LINEAR ALGEBRA

W W L CHEN

Chapter 7

EIGENVALUES AND EIGENVECTORS

7.1. Introduction

EXAMPLE 7.1.1. Consider a function $f : \mathbb{R}^2 \to \mathbb{R}^2$, defined for every $(x, y) \in \mathbb{R}^2$ by $f(x, y) = (s, t)$, where

$$\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Note that

$$\begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \end{pmatrix} = 2 \begin{pmatrix} 3 \\ -1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \end{pmatrix} = 6 \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

On the other hand, note that

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

form a basis for $\mathbb{R}^2$. It follows that every $\mathbf{u} \in \mathbb{R}^2$ can be written uniquely in the form $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, where $c_1, c_2 \in \mathbb{R}$, so that

$$A\mathbf{u} = A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1 A\mathbf{v}_1 + c_2 A\mathbf{v}_2 = 2c_1\mathbf{v}_1 + 6c_2\mathbf{v}_2.$$

Note that in this case, the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ can be described easily in terms of the two special vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ and the two special numbers 2 and 6. Let us now examine how these special vectors and numbers arise. We hope to find numbers $\lambda \in \mathbb{R}$ and non-zero vectors $\mathbf{v} \in \mathbb{R}^2$ such that

$$\begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix} \mathbf{v} = \lambda \mathbf{v}.$$

Since

$$\lambda \mathbf{v} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{v} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \mathbf{v},$$

we must have

$$\left(\begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) \mathbf{v} = \mathbf{0}.$$

In other words, we must have

$$\begin{pmatrix} 3-\lambda & 3 \\ 1 & 5-\lambda \end{pmatrix} \mathbf{v} = \mathbf{0}. \qquad (1)$$

In order to have non-zero $\mathbf{v} \in \mathbb{R}^2$, we must therefore ensure that

$$\det \begin{pmatrix} 3-\lambda & 3 \\ 1 & 5-\lambda \end{pmatrix} = 0.$$

Hence $(3-\lambda)(5-\lambda) - 3 = 0$, with roots $\lambda_1 = 2$ and $\lambda_2 = 6$. Substituting $\lambda = 2$ into (1), we obtain

$$\begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} \mathbf{v} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Substituting $\lambda = 6$ into (1), we obtain

$$\begin{pmatrix} -3 & 3 \\ 1 & -1 \end{pmatrix} \mathbf{v} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

DEFINITION. Suppose that

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \qquad (2)$$

is an $n \times n$ matrix with entries in $\mathbb{R}$. Suppose further that there exist a number $\lambda \in \mathbb{R}$ and a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} = \lambda \mathbf{v}$. Then we say that λ is an eigenvalue of the matrix A, and that $\mathbf{v}$ is an eigenvector corresponding to the eigenvalue λ.

Suppose that λ is an eigenvalue of the $n \times n$ matrix A, and that $\mathbf{v}$ is an eigenvector corresponding to the eigenvalue λ. Then $A\mathbf{v} = \lambda\mathbf{v} = \lambda I \mathbf{v}$, where I is the $n \times n$ identity matrix, so that $(A - \lambda I)\mathbf{v} = \mathbf{0}$. Since $\mathbf{v} \in \mathbb{R}^n$ is non-zero, it follows that we must have

$$\det(A - \lambda I) = 0. \qquad (3)$$

In other words, we must have

$$\det \begin{pmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{pmatrix} = 0.$$

Note that (3) is a polynomial equation. Solving this equation (3) gives the eigenvalues of the matrix A. On the other hand, for any eigenvalue λ of the matrix A, the set

$$\{\mathbf{v} \in \mathbb{R}^n : (A - \lambda I)\mathbf{v} = \mathbf{0}\} \qquad (4)$$

is the nullspace of the matrix $A - \lambda I$, a subspace of $\mathbb{R}^n$.

DEFINITION. The polynomial (3) is called the characteristic polynomial of the matrix A. For any root λ of (3), the space (4) is called the eigenspace corresponding to the eigenvalue λ.

