Section 5.2 The Characteristic Equation

Review:

$$A \quad \mathbf{x} \quad = \quad \lambda \quad \mathbf{x}$$

Find eigenvectors **x** by solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

How do we find the eigenvalues λ?

$\mathbf{x}$ must be nonzero

$$\Downarrow$$

$(A - \lambda I)\mathbf{x} = \mathbf{0}$ must have nontrivial solutions

$$\Downarrow$$

$(A - \lambda I)$ is not invertible

$$\Downarrow$$

$$\det(A - \lambda I) = 0$$

(called the *characteristic equation*)

$\boxed{\text{Solve } \det(A - \lambda I) = 0 \text{ for } \lambda \text{ to find the eigenvalues.}}$

Characteristic polynomial: $\det(A - \lambda I)$

Characteristic equation: $\det(A - \lambda I) = 0$

EXAMPLE: Find the eigenvalues of $A = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix}$.

Solution: Since

$$A - \lambda I = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -\lambda & 1 \\ -6 & 5 - \lambda \end{bmatrix},$$

the equation $\det(A - \lambda I) = 0$ becomes

$$-\lambda(5 - \lambda) + 6 = 0$$

$$\lambda^2 - 5\lambda + 6 = 0$$

Factor:

$$(\lambda - 2)(\lambda - 3) = 0.$$

So the eigenvalues are 2 and 3.

For a 3×3 matrix or larger, recall that a determinant can be computed by cofactor expansion.

EXAMPLE: Find the eigenvalues of $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -5 & 0 \\ 1 & 8 & 1 \end{bmatrix}$.

Solution:

$$A - \lambda I = \begin{bmatrix} 1 - \_\_\_\_ & 2 & 1 \\ 0 & -5 - \_\_\_\_ & 0 \\ 1 & 8 & 1 - \_\_\_\_ \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 & 1 \\ 0 & -5-\lambda & 0 \\ 1 & 8 & 1-\lambda \end{vmatrix} = (-5-\lambda)\begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix}$$

$$= (-5-\lambda)\left[(1-\lambda)^2 - 1\right] = (-5-\lambda)[1 - 2\lambda + \lambda^2 - 1]$$

$$= (-5-\lambda)[-2\lambda + \lambda^2] = -(5+\lambda)\lambda[-2+\lambda] = 0$$

$$\Rightarrow \lambda = -5, 0, 2$$

THEOREM (The Invertible Matrix Theorem - continued)

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

s. The number 0 is not an eigenvalue of A.

t. $\det A \neq 0$

Recall that if B is obtained from A by a sequence of row replacements or interchanges, but without scaling, then $\det A = (-1)^r \det B$, where r is the number of row interchanges.

Suppose the echelon form U is obtained from A by a sequence of row replacements or interchanges, but without scaling.

$$A \sim U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & u_{nn} \end{bmatrix}$$

The **determinant** of A, written $\det A$, is defined as follows:

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix} \text{product of} \\ \text{pivots in } U \end{pmatrix}, & \text{when } A \text{ is invertible} \\ 0, & \text{when } A \text{ is not invertible} \end{cases}$$

(r is the number of row interchanges)

EXAMPLE: Find the eigenvalues of $A = \begin{bmatrix} 3 & 2 & 3 \\ 0 & 6 & 10 \\ 0 & 0 & 2 \end{bmatrix}$.

Solution:

$$\det(A - \lambda I) = \det\begin{bmatrix} 3-\lambda & 2 & 3 \\ 0 & 6-\lambda & 10 \\ 0 & 0 & 2-\lambda \end{bmatrix}$$

Characteristic equation:

$$(\qquad\qquad)(\qquad\qquad)(\qquad\qquad) = 0.$$

eigenvalues: _____, _____, _____

The (**algebraic**) **multiplicity** of an eigenvalue is its multiplicity as a root of the characteristic equation.

EXAMPLE: Find the characteristic polynomial of

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ 9 & 1 & 3 & 0 \\ 1 & 2 & 5 & -1 \end{bmatrix}$$

and then find all the eigenvalues and the algebraic multiplicity of each eigenvalue.

Solution:

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 0 & 0 & 0 \\ 5 & 3-\lambda & 0 & 0 \\ 9 & 1 & 3-\lambda & 0 \\ 1 & 2 & 5 & -1-\lambda \end{vmatrix}$$

$$= (2-\lambda)(3-\lambda)(3-\lambda)(-1-\lambda) = 0$$

eigenvalues: _____, _____, _____

Similarity

Numerical methods for finding approximating eigenvalues are based upon Theorem 4 to be described shortly.

For $n \times n$ matrices A and B, we say the A is **similar** to B if there is an invertible matrix P such that

$$P^{-1}AP = B \qquad \text{or equivalently,} \qquad A = PBP^{-1}.$$

Theorem 4: If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Proof: If $B = P^{-1}AP$, then

$$\det(B - \lambda I) = \det[P^{-1}AP - P^{-1}\lambda IP] = \det[P^{-1}(A - \lambda I)P]$$

$$= \det P^{-1} \cdot \det(A - \lambda I) \cdot \det P = \det(A - \lambda I).$$

Application to Markov Chains

EXAMPLE Consider the migration matrix $M = \begin{bmatrix} .95 & .90 \\ .05 & .10 \end{bmatrix}$ and define $\mathbf{x}_{k+1} = M\mathbf{x}_k$. It can be shown that

$$\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$$

converges to a steady state vector $\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$. Why?

The answer lies in examining the corresponding eigenvectors.

First we find the eigenvalues:

$$\det(M - \lambda I) = \det\left(\begin{bmatrix} .95 - \lambda & .90 \\ .05 & .10 - \lambda \end{bmatrix}\right) = \lambda^2 - 1.05\lambda + 0.05$$

So solve

$$\lambda^2 - 1.05\lambda + 0.05 = 0$$

By factoring

$$\lambda = 0.05, \ \lambda = 1$$

It can be shown that the eigenspace corresponding to $\lambda = 1$ is span$\{\mathbf{v}_1\}$ where $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and the eigenspace corresponding to $\lambda = 0.05$ is span$\{\mathbf{v}_2\}$ where $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Note that

$$M\mathbf{v}_1 = \mathbf{v}_1,$$

and so $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$ is our steady state vector.

Then for a given vector $\mathbf{x}_0$,

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$$

$$\mathbf{x}_1 = M\mathbf{x}_0 = M(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1M\mathbf{v}_1 + c_2M\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2(0.05)\mathbf{v}_2$$

$$\mathbf{x}_2 = M\mathbf{x}_1 = M(c_1\mathbf{v}_1 + c_2(0.05)\mathbf{v}_2) = c_1M\mathbf{v}_1 + c_2(0.05)M\mathbf{v}_2 = c_1\mathbf{v}_1 +$$

and in general

$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2(0.05)^k\mathbf{v}_2$$

and so $\lim_{k\to\infty}\mathbf{x}_k = \lim_{k\to\infty}\left(c_1\mathbf{v}_1 + c_2(0.05)^k\mathbf{v}_2\right) = c_1\mathbf{v}_1$

and this is the steady state when $c_1 = \frac{1}{2}$.