THEORY AND METHODS IN DETERMINING THE EIGENVALUES
AND EIGENVECTORS OF A MATRIX

APPROVED:

Major Professor

Minor Professor

Director of the Department of Mathematics

Dean of the Graduate School
THEORY AND METHODS IN DETERMINING THE EIGENVALUES
AND EIGENVECTORS OF A MATRIX

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Jerry Herschel Waldon, B. S.
Denton, Texas
August, 1965
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Unitary Spaces</td>
<td></td>
</tr>
<tr>
<td>Linear Operators</td>
<td></td>
</tr>
<tr>
<td>Eigenvalues and Hermitian Operators</td>
<td></td>
</tr>
<tr>
<td>Matrices</td>
<td></td>
</tr>
<tr>
<td>Eigenvalues and Matrices</td>
<td></td>
</tr>
<tr>
<td>Diagonalization of Matrices</td>
<td></td>
</tr>
<tr>
<td>The Companion Matrix</td>
<td></td>
</tr>
<tr>
<td>Bordering Matrices</td>
<td></td>
</tr>
<tr>
<td>II. THE ESCALATOR METHOD</td>
<td>51</td>
</tr>
<tr>
<td>III. THE METHOD OF ORTHOGONALIZATION OF SUCCESSIVE ITERATIONS</td>
<td>92</td>
</tr>
<tr>
<td>IV. TRANSFORMATION OF SYMMETRIC MATRICES TO TRIDIAGONAL FORM BY MEANS OF ROTATION</td>
<td>109</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>123</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

In the numerous problems of matrix algebra, one finds the problem of determining the eigenvalues and eigenvectors of a matrix quite frequently. The theory and methods leading to the solution of the eigenvalue and eigenvector problem are of considerable interest. The relation between vector spaces, matrices, eigenvalues, and eigenvectors is to be considered in this chapter, with particular concentration directed toward eigenvalues and eigenvectors. Three methods for determining the eigenvalues and eigenvectors shall be developed in the following chapters with detailed examples of the methods.

Unitary Spaces

**Definition 1**: A set of elements $x, y, \ldots$, which shall be called vectors, satisfying the following properties, is called a vector space $V$.

I. If each of $x$ and $y$ is an element of $V$, there exists a unique element $x + y$ in $V$ called the sum of $x$ and $y$.

II. If each of $x$, $y$, and $z$ is an element of $V$ and each of $a$, $b$ is a complex number, there exist unique vectors $ax$, $bx$, and $ay$ in $V$ such that
1. $a(x + y) = ax + ay$
2. $(ab)x = a(bx)$
3. $(a + b)x = ax + ab$
4. $(1)x = x$, where $(1)$ is the complex number one.
5. $x + y = y + x$
6. $x + (y + z) = (x + y) + z$

III. If $x$ is an element of $V$, there exists an element $\theta$ in $V$ such that $x + \theta = \theta + x = x$; furthermore, if $x$ is an element of $V$, there exists an element $-x$ in $V$ such that $x + (-x)$ equals $\theta$. The expression $x - y$ shall mean the sum $x + (-y)$. Hence, one can write $\theta = x - x$ for any $x \in V$.

If, in addition, the vector space satisfies the following condition,

IV. If each of $x$ and $y$ is an element of $V$ and $a$ is a complex number with complex conjugate $\overline{a}$, there exists a uniquely defined complex number $(x, y)$, called the inner product of $x$ and $y$, which satisfies the following

1. $(x, y) = (y, x)$
2. $(ax, y) = \overline{a}(x, y)$
3. $(x, x) \geq 0$
4. $(x, y + z) = (x, y) + (x, z)$
5. $(x, x) = 0$ if and only if $x = \theta$

then $V$ is called a unitary space $U$. 
Note:

\[(x + y, z) = (x, z) + (y, z)\]
\[(x, ay) = a(x, y)\]

Proof: By part IV, property 4, one sees that

\[(z, x + y) = (z, x) + (z, y)\]
\[(x + y, z) = (x, z) + (y, z)\]

and by part IV, property 2,

\[(ay, x) = \overline{a}(y, x)\]
\[(ay, x) = \overline{a}(y, x)\]
\[(x, ay) = \overline{a}(y, x)\]

Definition 2: An element \(x\), which is an ordered \(n\)-tuple of elements from a field \(F=(a_1, a_2, \ldots, a_n)\), is a vector with \(n\) components \(a_1\).

Let \(x = (a_1, a_2, \ldots, a_n), y = (b_1, b_2, \ldots, b_n)\) be vectors with complex components \(a_i, b_i\) respectively. Then define

1. \(x + y = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)\)
2. \(gx = (ga_1, ga_2, \ldots, ga_n)\), where \(g\) is complex
3. \((x, y) = \sum_{i=1}^{n} a_i b_i\)
4. \(x = \sum_{k} (a_1, a_2, \ldots, a_n)\), where \(a_i = 0\) for \(i = 1, 2, \ldots, k-1, k+1, \ldots, n\) and \(a_k = 1\).
5. \(x = 0_v = (0, 0, 0, \ldots, 0)\)
6. $x = y$ if $a_i = b_i$ for $i = 1, 2, \ldots, n$

7. $-x = (-a_1, -a_2, \ldots, -a_n)$

**Example 1:** The set of ordered $n$-tuples of complex numbers is a unitary space $U_n$.

**Proof:** I. Let each of $x = (a_1, a_2, \ldots, a_n)$, $y = (b_1, b_2, \ldots, b_n)$ and $z = (e_1, e_2, \ldots, e_n)$ be an element of $U_n$. By I above, one sees that

$$x + y = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n) \in U_n.$$

Assume $x + y = (g_1, g_2, \ldots, g_n) \in U_n$, then $(g_1, g_2, \ldots, g_n)$ is equivalent to $(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$ and $g_i = a_i + b_i$ by property 6 above. Therefore, $x + y$ is unique since $a_i + b_i$ is unique, because the $+$ operation for complex numbers is unique.

**Proof:** II. Let each of $d, c$ be a complex number. Then

1. $d(x + y) = d(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$
   
   $= (d(a_1 + b_1), d(a_2 + b_2), \ldots, d(a_n + b_n))$
   
   $= (da_1 + db_1, da_2 + db_2, \ldots, da_n + db_n)$
   
   $= (da_1, da_2, \ldots, da_n) + (db_1, db_2, \ldots, db_n)$
   
   $= d(a_1, a_2, \ldots, a_n) + d(b_1, b_2, \ldots, b_n)$
   
   $= dx + dy$.

2. $(dc)x = dc(a_1, a_2, \ldots, a_n)$

   $= (dca_1, dca_2, \ldots, dca_n)$

   $= (d(ca_1), d(ca_2), \ldots, d(ca_n))$

   $= d(ca_1, ca_2, \ldots, ca_n)$

   $= d(c(a_1, a_2, \ldots, a_n))$

   $= d(cx)$. 
3. \((d + c)x = (d + c)(a_1, a_2, \ldots, a_n)\)
\[= ((d + c)a_1, (d + c)a_2, \ldots, (d + c)a_n)\]
\[= (da_1 + ca_1, da_2 + ca_2, \ldots, da_n + ca_n)\]
\[= (da_1, da_2, \ldots, da_n) + (ca_1, ca_2, \ldots, ca_n)\]
\[= d(a_1, a_2, \ldots, a_n) + c(a_1, a_2, \ldots, a_n)\]
\[= dx + cx.\]

4. \((l)x = (l)(a_1, a_2, \ldots, a_n)\)
\[= (la_1, la_2, \ldots, la_n)\]
\[= (a_1, a_2, \ldots, a_n)\]
\[= x.\]

5. \(x + y = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)\)
\[= (b_1 + a_1, b_2 + a_2, \ldots, b_n + a_n)\]
\[= y + x\]

6. \(x + (y + z) = (a_1, a_2, \ldots, a_n)\)
\[+ (b_1 + e_1, b_2 + e_2, \ldots, b_n + e_n)\]
\[= (a_1 + b_1 + e_1, a_2 + b_2 + e_2, \ldots, a_n + b_n + e_n)\]
\[= (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)\]
\[+ (e_1, e_2, \ldots, e_n)\]
\[= (x + y) + z\]

Proof: III. Let \(0_v = (0, 0, \ldots, 0)\), then
\(x + 0_v = (a_1, a_2, \ldots, a_n) + (0, 0, \ldots, 0)\)
\[= (a_1 + 0, a_2 + 0, \ldots, a_n + 0)\]
\[= (a_1, a_2, \ldots, a_n)\]
\[= x.\]
\[ x + (-x) = (a_1, a_2, \ldots, a_n) + (-a_1, -a_2, \ldots, -a_n) \]
\[ = (a_1 - a_1, a_2 - a_2, \ldots, a_n - a_n) \]
\[ = (0, 0, \ldots, 0) \]
\[ = 0_v \]

Therefore, \( x - x = 0_v \).

