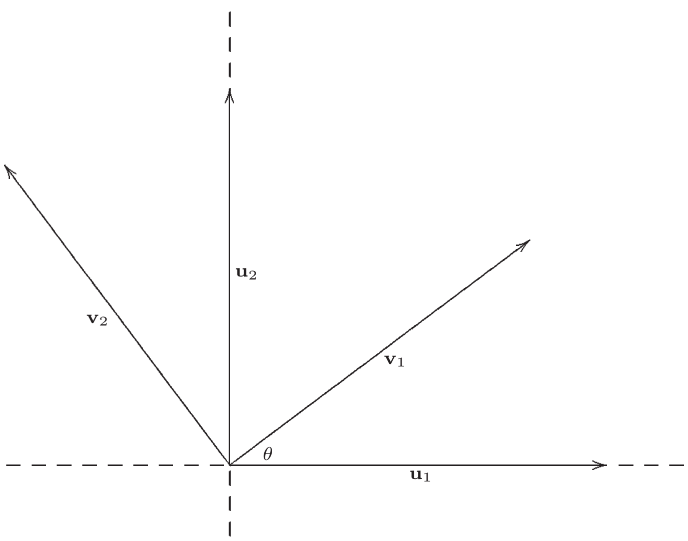
Chapter 10

ORTHOGONAL MATRICES

10.1. Introduction

DEFINITION. A square matrix A with real entries and satisfying the condition $A^{-1} = A^t$ is called an orthogonal matrix.

EXAMPLE 10.1.1. Consider the euclidean space $\mathbb{R}^2$ with the euclidean inner product. The vectors $\mathbf{u}_1 = (1,0)$ and $\mathbf{u}_2 = (0,1)$ form an orthonormal basis $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$. Let us now rotate $\mathbf{u}_1$ and $\mathbf{u}_2$ anticlockwise by an angle θ to obtain $\mathbf{v}_1 = (\cos\theta, \sin\theta)$ and $\mathbf{v}_2 = (-\sin\theta, \cos\theta)$. Then $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is also an orthonormal basis.



The transition matrix from the basis $\mathcal{C}$ to the basis $\mathcal{B}$ is given by

$$P = ([\mathbf{v}_1]_{\mathcal{B}} \quad [\mathbf{v}_2]_{\mathcal{B}}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Clearly

$$P^{-1} = P^t = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

In fact, our example is a special case of the following general result.

PROPOSITION 10A. *Suppose that $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are two orthonormal bases of a real inner product space V. Then the transition matrix P from the basis $\mathcal{C}$ to the basis $\mathcal{B}$ is an orthogonal matrix.*

EXAMPLE 10.1.2. The matrix

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

is orthogonal, since

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

Note also that the row vectors of A, namely $(1/3, -2/3, 2/3)$, $(2/3, -1/3, -2/3)$ and $(2/3, 2/3, 1/3)$ are orthonormal. So are the column vectors of A.

In fact, our last observation is not a coincidence.

PROPOSITION 10B. *Suppose that A is an $n \times n$ matrix with real entries. Then*
(a) A is orthogonal if and only if the row vectors of A form an orthonormal basis of $\mathbb{R}^n$ under the euclidean inner product; and
(b) A is orthogonal if and only if the column vectors of A form an orthonormal basis of $\mathbb{R}^n$ under the euclidean inner product.

PROOF. We shall only prove (a), since the proof of (b) is almost identical. Let $\mathbf{r}_1, \ldots, \mathbf{r}_n$ denote the row vectors of A. Then

$$AA^t = \begin{pmatrix} \mathbf{r}_1 \cdot \mathbf{r}_1 & \ldots & \mathbf{r}_1 \cdot \mathbf{r}_n \\ \vdots & & \vdots \\ \mathbf{r}_n \cdot \mathbf{r}_1 & \ldots & \mathbf{r}_n \cdot \mathbf{r}_n \end{pmatrix}.$$

It follows that $AA^t = I$ if and only if for every $i, j = 1, \ldots, n$, we have

$$\mathbf{r}_i \cdot \mathbf{r}_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

if and only if $\mathbf{r}_1, \ldots, \mathbf{r}_n$ are orthonormal. ○

PROPOSITION 10C. *Suppose that A is an $n \times n$ matrix with real entries. Suppose further that the inner product in $\mathbb{R}^n$ is the euclidean inner product. Then the following are equivalent:*
(a) A is orthogonal.
(b) For every $\mathbf{x} \in \mathbb{R}^n$, we have $\|A\mathbf{x}\| = \|\mathbf{x}\|$.
(c) For every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have $A\mathbf{u} \cdot A\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$.