EXAMPLE 7.1.2. The matrix

$$\begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix}$$

has characteristic polynomial $(3-\lambda)(5-\lambda)-3=0$; in other words, $\lambda^2-8\lambda+12=0$. Hence the eigenvalues are $\lambda_1=2$ and $\lambda_2=6$, with corresponding eigenvectors

$$\mathbf{v}_1=\begin{pmatrix} 3 \\ -1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{v}_2=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

respectively. The eigenspace corresponding to the eigenvalue 2 is

$$\left\{\mathbf{v}\in\mathbb{R}^2 : \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}\mathbf{v}=\mathbf{0}\right\}=\left\{c\begin{pmatrix} 3 \\ -1 \end{pmatrix} : c\in\mathbb{R}\right\}.$$

The eigenspace corresponding to the eigenvalue 6 is

$$\left\{\mathbf{v}\in\mathbb{R}^2 : \begin{pmatrix} -3 & 3 \\ 1 & -1 \end{pmatrix}\mathbf{v}=\mathbf{0}\right\}=\left\{c\begin{pmatrix} 1 \\ 1 \end{pmatrix} : c\in\mathbb{R}\right\}.$$

EXAMPLE 7.1.3. Consider the matrix

$$A=\begin{pmatrix} -1 & 6 & -12 \\ 0 & -13 & 30 \\ 0 & -9 & 20 \end{pmatrix}.$$

To find the eigenvalues of A, we need to find the roots of

$$\det\begin{pmatrix} -1-\lambda & 6 & -12 \\ 0 & -13-\lambda & 30 \\ 0 & -9 & 20-\lambda \end{pmatrix}=0;$$

in other words, $(\lambda+1)(\lambda-2)(\lambda-5)=0$. The eigenvalues are therefore $\lambda_1=-1$, $\lambda_2=2$ and $\lambda_3=5$. An eigenvector corresponding to the eigenvalue -1 is a solution of the system

$$(A+I)\mathbf{v}=\begin{pmatrix} 0 & 6 & -12 \\ 0 & -12 & 30 \\ 0 & -9 & 21 \end{pmatrix}\mathbf{v}=\mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_1=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

An eigenvector corresponding to the eigenvalue 2 is a solution of the system

$$(A-2I)\mathbf{v}=\begin{pmatrix} -3 & 6 & -12 \\ 0 & -15 & 30 \\ 0 & -9 & 18 \end{pmatrix}\mathbf{v}=\mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_2=\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$

An eigenvector corresponding to the eigenvalue 5 is a solution of the system

$$(A-5I)\mathbf{v}=\begin{pmatrix} -6 & 6 & -12 \\ 0 & -18 & 30 \\ 0 & -9 & 15 \end{pmatrix}\mathbf{v}=\mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_3=\begin{pmatrix} 1 \\ -5 \\ -3 \end{pmatrix}.$$

Note that the three eigenspaces are all lines through the origin. Note also that the eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent, and so form a basis for $\mathbb{R}^3$.

Example 7.1.4. Consider the matrix

$$A = \begin{pmatrix} 17 & -10 & -5 \\ 45 & -28 & -15 \\ -30 & 20 & 12 \end{pmatrix}.$$

To find the eigenvalues of A, we need to find the roots of

$$\det \begin{pmatrix} 17-\lambda & -10 & -5 \\ 45 & -28-\lambda & -15 \\ -30 & 20 & 12-\lambda \end{pmatrix} = 0;$$

in other words, $(\lambda+3)(\lambda-2)^2 = 0$. The eigenvalues are therefore $\lambda_1 = -3$ and $\lambda_2 = 2$. An eigenvector corresponding to the eigenvalue -3 is a solution of the system

$$(A+3I)\mathbf{v} = \begin{pmatrix} 20 & -10 & -5 \\ 45 & -25 & -15 \\ -30 & 20 & 15 \end{pmatrix} \mathbf{v} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}.$$

An eigenvector corresponding to the eigenvalue 2 is a solution of the system

$$(A-2I)\mathbf{v} = \begin{pmatrix} 15 & -10 & -5 \\ 45 & -30 & -15 \\ -30 & 20 & 10 \end{pmatrix} \mathbf{v} = \mathbf{0}, \qquad \text{with roots} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \qquad \text{and} \qquad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}.$$

Note that the eigenspace corresponding to the eigenvalue -3 is a line through the origin, while the eigenspace corresponding to the eigenvalue 2 is a plane through the origin. Note also that the eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent, and so form a basis for $\mathbb{R}^3$.

