Chapter 4: Eigenvalues and Eigenvectors

April 17, 2009

Week 11-12

1 Eigenvalues and eigenvectors

If a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ has the form

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

then we can easily see that $T(\boldsymbol{e}_i) = \lambda_i \boldsymbol{e}_i$ $(i = 1, 2, \ldots, n)$, i.e., each $\boldsymbol{e}_i$ is dilated λ_i times by T.

Example 1.1. The linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, defined by

$$T(\boldsymbol{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 3x_2 \end{bmatrix},$$

is to dilate the first coordinate two times and the second coordinate three times.

Example 1.2. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation defined by

$$T(\boldsymbol{x}) = \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

What can we say geometrically about T? Consider the basis $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ of $\mathbb{R}^2$, where

$$\boldsymbol{u}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Then

$$T(\boldsymbol{u}_1) = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 3\boldsymbol{u}_1,$$

$$T(\boldsymbol{u}_2) = \begin{bmatrix} -2 \\ -6 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = -2\boldsymbol{u}_2.$$

For any vector $\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2$, we have $[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, and

$$T(\boldsymbol{v}) = c_1 T(\boldsymbol{u}_1) + c_2 T(\boldsymbol{u}_2) = 3c_1\boldsymbol{u}_1 - 2c_2\boldsymbol{u}_2,$$

Thus

$$[T(\boldsymbol{v})]_{\mathcal{B}} = \begin{bmatrix} 3c_1 \\ -2c_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

If one uses the basis $\mathcal{B}$ to describe vector $\boldsymbol{v}$ with coordinate vector $[\boldsymbol{v}]_{\mathcal{B}}$, then the coordinate vector of $T(\boldsymbol{v})$ under the basis $\mathcal{B}$ simply described as

$$\left[T(\boldsymbol{v})\right]_{\mathcal{B}} = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \left[\boldsymbol{v}\right]_{\mathcal{B}}.$$

This means that the matrix of T relative to the basis $\mathcal{B}$ is as simple as a diagonal matrix.

The above discussion demonstrates that for a linear transformation $T : V \to V$, the nonzero vectors $\boldsymbol{v}$ satisfying the condition

$$T(\boldsymbol{v}) = \lambda \boldsymbol{v} \tag{1.1}$$

for some scalar λ is important in simplifying a linear transformation T.

Definition 1.1. Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with $T(\boldsymbol{x}) = A\boldsymbol{x}$. A nonzero vector $\boldsymbol{v}$ in $\mathbb{R}^n$ is called an **eigenvector** of T (the matrix A) if there exists a scalar λ such that

$$T(\boldsymbol{v}) = A\boldsymbol{v} = \lambda \boldsymbol{v}. \tag{1.2}$$

The scalar λ is called an **eigenvalue** of T (the matrix A) and the nonzero vector $\boldsymbol{v}$ is called an **eigenvector of T (the matrix A) corresponding to the eigenvalue λ**.

Example 1.3. Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$. Then $\boldsymbol{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ is an eigenvector of A. However, but $\boldsymbol{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ is not an eigenvector of A.

Proposition 1.2. *For any $n \times n$ matrix A, the value 0 is an eigenvalue of $A \iff \det A = 0$.*

Proof. Note that $\det A = 0 \iff A$ is not invertible $\iff \operatorname{Nul} A \neq \{\boldsymbol{0}\}$. The set of eigenvectors of A corresponding to the zero eigenvalue is the set $\operatorname{Nul} A - \{\boldsymbol{0}\}$. □

Theorem 1.3. *Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ be eigenvectors of a matrix A corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_p$, respectively. Then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are linearly independent.*

Proof. Let k be the largest positive integer such that $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are linearly independent. If $k = p$, nothing is to be proved. If $k < p$, then $\boldsymbol{v}_{k+1}$ is a linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$, i.e., there exist constants $c_1, c_2, \ldots, c_k$ such that

$$\boldsymbol{v}_{k+1} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k.$$

Applying the matrix A to both sides, on the one hand

$$\begin{aligned} A\boldsymbol{v}_{k+1} &= \lambda_{k+1}\boldsymbol{v}_{k+1} \\ &= \lambda_{k+1}(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k) \\ &= c_1\lambda_{k+1}\boldsymbol{v}_1 + \cdots + c_k\lambda_{k+1}\boldsymbol{v}_k. \end{aligned}$$

