§5.2 The characteristic Equation

Review:

1) **Def:** An eigenvector of an $n\times n$ matrix A is a nonzero vector $\vec{x}$ such that $A\vec{x} = \lambda\vec{x}$ for some scalar λ.

λ is called an eigenvalue of A

$\vec{x}$ is called an eigenvector corresponding to λ.

Def: The characteristic equation: A: $n\times n$ matrix

$\det(A-\lambda I)=0$ is called the characteristic equation of A.

$\det(A-\lambda I)$ is called the characteristic polynomial of A, a polynomial of degree n in the variable λ.

Thm: λ is an eigenvalue of A iff λ satisfies the characteristic equation $\det(A-\lambda I)=0$.

Review of the determinant:

A: $n\times n$ matrix; U: an echelon form of A by row replacements and row interchanges (without scaling). with r many interchanges

$$\det A = (-1)^r \det U = \begin{cases} (-1)^r \cdot (\text{product of pivots of } U), & U \text{ invertible} \\ 0, & \text{otherwise} \end{cases}$$

Thm: A: $n\times n$ matrix. Then A is invertible iff

1) The number 0 is not an eigenvalue of A

2) The determinant of A is not zero.

Thm: (Properties of Determinants)

(a) A is invertible iff $\det A \neq 0$.

(b) $\det(AB) = (\det A)\cdot(\det B)$

(c) $\det A^T = \det A$

(d) If A is triangular $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$, then

$\det A = a_{11} a_{22} \cdots a_{nn}$.

(e) A row replacement operation of A doesn't change the determinant. A row interchange changes the sign of the determinant. A row scaling also scales the determinant by the same scale factor.

Ex: Let $A = \begin{pmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

1) Find the characteristic equation of A.

2) Find all the eigenvalues of A.

Solution: 1) $\det(A-\lambda I) = \begin{vmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix}$

$= (5-\lambda)^2(3-\lambda)(1-\lambda)$

2) Let $\det(A - \lambda I) = 0$, we have

$$(5-\lambda)^2(3-\lambda)(1-\lambda) = 0$$

$$\lambda_1 = \lambda_2 = 5, \ \lambda_3 = 3, \ \lambda_4 = 1.$$

Ex: The characteristic polynomial of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Find the eigenvalues and their multiplicities.

Solution: $\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda^2 - 4\lambda - 12)$

$$= \lambda^4(\lambda - 6)(\lambda + 2)$$

eigenvalues: $\lambda = 0$ multiplicity 4

$\lambda = 6$ multiplicity 1

$\lambda = -2$ multiplicity 1

Exercise: Find the characteristic polynomial and the eigenvalues of the matrices

1) $A = \begin{pmatrix} 5 & -2 & 3 \\ 0 & 1 & 0 \\ 6 & 7 & -2 \end{pmatrix}$. 2) $A = \begin{pmatrix} 0 & 2 & 1 \\ 3 & 0 & 3 \\ 1 & 2 & 0 \end{pmatrix}$

Solution: 1) $\det(A - \lambda I) = \begin{vmatrix} 5-\lambda & -2 & 3 \\ 0 & 1-\lambda & 0 \\ 6 & 7 & -2-\lambda \end{vmatrix}$

$$= (1-\lambda)\cdot(-1)^{2+2}\cdot \begin{vmatrix} 5-\lambda & 3 \\ 6 & -2-\lambda \end{vmatrix}$$

$$= (1-\lambda)\left[(5-\lambda)(-2-\lambda) - 18\right]$$

$$= (1-\lambda)[-10+2\lambda-5\lambda+\lambda^2-18]$$

$$= (1-\lambda)(\lambda^2-3\lambda-28)$$

$$= (1-\lambda)(\lambda-7)(\lambda+4)$$

Let $\det(A-\lambda I)=0$, we get $\lambda_1=1$, $\lambda_2=7$, $\lambda_3=-4$.

So the characteristic polynomial is

$\det(A-\lambda I) = (1-\lambda)(\lambda-7)(\lambda+4)$.

The eigenvalues are $1, 7, -4$.

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

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

$$= -\begin{vmatrix} 1 & 2 & -\lambda \\ 0 & -\lambda-6 & 3+3\lambda \\ 0 & 2+2\lambda & 1-\lambda^2 \end{vmatrix}$$

$$= -1\cdot[(-\lambda-6)(1-\lambda^2)-(3+3\lambda)(2+2\lambda)]$$

$$= -[(\lambda-1)(\lambda+1)(\lambda+6)-6(1+\lambda)^2]$$

$$= -(\lambda+1)[(\lambda-1)(\lambda+6)-6(1+\lambda)]$$

$$= -(\lambda+1)[\lambda^2+5\lambda-6-6-6\lambda]$$

$$= -(\lambda+1)[\lambda^2-\lambda-12] = -(\lambda+1)(\lambda-4)(\lambda+3)$$

So the characteristic polynomial is $-(\lambda+1)(\lambda-4)(\lambda+3)$

Let $\det(A-\lambda I)=0$. That is $(\lambda+1)(\lambda-4)(\lambda+3)=0$

Eigenvalues : $\lambda_1=-1$, $\lambda_2=4$, $\lambda_3=-3$

* Similarity

Def: A, B $n\times n$ matrix

A is similar to B if there exists an invertible matrix P such that $P^{-1}AP = B$ or equivalently, $A = PBP^{-1}$.