Example 7.1.5. Consider the matrix

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

To find the eigenvalues of A, we need to find the roots of

$$\det \begin{pmatrix} 2-\lambda & -1 & 0 \\ 1 & 0-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{pmatrix} = 0;$$

in other words, $(\lambda-3)(\lambda-1)^2 = 0$. The eigenvalues are therefore $\lambda_1 = 3$ and $\lambda_2 = 1$. An eigenvector corresponding to the eigenvalue 3 is a solution of the system

$$(A-3I)\mathbf{v} = \begin{pmatrix} -1 & -1 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{v} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

An eigenvector corresponding to the eigenvalue 1 is a solution of the system

$$(A-I)\mathbf{v} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \mathbf{v} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Note that the eigenspace corresponding to the eigenvalue 3 is a line through the origin. On the other hand, the matrix

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

has rank 2, and so the eigenspace corresponding to the eigenvalue 1 is of dimension 1 and so is also a line through the origin. We can therefore only find two linearly independent eigenvectors, so that $\mathbb{R}^3$ does not have a basis consisting of linearly independent eigenvectors of the matrix A.

EXAMPLE 7.1.6. Consider the matrix

$$A = \begin{pmatrix} 3 & -3 & 2 \\ 1 & -1 & 2 \\ 1 & -3 & 4 \end{pmatrix}.$$

To find the eigenvalues of A, we need to find the roots of

$$\det \begin{pmatrix} 3-\lambda & -3 & 2 \\ 1 & -1-\lambda & 2 \\ 1 & -3 & 4-\lambda \end{pmatrix} = 0;$$

in other words, $(\lambda - 2)^3 = 0$. The eigenvalue is therefore $\lambda = 2$. An eigenvector corresponding to the eigenvalue 2 is a solution of the system

$$(A - 2I)\mathbf{v} = \begin{pmatrix} 1 & -3 & 2 \\ 1 & -3 & 2 \\ 1 & -3 & 2 \end{pmatrix} \mathbf{v} = \mathbf{0}, \qquad \text{with roots} \qquad \mathbf{v}_1 = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}.$$

Note now that the matrix

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

has rank 1, and so the eigenspace corresponding to the eigenvalue 2 is of dimension 2 and so is a plane through the origin. We can therefore only find two linearly independent eigenvectors, so that $\mathbb{R}^3$ does not have a basis consisting of linearly independent eigenvectors of the matrix A.

EXAMPLE 7.1.7. Suppose that λ is an eigenvalue of a matrix A, with corresponding eigenvector $\mathbf{v}$. Then

$$A^2\mathbf{v} = A(A\mathbf{v}) = A(\lambda\mathbf{v}) = \lambda(A\mathbf{v}) = \lambda(\lambda\mathbf{v}) = \lambda^2\mathbf{v}.$$

Hence λ^2 is an eigenvalue of the matrix A^2, with corresponding eigenvector $\mathbf{v}$. In fact, it can be proved by induction that for every natural number $k \in \mathbb{N}$, λ^k is an eigenvalue of the matrix A^k, with corresponding eigenvector $\mathbf{v}$.

EXAMPLE 7.1.8. Consider the matrix

$$\begin{pmatrix} 1 & 5 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}.$$

To find the eigenvalues of A, we need to find the roots of

$$\det \begin{pmatrix} 1-\lambda & 5 & 4 \\ 0 & 2-\lambda & 6 \\ 0 & 0 & 3-\lambda \end{pmatrix} = 0;$$

in other words, $(\lambda - 1)(\lambda - 2)(\lambda - 3) = 0$. It follows that the eigenvalues of the matrix A are given by the entries on the diagonal. In fact, this is true for all triangular matrices.