On the other hand,

$$\begin{aligned} A\boldsymbol{v}_{k+1} &= A(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k) \\ &= c_1A\boldsymbol{v}_1 + \cdots + c_kA\boldsymbol{v}_k \\ &= c_1\lambda_1\boldsymbol{v}_1 + \cdots + c_k\lambda_k\boldsymbol{v}_k. \end{aligned}$$

Thus

$$c_1(\lambda_{k+1} - \lambda_1)\boldsymbol{v}_1 + \cdots + c_k(\lambda_{k+1} - \lambda_k)\boldsymbol{v}_k = \mathbf{0}.$$

Since $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are linearly independent, we have

$$c_1(\lambda_{k+1} - \lambda_1) = \cdots = c_k(\lambda_{k+1} - \lambda_k) = 0.$$

Note that the eigenvalues are distinct. Hence

$$c_1 = \cdots = c_k = 0,$$

which implies that $\boldsymbol{v}_{k+1}$ is the zero vector $\mathbf{0}$. This is contradictory to that $\boldsymbol{v}_{k+1} \neq \mathbf{0}$. □

2 How to find eigenvectors?

To find an eigenvector of a matrix A, it is meant to find a nonzero vector $\boldsymbol{x}$ and scalar λ such that

$$A\boldsymbol{x} = \lambda\boldsymbol{x}. \tag{2.1}$$

Since $\lambda\boldsymbol{x} = \lambda I\boldsymbol{x}$, (2.1) is equivalent to $\lambda I\boldsymbol{x} - A\boldsymbol{x} = \mathbf{0}$, i.e.,

$$(\lambda I - A)\boldsymbol{x} = \mathbf{0}. \tag{2.2}$$

Since $\boldsymbol{x}$ is required to be nonzero, the system (2.3) is required to have nonzero solutions. We must have

$$\det(\lambda I - A) = 0. \tag{2.3}$$

Expanding the $\det(\lambda I - A)$, we see that

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

is a polynomial of degree n in λ, called the **characteristic polynomial** of A. To find eigenvalues of A, it is meant to find all roots of the polynomial $p(\lambda)$. The polynomial equation (2.3) about λ is called the **characteristic equation** of A. For each eigenvalue λ of A, the linear system

$$(\lambda I - A)\boldsymbol{x} = \mathbf{0}$$

is called the **eigensystem** for the eigenvalue λ. Its solution set $\operatorname{Nul}(\lambda I - A)$ is called the **eigenspace** corresponding to the eigenvalue λ.

Theorem 2.1. *The eigenvalues of a triangular matrix are the entries on its main diagonal.*

Example 2.1. The matrix

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 5 & 0 \\ 0 & -1 & 2 \end{bmatrix}.$$

has the characteristic polynomial

$$p(\lambda) = \begin{vmatrix} \lambda - 2 & 1 & 0 \\ 0 & \lambda - 5 & 0 \\ 0 & 1 & \lambda - 2 \end{vmatrix} = (\lambda - 2)^2(\lambda - 5).$$

Then there are two eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 5$.

For $\lambda_1 = 2$, the eigensystem

$$\begin{bmatrix} \lambda_1 - 2 & 1 & 0 \\ 0 & \lambda_1 - 5 & 0 \\ 0 & 1 & \lambda_1 - 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -3 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has two linearly independent eigenvectors

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = 5$, the eigensystem

$$\begin{bmatrix} \lambda_2 - 2 & 1 & 0 \\ 0 & \lambda_2 - 5 & 0 \\ 0 & 1 & \lambda_2 - 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has one linearly independent eigenvector

$$\boldsymbol{v}_3 = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}.$$

Example 2.2. The matrix

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 5 & 0 \\ -1 & -1 & 2 \end{bmatrix}$$

has the characteristic polynomial

$$p(\lambda) = \begin{vmatrix} \lambda - 2 & 1 & 0 \\ 0 & \lambda - 5 & 0 \\ 1 & 1 & \lambda - 2 \end{vmatrix} = (\lambda - 2)^2(\lambda - 5).$$

We obtain two eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 5$.