**Proof:** IV.

1. \((x, y) = \sum_{i=1}^{n} \overline{a_i} b_1\)
   \[= \overline{a_1} b_1 + \overline{a_2} b_2 + \ldots + \overline{a_n} b_n\]
   \[= a_1 \overline{b_1} + a_2 \overline{b_2} + \ldots + a_n \overline{b_n}\]
   \[= \overline{b_1} a_1 + \overline{b_2} a_2 + \ldots + \overline{b_n} a_n\]
   \[= \sum_{i=1}^{n} \overline{b_1} a_1\]
   \[= (y, x)\]

2. \((dx, y) = \sum_{i=1}^{n} \overline{d a_i} b_1\)
   \[= \sum_{i=1}^{n} \overline{d a_i} b_1\]
   \[= \overline{d} \sum_{i=1}^{n} a_i b_1\]
   \[= \overline{d}(x, y)\]

3. \((x, x) = \sum_{i=1}^{n} \overline{a_i} a_i\)
   \[= \sum_{i=1}^{n} |a_i|^2 \geq 0.\]
4. \((x, y + z) = \sum_{i=1}^{n} a_i (b_i + e_i)\)
   
   \[= \sum_{i=1}^{n} (a_i b_i + a_i e_i)\]
   
   \[= \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} a_i e_i\]
   
   \[= (x, y) + (x, z)\]

5. If \(x = 0\),

\[ (x, x) = \sum_{i=1}^{n} a_i a_i = \sum_{i=1}^{n} |a_i|^2 = 0. \]

If \((x, x) = 0\), \(\sum_{i=1}^{n} |a_i|^2 = 0. \)

Assume \(x \neq 0\), then \(\sum_{i=1}^{n} |a_i|^2 > 0\); but,

\[\sum_{i=1}^{n} |a_i|^2 = 0. \]

Hence, this is a contradiction.

Therefore, \(x = 0\).

**Definition 3**: Let each of \(x\) and \(y\) be elements of a unitary space \(U\). If \((x, y) = 0\), then \(x\) and \(y\) are orthogonal. The length of a vector \(x\) is \(\|x\| = \sqrt{(x, x)}\) and is always a non-negative real number. If \(\|x\| = \text{unitary}\), then \(x\) is normalized.

**Definition 4**: If \(S\) is a sequence of vectors, \(x_1, x_2, x_3, \ldots\), in a unitary space \(U\), satisfying the property that \((x_i, x_j) = 0\) for \(i \neq j\), \((i, j = 1, 2, 3, \ldots)\), then \(S\) is an orthogonal set. If, in addition, \(\|x_i\| = 1\), \((i = 1, 2, \ldots)\), \(S\) is an orthonormal set.
An alternate means of stating the definition of an orthogonal set \( S \) is \( (x_1, x_j) = \delta_{ij} \) where \( \delta_{ij} \), the Kronecker delta, is defined by

\[
\delta_{ij} = \begin{cases} 
1 & \text{if } i = j \\
0 & \text{if } i \neq j 
\end{cases}
\]

for \( x_i, x_j \in S \).

**Example 2:** The \( u_k \) of page three, number four, are orthogonal.

**Proof:**

\[
\begin{align*}
\mathbf{u}_1 &= (1, 0, 0, \ldots, 0) \\
\mathbf{u}_2 &= (0, 1, 0, \ldots, 0) \\
& \quad \vdots \\
\mathbf{u}_n &= (0, 0, 0, \ldots, 1).
\end{align*}
\]

Let \( a_{ip} \) be the \( p \)th component of \( \mathbf{u}_1 \), then

\[
(u_1, u_j) = \sum_{p=1}^{n} a_{ip} a_{jp} = a_{11}a_{1j} + a_{12}a_{2j} + \ldots + a_{1n}a_{nj}.
\]

If \( i = j \),

\[
(u_1, u_1) = a_{11}^2 + a_{12}^2 + \ldots + a_{1i}^2 + \ldots + a_{1n}^2
= 0 + 0 + \ldots + 1 + \ldots + 0
= 1.
\]

If \( i \neq j \),

\[
(u_1, u_j) = a_{11}a_{1j} + a_{12}a_{2j} + \ldots + a_{1i}a_{ij} + \ldots + a_{ij}a_{jj} + \ldots + a_{1n}a_{nj}
= 0*0 + 0*0 + \ldots + 1*0 + \ldots + \\
0*1 + \ldots + 0*0
= 0.
\]

**Definition 5:** Let \( x_1, x_2, \ldots, x_n \) be a set of vectors. The vectors \( x_1, x_2, \ldots, x_n \) are linearly dependent if
\[ a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = 0 \]

where \( a_i \) is a complex constant and \( a_i \neq 0 \) for some \( i = 1, 2, \ldots, n \). If \( \sum_{i=1}^{n} a_i x_i = 0 \) only when \( a_i = 0 \) for \( i = 1, 2, \ldots, n \), then the vectors are linearly independent.

Let \( x_1, x_2, \ldots, x_n \) be linearly independent. If one wishes to transform the set of vectors \( x_1, x_2, \ldots, x_n \) into a new set \( y_1, y_2, \ldots, y_n \) having the properties

1) \( (y_i, y_j) = \delta_{ij} \) and

2) each \( y_i \) is a linear combination of \( x_j \), where \( j = 1, 2, \ldots, n \); i.e., if each \( y_i = \sum_{i=1}^{n} a_j x_j \) for some choice of \( a_j \) with each \( a_j \) complex, one may do so by the Gram Schmidt process (3, p. 6).

Let \( y_1 = x_1/||x_1|| \), then \( ||y_1|| = 1 \). Next, assume that \( y_2' = x_2 - \lambda_1 y_1 \) and determine \( \lambda_1 \), such that \( (y_2', y_1) = 0 \), i.e., \( \lambda_1 = (y_1, x_2) \). Since \( x_1, x_2 \) are linearly independent, \( y_2' \neq 0 \) and one sets \( y_2 = y_2'/||y_2'|| \) using \( \lambda_1 = (y_1, x_2) \). In general, if \( y_1, y_2, \ldots, y_k \) have been constructed, write

\[ y_{k+1}' = x_{k+1} - \sigma_1 y_1 - \cdots - \sigma_k y_k \]

and determine \( \sigma_1, \sigma_2, \ldots, \sigma_k \) so that \( (y_{k+1}', y_j) = 0 \) for \( j = 1, 2, \ldots, k \), i.e., choose \( \sigma_j = (y_j, x_{k+1}) \). As before, \( y_{k+1}' \neq 0 \) and \( y_{k+1} = y_{k+1}'/||y_{k+1}'|| \).

Since \( y_{k+1}' = y_{k+1}'/||y_{k+1}'|| \), \( y_{k+1}/||y_{k+1}'|| = y_{k+1} \); and since \( (y_{k+1}', y_j) = 0 \) was constructed,
so that \((y_{k+1}, y_j) = 0\) for \(k = 1, 2, \ldots, n-1\) and \(j = 1, 2, \ldots, k\). Therefore, property 1) is satisfied for \(i \neq j\).

If \(i = j\),

\[
(y_i, y_i) = (y_i / \|y_i\|, y_i / \|y_i\|) = (1/\|y_i\|)(y_i, y_i / \|y_i\|)
\]

\[
= (1/\|y_i\|)(1/\|y_i\|)(y_i, y_i) = (1/\|y_i\|^2)(y_i, y_i)
\]

\[
= (1/(y_i, y_i))(y_i, y_i)
\]

\[
= 1
\]

and property 1) is satisfied.

Since \(y_1\) was constructed as a linear combination of \(x_1\), \(y_2\) was constructed as a linear combination of \(x_2\) and \(y_1\), hence, \(x_2\) and \(x_1\), and, in general, \(y_{k+1}\) was constructed as a linear combination of \(x_{k+1}, y_1, y_2, \ldots, y_k\), hence \(x_{k+1}, x_k, x_{k-1}, \ldots, x_2, x_1\), each \(y\) is a linear combination of \(x_j\) and property 2) is satisfied for \(j = 1, 2, \ldots, n\).

Example 3: Let \(x_1 = (1, 5, -1)\), \(x_2 = (0, 2i, 5 - i)\), and \(x_3 = (-1, 7 - i, 6 + i)\) be a set of vectors from \(U_3\) where \(i = \sqrt{-1}\). Show that \(x_1, x_2, \) and \(x_3\) are linearly independent. Transform \(x_1, x_2, \) and \(x_3\) into \(y_1, y_2, \) and \(y_3\) by use of the Gram Schmidt process.