7.2. The Diagonalization Problem

EXAMPLE 7.2.1. Let us return to Examples 7.1.1 and 7.1.2, and consider again the matrix

$$A = \begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix}.$$

We have already shown that the matrix A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 6$, with corresponding eigenvectors

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

respectively. Since the eigenvectors form a basis for $\mathbb{R}^2$, every $\mathbf{u} \in \mathbb{R}^2$ can be written uniquely in the form

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2, \quad \text{where } c_1, c_2 \in \mathbb{R}, \tag{5}$$

and

$$A\mathbf{u} = 2c_1\mathbf{v}_1 + 6c_2\mathbf{v}_2. \tag{6}$$

Write

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \qquad \mathbf{u} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad A\mathbf{u} = \begin{pmatrix} s \\ t \end{pmatrix}.$$

Then (5) and (6) can be rewritten as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{7}$$

and

$$\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2c_1 \\ 6c_2 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{8}$$

respectively. If we write

$$P = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix},$$

then (7) and (8) become $\mathbf{u} = P\mathbf{c}$ and $A\mathbf{u} = PD\mathbf{c}$ respectively, so that $AP\mathbf{c} = PD\mathbf{c}$. Note that $\mathbf{c} \in \mathbb{R}^2$ is arbitrary. This implies that $(AP - PD)\mathbf{c} = \mathbf{0}$ for every $\mathbf{c} \in \mathbb{R}^2$. Hence we must have $AP = PD$. Since P is invertible, we conclude that

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

Note here that

$$P = (\,\mathbf{v}_1 \quad \mathbf{v}_2\,) \qquad \text{and} \qquad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Note also the crucial point that the eigenvectors of A form a basis for $\mathbb{R}^2$.

We now consider the problem in general.

PROPOSITION 7A. *Suppose that A is an $n \times n$ matrix, with entries in $\mathbb{R}$. Suppose further that A has eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, not necessarily distinct, with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Then*

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

where

$$P = (\,\mathbf{v}_1 \quad \ldots \quad \mathbf{v}_n\,) \qquad \text{and} \qquad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Proof. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, they form a basis for $\mathbb{R}^n$, so that every $\mathbf{u} \in \mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n, \quad \text{where } c_1, \ldots, c_n \in \mathbb{R}, \tag{9}$$

and

$$A\mathbf{u} = A(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1A\mathbf{v}_1 + \ldots + c_nA\mathbf{v}_n = \lambda_1c_1\mathbf{v}_1 + \ldots + \lambda_nc_n\mathbf{v}_n. \tag{10}$$

Writing

$$\mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

we see that (9) and (10) can be rewritten as

$$\mathbf{u} = P\mathbf{c} \qquad \text{and} \qquad A\mathbf{u} = P\begin{pmatrix} \lambda_1c_1 \\ \vdots \\ \lambda_nc_n \end{pmatrix} = PD\mathbf{c}$$

respectively, so that

$$AP\mathbf{c} = PD\mathbf{c}.$$

Note that $\mathbf{c} \in \mathbb{R}^n$ is arbitrary. This implies that $(AP - PD)\mathbf{c} = \mathbf{0}$ for every $\mathbf{c} \in \mathbb{R}^n$. Hence we must have $AP = PD$. Since the columns of P are linearly independent, it follows that P is invertible. Hence $P^{-1}AP = D$ as required. ◯

Example 7.2.2. Consider the matrix

$$A = \begin{pmatrix} -1 & 6 & -12 \\ 0 & -13 & 30 \\ 0 & -9 & 20 \end{pmatrix},$$

as in Example 7.1.3. We have $P^{-1}AP = D$, where

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & -5 \\ 0 & 1 & -3 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Example 7.2.3. Consider the matrix

$$A = \begin{pmatrix} 17 & -10 & -5 \\ 45 & -28 & -15 \\ -30 & 20 & 12 \end{pmatrix},$$

as in Example 7.1.4. We have $P^{-1}AP = D$, where

$$P = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Definition. Suppose that A is an $n \times n$ matrix, with entries in $\mathbb{R}$. We say that A is diagonalizable if there exists an invertible matrix P, with entries in $\mathbb{R}$, such that $P^{-1}AP$ is a diagonal matrix, with entries in $\mathbb{R}$.