PROOF. ((a)⇒(b)) Suppose that A is orthogonal, so that $A^tA = I$. It follows that for every $\mathbf{x} \in \mathbb{R}^n$, we have

$$\|A\mathbf{x}\|^2 = A\mathbf{x} \cdot A\mathbf{x} = \mathbf{x}^t A^t A\mathbf{x} = \mathbf{x}^t I\mathbf{x} = \mathbf{x}^t\mathbf{x} = \mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2.$$

((b)⇒(c)) Suppose that $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for every $\mathbf{x} \in \mathbb{R}^n$. Then for every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have

$$A\mathbf{u} \cdot A\mathbf{v} = \tfrac{1}{4}\|A\mathbf{u}+A\mathbf{v}\|^2 - \tfrac{1}{4}\|A\mathbf{u}-A\mathbf{v}\|^2 = \tfrac{1}{4}\|A(\mathbf{u}+\mathbf{v})\|^2 - \tfrac{1}{4}\|A(\mathbf{u}-\mathbf{v})\|^2 = \tfrac{1}{4}\|\mathbf{u}+\mathbf{v}\|^2 - \tfrac{1}{4}\|\mathbf{u}-\mathbf{v}\|^2 = \mathbf{u} \cdot \mathbf{v}.$$

((c)⇒(a)) Suppose that $A\mathbf{u} \cdot A\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ for every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Then

$$I\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{v} = A\mathbf{u} \cdot A\mathbf{v} = \mathbf{v}^t A^t A\mathbf{u} = A^t A\mathbf{u} \cdot \mathbf{v},$$

so that

$$(A^tA - I)\mathbf{u} \cdot \mathbf{v} = 0.$$

In particular, this holds when $\mathbf{v} = (A^tA - I)\mathbf{u}$, so that

$$(A^tA - I)\mathbf{u} \cdot (A^tA - I)\mathbf{u} = 0,$$

whence

$$(A^tA - I)\mathbf{u} = \mathbf{0}, \tag{1}$$

in view of Proposition 9A(d). But then (1) is a system of n homogeneous linear equations in n unknowns satisfied by every $\mathbf{u} \in \mathbb{R}^n$. Hence the coefficient matrix $A^tA - I$ must be the zero matrix, and so $A^tA = I$. ○

PROOF OF PROPOSITION 10A. For every $\mathbf{u} \in V$, we can write

$$\mathbf{u} = \beta_1\mathbf{u}_1 + \ldots + \beta_n\mathbf{u}_n = \gamma_1\mathbf{v}_1 + \ldots + \gamma_n\mathbf{v}_n, \quad \text{where } \beta_1, \ldots, \beta_n, \gamma_1, \ldots, \gamma_n \in \mathbb{R},$$

and where $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are two orthonormal bases of V. Then

$$\begin{aligned}\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle = \langle \beta_1\mathbf{u}_1 + \ldots + \beta_n\mathbf{u}_n, \beta_1\mathbf{u}_1 + \ldots + \beta_n\mathbf{u}_n \rangle &= \sum_{i=1}^n \sum_{j=1}^n \beta_i\beta_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{i=1}^n \beta_i^2 \\ &= (\beta_1, \ldots, \beta_n) \cdot (\beta_1, \ldots, \beta_n).\end{aligned}$$

