9 Eigenvalues and Eigenvectors

9.1 Eigenvalues and Characteristic Polynomials

Definition 9.1 [page 285] An *eigenvector*（固有ベクトル） of an $n \times n$ matrix A is a nonzero vector $\boldsymbol{x} \in \mathbb{R}^n$ such that

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

for some scalar λ. A scalar λ is called an *eigenvalue*（固有値） of A, if there is a nontrivial solution $\boldsymbol{x}$ of $A\boldsymbol{x} = \lambda\boldsymbol{x}$; such an $\boldsymbol{x}$ is called an *eigenvector corresponding to* λ.

$$\exists \boldsymbol{x} \neq \mathbf{0}, A\boldsymbol{x} = \lambda\boldsymbol{x} \Leftrightarrow \exists \boldsymbol{x} \neq \mathbf{0}, (A - \lambda I)\boldsymbol{x} = \mathbf{0} \Leftrightarrow \det(A - \lambda I) = 0.$$

Definition 9.2 [page 294] The determinant $\det(A - xI)$ is a polynomial of degree n in x. It is called the *characteristic polynomial*（固有（特性）多項式） of A, and $\det(A - xI) = 0$ the *characteristic equation*（固有方程式） of A. The (algebraic) *multiplicity*（重複度） of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

Theorem 9.1 (Theorem 2 in page 288) *If $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots \boldsymbol{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ of an $n \times n$ matrix A. Then the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ is linearly independent.*

9.2 Diagonalization

Definition 9.3 [page 300] If A and B are $n \times n$ matrices, then A is *similar*（相似） to B if there is an invertible matrix P such that $P^{-1}AP = B$, or equivalently $A = PBP^{-1}$.

Theorem 9.2 (Theorem 4 in page 295) *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.*

Definition 9.4 [page 295] A square matrix A is said to be *diagonalizable*（対角化可能） if A is similar to a diagonal matrix, i.e., there is an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Theorem 9.3 (Theorem 5 in page 300) *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. In fact $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal eintries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.*

In particular, an $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Proof. Suppose A is diagonalizable. Then there is an invertible matrix $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ and a diagonal matrix D such that $A = PDP^{-1}$. Let $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, where λ_i

is the ith diagonal entry. Then

$$\begin{aligned} [A\boldsymbol{v}_1, A\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n] &= A[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] \\ &= AP \\ &= PD \\ &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ &= [\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2, \ldots, \lambda_n\boldsymbol{v}_n]. \end{aligned}$$

Hence $A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1$, $A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n$.

Suppose there are n linearly independent eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$. Then $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ is invertible by Theorem 8 in Chapter 2. Let $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then

$$\begin{aligned} AP &= A[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] \\ &= [A\boldsymbol{v}_1, A\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n] \\ &= [\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2, \ldots, \lambda_n\boldsymbol{v}_n] \\ &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ &= PD. \end{aligned}$$

Since P is invertible, $A = PDP^{-1}$. ■

Example 9.1

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 6 & 1 & 3 \\ 0 & 4 & 3 \end{bmatrix}, \quad \boldsymbol{u} = \begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \quad \boldsymbol{w} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}.$$

Then

$$A\boldsymbol{u} = \begin{bmatrix} 0 & 1 & 0 \\ 6 & 1 & 3 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 6 \\ 36 \\ 48 \end{bmatrix} = 6 \cdot \begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix}.$$

$$\begin{aligned} \det(A - xI) &= \begin{vmatrix} -x & 1 & 0 \\ 6 & 1-x & 3 \\ 0 & 4 & 3-x \end{vmatrix} \\ &= -x(((x-1)(x-3) - 12) - (-1)(-6)(x-3)) = -(x^3 - 4x^2 - 15x + 18) \\ &= -(x-6)(x+3)(x-1). \end{aligned}$$

Find nontrivial solutions of $(A - 6I)\boldsymbol{x} = \boldsymbol{0}$, $(A - (-3)I)\boldsymbol{x} = \boldsymbol{0}$ and $(A - I)\boldsymbol{x} = \boldsymbol{0}$. They are $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$. Let $T = [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]$. Then

$$\begin{aligned} AT &= \begin{bmatrix} 0 & 1 & 0 \\ 6 & 1 & 3 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 6 & -3 & 1 \\ 8 & 2 & -2 \end{bmatrix} = \begin{bmatrix} 6 & -3 & 1 \\ 36 & 9 & 1 \\ 48 & -6 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 6 & -3 & 1 \\ 8 & 2 & -2 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = TD. \end{aligned}$$

$$T^{-1}AT = D = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } A = TDT^{-1}.$$

Example 9.2 [Theorem 1 in page 291 (269)] If A is an $n \times n$ triangular matrix, then the egenvalues of A are the entries on the main diagonal of A.

Example 9.3 Let

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}. \text{ Then } \det(A - xI) = \begin{vmatrix} -x & 1 & 0 & 1 \\ 1 & -x & 1 & 0 \\ 0 & 1 & -x & 1 \\ 1 & 0 & 1 & -x \end{vmatrix} = x^2(x-2)(x+2).$$

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{v}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \; P = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}.$$

$$AP = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix} = PD.$$

9.3 Extra Results without a Proof (Not to be included in Final)

Theorem 9.4 (Cayley-Hamilton (See page 344 Exercise 7)) *Let A be an $n \times n$ matrix and $p(x) = \det(A - xI)$ is the characteristic polynomial of A. Then $p(A) = O$.*

Proof. The proof is complicated. So we prove only when A is diagonalizable. If $A = D$ is a diagonal matrix, this is obvious. For the general case, suppose $P^{-1}AP = D$. Then $A = PDP^{-1}$ and $p(A) = Pp(D)P^{-1}$. By Theorem 9.2, the characteristic polynomial of D is equal to $p(x)$. Now clearly $p(D) = O$. ■

Theorem 9.5 (Theorems 2, 3 in Section 7.1) *Let A be an $n \times n$ matrix. Then the following are equivalent.*

(i) *There is a matrix P such that $P^{-1} = P^\top$* [9] *and $P^\top AP$ is diagonal.*

(ii) *There are n eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ such that $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \delta_{i,j}$.* [10]

(iii) $A = A^\top$.

Theorem 9.6 (Triangulation（三角化可能）) (1) *If A is a square matrix such that all eigenvalues are real. Then there is an invertible matrix P such that $P^{-1}AP$ is an upper triangular matrix.*
(2) *If A is a square matrix such that all entries are complex numbers. Then there is an invertible matrix P such that $P^{-1}AP$ is an upper triangular matrix.*

---

[9] A square matrix with the property $P^{-1} = P^\top$ is called an orthogonal matrix（直交行列）.

[10] The set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ with the property $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \delta_{i,j}$ is called orthonormal（正規直交）.