Math 2331 – Linear Algebra

5.2 The Characteristic Equation

Jiwen He

Department of Mathematics, University of Houston

jiwenhe@math.uh.edu
math.uh.edu/∼jiwenhe/math2331

5.2 The Characteristic Equation

- The Characteristic Equation: Definition and Examples
- The Invertible Matrix Theorem (continued)
- Row Reductions and Determinants
- Similarity
- Application to Markov Chains

The Characteristic Equation

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

Find eigenvectors $\mathbf{x}$ 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*)

Solve $\det(A - \lambda I) = 0$ for λ to find the eigenvalues.

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

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

The Characteristic Equation: Example

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

$$-\lambda(5-\lambda) + 6 = 0 \quad \Longrightarrow \quad \lambda^2 - 5\lambda + 6 = 0$$

Factor:

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

So the eigenvalues are 2 and 3.

The Characteristic Equation: Example

Example

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

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

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

$$= (-5-\lambda)\begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = (-5-\lambda)\left[(1-\lambda)^2 - 1\right]$$

$$= (-5-\lambda)\left[1 - 2\lambda + \lambda^2 - 1\right] = -(5+\lambda)\,\lambda\,[-2+\lambda] = 0$$

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

The Invertible Matrix Theorem - continued

Theorem (IMT (cont.))

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. $\det A \neq 0$

Algebraic Multiplicity

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

Row Reductions and Determinants

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 r 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:

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

Row Reductions and Determinants: Example

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: _____, _____, _____

Row Reductions and Determinants: Example

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 their algebraic multiplicity.

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.

Similarity

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.

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

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

Application to Markov Chains: Example

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

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

Application to Markov Chains: Example (cont.)

The eigenvector corresponding to $\lambda = 1$ is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and the eigenvector corresponding to $\lambda = 0.05$ is $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

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

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

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

$$\mathbf{x}_2 = M\mathbf{x}_1 = M\left(c_1\mathbf{v}_1 + c_2\left(0.05\right)\mathbf{v}_2\right) = c_1 M\mathbf{v}_1 + c_2\left(0.05\right)M\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2\left(0.05\right)^2\mathbf{v}_2$$

and in general $\quad \mathbf{x}_k = c_1\mathbf{v}_1 + c_2\left(0.05\right)^k\mathbf{v}_2$

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

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