For $\lambda_1 = 2$ (though it is of multiplicity 2), the eigensystem

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -3 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has only one linearly independent eigenvector

$$\boldsymbol{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = 5$, eigen-system

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has one linearly independent eigenvector

$$\boldsymbol{v}_2 = \begin{bmatrix} 3 \\ -9 \\ 2 \end{bmatrix}.$$

Example 2.3. Find the eigenvalues and the eigenvectors for the matrix

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ 3 & 3 & 1 \end{bmatrix}.$$

The characteristic equation of A is

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} \lambda-1 & -3 & -3 \\ -3 & \lambda-1 & -3 \\ -3 & -3 & \lambda-1 \end{vmatrix} \quad (R_2 - R_3) \\ &= \begin{vmatrix} \lambda-1 & -3 & -3 \\ 0 & \lambda+2 & -(\lambda+2) \\ -3 & -3 & \lambda-1 \end{vmatrix} \\ &= (\lambda-1)\begin{vmatrix} \lambda+2 & -(\lambda+2) \\ -3 & \lambda-1 \end{vmatrix} - 3\begin{vmatrix} -3 & -3 \\ \lambda+2 & -(\lambda+2) \end{vmatrix} \\ &= (\lambda-1)(\lambda+2)(\lambda-4) - 18(\lambda+2) = (\lambda+2)^2(\lambda-7). \end{aligned}$$

Then A has two eigenvalues $\lambda_1 = -2$ and $\lambda_2 = 7$.

For $\lambda_1 = -2$ (its multiplicity is 2), the eigen-system

$$\begin{bmatrix} -3 & -3 & -3 \\ -3 & -3 & -3 \\ -3 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has two linearly independent eigenvectors

$$\boldsymbol{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

For $\lambda = 7$, the eigen-system

$$\begin{bmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has one linearly independent eigenvector

$$\boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Theorem 2.2. *Let λ, μ and ν be distinct eigenvalues of a matrix A. Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ be linearly independent eigenvectors for the eigenvalue λ; $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_q$ be linearly independent eigenvectors for the eigenvalue μ; and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ be linearly independent eigenvectors for the eigenvalue ν. Then the vectors*

$$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p,\ \boldsymbol{v}_1, \ldots, \boldsymbol{v}_q,\ \boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$$

are linearly independent.

Proof. Suppose there are scalars $a_1, \ldots, a_p, b_1, \ldots, b_q, c_1, \ldots, c_r$ such that

$$(a_1\boldsymbol{u}_1 + \cdots + a_p\boldsymbol{u}_p) + (b_1\boldsymbol{v}_1 + \cdots + b_q\boldsymbol{v}_q) + (c_1\boldsymbol{w}_1 + \cdots + c_r\boldsymbol{w}_r) = \boldsymbol{0}. \tag{2.4}$$

It suffices to show that all the scalars $a_1, \ldots, a_p, b_1, \ldots, b_q, c_1, \ldots, c_r$ are 0. Set

$$\boldsymbol{u} = a_1\boldsymbol{u}_1 + \cdots + a_p\boldsymbol{u}_p, \quad \boldsymbol{v} = b_1\boldsymbol{v}_1 + \cdots + b_q\boldsymbol{u}_q, \quad \boldsymbol{w} = c_1\boldsymbol{w}_1 + \cdots + c_r\boldsymbol{w}_r.$$

Note that

$$A\boldsymbol{u} = a_1A\boldsymbol{u}_1 + \cdots + a_pA\boldsymbol{u}_p = a_1\lambda\boldsymbol{u}_1 + \cdots a_p\lambda\boldsymbol{u}_p = \lambda\boldsymbol{u}.$$

Similarly, $A\boldsymbol{v} = \mu\boldsymbol{v}$ and $A\boldsymbol{w} = \nu\boldsymbol{w}$. If $\boldsymbol{u} = \boldsymbol{0}$, then the linear independence of $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ implies that

$$a_1 = \cdots = a_p = 0.$$

Similarly, $\boldsymbol{v} = \boldsymbol{0}$ implies $b_1 = \cdots = b_q = 0$, and $\boldsymbol{w} = \boldsymbol{0}$ implies $c_1 = \cdots = c_r = 0$.

Now we claim that $\boldsymbol{u} = \boldsymbol{v} = \boldsymbol{w} = \boldsymbol{0}$. If not, there are following three types.