From Definition 5, \(x_1, x_2, \) and \(x_3\) are linearly independent if \(\sum_{i=1}^{3} a_i x_i = 0\) only when \(a_i = 0\) for \(i = 1, 2, 3\).
Hence, one needs to determine the $a_i$ for $i = 1, 2, 3$ to test for linear independence. It follows that

$$a_1 x_1 + a_2 x_2 + a_3 x_3 = 0 \iff (0, 0, 0)$$

$$a_1(i, 5, -1) + a_2(0, 2i, 5-i) + a_3(-1, 7-i, 6+i) = (0, 0, 0)$$

$$(a_1 + a_2 + (7-i)a_3, 5a_1 + 2a_2 + (7-i)a_3, -a_1 + (5-i)a_2 + (6+i)a_3) = (0, 0, 0).$$

Now,

$$a_1 - a_3 = 0$$
$$5a_1 + 2a_2 + (7-i)a_3 = 0$$
$$-a_1 + (5-i)a_2 + (6+i)a_3 = 0$$

Using $a_1 = -a_3$ from the first equation in the second and third equation and solving simultaneously, it can be shown that $a_3 = -2a_2/(7-6i)$ so that $a_1 = -2a_2/(7-6i)$. Using these values of $a_1$ and $a_3$ in the third equation, one finds

$$2a_2/(7-6i) + (5-i)a_2 - (6+i)(2a_2)/(7-6i) = 0$$
$$2a_2 + (5-i)(7-6i)a_2 - 2(6+i)a_2 = 0$$
$$4a_2 + 29a_2 - 37a_2 - 12a_2 = 0$$
$$a_2(33 - 49i) = 0.$$ 

Since $33 - 49i \neq 0$, $a_2$ must be 0. Therefore, $a_3 = 0$ and $a_1 = 0$. Hence, $x_1$, $x_2$, and $x_3$ are linearly independent.

Using the Gram-Schmidt process, 

$$y_1 = x_1 / \|x_1\| = x_1 / \sqrt{\sum_{i=1}^{3} a_{ii} a_{i1}}$$

where $a_{ii}$ is the $i$th component of $x_1$. Hence,

$$y_1 = x_1 / \sqrt{1+25+1} = (1/\sqrt{27})(i, 5, -1).$$
Next, assume \( y_2' = x_2 - \lambda_1 y_1 \) and determine \( \lambda_1 \) so that
\[
\lambda_1 = (y_1, x_2) = (x_1^*/\sqrt{27}, x_2) = (1/\sqrt{27}) \sum_{i=1}^{3} \overline{a_{i1}} a_{2i}
\]
\[
= (1/\sqrt{27})(10i - 5 + i) = (-1/\sqrt{27})(5 - 11i).
\]
Therefore,
\[
y_2' = (0, 2i, 5 - i) + (1/\sqrt{27})(5 - 11i)(x_1^*/\sqrt{27})
\]
\[
= (0, 2i, 5 - i) + ((5 - 11i)/27)(i, 5, -1)
\]
\[
= (1/27)(11 + 5i, 25 - i, 130 - 161)
\]
and
\[
y_2 = \frac{y_2'/||y_2'||}{\sqrt{\sum_{i=1}^{3} b_{2i}^2 b_{3i}^2}}
\]
where \( b_{2i} \) is the ith component of \( y_2' \). Then
\[
y_2 = 27y_2'/\sqrt{(146 + 626 + 17156)} = 27y_2'/\sqrt{17928}
\]
\[
= (1/\sqrt{17928})(11 + 5i, 25 - i, 130 - 161).
\]
One now assumes \( y_j = x_3 - \sigma_1 y_1 - \sigma_2 y_2 \) and determines \( \sigma_1 \) and \( \sigma_2 \) so that \( \sigma_1 = (y_1, x_3) \) and \( \sigma_2 = (y_2, x_3) \).
\[
\sigma_1 = (x_1^*/\sqrt{27}, x_3) = (1/\sqrt{27})(x_1, x_3) = (1/\sqrt{27}) \sum_{i=1}^{3} \overline{a_{i1}} a_{3i}
\]
\[
\sigma_1 = ((1/\sqrt{27})(1 + 35 - 51 - 6 - i) = (1/\sqrt{27})(29 - 51).
\]
\[
\sigma_2 = (1/\sqrt{17928}) \sum_{i=1}^{3} b_{2i} a_{3i}
\]
\[
\sigma_2 = (1/\sqrt{17928})(-11 + 51 + 176 - 181 + 764 + 2261)
\]
\[
\sigma_2 = (1/\sqrt{17928})(929 + 2131).
\]
Therefore,
\[
y_j = (-1, 7 - i, 6 + i) - (1/\sqrt{27})(29 - 51)(1/\sqrt{27})(i, 5, -1)
\]
\[
- (1/\sqrt{17928})(929 + 2131)(1/\sqrt{17928})(11 + 5i, 25 - i, 130 - 161)
\]
\[ y_3 = (1/17928)(-30402 - 262441, 5778 - 5724i, 2646 + 17821). \]

so that

\[ y_3 = \frac{y_3}{\|y_3\|} = \frac{17928y_3}{\sqrt{2214235440}} = (1/\sqrt{2214235440})(-30402 - 262441, 5778 - 5724i, 2646 + 17821). \]

Checking \((y_j, y_k)\) as described, one finds that

\[(y_1, y_1) = (1/27)(1 + 25 + 1) = 1 \]
\[(y_2, y_2) = (1/17928)(121 + 26 + 625 + 1 + 16900 + 256) = 1 \]
\[(y_3, y_3) = (1/2214235440)(924281604 + 1213627536 + 33385284 + 32764176 + 7001316 + 3175524) = 1 \]
\[(y_1, y_2) = (1/\sqrt{27})(1/\sqrt{17928})(-111 + 5 + 125 - 51 - 130 + 161) = 0 \]
\[(y_1, y_3) = (1/\sqrt{27})(1/\sqrt{2214235440})(304021 - 26244 + 28890 - 286201 - 2646 - 17821) = 0 \]
\[(y_2, y_3) = (1/\sqrt{17928})(1/\sqrt{2214235440})(-465642 - 1366741 + 150174 - 1373221 + 315468 + 273996i) = 0 \]

so that property 2) of the Gram-Schmidt process is satisfied.

**Definition 6:** If \( S \) is a set of vectors \( x_1, x_2, \ldots, x_n \), \( S \) spans a vector space \( V \) if every vector of \( V \) is a linear combination of \( x_1, x_2, \ldots, x_n \). \( S \) forms a basis of \( V \) if \( S \) spans \( V \) and \( S \) is linearly independent.

**Linear Operators**

**Definition 7:** A linear operator \( T \) on a unitary space \( U \) is a mapping of each vector \( x \) of \( U \) to a unique vector \( Tx \) of
U so that $T(\alpha x + y) = \alpha Tx + Ty$ for every pair of vectors $x$, $y$ in $U$ and every complex number $\alpha$.

**Definition 7':** An alternate definition would be $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$ for each pair of vectors $x$, $y$ in $U$ and every complex number $\alpha$ and $\beta$.

**Proof:**

$$T(\alpha x + \beta y) = T(\beta (\frac{\alpha}{\beta} x + y)),$$ if $\beta \neq 0$

$$= \beta (T(\frac{\alpha}{\beta} x + y)) \text{ by Definition 7}$$

$$= \beta (\frac{\alpha}{\beta} Tx + Ty) \text{ by Definition 7}$$

$$= \alpha Tx + \beta Ty$$

Therefore, Definition 7 and Definition 7' are equivalent.

**Definition 8:** Let $x$ be an element of a vector space $V$. The linear operator $I$ which maps each vector $x$ to the vector $x$ itself, $Ix = x$, is called the identity operator. The zero operator, $0$, is the operator which maps each $x$ to $0$, $0x = 0$.

**Definition 9:** If each $T$ and $W$ is a linear operator on a unitary space $U$, then $T = W$, i.e., $T$ and $W$ are called equal operators if $Tx = Wx$ for each $x$ in $U$.

**Definition 10:** If $T$ and $W$ are linear operators on a vector space $V$ and if $\alpha$ is a complex constant, then $(T + W)x = Dx Tx + Wx$ and $(\alpha T)x = \alpha (Tx)$.

**Theorem 1:** If $V$ is a vector space and if $V_\perp = \{ T_\perp | \ T_\perp$ is a linear operator on $V \}$, then $V_\perp$ with the operations in Definition 10 is a vector space.
Proof: I. If each of $T_i$ and $T_j$ is an element of $V$, and $x \in V$, then

$$T_i(x) + T_j(x) = (T_i + T_j)(x).$$

Assume $T_i(x) + T_j(x) = T_k(x) \in V$ then

$$T_k(x) = (T_i + T_j)(x)$$

$$T_k = T_i + T_j.$$ 

Now,

$$(T_i + T_j)(\alpha x + y) = T_i(\alpha x + y) + T_j(\alpha x + y)$$

$$= \alpha T_i x + T_i y + \alpha T_j x + T_j y$$

$$= \alpha(T_i + T_j)x + (T_i + T_j)y$$

so that $T_i + T_j$ is a unique linear operator.

