c_0+c_1\lambda$ does not yield enough independent equations to form a system for determining the coefficients c_0, c_1. Use the derivative (with respect to λ) of the equation evaluated at λ_1 as the extra needed equation to form a system. Compute A^m and use this result to compute the indicated power of the matrix A

$A=\begin{pmatrix}7 & 3\\ -3 & 1\end{pmatrix}$; $m=6$ (page 403, Problem 11)

$\det(A-\lambda I)=\begin{vmatrix}7-\lambda & 3\\ -3 & 1-\lambda\end{vmatrix}=\lambda^2-8\lambda+16=0$

➔ $\lambda=4,\ 4$

➔ $A^2 - 8A + 16I = 0$

➔ $A^m = c_0 I + c_1 A$ and $\lambda^m = c_0 + c_1 \lambda$

For $\lambda = 4$ ➔ $4^m = c_0 + c_1 4$

On the other hand, take the derivative (with respect to λ) of the equation evaluated at λ_1

➔ $m\lambda^{m-1} = c_1$

➔ $c_1 = m\lambda^{m-1} = m4^{m-1}$, $c_0 = 4^m - c_1 4 = 4^m - m4^{m-1}4 = (1-m)4^m$

➔ $$A^m = c_0 I + c_1 A = \begin{bmatrix} (1-m)4^m & 0 \\ 0 & (1-m)4^m \end{bmatrix} + (m4^{m-1}) \begin{bmatrix} 7 & 3 \\ -3 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 7m4^{m-1} + (1-m)4^m & 3m4^{m-1} \\ -3m4^{m-1} & m4^{m-1} + (1-m)4^m \end{bmatrix}$$

For $m = 6$

➔ $$A^m = \begin{bmatrix} 22528 & 18432 \\ -18432 & -14336 \end{bmatrix}$$

Note that, you should extend this example to (page 403, Problem 12) for the **Final**.

Proof of Cayley-Hamilton Theorem

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\quad B=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$\begin{aligned} AB &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{bmatrix} \end{aligned}$$

where column 1 is equal to

$$\begin{Bmatrix} a_{11}b_{11}+a_{12}b_{21} \\ a_{21}b_{11}+a_{22}b_{21} \end{Bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{Bmatrix} b_{11} \\ b_{21} \end{Bmatrix} = A\begin{Bmatrix} b_{11} \\ b_{21} \end{Bmatrix}$$

and column 2 is equal to

$$\begin{Bmatrix} a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{12}+a_{22}b_{22} \end{Bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{Bmatrix} b_{12} \\ b_{22} \end{Bmatrix} = A\begin{Bmatrix} b_{12} \\ b_{22} \end{Bmatrix}$$

Thus, we can rewrite

$$\begin{aligned} AB &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} AX_1 & AX_2 \end{bmatrix} \end{aligned}$$

where $X_1=\begin{Bmatrix} b_{11} \\ b_{21} \end{Bmatrix},\quad X_2=\begin{Bmatrix} b_{12} \\ b_{22} \end{Bmatrix}$

In general, for $A_{n\times n}$, $B_{n\times n}$

$$AB = A\begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix} = \begin{bmatrix} AX_1 & AX_2 & \dots & AX_n \end{bmatrix}$$

where $X_1, \quad X_2, \quad \dots \quad, X_n$ are the columns of $B_{n\times n}$.

For $A_{n\times n}$, can we find an n×n nonsingular matrix P such that

$P^{-1}A_{n\times n}P = D$ is a diagonal matrix ?

If such a matrix P can be found, then we say that the matrix $A_{n\times n}$ can be diagonalized, or is diagonalizable, and that P diagonalizes $A_{n\times n}$.

For simplicity, let us assume that $A_{3\times 3}$ is a diagonalizable matrix. Then there exists a 3×3 nonsingular matrix P such that $P^{-1}A_{3\times 3}P = D$ or $A_{3\times 3}P = PD$, where D is a diagonal matrix $D = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{pmatrix}$.

If $P_{1,}\ P_{2,}\ P_3$ denote the columns of P

➔ $A_{3\times 3}P = PD$ ➔ $\left(AP_1 \quad AP_2 \quad AP_3\right) = \left(d_{11}P_1 \quad d_{22}P_2 \quad d_{33}P_3\right)$

Or $AP_1 = d_{11}P_1, \quad AP_2 = d_{22}P_2, \quad AP_3 = d_{33}P_3$

$P_1 \leftrightarrow d_{11}, \quad P_2 \leftrightarrow d_{22}, \quad P_3 \leftrightarrow d_{33}$ (eigenvector <—> eigenvalue)

➔ $P^{-1}AP = D$

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad D^2 = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1^2 & 0 & \dots & 0 \\ 0 & \lambda_2^2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n^2 \end{pmatrix}$$

➔ $$D^n = \begin{pmatrix} \lambda_1^n & 0 & \dots & 0 \\ 0 & \lambda_2^n & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n^n \end{pmatrix}$$

➔ $D^2 = (P^{-1}AP)(P^{-1}AP) = P^{-1}APP^{-1}AP = P^{-1}A^2P$ ➔ $D^n = P^{-1}A^nP$

characteristic equation ➔ $(-1)^n \lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0 = 0$

➔ $(-1)^n D^n + c_{n-1}D^{n-1} + \ldots + c_1 D + c_0 I = 0$

➔ $(-1)^n P^{-1}A^n P + c_{n-1}P^{-1}A^{n-1}P + \ldots + c_1 P^{-1}AP + c_0 I = 0$

➔ $P^{-1}\left[(-1)^n A^n + c_{n-1}A^{n-1} + \ldots + c_1 A + c_0 I\right]P = 0$ (NOTE $I = P^{-1}IP$)

➔ $(-1)^n A^n + c_{n-1}A^{n-1} + \ldots + c_1 A + c_0 I = 0$ (**when *A* is diagonalizable**).