Type 1: $\boldsymbol{u} \neq \boldsymbol{0}$, $\boldsymbol{v} = \boldsymbol{w} = \boldsymbol{0}$. Since $\boldsymbol{v} = \boldsymbol{w} = \boldsymbol{0}$, it follows from (2.4) that $\boldsymbol{u} = \boldsymbol{0}$, a contradiction.

Type 2: $\boldsymbol{u} \neq 0$, $\boldsymbol{v} \neq \boldsymbol{0}$, $\boldsymbol{w} = \boldsymbol{0}$. Then $\boldsymbol{u}$ is the eigenvector of A for the eigenvalue λ and $\boldsymbol{v}$ the eigenvector of A for the eigenvalue μ; they are eigenvectors for distinct eigenvalues. So $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly independent. But (2.4) shows that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{0}$, which means that $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly dependent, a contradiction.

Type 3: $\boldsymbol{u} \neq 0$, $\boldsymbol{v} \neq \boldsymbol{0}$, $\boldsymbol{w} \neq \boldsymbol{0}$. This means that $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are eigenvectors of A for distinct eigenvalues λ, μ, ρ respectively. So they are linearly independent. However, (2.4) shows that $\boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w} = \boldsymbol{0}$, which means that $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are linearly dependent, a contradiction again. □

Note 1. The above theorem is also true for more than three distinct eigenvalues.

3 Diagonalization

Definition 3.1. An $n \times n$ matrix A is said to be **similar** to an $n \times n$ matrix B if there is an invertible matrix P such that

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

Theorem 3.2. *Similar matrices have the same characteristic polynomial and hence have the same eigenvalues.*

Note. Similar matrices may have different eigenvectors. For instance, the matrices

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 5 & 0 \\ -1 & -1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 5 & 0 \\ 0 & -1 & 2 \end{bmatrix}$$

have the same eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 5$; but A and B have different eigenvectors.

An square matrix D is said to be **diagonal** if all non-diagonal entries are zero, that is,

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}.$$

It is easy to see that for any k,

$$D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}.$$

Definition 3.3. A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, there exists an invertible matrix P and a diagonal matrix D such that

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

Theorem 3.4 (Diagonalization)**.** *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.*

Proof. We demonstrate the proof for the case $n = 3$.

If A is diagonalizable, there exist an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$, where

$$P = [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}, \quad D = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{bmatrix}.$$

Note that $P^{-1}AP = D$ is equivalent to $AP = PD$. Since $AP = A[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = [A\boldsymbol{u}, A\boldsymbol{v}, A\boldsymbol{w}]$ and

$$PD = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{bmatrix} = \begin{bmatrix} \lambda u_1 & \mu v_1 & \nu w_1 \\ \lambda u_2 & \mu v_2 & \nu w_2 \\ \lambda u_3 & \mu v_3 & \nu w_3 \end{bmatrix} = [\lambda \boldsymbol{u}, \mu \boldsymbol{v}, \nu \boldsymbol{w}],$$

we have $[A\boldsymbol{u}, A\boldsymbol{v}, A\boldsymbol{w}] = [\lambda \boldsymbol{u}, \mu \boldsymbol{v}, \nu \boldsymbol{w}]$, i.e.,

$$A\boldsymbol{u} = \lambda \boldsymbol{u}, \quad A\boldsymbol{v} = \mu \boldsymbol{v}, \quad A\boldsymbol{w} = \nu \boldsymbol{w}.$$

Since P is invertible, then the vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly independent. This means that $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are three linearly independent eigenvectors of A.

Conversely, if A has three linear independent eigenvectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ corresponding to the eigenvalues λ, μ, ν respectively. Then $A\boldsymbol{u} = \lambda \boldsymbol{u}$, $A\boldsymbol{v} = \mu \boldsymbol{v}$, $A\boldsymbol{w} = \nu \boldsymbol{w}$. Let

$$P = [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}, \quad D = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{bmatrix}.$$

Then

$$\begin{aligned} AP &= A[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = [A\boldsymbol{u}, A\boldsymbol{v}, A\boldsymbol{w}] = [\lambda \boldsymbol{u}, \mu \boldsymbol{v}, \nu \boldsymbol{w}] \\ &= \begin{bmatrix} \lambda u_1 & \mu v_1 & \nu w_1 \\ \lambda u_2 & \mu v_2 & \nu w_2 \\ \lambda u_3 & \mu v_3 & \nu w_3 \end{bmatrix} \\ &= \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{bmatrix} = PD. \end{aligned}$$