Proof: II. If each of $T_i$, $T_j$ and $T_k$ is an element of $V$ and each of $\alpha, \beta$, is a complex number, then

1. $\alpha(T_i + T_j)(x) = \alpha(T_i(x) + T_j(x))$

$$= \alpha T_i(x) + \alpha T_j(x)$$

2. $(\alpha \beta)T_i(x) = (\alpha \beta)T_i(x) + (\alpha(\beta T_i(x)))$

$$= \alpha(\beta T_i(x))$$

3. $(\alpha + \beta)T_i(x) = ((\alpha + \beta)T_i)(x)$

$$= (\alpha T_i + \beta T_i)(x)$$

$$= \alpha T_i(x) + \beta T_i(x)$$

4. $(1)T_i(x) = (1T_i)(x) = T_i(x)$

5. $(T_i + T_j)x = T_i(x) + T_j(x)$

$$= T_j(x) + T_i(x)$$

$$= (T_j + T_i)(x)$$
6. \( T_i(x) + (T_j(x) + T_k(x)) = T_i(x) + (T_j + T_k)(x) \)
\[ = (T_i + T_j + T_k)(x) \]
\[ = (T_i + T_j)(x) + T_k(x) \]

Proof: III. Let \( \Theta \) be the operator of Definition 8 which maps each \( x \) in \( V \) to \( \Theta \), then
\[ (T_i + \Theta)(x) = T_i(x) + \Theta(x) = T_i(x). \]

Now,
\[ T_i(\alpha x + \beta y) = \alpha T_i(x) + \beta T_i(y). \]

Letting \( \alpha = 1, \beta = -1, \) and \( y = x, \) then
\[ T_i(\alpha x + \beta y) = T_i(x - x) = T_i(x) + (-T_i(x)) \]
\[ = 0 \]
\[ = T_i(x) + (-T_i(x)) \]

Definition 11: If \( x \) is an element of a vector space \( V \) and each of \( T \) and \( W \) is a linear operator, the product \( TW \) is defined by \( (TW)(x) = T(W(x)). \) If \( TW = WT \), \( T \) commutes with \( W \); but, in general, \( TW \neq WT \). In any case, the commutator
\[ [T, W] = TW - WT. \]
Obviously, \( T \) commutes with \( W \) if and only if \( [T, W] = 0. \)

Definition 12: Let \( T \) be a linear operator on a vector space \( V \). If there exists a linear operator \( W \) on \( V \) so that
\[ WT = TW = I, \]
\( W \) is called the inverse operator of \( T \). \( T \) has at most one inverse operator; since, if \( Z \) is also an inverse operator of \( T \), \( Z(TW) = Z(I) = Z = (ZT) W = IW = W. \) Therefore, if \( T \) has an inverse, it shall be denoted by \( W = T^{-1}. \) Therefore,
\[ T^{-1}T = TT^{-1} = I. \]

Although \( T^{-1} \) is defined, \( T^{-1} \) may not exist. If \( T^{-1} \) exists, then \( T^{-1} \) "undoes" what \( T \) has done, i.e.,
\[ T^{-1}(T(x)) = (T^{-1}T)(x) = I(x) = x \]

for every \( x \in V \).

**Definition 13:** If \( T \) is a linear operator on the vector space \( V \) and \( T \) has an inverse \( T^{-1} \), \( T \) is nonsingular; otherwise, \( T \) is singular.

**Theorem 2:** If each of \( T \) and \( W \) is a nonsingular operator, the inverse of the product is the product of the inverses in reverse order, i.e., \( (TW)^{-1} = W^{-1}T^{-1} \).

**Proof:**

\[
TW(W^{-1}T^{-1}) = T(WW^{-1})T^{-1} = TIT^{-1} = TT^{-1} = I \quad \text{and} \\
(W^{-1}T^{-1})TW = W^{-1}(T^{-1}T)W = W^{-1}IW = W^{-1}W = I.
\]

But, \( (TW)^{-1} \) is that operator such that \( (TW)^{-1}TW = I = TW(TW)^{-1} \) and is unique. Therefore, \( W^{-1}T^{-1} = (TW)^{-1} \).

**Eigenvalues and Hermitian Operators**

For the present discussion the word *space* shall stand for unitary space.

**Definition 14:** Let \( T \) be a linear operator on a space \( U \). If there exists a nonzero vector \( x \in U \) and a complex number \( \lambda \) such that

\[ Tx = \lambda x \]

then the nonzero vector \( x \) is called an eigenvector (proper vector, characteristic vector, latent vector) of the operator \( T \). For any such \( x \), the number \( \lambda \) is called the eigenvalue (proper root, characteristic value, characteristic root, proper value, latent root, latent value, latent number) of \( T \) corresponding to the eigenvector \( x \).
Remark: Intuitively, if there exists a nonzero vector which, when operated on by $T$, does not have its direction changed, then the vector is an eigenvector of $T$.

**Definition 15:** Let $T$ be a linear operator. If there exists a linear operator $T^*$ having the property that

$$(x, Ty) = (T^*x, y)$$

for every pair of vectors $x, y$ in $U$, then $T^*$ is called an adjoint operator of $T$.

**Theorem 3:** There can be at most one adjoint operator for $T$.

**Proof:** If $T^*$ exists and $(x, Ty) = (T^*x, y)$ and there is another operator $Z^*$ such that $(x, Ty) = (Z^*x, y)$ for every pair of vectors $x, y$ in $U$, then

$$(T^*x, y) = (Z^*x, y)$$

$T^*x = Z^*x$

and from Definition 9, $T^*$ and $Z^*$ are equal operators.

**Note:** If $T^*$ exists, then $(T^*)^*$ exists and $(T^*)^* = T$.

**Definition 16:** Let $T$ be a linear operator on a space $U$. $T$ is Hermitian, or self-adjoint, if $T^* = T$, or equivalently, if $(x, Ty) = (Tx, y)$ for every $x, y$ in $U$.

**Theorem 4:** Let each of $T$ and $W$ be a linear operator that possesses an adjoint $T^*$ and $W^*$ respectively. Then the adjoint of $TW$ exists and is $W^*T^*$.

**Proof:** Let each of $x$ and $y$ be an arbitrary vector of $U$, then
\[(x, Ty) = (T^*x, y)\]
\[(x, Wy) = (W^*x, y)\]

Each of \(Ty, Wy, T^*x, \) and \(W^*x\) is a vector in \(U\) so that
\[(x, TWy) = (x, T(Wy))\]
\[= (T^*x, Wy)\]
\[= (W^*T^*x, y)\]

Therefore, the adjoint of \(TW\) exists and is \(W^*T^*\) by definition.

**Theorem 5:** Let \(T\) be a self-adjoint operator and \(x\) an arbitrary vector of \(U\), then \((x, Tx)\) is a real number.

**Proof:**
\[(x, Tx) = (T^*x, x)\] by Definition 15
\[= (Tx, x)\] by Definition 16
\[= (x, Tx).\]

Hence, \((x, Tx)\) is real since it equals its complex conjugate.

**Theorem 6:** The eigenvalues of a Hermitian operator are real.

**Proof:** Let \(H\) be a Hermitian operator, \(x\) be an eigenvector of \(H\), and \(\lambda\) be an eigenvalue of \(H\). If
\[Hx = \lambda x,\]
\[(x, Hx) = (x, \lambda x)\]
\[= \lambda (x, x).\]

\((x, Hx)\) is real by Theorem 5 and \((x, x)\) is real by Definition 1, IV, 3. Therefore, \(\lambda\) is real since if \(\lambda\) were complex, \((x, Hx)\) would be complex; but, \((x, Hx)\) is real.

**Theorem 7:** Let each of \(x\) and \(y\) be eigenvectors of a Hermitian operator belonging to distinct eigenvalues \(\lambda_1\), and \(\lambda_2\) respectively. Then \(x\) and \(y\) are orthogonal. In other
words, given that $Hx = \lambda_1 x$, $Hy = \lambda_2 y$, $\lambda_1 \neq \lambda_2$, and $H = H^*$, prove $(x, y) = 0$.

**Proof:** Taking the inner products $(y, Hx)$ and $(x, Hy)$,

$$(y, Hx) = (y, \lambda_1 x) = \lambda_1 (y, x)$$

$$(x, Hy) = (x, \lambda_2 y) = \lambda_2 (x, y).$$

Also,

$$(y, Hx) = (H^* y, x) = (Hy, x)$$

$$(x, Hy) = (H^* x, y) = (Hx, y).$$

Hence,

$$\lambda_2 (x, y) = (Hx, y)$$

$$= (y, Hx)$$

$$= \overline{\lambda_1 (y, x)}$$

$$= \lambda_1 (x, y)$$

and $(x, y) = 0$ since $\lambda_1 \neq \lambda_2$.