Similarly,

$$\begin{aligned}\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle = \langle \gamma_1\mathbf{v}_1 + \ldots + \gamma_n\mathbf{v}_n, \gamma_1\mathbf{v}_1 + \ldots + \gamma_n\mathbf{v}_n \rangle &= \sum_{i=1}^n \sum_{j=1}^n \gamma_i\gamma_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^n \gamma_i^2 \\ &= (\gamma_1, \ldots, \gamma_n) \cdot (\gamma_1, \ldots, \gamma_n).\end{aligned}$$

It follows that in $\mathbb{R}^n$ with the euclidean norm, we have $\|[\mathbf{u}]_\mathcal{B}\| = \|[\mathbf{u}]_\mathcal{C}\|$, and so $\|P[\mathbf{u}]_\mathcal{C}\| = \|[\mathbf{u}]_\mathcal{C}\|$ for every $\mathbf{u} \in V$. Hence $\|P\mathbf{x}\| = \|\mathbf{x}\|$ holds for every $\mathbf{x} \in \mathbb{R}^n$. It now follows from Proposition 10C that P is orthogonal. ○

10.2. Eigenvalues and Eigenvectors

In this section, we give a brief review on eigenvalues and eigenvectors first discussed in Chapter 7.

Suppose that

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

is an $n \times n$ matrix with real entries. 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. In this case, we have $A\mathbf{v} = \lambda\mathbf{v} = \lambda I\mathbf{v}$, 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

$$\det(A - \lambda I) = 0. \tag{2}$$

In other words, we must have

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

Note that (2) is a polynomial equation. The polynomial $\det(A - \lambda I)$ is called the characteristic polynomial of the matrix A. Solving this equation (2) gives the eigenvalues of the matrix A.

On the other hand, for any eigenvalue λ of the matrix A, the set

$$\{\mathbf{v} \in \mathbb{R}^n : (A - \lambda I)\mathbf{v} = \mathbf{0}\} \tag{3}$$

is the nullspace of the matrix $A - \lambda I$, and forms a subspace of $\mathbb{R}^n$. This space (3) is called the eigenspace corresponding to the eigenvalue λ.

Suppose now that A has eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, not necessarily distinct, with corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$, and that $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent. Then it can be shown that

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

where

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

In fact, we say that A is diagonalizable if there exists an invertible matrix P with real entries such that $P^{-1}AP$ is a diagonal matrix with real entries. It follows that A is diagonalizable if its eigenvectors form a basis of $\mathbb{R}^n$. In the opposite direction, one can show that if A is diagonalizable, then it has n linearly independent eigenvectors in $\mathbb{R}^n$. It therefore follows that the question of diagonalizing a matrix A with real entries is reduced to one of linear independence of its eigenvectors.

We now summarize our discussion so far.

DIAGONALIZATION PROCESS. *Suppose that A is an $n \times n$ matrix with real entries.*
(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 $P^{-1}AP = D$.

In particular, it can be shown that if 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. It follows that all such matrices A are diagonalizable.

10.3. Orthonormal Diagonalization

We now consider the euclidean space $\mathbb{R}^n$ an as inner product space with the euclidean inner product. Given any $n \times n$ matrix A with real entries, we wish to find out whether there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

Recall that in the Diagonalization process discussed in the last section, the columns of the matrix P are eigenvectors of A, and these vectors form a basis of $\mathbb{R}^n$. It follows from Proposition 10B that this basis is orthonormal if and only if the matrix P is orthogonal.

Definition. An $n \times n$ matrix A with real entries is said to be orthogonally diagonalizable if there exists an orthogonal matrix P with real entries such that $P^{-1}AP = P^tAP$ is a diagonal matrix with real entries.

First of all, we would like to determine which matrices are orthogonally diagonalizable. For those that are, we then need to discuss how we may find an orthogonal matrix P to carry out the diagonalization.

To study the first question, we have the following result which gives a restriction on those matrices that are orthogonally diagonalizable.

PROPOSITION 10D. *Suppose that A is a orthogonally diagonalizable matrix with real entries. Then A is symmetric.*

Proof. Suppose that A is orthogonally diagonalizable. Then there exists an orthogonal matrix P and a diagonal matrix D, both with real entries and such that $P^tAP = D$. Since $PP^t = P^tP = I$ and $D^t = D$, we have

$$A = PDP^t = PD^tP^t,$$

so that

$$A^t = (PD^tP^t)^t = (P^t)^t(D^t)^tP^t = PDP^t = A,$$

whence A is symmetric. ○

Our first question is in fact answered by the following result which we state without proof.