Since $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly independent, the matrix P is invertible. Thus

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

This means that A is diagonalizable. □

Example 3.1. Diagonalize the matrix

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

and compute A^8. The characteristic polynomial of A is

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda-3 & -1 & -1 \\ -1 & \lambda-3 & -1 \\ 1 & 1 & \lambda-1 \end{vmatrix} \begin{matrix} R_2 + R_3 \\ R_1 - (\lambda-3)R_3 \end{matrix} \\ &= \begin{vmatrix} 0 & -(\lambda-2) & -(\lambda-2)^2 \\ 0 & \lambda-2 & \lambda-2 \\ 1 & 1 & \lambda-1 \end{vmatrix} = (\lambda-2)^2(\lambda-3). \end{aligned}$$

There are two eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 3$.

For $\lambda_1 = 2$, the eigensystem

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has two linearly independent eigenvectors

$$\boldsymbol{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = 3$, the eigensystem

$$\begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has one linearly independent eigenvector

$$\boldsymbol{v}_3 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

Set

$$P = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Then

$$P^{-1}AP = \begin{bmatrix} -1 & 0 & -1 \\ 1 & 1 & 2 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = D.$$

Equivalently,

$$PDP^{-1} = A.$$

Thus

$$\begin{aligned}
A^8 &= \underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{8} = PD^8P^{-1} \\
&= \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2^8 & 0 & 0 \\ 0 & 2^8 & 0 \\ 0 & 0 & 3^8 \end{bmatrix} \begin{bmatrix} -1 & 0 & -1 \\ 1 & 1 & 2 \\ -1 & -1 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 3^8 & 3^8-2^8 & 3^8-2^8 \\ 3^8-2^8 & 3^8 & 3^8-2^8 \\ 2^8-3^8 & 2^8-3^8 & 2^9-3^8 \end{bmatrix}.
\end{aligned}$$

Example 3.2. Compute the matrix A^8, where

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

The characteristic polynomial of A is

$$\begin{aligned}
|\lambda I - A| &= \begin{vmatrix} \lambda-1 & -1 & -1 \\ 0 & \lambda-3 & -3 \\ 2 & -1 & \lambda-1 \end{vmatrix} \\
&= (\lambda-1)\begin{vmatrix} \lambda-3 & -3 \\ -1 & \lambda-1 \end{vmatrix} + 2\begin{vmatrix} -1 & -1 \\ \lambda-3 & -3 \end{vmatrix} \\
&= (\lambda-1)(\lambda^2-4\lambda)+2\lambda \\
&= \lambda(\lambda-2)(\lambda-3).
\end{aligned}$$

We have eigenvalues $\lambda_1 = 0$, $\lambda_2 = 2$, and $\lambda_3 = 3$.

For $\lambda_1 = 0$,

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & -3 & -3 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \boldsymbol{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = 2$,

$$\begin{bmatrix} 1 & -1 & -1 \\ 0 & -1 & -3 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix}.$$

For $\lambda_3 = 3$,

$$\begin{bmatrix} 2 & -1 & -1 \\ 0 & 0 & -3 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}.$$

Set

$$P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3] = \begin{bmatrix} 0 & -2 & 1/2 \\ -1 & -3 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Then

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = D.$$

Thus

$$\begin{aligned}
A^8 &= \underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{8} = P\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^8 P^{-1} \\
&= \begin{bmatrix} 0 & -2 & 1/2 \\ -1 & -3 & 1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2^8 & 0 \\ 0 & 0 & 3^8 \end{bmatrix}\begin{bmatrix} 1 & -1/2 & 1/2 \\ -1 & 1/2 & 1/2 \\ -2 & 2 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 2^8 & 0 & 0 \\ 2^8 & 2^7 & 2^7 \\ -2^8 & 2^7 & 2^7 \end{bmatrix}.
\end{aligned}$$