**Definition 17:** Let $\gamma$ be a linear operator on $U$. If $\gamma^{-1}$ exists, if $\gamma^*$ exists, and if $\gamma^{-1} = \gamma^*$, then $\gamma$ is called a unitary operator and $\gamma \gamma^* = \gamma^* \gamma = I$.

**Definition 18:** Let $\gamma$ be a linear operator on $U$. If $\gamma$ preserves all inner products, i.e., $(x, y) = (\gamma x, \gamma y)$ for all $x, y$ in $U$, then $\gamma$ is called an isometric operator or isometry.

**Note:** An isometric operator preserves the length of every vector, since $\|Ux\|^2 = (Ux, Ux) = (x, x) = \|x\|^2$. Thus an isometry may be thought of as a generalized rotation of the unitary space $U$. 
Theorem 8: If $\gamma^*$ exists, then $\gamma$ is isometric if and only if it is unitary.

Proof: If $\gamma$ is unitary and each of $x$ and $y$ is a vector of $U$, then

$$(\gamma x, \gamma y) = (x, \gamma^* \gamma y) = (x, y).$$

Hence, $\gamma$ is isometric. If $\gamma$ is isometric, then

$$(\gamma x, \gamma y) = (\gamma^* \gamma x, y) = (x, y)$$

for every $x, y$ in $U$ and

$$(\gamma^* \gamma - I)x, y) = (\gamma^* \gamma x, y) - (Ix, y)$$

$$= (\gamma^* \gamma x, y) - (x, y)$$

$$= 0.$$

Since this is true for every $y$ in $U$, it is true in particular for $y = (\gamma^* \gamma - I)x$ so that

$$(((\gamma^* \gamma - I)x, (\gamma^* \gamma - I)x) = 0.$$

Hence, $(\gamma^* \gamma - I)x = 0$. Since $x$ was an arbitrary vector of $U$,

$$\gamma^* \gamma - I = 0$$

$$\gamma^* \gamma = I$$

and similarly, $\gamma \gamma^* = I$, so that $\gamma$ is unitary.

Matrices

Definition 19: Euclidean $n$-space is the space of vectors $x$ that satisfy Definition 2 and have the properties 1 through 7 and will be denoted by $E_n$.

Note: The symbol $(x)_i$ will be the $i$th component of $x$.

Let $E_n$ be an Euclidean $n$-space and $E_m$ be an Euclidean $m$-space and let $T$ be a linear operator which associates with each $x \in E_m$ a unique element $y \in E_n$ such that $y = Tx$. Let
e_1, e_2, ..., e_m and f_1, f_2, ..., f_n be a basis of E_m and E_n respectively. The vectors T e_j, (j = 1, 2, ..., m), are in E_n and are a linear combination of the f_i, (i = 1, 2, ..., n), i.e., T e_j = \sum_{i=1}^{n} t_{ij} f_i. If x = (x_1, x_2, ..., x_m) \in E_m, x = \sum_{j=1}^{m} \alpha_j e_j. Therefore, T x = \sum_{j=1}^{m} \alpha_j T e_j and T x = \sum_{j=1}^{m} \alpha_j \sum_{i=1}^{n} t_{ij} f_i.

Now, T x = \sum_{i=1}^{n} \left( \sum_{j=1}^{m} t_{ij} \alpha_j \right) f_i. Hence, (T x)_i = \sum_{j=1}^{m} t_{ij} \alpha_j.

Therefore,

(T x)_1 = t_{11} x_1 + t_{12} x_2 + t_{13} x_3 + ... + t_{1m} x_m

(T x)_2 = t_{21} x_1 + t_{22} x_2 + t_{23} x_3 + ... + t_{2m} x_m

\vdots

(T x)_n = t_{n1} x_1 + t_{n2} x_2 + t_{n3} x_3 + ... + t_{nm} x_m.

Definition 20: Consider the numbers t_{ij} arranged in a rectangular array having n rows and m columns,

\[\begin{array}{cccc}
t_{11} & t_{12} & t_{13} & \cdots & t_{1m} \\
t_{21} & t_{22} & t_{23} & \cdots & t_{2m} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
t_{n1} & t_{n2} & t_{n3} & \cdots & t_{nm}
\end{array}\]

then this array is called an n \times m matrix associated with the operator T. Since the action of operator T is fully described if one knows the numbers t_{ij}, (i = 1, 2, ..., n; j = 1, 2, ..., m), one uses T to denote the matrix. The equation T = (t_{ij}), (i = 1, 2, ..., n; j = 1, 2, ..., m),
means that $T$ is the matrix which has the number $t_{ij}$ in row $i$ and column $j$.

Let each of $T$ and $W$ be a linear operator which carries $E_m$ into $E_n$ and let $(t_{ij})$, $(w_{ij})$ be the matrix which represents $T$, $W$ respectively. If $x$ is a vector of $E_m$, then

$$(T + W)x = Tx + Wx$$

$$((T + W)x)_i = (Tx)_i + (Wx)_i$$

$$= \sum_{j=1}^{m} t_{ij} \alpha_j + \sum_{j=1}^{m} w_{ij} \alpha_j$$

$$= \sum_{j=1}^{m} (t_{ij} + w_{ij}) \alpha_j$$

Definition 21: In view of Definition 20 and the fact that $((T + W)x)_i = \sum_{i=1}^{m} (t_{ij} + w_{ij}) \alpha_j$, one sees immediately that the sum of two operators can be represented by matrix $(t_{ij} + w_{ij})$, or the sum of two matrices $T + W = (t_{ij} + w_{ij})$.

In order to define a meaningful product of two matrices, some restrictions must be made in the definition of the operators $T$ and $W$ above. As it is, $TW$ would be meaningless since if $x$ is in $E_m$, $Wx$ is in $E_n$ and $T$ is not defined on the vector $Wx$. Therefore, let $T$ carry $E_m$ into $E_p$, $W$ carry $E_p$ into $E_n$ and note that $WT$ (not $TW$) is meaningful and carries $E_m$ into $E_n$, i.e., $E_m \xrightarrow{T} E_p \xrightarrow{W} E_n$ or $x \rightarrow Tx \rightarrow WTx$ where $x \in E_m$, $Tx \in E_p$ and $WTx \in E_n$. WT should be representable by a matrix $WT$ of $n$ rows and $m$ columns.
If each of $T = (t_{ij})$ and $W = (w_{ij})$ is the matrix representation of the operator $T$ and $W$ respectively and $x = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ is a vector of $E_m$, then

$$(Tx)_i = \sum_{j=1}^{m} t_{ij} \alpha_j.$$  

Applying $W$ to $Tx \in E_p$, 

$$(W(Tx))_i = \sum_{k=1}^{p} w_{ik} (Tx)_k = \sum_{k=1}^{p} w_{ik} \sum_{j=1}^{m} t_{kj} \alpha_j = \sum_{j=1}^{m} \left( \sum_{k=1}^{p} w_{ik} t_{kj} \right) \alpha_j.$$  

**Definition 22:** In view of Definition 20 and the fact that $(W(Tx))_i = \sum_{j=1}^{m} \sum_{k=1}^{p} w_{ik} t_{kj} \alpha_j$, one sees that the product operator $WT$ can be represented by a matrix $\left( \sum_{k=1}^{p} w_{ik} t_{kj} \right)$ or the product of two matrices $WT = \left( \sum_{k=1}^{p} w_{ik} t_{kj} \right)$, $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$. Obviously the product of an $n \times p$ matrix and a $p \times m$ matrix is an $n \times m$ matrix.

Let $T = (t_{ij})$, $(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$, be an $m \times n$ matrix. Consider the $n \times m$ matrix $W = (w_{ij})$ where $w_{ij} = \overline{t}_{ji}$, $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$. Let $x = (\alpha_1, \alpha_2, \ldots, \alpha_m)$, $y = (\beta_1, \beta_2, \ldots, \beta_n)$ be arbitrary vectors in $E_m$ and $E_n$ respectively. Then
\[(x, Ty) = \sum_{i=1}^{m} \bar{\alpha}_i (Ty)_i\]
\[= \sum_{i=1}^{m} \bar{\alpha}_i \sum_{k=1}^{n} t_{ik} \beta_k\]
\[= \sum_{i=1}^{m} \sum_{k=1}^{n} \bar{\alpha}_i t_{ik} \beta_k\]

\[(Wx, y) = \sum_{i=1}^{n} (Wx)_i \beta_i\]
\[= \sum_{i=1}^{n} (\sum_{k=1}^{m} t_{ki} \alpha_k) \beta_i\]
\[= \sum_{i=1}^{n} \sum_{k=1}^{m} t_{ki} \alpha_k \beta_i\]
\[= \sum_{i=1}^{m} \sum_{k=1}^{n} t_{ik} \alpha_i \beta_k\]
\[= (x, Ty)\]

**Definition 23:** Since, by Definition 15, \(W\) has the property of the adjoint operator of \(T\), i.e., \(W = T^*\), \(W\) shall be called the adjoint matrix of \(T\). Symbolically, \((T^*)_ij = (T)_{ji}\). The adjoint matrix is sometimes referred to as the conjugate transpose matrix or the Hermitian conjugate matrix.