PROPOSITION 10E. *Suppose that A is an $n \times n$ matrix with real entries. Then it is orthogonally diagonalizable if and only if it is symmetric.*

The remainder of this section is devoted to finding a way to orthogonally diagonalize a symmetric matrix with real entries. We begin by stating without proof the following result. The proof requires results from the theory of complex vector spaces.

PROPOSITION 10F. *Suppose that A is a symmetric matrix with real entries. Then all the eigenvalues of A are real.*

Our idea here is to follow the Diagonalization process discussed in the last section, knowing that since A is diagonalizable, we shall find a basis of $\mathbb{R}^n$ consisting of eigenvectors of A. We may then wish to orthogonalize this basis by the Gram-Schmidt process. This last step is considerably simplified in view of the following result.

PROPOSITION 10G. *Suppose that $\mathbf{u}_1$ and $\mathbf{u}_2$ are eigenvectors of a symmetric matrix A with real entries, corresponding to distinct eigenvalues λ_1 and λ_2 respectively. Then $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$. In other words, eigenvectors of a symmetric real matrix corresponding to distinct eigenvalues are orthogonal.*

PROOF. Note that if we write $\mathbf{u}_1$ and $\mathbf{u}_2$ as column matrices, then since A is symmetric, we have

$$A\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{u}_2^t A\mathbf{u}_1 = \mathbf{u}_2^t A^t \mathbf{u}_1 = (A\mathbf{u}_2)^t \mathbf{u}_1 = \mathbf{u}_1 \cdot A\mathbf{u}_2.$$

It follows that

$$\lambda_1 \mathbf{u}_1 \cdot \mathbf{u}_2 = A\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{u}_1 \cdot A\mathbf{u}_2 = \mathbf{u}_1 \cdot \lambda_2 \mathbf{u}_2,$$

so that $(\lambda_1 - \lambda_2)(\mathbf{u}_1 \cdot \mathbf{u}_2) = 0$. Since $\lambda_1 \neq \lambda_2$, we must have $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$. ◯

We can now follow the procedure below.

ORTHOGONAL DIAGONALIZATION PROCESS. *Suppose that A is a symmetric $n \times n$ matrix with real entries.*

(1) Determine the n real roots $\lambda_1, \ldots, \lambda_n$ of the characteristic polynomial $\det(A - \lambda I)$, and find n linearly independent eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of A corresponding to these eigenvalues as in the Diagonalization process.

(2) Apply the Gram-Schmidt orthogonalization process to the eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ to obtain orthogonal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of A, noting that eigenvectors corresponding to distinct eigenvalues are already orthogonal.

(3) Normalize the orthogonal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ to obtain orthonormal eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$ of A. These form an orthonormal basis of $\mathbb{R}^n$. Furthermore, write

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

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

REMARK. Note that if we apply the Gram-Schmidt orthogonalization process to eigenvectors corresponding to the same eigenvalue, then the new vectors that result from this process are also eigenvectors corresponding to this eigenvalue. Why?

EXAMPLE 10.3.1. Consider the matrix

$$A=\begin{pmatrix}2&2&1\\2&5&2\\1&2&2\end{pmatrix}.$$

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

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

in other words, $(\lambda-7)(\lambda-1)^2=0$. The eigenvalues are therefore $\lambda_1=7$ and (double root) $\lambda_2=\lambda_3=1$. An eigenvector corresponding to $\lambda_1=7$ is a solution of the system

$$(A-7I)\mathbf{u}=\begin{pmatrix}-5&2&1\\2&-2&2\\1&2&-5\end{pmatrix}\mathbf{u}=\mathbf{0},\qquad\text{with root}\qquad \mathbf{u}_1=\begin{pmatrix}1\\2\\1\end{pmatrix}.$$

