The Cayley-Hamilton Theorem for 2 x 2 Matrices

Characteristic polynomial and the Cayley-Hamilton theorem

Suppose that a 2×2 matrix $\mathbf{A}$ has characteristic equation

$a\lambda^2 + b\lambda + c = 0$

To say that λ_1 and λ_2 are eigenvalues of $\mathbf{A}$ means that they are solutions to this equation,

Hence

$a\lambda_1^{\ 2} + b\lambda_1 + c = 0 \qquad a\lambda_2^{\ 2} + b\lambda_2 + c = 0$

Let $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ then

$$
\begin{aligned}
a\mathbf{D}^2 + b\mathbf{D} + c\mathbf{I} &= a\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 + b\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + c\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} a\lambda_1^{\ 2} + b\lambda_1 + c \\ a\lambda_2^{\ 2} + b\lambda_2 + c \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

Now suppose that $\mathbf{A} = \mathbf{XDX}^{-1}$ then

$$
\begin{aligned}
a\mathbf{A}^2 + b\mathbf{A} + c\mathbf{I} &= a\left(\mathbf{XDX}^{-1}\right)^2 + b\left(\mathbf{XDX}^{-1}\right) + c\mathbf{I} \\
&= a\left(\mathbf{XDX}^{-1}\right)\left(\mathbf{XDX}^{-1}\right) + b\left(\mathbf{XDX}^{-1}\right) + c\mathbf{I} \\
&= a\left(\mathbf{XDX}^{-1}\mathbf{XDX}^{-1}\right) + b\left(\mathbf{XDX}^{-1}\right) + c\mathbf{IXX}^{-1} \\
&= a\left(\mathbf{XD}^2\mathbf{X}^{-1}\right) + b\left(\mathbf{XDX}^{-1}\right) + c\mathbf{XIX}^{-1} \\
&= \mathbf{X}\left(a\mathbf{D}^2\mathbf{X}^{-1} + b\mathbf{DX}^{-1} + c\mathbf{IX}^{-1}\right) \\
&= \mathbf{X}\left(\left(a\mathbf{D}^2 + b\mathbf{D} + c\mathbf{I}\right)\mathbf{X}^{-1}\right) \\
&= \mathbf{X}\left(a\mathbf{D}^2 + b\mathbf{D} + c\mathbf{I}\right)\mathbf{X}^{-1}
\end{aligned}
$$

In this argument we have made use of the fact that $\mathbf{I} = \mathbf{XX}^{-1}$ and that $\mathbf{ID} = \mathbf{DI}$ for any matrix, $\mathbf{D}$.

Hence

$$a\mathbf{A}^2 + b\mathbf{A} + c\mathbf{I} = \mathbf{X}^{-1}\left(a\mathbf{D}^2 + b\mathbf{D} + c\mathbf{I}\right)\mathbf{X} = \mathbf{0}$$

So a 2×2 matrix satisfies its own characteristic equation, which is the Cayley-Hamilton theorem.

Cayley-Hamilton theorem

Every matrix A is a root of its own characteristic polynomial.

Remark

The Cayley-Hamilton theorem can be established for square matrices of any dimension. This theorem is used as a computational device for finding inverses of matrices and higher powers.

Example

You are given the matrix $\mathbf{M} = \begin{pmatrix} 3 & 2 & 2 \\ -2 & -2 & -2 \\ 1 & 2 & 2 \end{pmatrix}$

(a) Show that 0 and 1 are eigenvalues of $\mathbf{M}$, and find the other eigenvalue.

(b) Using the Cayley Hamilton theorem, or otherwise

(i) Show that $\mathbf{M}^4 = 11\mathbf{M}^2 - 6\mathbf{M}$ and verify that the matrix **M** satisfies its own characteristic equation

(iii) Explain why it is not possible to obtain a value for $\mathbf{M}^{-1}$ by cancelling down and rearranging the characteristic equation

Solution

(a) The characteristic equation is

$$\begin{vmatrix} 3-\lambda & 2 & 2 \\ -2 & -2-\lambda & -2 \\ 1 & 2 & 2-\lambda \end{vmatrix} = 0$$

If $\lambda_1 = 0$ we obtain

$$\begin{vmatrix} 3 & 2 & 2 \\ -2 & -2 & -2 \\ 1 & 2 & 2 \end{vmatrix} = 3\begin{vmatrix} -2 & -2 \\ 2 & 2 \end{vmatrix} + 2\begin{vmatrix} -2 & -2 \\ 2 & 1 \end{vmatrix} + 2\begin{vmatrix} -2 & -2 \\ 1 & 2 \end{vmatrix} = 3(-4+4) + 2(2+4) + 2(-4+2) = 0$$

If $\lambda_2 = 2$ we obtain

$$\begin{vmatrix} 1 & 2 & 2 \\ -2 & -4 & -2 \\ 1 & 2 & 0 \end{vmatrix} = 1\begin{vmatrix} -4 & -2 \\ 2 & 0 \end{vmatrix} + 2\begin{vmatrix} -2 & -2 \\ 0 & 1 \end{vmatrix} + 2\begin{vmatrix} -2 & -4 \\ 1 & 2 \end{vmatrix} = 4 - 4 + 0 = 0$$

If $\lambda_1 = 0$ or $\lambda_2 = 2$, we observe that the above equation holds.

Therefore, $\lambda_1 = 0$ and $\lambda_2 = 2$ are eigenvalues.

