Section 5.2 – The Characteristic Equation

Main Ideas in this section:

- Characteristic Equation, Finding Eigenvalues
- Similar matrices
- Stochastic Matrices, steady state probability vectors

In the last section, we found eigenvectors of a matrix corresponding to given eigenvalues. In this section, we learn how to determine these eigenvalues. The method relies in part on some results obtained in section 5.1:

Theorem (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. The determinant of A is *not* zero.

How does this help us find eigenvalues?? Read on...

To find eigenvalues, given A:

$$\begin{gathered}\mathbf{x} \text{ must be nonzero}\\ \Downarrow \\ (A-\lambda I)\mathbf{x}=\mathbf{0} \text{ must have nontrivial solutions}\\ \Downarrow \\ (A-\lambda I) \text{ is not invertible}\\ \Downarrow \\ \det(A-\lambda I)=0.\end{gathered}$$

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

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

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

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

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

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

$$\begin{aligned} -\lambda(5-\lambda)+6 &= 0 \\ \lambda^2-5\lambda+6 &= 0 \\ (\lambda-2)(\lambda-3) &= 0 \end{aligned}$$

Eigenvalues are 2 and 3.

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

Exercise Prove that if A is an upper triangular matrix, the diagonal entries are the eigenvalues of A

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

Exercise Find the eigenvalues, along with multiplicities, for

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

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

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

Similar Matrices Given two $n \times n$ matrices A and B, we say that A is similar to B if there is an invertible matrix P such that

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

Example Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $P = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$. Then compute

$$P^{-1}AP = \begin{bmatrix} -5/2 & 3/2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 8 & 11 \\ -2 & -3 \end{bmatrix} = B.$$

A and B are similar.

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.

Exercise: find characteristic polynomials for A and B from the previous example.

Warning: Similarity is not the same as row equivalence. (A is row equivalent to B if $B = EA$ for some invertible matrix E.) Row operations on a matrix usually change its eigenvalues.

Stochastic Matrices

Definitions A vector with nonnegative entries that add up to 1 is called a **probability vector**. A **stochastic matrix** is a square matrix whose columns are probability vectors. A **Markov chain** is a sequence of probability vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$, together with a stochastic matrix P, such that $\mathbf{x}_{k+1} = P\mathbf{x}_k$ for $k = 0, 1, 2, \ldots$

Meaningful applications: migration models.

- entries in the probability vector $\mathbf{x} \in \mathbb{R}^n$ list the probability that a system is in each of n possible states (or they list proportion of entire population in each state).
- $n \times n$ migration matrix describes state changes that occur from one time period to the next.

Example Consider 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 the Markov chain vectors converge to a steady state vector $\mathbf{x} = \begin{bmatrix} 18/19 \\ 1/19 \end{bmatrix}$.

Why??? The answer lies in the eigenvalues and eigenvectors.

- First find the eigenvalues:

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

$$= (\lambda - .05)(\lambda - 1) \quad \Rightarrow \quad \lambda = .05, \lambda = 1.$$

- It can be shown that the eigenspace corresponding to $\lambda = 1$ is span$\{\mathbf{v}_1\}$ with $\mathbf{v}_1 = \begin{bmatrix} 18 \\ 1 \end{bmatrix}$, and the eigenspace corresponding to $\lambda = .05$ is span$\{\mathbf{v}_2\}$ with $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

So for any choice of probability vector $\mathbf{x}_0, \ldots$

Summary

- Find eigenvalues by solving $\det(A - \lambda I) = 0$.
- Similar matrices have the same eigenvalues, row-equivalent matrices typically do not have the same eigenvalues.
- Dominant eigenvalue/vector term dominates when computing $\mathbf{x}_{k+1} = M\mathbf{x}_k$ as k gets large (M stochastic, $\mathbf{x}_k$ a probability vector).