It follows from Proposition 7A that an $n \times n$ matrix A with entries in $\mathbb{R}$ is diagonalizable if its eigenvectors form a basis for $\mathbb{R}^n$. In the opposite direction, we establish the following result.

PROPOSITION 7B. *Suppose that A is an $n \times n$ matrix, with entries in $\mathbb{R}$. Suppose further that A is diagonalizable. Then A has n linearly independent eigenvectors in $\mathbb{R}^n$.*

Proof. Suppose that A is diagonalizable. Then there exists an invertible matrix P, with entries in $\mathbb{R}$, such that $D = P^{-1}AP$ is a diagonal matrix, with entries in $\mathbb{R}$. Denote by $\mathbf{v}_1, \ldots, \mathbf{v}_n$ the columns of P; in other words, write

$$P = (\,\mathbf{v}_1 \quad \ldots \quad \mathbf{v}_n\,).$$

Also write

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Clearly we have $AP = PD$. It follows that

$$(\,A\mathbf{v}_1 \quad \ldots \quad A\mathbf{v}_n\,) = A\,(\,\mathbf{v}_1 \quad \ldots \quad \mathbf{v}_n\,) = (\,\mathbf{v}_1 \quad \ldots \quad \mathbf{v}_n\,) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = (\,\lambda_1\mathbf{v}_1 \quad \ldots \quad \lambda_n\mathbf{v}_n\,).$$

Equating columns, we obtain

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \qquad \ldots, \qquad A\mathbf{v}_n = \lambda_n\mathbf{v}_n.$$

It follows that A has eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$. Since P is invertible and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are the columns of P, it follows that the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. ○

In view of Propositions 7A and 7B, the question of diagonalizing a matrix A with entries in $\mathbb{R}$ is reduced to one of linear independence of its eigenvectors.

PROPOSITION 7C. *Suppose that A is an $n \times n$ matrix, with entries in $\mathbb{R}$. Suppose further that A has distinct eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.*

Proof. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent. Then there exist $c_1, \ldots, c_n \in \mathbb{R}$, not all zero, such that

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}. \tag{11}$$

Then

$$A(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1A\mathbf{v}_1 + \ldots + c_nA\mathbf{v}_n = \lambda_1c_1\mathbf{v}_1 + \ldots + \lambda_nc_n\mathbf{v}_n = \mathbf{0}. \tag{12}$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are all eigenvectors and hence non-zero, it follows that at least two numbers among $c_1, \ldots, c_n$ are non-zero, so that $c_1, \ldots, c_{n-1}$ are not all zero. Multiplying (11) by λ_n and subtracting from (12), we obtain

$$(\lambda_1 - \lambda_n)c_1\mathbf{v}_1 + \ldots + (\lambda_{n-1} - \lambda_n)c_{n-1}\mathbf{v}_{n-1} = \mathbf{0}.$$

Note that since $\lambda_1, \ldots, \lambda_n$ are distinct, the numbers $\lambda_1 - \lambda_n, \ldots, \lambda_{n-1} - \lambda_n$ are all non-zero. It follows that $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ are linearly dependent. To summarize, we can eliminate one eigenvector and the remaining ones are still linearly dependent. Repeating this argument a finite number of times, we arrive at a linearly dependent set of one eigenvector, clearly an absurdity. ◯

We now summarize our discussion in this section.

DIAGONALIZATION PROCESS. *Suppose that A is an $n \times n$ matrix with entries in $\mathbb{R}$.*
(1) Determine whether the n roots of the characteristic polynomial $\det(A - \lambda I)$ are real.
(2) If not, then A is not diagonalizable. If so, then find the eigenvectors corresponding to these eigenvalues. Determine whether we can find n linearly independent eigenvectors.
(3) If not, then A is not diagonalizable. If so, then write

$$P = (\,\mathbf{v}_1 \quad \ldots \quad \mathbf{v}_n\,) \qquad \text{and} \qquad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are the eigenvalues of A and where $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ are respectively their corresponding eigenvectors. Then $P^{-1}AP = D$.