The characteristic equation yields

$$\begin{vmatrix} 3-\lambda & 2 & 2 \\ -2 & -2-\lambda & -2 \\ 1 & 2 & 2-\lambda \end{vmatrix} = 0$$

$$(3-\lambda)\begin{vmatrix} -2-\lambda & -2 \\ 2 & 2-\lambda \end{vmatrix} + 2\begin{vmatrix} -2 & -2 \\ 2-\lambda & 1 \end{vmatrix} + 2\begin{vmatrix} -2 & -2-\lambda \\ 1 & 2 \end{vmatrix} = 0$$

$$(3-\lambda)\left((-2-\lambda)(2-\lambda)+4\right)+2\left(-2+2(2-\lambda)\right)+2\left(-4-(-2-\lambda)\right)=0$$

$$-(3-\lambda)(2+\lambda)(2-\lambda)+4(3-\lambda)-4+4(2-\lambda)-8+2(2+\lambda)=0$$

$$-\left(6+\lambda-\lambda^2\right)(2-\lambda)+12-4\lambda-4+8-4\lambda-8+4+2\lambda=0$$

$$-\left(12-4\lambda-3\lambda^2+\lambda^3\right)+12-6\lambda=0$$

$$-2\lambda+3\lambda^2-\lambda^3=0$$

$$\lambda^3-3\lambda^2+2\lambda=0$$

$$\lambda\left(\lambda^2-3\lambda+2\right)=0$$

$$\lambda(\lambda-1)(\lambda-2)=0$$

So the three eigenvalues are

$\lambda_1 = 0 \quad \lambda_2 = 1 \quad \lambda_3 = 2$

(b) (i) The characteristic equation is $\lambda^3-3\lambda^2+2\lambda=0$

Therefore, from Cayley-Hamilton theorem, we have

$$\mathbf{M}^3-3\mathbf{M}^2+2\mathbf{M}=\mathbf{0}$$

Hence

$$\mathbf{M}^4-3\mathbf{M}^3+2\mathbf{M}^2=\mathbf{0}$$

$$\mathbf{M}^4=3\mathbf{M}^3+2\mathbf{M}^2$$

$$=3\left(3\mathbf{M}^2-2\mathbf{M}\right)+2\mathbf{M}^2$$

$$=11\mathbf{M}^2-6\mathbf{M}$$

(ii) The characteristic equation is $\lambda^3-3\lambda^2+2\lambda=0$

and we have $\mathbf{M}^3-3\mathbf{M}^2+2\mathbf{M}=\mathbf{0}$

To verify this equation we compute

$$\mathbf{M}^2=\begin{pmatrix} 3 & 2 & 2 \\ -2 & -2 & -2 \\ 1 & 2 & 2 \end{pmatrix}\begin{pmatrix} 3 & 2 & 2 \\ -2 & -2 & -2 \\ 1 & 2 & 2 \end{pmatrix}=\begin{pmatrix} 7 & 6 & 6 \\ -4 & -4 & -4 \\ 1 & 2 & 2 \end{pmatrix}$$

$$\mathbf{M}^3=\begin{pmatrix} 7 & 6 & 6 \\ -4 & -4 & -4 \\ 1 & 2 & 2 \end{pmatrix}\begin{pmatrix} 3 & 2 & 2 \\ -2 & -2 & -2 \\ 1 & 2 & 2 \end{pmatrix}=\begin{pmatrix} 15 & 14 & 14 \\ -8 & -8 & -8 \\ 1 & 2 & 2 \end{pmatrix}$$

Hence

$$\mathbf{M}^3 - 3\mathbf{M}^2 + 2\mathbf{M} = \begin{pmatrix} 15 & 14 & 14 \\ -8 & -8 & -8 \\ 1 & 2 & 2 \end{pmatrix} - 3\begin{pmatrix} 7 & 6 & 6 \\ -4 & -4 & -4 \\ 1 & 2 & 2 \end{pmatrix} + 2\begin{pmatrix} 3 & 2 & 2 \\ -2 & -2 & -2 \\ 1 & 2 & 2 \end{pmatrix} = \mathbf{0}$$

So we have verified the equation.

(iii) To obtain an inverse we would like to argue as follows

$\mathbf{M}\left(\mathbf{M}^2 - 3\mathbf{M}^2 + 2\mathbf{I}\right) = \mathbf{0}$ (1)

$\mathbf{M}^2 - 3\mathbf{M} + 2\mathbf{I} = \mathbf{0}$ (2)

$\mathbf{M}^2 = 3\mathbf{M} - 2\mathbf{I}$

and so forth

This is not valid!

The cancellation law from step (1) to step (2) does not apply to matrices.

You can have two non-zero matrices that multiply together to give the zero matrix, as precisely in this case! In the line $\mathbf{M}\left(\mathbf{M}^2 - 3\mathbf{M}^2 + 2\mathbf{I}\right) = \mathbf{0}$ it does not follow that since $\mathbf{M} \neq \mathbf{0}$ that $\mathbf{M}^2 - 3\mathbf{M} + 2\mathbf{I} = \mathbf{0}$ and in fact, in this case $\mathbf{M}^2 - 3\mathbf{M} + 2\mathbf{I} \neq \mathbf{0}$

Thus we cannot obtain a rearrangement of the equation in terms of the identity matrix $\mathbf{I}$; hence we cannot multiply through by $\mathbf{M}^{-1}$ to obtain and expression for $\mathbf{M}^{-1}$ - the inverse of $\mathbf{M}$