Eigenvectors corresponding to $\lambda_2=\lambda_3=1$ are solutions of the system

$$(A-I)\mathbf{u}=\begin{pmatrix}1&2&1\\2&4&2\\1&2&1\end{pmatrix}\mathbf{u}=\mathbf{0},\qquad\text{with roots}\qquad \mathbf{u}_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}\qquad\text{and}\qquad \mathbf{u}_3=\begin{pmatrix}2\\-1\\0\end{pmatrix}$$

which are linearly independent. Next, we apply the Gram-Schmidt orthogonalization process to $\mathbf{u}_2$ and $\mathbf{u}_3$, and obtain

$$\mathbf{v}_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}\qquad\text{and}\qquad \mathbf{v}_3=\begin{pmatrix}1\\-1\\1\end{pmatrix}$$

which are now orthogonal to each other. Note that we do not have to do anything to $\mathbf{u}_1$ at this stage, in view of Proposition 10G. We now conclude that

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

form an orthogonal basis of $\mathbb{R}^3$. Normalizing each of these, we obtain respectively

$$\mathbf{w}_1=\begin{pmatrix}1/\sqrt{6}\\2/\sqrt{6}\\1/\sqrt{6}\end{pmatrix},\qquad \mathbf{w}_2=\begin{pmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{pmatrix},\qquad \mathbf{w}_3=\begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}.$$

We now take

$$P=(\,\mathbf{w}_1\quad\mathbf{w}_2\quad\mathbf{w}_3\,)=\begin{pmatrix}1/\sqrt{6}&1/\sqrt{2}&1/\sqrt{3}\\2/\sqrt{6}&0&-1/\sqrt{3}\\1/\sqrt{6}&-1/\sqrt{2}&1/\sqrt{3}\end{pmatrix}.$$

Then

$$P^{-1}=P^t=\begin{pmatrix}1/\sqrt{6}&2/\sqrt{6}&1/\sqrt{6}\\1/\sqrt{2}&0&-1/\sqrt{2}\\1/\sqrt{3}&-1/\sqrt{3}&1/\sqrt{3}\end{pmatrix}\qquad\text{and}\qquad P^tAP=\begin{pmatrix}7&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

EXAMPLE 10.3.2. 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 $\lambda_1=-1$ is a solution of the system

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

An eigenvector corresponding to $\lambda_2=2$ is a solution of the system

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

An eigenvector corresponding to $\lambda_3=5$ is a solution of the system

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

Note that while $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ correspond to distinct eigenvalues of A, they are not orthogonal. The matrix A is not symmetric, and so Proposition 10G does not apply in this case.

EXAMPLE 10.3.3. Consider the matrix

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

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

$$\det \begin{pmatrix} 5-\lambda & -2 & 0 \\ -2 & 6-\lambda & 2 \\ 0 & 2 & 7-\lambda \end{pmatrix} = 0;$$

in other words, $(\lambda-3)(\lambda-6)(\lambda-9)=0$. The eigenvalues are therefore $\lambda_1=3$, $\lambda_2=6$ and $\lambda_3=9$. An eigenvector corresponding $\lambda_1=3$ is a solution of the system

$$(A-3I)\mathbf{u} = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 3 & 2 \\ 0 & 2 & 4 \end{pmatrix} \mathbf{u} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{u}_1 = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}.$$

An eigenvector corresponding to $\lambda_2=6$ is a solution of the system

$$(A-6I)\mathbf{u} = \begin{pmatrix} -1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix} \mathbf{u} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{u}_2 = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}.$$

An eigenvector corresponding to $\lambda_3 = 9$ is a solution of the system

$$(A - 9I)\mathbf{u} = \begin{pmatrix} -4 & -2 & 0 \\ -2 & -3 & 2 \\ 0 & 2 & -2 \end{pmatrix} \mathbf{u} = \mathbf{0}, \qquad \text{with root} \qquad \mathbf{u}_3 = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}.$$