Changing A into $P^{-1}AP$ is called a similarity transformation

Thm: 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

$$\begin{aligned} B - \lambda I &= P^{-1}AP - \lambda I \\ &= P^{-1}AP - P^{-1}(\lambda I)P \\ &= P^{-1}(A-\lambda I)P \end{aligned}$$

$$\begin{aligned} \det(B-\lambda I) &= \det(P^{-1}(A-\lambda I)P) \\ &= \det P^{-1} \cdot \det(A-\lambda I) \det P \\ &= \det(A-\lambda I) \det P^{-1} \cdot \det P \\ &= \det(A-\lambda I) \det(P^{-1} \cdot P) \\ &= \det(A-\lambda I) \det I_n \\ &= \det(A-\lambda I) \end{aligned}$$

Warnings: 1) The matrices $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ and $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ are not similar even though they have the same eigenvalues

2) Similarity is not the same as row equivalence. Row operations on a matrix usually change its eigenvalues.

* Application to Dynamical Systems

Example: Let $A = \begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix}$. Analyze the long-term behavior of the dynamical system defined by $\vec{x}_{k+1} = A\vec{x}_k$ $(k=0,1,2,\cdots)$ with $\vec{x}_0 = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$.

Solution: We first find the eigenvalues of A and a basis for each eigenspace.

$$\det \begin{pmatrix} 0.95-\lambda & 0.03 \\ 0.05 & 0.97-\lambda \end{pmatrix}$$

$$= (0.95-\lambda)(0.97-\lambda) - (0.03)\cdot(0.05)$$

$$= \lambda^2 - 1.92\lambda + 0.92 = 0$$

By the quadratic formula

$$\lambda = \frac{1.92 \pm \sqrt{(1.92)^2 - 4\cdot 0.92}}{2}$$

$$= \frac{1.92 \pm 0.08}{2}$$

$\lambda_1 = 1$, $\lambda_2 = 0.92$

For $\lambda_1 = 1$, $\begin{pmatrix} 0.95-1 & 0.03 \\ 0.05 & 0.97-1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\begin{pmatrix} -0.05 & 0.03 \\ 0.05 & -0.03 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$-5x_1 + 3x_2 = 0$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = c\begin{pmatrix} 3 \\ 5 \end{pmatrix} = c\vec{v}_1, \quad \text{where } \vec{v}_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

For $\lambda_2 = 0.92$ $\begin{pmatrix} 0.95-0.92 & 0.03 \\ 0.05 & 0.97-0.92 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\begin{pmatrix} 0.03 & 0.03 \\ 0.05 & 0.05 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = c\begin{pmatrix} 1 \\ -1 \end{pmatrix} = c\vec{v}_2, \quad \text{where } \vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Then $\{\vec{v}_1, \vec{v}_2\}$ is obviously a basis for $\mathbb{R}^2$ since the $\vec{v}_1$ and $\vec{v}_2$ are linearly independent and $\dim \mathbb{R}^2 = 2$. So there exist weight c_1 and c_2 such that

$$\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2 = [\vec{v}_1 \ \ \vec{v}_2]\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

In fact, $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = [v_1 \ \ v_2]^{-1}\vec{x}_0$

$$= \begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}^{-1}\begin{pmatrix} 0.60 \\ 0.40 \end{pmatrix}$$

$$= \frac{1}{-8}\begin{pmatrix} -1 & -1 \\ -5 & 3 \end{pmatrix}\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = \begin{pmatrix} 0.125 \\ 0.225 \end{pmatrix}$$

Since $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors of A with $A\vec{v}_1 = \vec{v}_1$ and $A\vec{v}_2 = 0.92\vec{v}_2$, we have

$$\vec{x}_1 = A\vec{x}_0 = c_1 A\vec{v}_1 + c_2 A\vec{v}_2$$
$$= c_1\vec{v}_1 + c_2(0.92)\vec{v}_2$$
$$\vec{x}_2 = A\vec{x}_1 = c_1 A\vec{v}_1 + c_2(0.92)\cdot A\vec{v}_2$$
$$= c_1\vec{v}_1 + c_2(0.92)^2\vec{v}_2$$

...

In general, $\vec{x}_k = c_1\vec{v}_1 + c_2(0.92)^k\vec{v}_2$, $k = 0, 1, 2, \cdots$
$$= 0.125\begin{pmatrix}3\\5\end{pmatrix} + 0.225(0.92)^k\begin{pmatrix}1\\-1\end{pmatrix}$$

As $k \to \infty$, $(0.92)^k$ tends to zero and $\vec{x}_k$ tends to
$$0.125\cdot\begin{pmatrix}3\\5\end{pmatrix} = \begin{pmatrix}0.375\\0.625\end{pmatrix}$$

Exercise 1: Show that if $A = QR$ with Q invertible, then A is similar to $A_1 = RQ$

Proof: Let $P = Q$. Then
$$P^{-1}AP = Q^{-1}QR\cdot Q = RQ = A_1$$
So A is similar to A_1.