7.3. Some Remarks

In all the examples we have discussed, we have chosen matrices A such that the characteristic polynomial $\det(A - \lambda I)$ has only real roots. However, there are matrices A where the characteristic polynomial has non-real roots. If we permit $\lambda_1, \ldots, \lambda_n$ to take values in $\mathbb{C}$ and permit "eigenvectors" to have entries in $\mathbb{C}$, then we may be able to "diagonalize" the matrix A, using matrices P and D with entries in $\mathbb{C}$. The details are similar.

Example 7.3.1. Consider the matrix

$$A = \begin{pmatrix} 1 & -5 \\ 1 & -1 \end{pmatrix}.$$

To find the eigenvalues of A, we need to find the roots of

$$\det\begin{pmatrix} 1-\lambda & -5 \\ 1 & -1-\lambda \end{pmatrix} = 0;$$

in other words, $\lambda^2+4=0$. Clearly there are no real roots, so the matrix A has no eigenvalues in $\mathbb{R}$. Try to show, however, that the matrix A can be "diagonalized" to the matrix

$$D = \begin{pmatrix} 2\mathrm{i} & 0 \\ 0 & -2\mathrm{i} \end{pmatrix}.$$

We also state without proof the following useful result which will guarantee many examples where the characteristic polynomial has only real roots.

PROPOSITION 7D. *Suppose that A is an $n \times n$ matrix, with entries in $\mathbb{R}$. Suppose further that A is symmetric. Then the characteristic polynomial $\det(A-\lambda I)$ has only real roots.*

We conclude this section by discussing an application of diagonalization. We illustrate this by an example.

EXAMPLE 7.3.2. Consider the matrix

$$A = \begin{pmatrix} 17 & -10 & -5 \\ 45 & -28 & -15 \\ -30 & 20 & 12 \end{pmatrix},$$

as in Example 7.2.3. Suppose that we wish to calculate A^{98}. Note that $P^{-1}AP=D$, where

$$P = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

It follows that $A = PDP^{-1}$, so that

$$A^{98} = \underbrace{(PDP^{-1})\ldots(PDP^{-1})}_{98} = PD^{98}P^{-1} = P\begin{pmatrix} 3^{98} & 0 & 0 \\ 0 & 2^{98} & 0 \\ 0 & 0 & 2^{98} \end{pmatrix}P^{-1}.$$

This is much simpler than calculating A^{98} directly.

7.4. An Application to Genetics

In this section, we discuss very briefly the problem of autosomal inheritance. Here we consider a set of two genes designated by G and g. Each member of the population inherits one from each parent, resulting in possible genotypes GG, Gg and gg. Furthermore, the gene G dominates the gene g, so that in the case of human eye colours, for example, people with genotype GG or Gg have brown eyes while people with genotype gg have blue eyes. It is also believed that each member of the population has equal probability of inheriting one or the other gene from each parent. The table below gives these peobabilities in detail. Here the genotypes of the parents are listed on top, and the genotypes of the offspring are listed on the left.

	$GG-GG$	$GG-Gg$	$GG-gg$	$Gg-Gg$	$Gg-gg$	$gg-gg$
GG	1	$\frac{1}{2}$	0	$\frac{1}{4}$	0	0
Gg	0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0
gg	0	0	0	$\frac{1}{4}$	$\frac{1}{2}$	1

EXAMPLE 7.4.1. Suppose that a plant breeder has a large population consisting of all three genotypes. At regular intervals, each plant he owns is fertilized with a plant known to have genotype GG, and is then disposed of and replaced by one of its offsprings. We would like to study the distribution of the three genotypes after n rounds of fertilization and replacements, where n is an arbitrary positive integer. Suppose that $GG(n)$, $Gg(n)$ and $gg(n)$ denote the proportion of each genotype after n rounds of fertilization and replacements, and that $GG(0)$, $Gg(0)$ and $gg(0)$ denote the initial proportions. Then clearly we have

$$GG(n) + Gg(n) + gg(n) = 1 \quad \text{for every } n = 0, 1, 2, \ldots.$$

On the other hand, the left hand half of the table above shows that for every $n = 1, 2, 3, \ldots$, we have

$$\begin{aligned} GG(n) &= GG(n-1) + \tfrac{1}{2} Gg(n-1), \\ Gg(n) &= \tfrac{1}{2} Gg(n-1) + gg(n-1), \end{aligned}$$

and

$$gg(n) = 0,$$

so that

$$\begin{pmatrix} GG(n) \\ Gg(n) \\ gg(n) \end{pmatrix} = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} GG(n-1) \\ Gg(n-1) \\ gg(n-1) \end{pmatrix}.$$