**Definition 24:** If \(T\) is a square \((n \times n)\) matrix, where \(T = (t_{ij})\), and if \(t_{ij} = \overline{t_{ji}}\), then \(T\) is Hermitian.

**Definition 25:** If \(T\) is a \(n \times m\) matrix, the transpose of \(T\), \(T^T\), is given by \((T^T)_{ij} = (T)_{ji}\), \((j = 1, 2, \ldots, n; i = 1, 2, \ldots, m)\).

**Note:** \((T^*)_ij = (T^T)_{ij}\).
Theorem 9: If $T$ is an $n \times m$ matrix and $W$ is an $m \times k$ matrix, the transpose of $TW$ is the product of $T^T$ and $W^T$ in reverse order, i.e., $(TW)^T = W^T T^T$.

**Proof:** Let $T = (t_{ij})$ and $W = (w_{jp})$, $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m; p = 1, 2, \ldots, k)$, then

$$TW = (\delta_{ip})$$

$$\delta_{ip} = \sum_{j=1}^{m} t_{ij} w_{jp}$$

$$(TW)^T = (\delta_{ip})^T$$

$$= (\delta_{pi})$$

$$W^T T = (\gamma_{pi})$$

Now $\gamma_{pi}$ is the component of the matrix $W^T T$ which is formed by multiplying row $p$ of $W^T$ by column $i$ of $T^T$ or

$$\gamma_{pi} = \sum_{j=1}^{m} w_{jp} t_{ij}$$

$$= \sum_{j=1}^{m} t_{ij} w_{jp}$$

$$= \delta_{ip} = \epsilon_{pi}$$

$$(\gamma_{pi}) = (\epsilon_{pi})$$

$$W^T T = (TW)^T$$

**Definition 26:** If $T$ is a square, $n \times n$, matrix, then $T$ is symmetric if $T = T^T$.

**Note:** If $T$ is a square matrix with real elements and $T$ is symmetric, then $T$ is Hermitian since $t_{ij} = \overline{t_{ij}} = \overline{t_{ji}}$. 
Theorem 10: Let each of $T$ and $W$ be a square Hermitian matrix. In order that $TW$ be Hermitian it is necessary and sufficient that $T$ and $W$ commute.

Proof:

$$(T)_{ij} = (T^*)_{ij} = (T^T)_{ij}$$
$$(W)_{ij} = (W^*)_{ij} = (W^T)_{ij}$$

If $TW$ is Hermitian,

$$(TW)^* = W^T_T^* = WT$$
$$(TW)^* = TW.$$ Therefore, $WT = TW$, i.e., $T$ and $W$ commute. If $T$ and $W$ commute,

$$TW = WT$$
$$(TW)^* = (WT)^* = T^*W^* = TW.$$ Hence, $TW$ is Hermitian.

Definition 27: If $T$ is an $n \times n$ matrix, $(t_{ij})$, $(i = 1, 2, \ldots, n)$, the adjugate of $T$, $(\text{adj } T)$, is the $n \times n$ matrix formed by the cofactor of each element $t_{ij}$ of $T$, i.e., the element $t_{ij} \in (\text{adj } T)$ is the number formed by finding the determinant of $T$ after having deleted the $i$th row and the $j$th column and multiplying by $(-1)^{i+j}$.

Example 4: Let $T$ be the $3 \times 3$ matrix

$$T = \begin{bmatrix} 1 & 7 & 2 \\ -1 & 3 & 0 \\ 9 & 4 & 2 \end{bmatrix}$$
\[
\begin{align*}
\text{t}^{11} &= (-1)^{1+1} \begin{vmatrix} 3 & 0 \\ 4 & 2 \end{vmatrix} = 6 \\
\text{t}^{12} &= (-1)^{1+2} \begin{vmatrix} -1 & 0 \\ 9 & 2 \end{vmatrix} = 2 \\
\text{t}^{13} &= (-1)^{1+3} \begin{vmatrix} -1 & 3 \\ 9 & 4 \end{vmatrix} = -31 \\
\text{t}^{21} &= (-1)^{2+1} \begin{vmatrix} 7 & 2 \\ 4 & 2 \end{vmatrix} = -6 \\
\text{t}^{22} &= (-1)^{2+2} \begin{vmatrix} 1 & 2 \\ 9 & 2 \end{vmatrix} = -16 \\
\text{t}^{23} &= (-1)^{2+3} \begin{vmatrix} 1 & 7 \\ 9 & 4 \end{vmatrix} = 59 \\
\text{t}^{31} &= (-1)^{3+1} \begin{vmatrix} 7 & 2 \\ 3 & 0 \end{vmatrix} = -6 \\
\text{t}^{32} &= (-1)^{3+2} \begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} = -2 \\
\text{t}^{33} &= (-1)^{3+3} \begin{vmatrix} 1 & 7 \\ -1 & 3 \end{vmatrix} = 10 \\
\text{(adj } T) &= \begin{bmatrix} 6 & 2 & -31 \\ -6 & -16 & 59 \\ -6 & -2 & 10 \end{bmatrix}.
\end{align*}
\]
The Laplace expansion of the determinant of a square matrix $T$ by cofactors has the form

$$\sum_{j=1}^{n} t_{ij} t_{kj} = \det (T), \quad (i = 1, 2, \ldots, n).$$

Let $i \neq k$ and consider $\sum_{j=1}^{n} t_{ij} t_{kj}$, which is the sum of products of the elements of one row by the cofactors of another row. One sees immediately that this is the determinant of the matrix $T$ with the $k$th row deleted and the $i$th row substituted in its place. But this matrix has two rows that are identical. Hence, $\det (T) = 0$. Symbolically,

$$\sum_{j=1}^{n} t_{ij} t_{kj} = 0$$

if $i \neq k$. Therefore,

$$\sum_{j=1}^{n} t_{ij} t_{kj} = \delta_{ik} (\det (T))$$

for $i, k = 1, 2, \ldots, n$.

If $\det (T) \neq 0$, one may define the matrix

$$(T^{-1})_{ij} = (t_{ji})/\det (T)$$

and

$$(TT^{-1})_{ij} = \sum_{k=1}^{n} t_{ik} (T^{-1})_{kj}$$

$$(TT^{-1})_{ij} = \sum_{k=1}^{n} t_{ik} (t_{jk})/\det (T) = \delta_{ij}$$

or

$$(TT^{-1})_{11} = 1$$

$$(TT^{-1})_{12} = 0$$
\[(TT^{-1})_{13} = 0 \]
\[\vdots\]
\[(TT^{-1})_{1n} = 0\]
\[(TT^{-1})_{21} = 0\]
\[(TT^{-1})_{22} = 1\]
\[\vdots\]
\[(TT^{-1})_{nn} = 1\]

so that

\[
TT^{-1} = \begin{bmatrix}
1 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1
\end{bmatrix}
\]

**Definition 28:** If \( I \) is an \( n \times n \) matrix which has ones on the diagonal and zeroes in all other positions, then \( I \) is called the \( n \times n \) unit matrix or the \( n \times n \) product identity matrix since if \( T \) is \( p \times n \),

\[
TI = \begin{bmatrix}
t_{11} & t_{12} & \cdots & t_{1n} \\
t_{21} & t_{22} & \cdots & t_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
t_{p1} & t_{p2} & \cdots & t_{pn}
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1
\end{bmatrix}
= \begin{bmatrix}
t_{11} & t_{12} & \cdots & t_{1n} \\
t_{21} & t_{22} & \cdots & t_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
t_{p1} & t_{p2} & \cdots & t_{pn}
\end{bmatrix} \begin{bmatrix}
t_{11} & t_{12} & \cdots & t_{1n} \\
t_{21} & t_{22} & \cdots & t_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
t_{p1} & t_{p2} & \cdots & t_{pn}
\end{bmatrix} \cdot \begin{bmatrix}
t_{11} & t_{12} & \cdots & t_{1n} \\
t_{21} & t_{22} & \cdots & t_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
t_{p1} & t_{p2} & \cdots & t_{pn}
\end{bmatrix}.
In view of Definition 28, \( TT^{-1} = I \). With a simple modification, one can immediately see that \( T^{-1}T = I \) so that \( T^{-1} \) plays the role of an inverse of \( T \).