Exercise 2: Show that if A and B are similar, then $\det A = \det B$.

Proof: Since A and B are similar, there exists an invertible matrix P such that

$P^{-1}AP = B.$

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1})\cdot\det(A)\det(P)$$
$$= \det(A)\det(P^{-1}P)$$
$$= \det(A)\det(I_n)$$
$$= \det(A).$$

Exercise: Mark each statement True or False. Justify each answer.

1) The determinant of A is the product of the diagonal entries in A.
2) An elementary row operation on A does not change the determinant.
3) $(\det A)\cdot(\det B) = \det(AB)$
4) If $\lambda+5$ is a factor of the characteristic polynomial of A, then 5 is an eigenvalue of A.

Solution: 1). False

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}$$

Then $\det A = 1\cdot(-1) - 2\cdot 3 = -7 \neq 1\cdot(-1)$

2) False

Interchange and scaling will change the determinant

3) True

4) False

$\lambda = -5$ is an eigenvalue of A.

Exercise: Show that A and A^T have the same characteristic polynomial.

Solution: $\det(A^T - \lambda I) = \det((A-\lambda I)^T)$

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

Exercise: Let $A = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{pmatrix}$, $\vec{v}_1 = \begin{pmatrix} 0.3 \\ 0.6 \\ 0.1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$

$\vec{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, and $\vec{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

a) Show that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are eigenvectors of A.

b) Let $\vec{x}_0$ be any vector in $\mathbb{R}^3$ with nonnegative entries whose sum is 1. Explain why there are constants c_1, c_2 and c_3 such that $\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3$. Compute $\vec{w}^T\vec{x}_0$, and deduce that $c_1 = 1$.

Solution: a). $A\vec{v}_1 = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{pmatrix}\begin{pmatrix} 0.3 \\ 0.6 \\ 0.1 \end{pmatrix}$

$= \begin{pmatrix} 0.5\times 0.3 + 0.2\times 0.6 + 0.3\times 0.1 \\ 0.3\times 0.3 + 0.8\times 0.6 + 0.3\times 0.1 \\ 0.2\times 0.3 + 0\times 0.6 + 0.4\times 0.1 \end{pmatrix}$

$$= \begin{pmatrix} 0.3 \\ 0.6 \\ 0.1 \end{pmatrix} = \vec{v}_1$$

$$A\vec{v}_2 = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5\times 1 + 0.2\times(-3) + 0.3\times 2 \\ 0.3\times 1 + 0.8\times(-3) + 0.3\times 2 \\ 0.2\times 1 + 0\times(-3) + 0.4\times 2 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5 \\ -1.5 \\ 1 \end{pmatrix} = \frac{1}{2}\vec{v}_2$$

$$A\vec{v}_3 = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5\times(-1) + 0.2\times 0 + 0.3\times 1 \\ 0.3\times(-1) + 0.8\times 0 + 0.3\times 1 \\ 0.2\times(-1) + 0\times 0 + 0.4\times 1 \end{pmatrix}$$

$$= \begin{pmatrix} -0.2 \\ 0 \\ 0.2 \end{pmatrix} = 0.2\,\vec{v}_3$$

So $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are eigenvectors of A.

b) Since $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are eigenvectors corresponding to distict eigenvalue $1, \frac{1}{2}, 0.2$, they are linearly independent and form a basis in $\mathbb{R}^3$.

So there are constants c_1, c_2, c_3 such that
$\vec{x}_0 = c_1 \vec{v}_1 + c_2 \vec{v}_2 + c_3 \vec{v}_3$.

$$\vec{w}^T \vec{x}_0 = c_1 \vec{w}^T \vec{v}_1 + c_2 \vec{w}^T \vec{v}_2 + c_3 \vec{w}^T \vec{v}_3$$
$$= c_1 (0.3+0.6+0.1) + c_2 (1-3+2) + c_3 (-1+0+1)$$
$$= c_1$$

Since $\vec{w}^T x_0 = x_{01} + x_{02} + x_{03} = 1$

we have $c_1 = 1$.

c). $\vec{x}_k = A^k \vec{x}_0 = A^k (c_1 \vec{v}_1 + c_2 \vec{v}_2 + c_3 \vec{v}_3)$
$$= c_1 A^k \vec{v}_1 + c_2 A^k \vec{v}_2 + c_3 A^k \vec{v}_3$$
$$= c_1 v_1 + c_2 \cdot \left(\tfrac{1}{2}\right)^k \vec{v}_2 + c_3 (0.2)^k \vec{v}_3$$
$$= v_1 + c_2 \left(\tfrac{1}{2}\right)^k \vec{v}_2 + c_3 (0.2)^k \vec{v}_3$$
$$\longrightarrow v_1 \quad \text{as } k \to \infty.$$