It follows that

$$\begin{pmatrix} GG(n) \\ Gg(n) \\ gg(n) \end{pmatrix} = A^n \begin{pmatrix} GG(0) \\ Gg(0) \\ gg(0) \end{pmatrix} \quad \text{for every } n = 1, 2, 3, \ldots,$$

where the matrix

$$A = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

has eigenvalues $\lambda_1 = 1$, $\lambda_2 = 0$, $\lambda_3 = 1/2$, with respective eigenvectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

We therefore write

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1/2 \end{pmatrix}, \qquad \text{with} \qquad P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & -1 & -2 \end{pmatrix}.$$

Then $P^{-1}AP = D$, so that $A = PDP^{-1}$, and so

$$\begin{aligned} A^n = PD^nP^{-1} &= \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1/2^n \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & -1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 - 1/2^n & 1 - 1/2^{n-1} \\ 0 & 1/2^n & 1/2^{n-1} \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

It follows that

$$\begin{aligned}
\begin{pmatrix} GG(n) \\ Gg(n) \\ gg(n) \end{pmatrix} &= \begin{pmatrix} 1 & 1-1/2^n & 1-1/2^{n-1} \\ 0 & 1/2^n & 1/2^{n-1} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} GG(0) \\ Gg(0) \\ gg(0) \end{pmatrix} \\
&= \begin{pmatrix} GG(0)+Gg(0)+gg(0)-Gg(0)/2^n-gg(0)/2^{n-1} \\ Gg(0)/2^n+gg(0)/2^{n-1} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 1-Gg(0)/2^n-gg(0)/2^{n-1} \\ Gg(0)/2^n+gg(0)/2^{n-1} \\ 0 \end{pmatrix} \to \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{as } n \to \infty.
\end{aligned}$$

This means that nearly the whole crop will have genotype GG.

Problems for Chapter 7

1. For each of the following 2×2 matrices, find all eigenvalues and describe the eigenspace of the matrix; if possible, diagonalize the matrix:

   a) $\begin{pmatrix} 3 & 4 \\ -2 & -3 \end{pmatrix}$ b) $\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$

2. For each of the following 3×3 matrices, find all eigenvalues and describe the eigenspace of the matrix; if possible, diagonalize the matrix:

   a) $\begin{pmatrix} -2 & 9 & -6 \\ 1 & -2 & 0 \\ 3 & -9 & 5 \end{pmatrix}$ b) $\begin{pmatrix} 2 & -1 & -1 \\ 0 & 3 & 2 \\ -1 & 1 & 2 \end{pmatrix}$

   c) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

3. Consider the matrices $A = \begin{pmatrix} -10 & 6 & 3 \\ -26 & 16 & 8 \\ 16 & -10 & -5 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & -6 & -16 \\ 0 & 17 & 45 \\ 0 & -6 & -16 \end{pmatrix}$.

   a) Show that A and B have the same eigenvalues.
   b) Reduce A and B to the same disgonal matrix.
   c) Explain why there is an invertible matrix R such that $R^{-1}AR = B$.

4. Find A^8 and B^8, where A and B are the two matrices in Problem 3.

5. Suppose that $\theta \in \mathbb{R}$ is not an integer multiple of π. Show that the matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ does not have an eigenvector in $\mathbb{R}^2$.

6. Consider the matrix $A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$, where $\theta \in \mathbb{R}$.

   a) Show that A has an eigenvector in $\mathbb{R}^2$ with eigenvalue 1.
   b) Show that any vector $\mathbf{v} \in \mathbb{R}^2$ perpendicular to the eigenvector in part (a) must satisfy $A\mathbf{v} = -\mathbf{v}$.

7. Let $a \in \mathbb{R}$ be non-zero. Show that the matrix $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ cannot be diagonalized.