**Definition 29:** In view of the above, if \( T \) is a square matrix and if there exists a matrix \( T^{-1} \) such that \( TT^{-1} = I = T^{-1}T \), then \( T^{-1} \) is the inverse of \( T \) and

\[
T^{-1} = (1/\det(T))(\text{adj } T)^T
\]

**Example 5:** Let \( T \) be the matrix of Example 4. Then

\[
T = \begin{bmatrix} 1 & 7 & 2 \\ -1 & 3 & 0 \\ 9 & 4 & 2 \end{bmatrix}
\]

and

\[
\text{adj } (T) = \begin{bmatrix} 6 & 2 & -31 \\ -6 & -16 & 59 \\ -6 & -2 & 10 \end{bmatrix}
\]

By the method of pivotal condensation (2, pp. 121-124),

\[
\det(T) = (1/1) \begin{vmatrix} 10 & 2 \\ -59 & -16 \end{vmatrix} = -42.
\]

Therefore,

\[
T^{-1} = -(1/42) \begin{bmatrix} 6 & 2 & -31 \\ -6 & -16 & 59 \\ -6 & -2 & 10 \end{bmatrix}
\]

and

\[
TT^{-1} = -(1/42) \begin{bmatrix} 1 & 7 & 2 \\ -1 & 3 & 0 \\ 9 & 4 & 2 \end{bmatrix} \begin{bmatrix} 6 & -6 & -6 \\ -1 & 3 & 0 \\ -31 & 59 & 10 \end{bmatrix} = \begin{bmatrix} 6 & -6 & -6 \\ -1 & 3 & 0 \\ -31 & 59 & 10 \end{bmatrix}
\]
In light of the above discussion and Definition 29, one may characterize nonsingular matrices as follows.

**Theorem 11:** If $T$ is a square matrix, it is necessary and sufficient for the $\det(T)$ to be nonzero in order for $T$ to be nonsingular.

**Proof:** Remembering that $\det(AB) = \det(A) \times \det(B)$ and $I = TT^{-1}$, then if $\det(T) = 0$ and if $T^{-1}$ exists,

$$1 = \det(I) = \det(TT^{-1})$$

$$= \det(T) \times \det(T^{-1})$$

$$= (\det(T))(0)$$

$$= 0.$$
Theorem 12: If $T^{-1}$ exists, then $T^{-1}$ is unique.

Proof: Assume $W$ is also an inverse of $T$, then

\[ TT^{-1} = TW = I \]
\[ T^{-1}TW = T^{-1}I \]
\[ IW = T^{-1} \]
\[ W = T^{-1} \]

Eigenvalues of Matrices

In this and all following sections, all matrices will be $n \times n$ unless otherwise specified, and all vectors will be column vectors in order to have a meaningful product.

Suppose $x$ is an eigenvector of $T$ corresponding to the eigenvalue $\lambda$. Then $Tx = \lambda x$ or

\[ \sum_{j=1}^{n} t_{ij} x_j = \lambda x_i, \quad (i = 1, 2, \ldots, n), \]

or equivalently,

\[ \sum_{j=1}^{n} (t_{ij} - \lambda \delta_{ij}) x_j = 0, \quad (i = 1, 2, \ldots, n). \]

One sees immediately that this is a system of linear, algebraic, homogeneous equations with $n$ unknowns

\[ (t_{11} - \lambda)x_1 + t_{12}x_2 + \cdots + t_{1n}x_n = 0 \]
\[ t_{21}x_1 + (t_{22} - \lambda)x_2 + \cdots + t_{2n}x_n = 0 \]
\[ \vdots \]
\[ t_{n1}x_1 + t_{n2}x_2 + \cdots + (t_{nn} - \lambda)x_n = 0. \]

If $x = 0$, then $x_1 = x_2 = \ldots = x_n = 0$ is obviously a trivial solution.
In order to get a nontrivial solution, it is necessary that the determinant of the coefficients vanish, i.e.,

\[
\begin{vmatrix}
  t_{11} - \lambda & t_{12} & \cdots & t_{1n} \\
  t_{21} & t_{22} - \lambda & \cdots & t_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  t_{n1} & t_{n2} & \cdots & t_{nn} - \lambda
\end{vmatrix} = 0
\]

or \( \det (T - \lambda I) = 0 \).

**Definition 30:** The polynomial equation, \( \det (T - \lambda I) = 0 \), is called the characteristic equation of \( T \). The polynomial \( \phi(\lambda) = \det (T - \lambda I) \) is called the characteristic polynomial of \( T \).

**Definition 31:** If \( \lambda \) is any one of \( \lambda_1, \lambda_2, \ldots, \lambda_n \), then \( \lambda \) is an eigenvalue of \( T \) and conversely.

**Example 6:** Let \( T \) be the \( 3 \times 3 \) matrix

\[
T = \begin{bmatrix}
  1 & 2 & -1 \\
  0 & 3 & -1 \\
  2 & 0 & 5
\end{bmatrix}
\]

\[
\det (T - \lambda I) = \begin{vmatrix}
  1 - \lambda & 2 & -1 \\
  0 & 3 - \lambda & -1 \\
  2 & 0 & 5 - \lambda
\end{vmatrix}
\]
\[
\text{det} \ (T - \lambda I) = 15 - 23\lambda + 9\lambda^2 - \gamma^3 - 4 + 2(3 - \lambda) \\
= 17 - 25\lambda + 9\lambda^2 - \gamma^3 \\
= (1 - \lambda)(\gamma - 4 - i)(\gamma - 4 + i).
\]

Therefore, \(\lambda_1 = 1, \lambda_2 = 4 + i, \lambda_3 = 4 - i\). One needs to determine a vector, \(x_i = (\alpha_{i1}, \alpha_{i2}, \alpha_{i3})\), such that

\[
T x_i = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 2 & 0 & 5 \end{bmatrix} \begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \alpha_{i3} \end{bmatrix} = \lambda_i \begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \alpha_{i3} \end{bmatrix}, \ (i = 1, 2, 3).
\]

\[
T x_i = \begin{bmatrix} \alpha_{i1} + 2\alpha_{i2} - \alpha_{i3} \\ 3\alpha_{i2} - \alpha_{i3} \\ 2\alpha_{i1} + 5\alpha_{i3} \end{bmatrix} = \lambda_i \begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \alpha_{i3} \end{bmatrix}.
\]

If \(i = 1\),

\begin{align*}
\alpha_{11} + 2\alpha_{12} - \alpha_{13} &= \alpha_{11} \\
3\alpha_{12} - \alpha_{13} &= \alpha_{12} \\
2\alpha_{11} + 5\alpha_{13} &= \alpha_{13}
\end{align*}

\[
2\alpha_{12} = \alpha_{13} \\
2\alpha_{11} = -4\alpha_{13}
\]

and \(x_1 = (-2\alpha_{13}, (1/2)\alpha_{13}, \alpha_{13})\) where \(\alpha_{13}\) is arbitrary.

If \(i = 2\),

\begin{align*}
\alpha_{21} + 2\alpha_{22} - \alpha_{23} &= (4 + i)\alpha_{21} \\
3\alpha_{22} - \alpha_{23} &= (4 + i)\alpha_{22} \\
2\alpha_{21} + 5\alpha_{23} &= (4 + i)\alpha_{23}
\end{align*}
\[-(1 + i)\alpha_{22} = \alpha_{23}\]
\[2\alpha_{21} = -(1 + i)\alpha_{23}\]

and \(x_2 = ((-1 + i)/2)\alpha_{23}, -(1/(1 + i))\alpha_{23}, \alpha_{23}\) where \(\alpha_{23}\)

is arbitrary. If \(i = 3\),
\[
\alpha_{31} + 2\alpha_{32} - \alpha_{33} = (4 - i)\alpha_{31}
\]
\[3\alpha_{32} - \alpha_{33} = (4 - i)\alpha_{32}\]
\[2\alpha_{31} + 5\alpha_{33} = (4 - i)\alpha_{33}\]
\[-(1 + i)\alpha_{32} = \alpha_{33}\]
\[2\alpha_{31} = -(1 + i)\alpha_{33}\]

and \(x_3 = ((-1 + i)/2)\alpha_{33}, (1/(1 + i))\alpha_{33}, \alpha_{33}\) where \(\alpha_{33}\)

is arbitrary.