Note now that the eigenvalues are distinct, so it follows from Proposition 10G that $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are orthogonal, so we do not have to apply Step (2) of the Orthogonal diagonalization process. Normalizing each of these vectors, we obtain respectively

$$\mathbf{w}_1 = \begin{pmatrix} 2/3 \\ 2/3 \\ -1/3 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 2/3 \\ -1/3 \\ 2/3 \end{pmatrix}, \qquad \mathbf{w}_3 = \begin{pmatrix} -1/3 \\ 2/3 \\ 2/3 \end{pmatrix}.$$

We now take

$$P = (\, \mathbf{w}_1 \quad \mathbf{w}_2 \quad \mathbf{w}_3 \,) = \begin{pmatrix} 2/3 & 2/3 & -1/3 \\ 2/3 & -1/3 & 2/3 \\ -1/3 & 2/3 & 2/3 \end{pmatrix}.$$

Then

$$P^{-1} = P^t = \begin{pmatrix} 2/3 & 2/3 & -1/3 \\ 2/3 & -1/3 & 2/3 \\ -1/3 & 2/3 & 2/3 \end{pmatrix} \qquad \text{and} \qquad P^t A P = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Problems for Chapter 10

1. Prove Proposition 10B(b).

2. Let

$$A = \begin{pmatrix} a+b & b-a \\ a-b & b+a \end{pmatrix},$$

where $a, b \in \mathbb{R}$. Determine when A is orthogonal.

3. Suppose that A is an orthogonal matrix with real entries. Prove that
 a) A^{-1} is an orthogonal matrix; and
 b) $\det A = \pm 1$.

4. Suppose that A and B are orthogonal matrices with real entries. Prove that AB is orthogonal.

5. Verify that for every $a \in \mathbb{R}$, the matrix

$$A = \frac{1}{1+2a^2} \begin{pmatrix} 1 & -2a & 2a^2 \\ 2a & 1-2a^2 & -2a \\ 2a^2 & 2a & 1 \end{pmatrix}$$

is orthogonal.

6. Suppose that λ is an eigenvalue of an orthogonal matrix A with real entries. Prove that $1/\lambda$ is also an eigenvalue of A.

7. Suppose that

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is an orthogonal matrix with real entries. Explain why $a^2 + b^2 = c^2 + d^2 = 1$ and $ac + bd = 0$, and quote clearly any result that you use. Deduce that A has one of the two possible forms

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad \text{or} \qquad A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix},$$

where $\theta \in [0, 2\pi)$.

8. Consider the matrix

$$A = \begin{pmatrix} 1 & -\sqrt{6} & \sqrt{3} \\ -\sqrt{6} & 2 & \sqrt{2} \\ \sqrt{3} & \sqrt{2} & 3 \end{pmatrix}.$$

 a) Find the characteristic polynomial of A and show that A has eigenvalues 4 (twice) and -2.
 b) Find an eigenvector of A corresponding to the eigenvalue -2.
 c) Find two orthogonal eigenvectors of A corresponding to the eigenvalue 4.
 d) Find an orthonormal basis of $\mathbb{R}^3$ consisting of eigenvectors of A.
 e) Using the orthonormal basis in part (d), find a matrix P such that $P^t A P$ is a diagonal matrix.

9. Apply the Orthogonal diagonalization process to each of the following matrices:

a) $A = \begin{pmatrix} 5 & 0 & 6 \\ 0 & 11 & 6 \\ 6 & 6 & -2 \end{pmatrix}$ b) $A = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

c) $A = \begin{pmatrix} 1 & -4 & 2 \\ -4 & 1 & -2 \\ 2 & -2 & -2 \end{pmatrix}$ d) $A = \begin{pmatrix} 2 & 0 & 36 \\ 0 & 3 & 0 \\ 36 & 0 & 23 \end{pmatrix}$

e) $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ f) $A = \begin{pmatrix} -7 & 24 & 0 & 0 \\ 24 & 7 & 0 & 0 \\ 0 & 0 & -7 & 24 \\ 0 & 0 & 24 & 7 \end{pmatrix}$

10. Suppose that B is an $m \times n$ matrix with real entries. Prove that the matrix $A = B^t B$ has an orthonormal set of n eigenvectors.