4 Complex eigenvalues

Theorem 4.1. *For a 2×2 matrix, if one of the eigenvalues of A is not a real number, then the other eigenvalue must be conjugate to this complex eigenvalue. Let*

$$\lambda = a - b\imath \quad \text{with} \quad b \neq 0$$

be a complex eigenvalue of A and let $\boldsymbol{x} = \boldsymbol{u} + \imath\boldsymbol{v}$ be a complex eigenvector of A for λ, that is,

$$A(\boldsymbol{u} + \imath\boldsymbol{v}) = (a - b\imath)(\boldsymbol{u} + \imath\boldsymbol{v}).$$

Let $P = [\boldsymbol{u}, \boldsymbol{v}]$. Then

$$P^{-1}AP = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Proof.

$$\begin{aligned}
A\boldsymbol{x} &= A\boldsymbol{u} + \imath A\boldsymbol{v}, \\
A\boldsymbol{x} &= \lambda\boldsymbol{x} = (a - b\imath)(\boldsymbol{u} + \imath\boldsymbol{v}) = (a\boldsymbol{u} + b\boldsymbol{v}) + \imath(-b\boldsymbol{u} + a\boldsymbol{v}).
\end{aligned}$$

It follows that

$$A\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{v}, \qquad A\boldsymbol{v} = -b\boldsymbol{u} + a\boldsymbol{v}.$$

Thus

$$AP = A[\boldsymbol{u}, \boldsymbol{v}] = [\boldsymbol{u}, \boldsymbol{v}]\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = P\begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

□

Example 4.1. Let $A = \begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$. Find a matrix P such that $P^{-1}AP$ is diagonal or antisymmetric. Set

$$\begin{vmatrix} \lambda - 5 & 2 \\ -1 & \lambda - 3 \end{vmatrix} = \lambda^2 - 8\lambda + 17 = 0.$$

The eigenvalues of A are complex numbers

$$\lambda = \frac{-8 \pm \sqrt{64 - 4\cdot 17}}{2} = 4 \pm \imath.$$

For $\lambda = 4 - \imath$, we have the eigensystem

$$\begin{bmatrix} -1-\imath & 2 \\ -1 & 1-\imath \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solving the complex system of linear equations, we have the eigenvector

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1-\imath \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \imath \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Let $P = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$. Then $P^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

$$\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ 1 & 4 \end{bmatrix}$$

5 Eigenvalues of symmetric matrices

A square matrix A is said to be **symmetric** if $A^T = A$. One can view a real matrix as a complex matrix whose entries are real numbers. So one can consider complex eigenvalues and complex eigenvectors. A nonzero complex vector $\boldsymbol{v}$ is called an **eigenvector of a square complex matrix** A if there exists a complex number λ such that

$$A\boldsymbol{v} = \lambda\boldsymbol{v}.$$

The scalar λ is called a **complex eigenvalue** of A, and the vector $\boldsymbol{v}$ is called a **complex eigenvector of A corresponding to the eigenvalue** λ.

Theorem 5.1. *If A is real symmetric matrix, then the eigenvalues of A must be real numbers.*

Proof. Let $\boldsymbol{v}$ be an eigenvector of A corresponding to an eigenvalue λ of A, i.e., $A\boldsymbol{v} = \lambda\boldsymbol{v}$. Then

$$\bar{\boldsymbol{v}}^T A\boldsymbol{v} = \bar{\boldsymbol{v}}^T \lambda\boldsymbol{v} = \lambda\bar{\boldsymbol{v}}^T\boldsymbol{v} = \lambda(\bar{v}_1 v_1 + \bar{v}_2 v_2 + \cdots + \bar{v}_n v_n).$$

On the other hand,

$$\bar{\boldsymbol{v}}^T A\boldsymbol{v} = \bar{\boldsymbol{v}}^T \bar{A}\boldsymbol{v} = \overline{\boldsymbol{v}^T A}\boldsymbol{v} = \overline{\boldsymbol{v}^T A^T}\boldsymbol{v} = \overline{(A\boldsymbol{v})^T}\boldsymbol{v} = \overline{(\lambda\boldsymbol{v})^T}\boldsymbol{v} = \bar{\lambda}\bar{\boldsymbol{v}}^T\boldsymbol{v} = \lambda(\bar{v}_1 v_1 + \bar{v}_2 v_2 + \cdots + \bar{v}_n v_n).$$

Since $\bar{v}_1 v_1 + \bar{v}_2 v_2 + \cdots + \bar{v}_n v_n \neq 0$, it follows that $\lambda = \bar{\lambda}$. This means that λ is a real number. □