**Definition 32:** If \(T\) is an \(n \times n\) matrix, the trace or spur of \(T\), \(Tr (T)\), is the sum of the diagonal elements of \(T\),

\[Tr (T) = \sum_{i=1}^{n} t_{ii}.\]

**Theorem 13:** Let the eigenvalues of a matrix \(T\) be \(\lambda_1, \lambda_2, \ldots, \lambda_n\), then

\[\lambda_1 \lambda_2 \ldots \lambda_n = \det (T)\]
\[\lambda_1 + \lambda_2 + \ldots + \lambda_n = Tr (T).\]

**Proof:** By expanding the characteristic polynomial \(\varphi(\lambda)\),
\[
\varphi(\lambda) = (-1)^n \left[ \lambda^n - (t_{11} + t_{12} + \ldots + t_{1n})\lambda^{n-1} + \ldots \right.
\]
\[\left. + (-1)^n\det (T) \right]\]
\[= (-1)^n \left[ (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) \ldots (\lambda - \lambda_n) \right].\]
Expanding the latter, one finds that
\[ \phi'(\lambda) = (-1)^n \left[ x^n - (\lambda_1 + \lambda_2 + \ldots + \lambda_n) x^{n-1} + (\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \ldots + \lambda_1 \lambda_n) x^{n-2} + \ldots + (-1)^n \lambda_1 \lambda_2 \lambda_3 \ldots \lambda_n \right]. \]
Equating \( \phi'(\lambda) \) and \( \phi(\lambda) \),
\[ \lambda_1 + \lambda_2 + \lambda_3 + \ldots + \lambda_n = t_{11} + t_{22} + t_{33} + \ldots + t_{nn} = \text{Tr} (T) \]
and
\[ \lambda_1 \lambda_2 \lambda_3 \ldots \lambda_n = \det (T). \]

Diagonalization of Matrices

Let \( T \) be a matrix with eigenvectors \( x_1, x_2, \ldots, x_n \) corresponding to eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \) respectively, then \( Tx_i = \lambda_i x_i \) for \( i = 1, 2, \ldots, n \).

**Definition 33:** Let \( P \) be the matrix formed by using the eigenvectors of \( T \) as columns for \( P \), i.e.,
\[
P = \begin{bmatrix}
\alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\
\alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn}
\end{bmatrix}
\]
where \( \alpha_{ij} \) is the \( i \)th component of \( x_j \). \( P \) is called the polar matrix of \( T \) or the modal matrix of \( T \).

The matrix \( P \) shall be denoted by \( (P)_{ij} = (x_j)_i \), \( (i, j = 1, 2, \ldots, n) \). Furthermore, define a diagonal matrix \( \Lambda \) by
placing the eigenvalues of $T$ on the main diagonal. Hence,

$$
\begin{bmatrix}
\lambda_1 & 0 & 0 & \ldots & 0 \\
0 & \lambda_2 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & \lambda_n
\end{bmatrix}
$$

The matrix $\Lambda$ shall also be denoted by $\Lambda = \text{diag} (\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $(\Lambda)_{ij} = \lambda_i \delta_{ij}$, $(i, j = 1, 2, \ldots, n)$. Now,

$$
(TP)_{ij} = \sum_{k=1}^{n} t_{ik}(P)_{kj} = \sum_{k=1}^{n} t_{ik}(x_j)_k
$$

$$
= \lambda_j x_j_i
$$

and

$$
(P\Lambda)_{ij} = \sum_{k=1}^{n} (P)_{ik}(\Lambda)_{kj} = \sum_{k=1}^{n} (x_k)_i \lambda_k \delta_{kj}
$$

$$
= \lambda_j x_j_i = (TP)_{ij}
$$

so that

$$
TP = P\Lambda.
$$

If the eigenvectors are linearly independent, the columns of $P$ are linearly independent and $\det (P) \neq 0$ so that $P^{-1}$ exists and

$$
TPP^{-1} = P\Lambda P^{-1}
$$

$$
T = P\Lambda P^{-1}.
$$

**Definition 34**: If $T$ is a matrix which can be represented in the form $T = P\Lambda P^{-1}$, finding the matrices $P$ and $\Lambda$ is called diagonalizing $T$. A matrix which has $n$ linearly independent eigenvectors is said to be diagonalizable.
**Definition 35:** Let each of $T$ and $W$ be a matrix. $T$ and $W$ are similar if there exists a nonsingular matrix $P$ such that $T = P^{-1}WP$.

**Theorem 14:** Similar matrices have the same eigenvalues.

**Proof:** If $T$ and $W$ are similar matrices, then there exists a matrix $P$ such that $T = P^{-1}WP$. If $\phi(\lambda)$ and $\psi(\lambda)$ are the characteristic polynomials of $T$ and $W$ respectively, then

\[
\phi(\lambda) = \det (T - \lambda I) = \det (P^{-1}WP - \lambda I) \\
= \det (P^{-1}WP - \lambda P^{-1}P) \\
= \det (P^{-1}(W - \lambda I)P) \\
= \det (P^{-1}) \det (W - \lambda I) \det (P) \\
= \det (W - \lambda I) \\
= \psi(\lambda).
\]

The following theorem has been proven.

**Theorem 15:** An arbitrary diagonalizable matrix $T$ is similar to a diagonal matrix $\Lambda$.

Consider the real symmetric matrix

\[
T_2 = \begin{bmatrix}
t_{11} & t_{12} \\
t_{21} & t_{22}
\end{bmatrix}
\]

and let $\lambda_1$ and $x_1$ be an associated characteristic root and characteristic vector where $x_1$ is normalized. Now form an orthogonal matrix $Q_2$, i.e., the columns of $Q_2$ are mutually orthogonal, with $x_1$ as one of its columns and designate the other column as $x_2$. Since

\[
T_2 x_1 = \lambda_1 x_1
\]
\[ Q_2^{\top}T_2Q_2 = Q_2^{\top}\begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}\begin{bmatrix} (x_1)_1 \\ (x_2)_1 \end{bmatrix} \begin{bmatrix} (x_1)_2 \\ (x_2)_2 \end{bmatrix} \]

\[ = Q_2^{\top}\begin{bmatrix} t_{11}(x_1)_1 + t_{12}(x_1)_2 & t_{11}(x_2)_1 + t_{12}(x_2)_2 \\ t_{21}(x_1)_1 + t_{22}(x_1)_2 & t_{21}(x_2)_1 + t_{22}(x_2)_2 \end{bmatrix} \]

\[ = Q_2^{\top}\begin{bmatrix} \lambda_1(x_1)_1 & \lambda_1(x_1)_2 \\ \lambda_1(x_2)_1 & \lambda_1(x_2)_2 \end{bmatrix}\begin{bmatrix} t_{11}(x_1)_1 + t_{12}(x_2)_2 \\ t_{21}(x_2)_1 + t_{22}(x_2)_2 \end{bmatrix} \]

\[ = \begin{bmatrix} \lambda_1(x_1)_1^2 + \lambda_1(x_1)_2^2 & b_{21} \\ \lambda_1(x_2)_1(x_1)_1 + \lambda_1(x_2)_2(x_1)_2 & b_{22} \end{bmatrix} \]

where \( b_{21}, b_{22} \) can be determined and \( (x_2)_1(x_1)_1 + (x_2)_2(x_1)_2 = 0 \) since \( x_1 \) and \( x_2 \) are orthogonal and \( (x_1)_1^2 + (x_1)_2^2 = 1 \) since \( x_1 \) is normalized. Hence,

\[ Q_2^{\top}T_2Q_2 = \begin{bmatrix} \lambda_1 & b_{21} \\ 0 & b_{22} \end{bmatrix}. \]

Now,

\[ (Q_2^{\top}T_2Q_2)^T = (Q_2)^T(Q_2^{\top}T_2)^T = Q_2^{\top}T_2(Q_2)^T \]

\[ = Q_2^{\top}T_2Q_2 = Q_2^{\top}T_2Q_2 \]

so that \( Q_2^{\top}T_2Q_2 \) is symmetric. Hence, \( b_{21} = 0 \). Since \( Q_2 \) is orthogonal, \( Q_2^{\top}Q_2 = I \) so that \( Q_2^{\top} = Q_2^{-1} \) and by Theorem 16,
\[ b_{22} = \lambda_2 \text{ or} \]
\[
Q_2^T T_2 Q_2 = \begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix}
\]

where
\[
n_2 = \left[ t_{11}(x_2)_1 + t_{12}(x_2)_2 \right] (x_2)_1 + \left[ t_{21}(x_2)_1 + t_{22}(x_2)_2 \right] (x_2)_2
\]
\[
= t_{11}(x_2)_1^2 + t_{22}(x_2)_2^2 + 2t_{12}(x_2)_2(x_2)_1.
\]

If one proceeds inductively, one assumes that for each 
\( k, (k = 1, 2, \ldots, n), \) one can determine an orthogonal matrix 
\( Q_k \) which reduces a real symmetric \( T_k = (t_{ij}), (i, j = 1, 2, \ldots, k), \) to diagonal form

\[ Q_k^T T_k Q_k \begin{bmatrix} T \
0 \\
\end{bmatrix} \begin{bmatrix} \lambda \end{bmatrix} \]

where \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the characteristic roots of \( T_k. \)

One now needs to show the reduction for a matrix \( T_{n+1} = (t_{ij}), (i, j = 1, 2, \ldots, n + 1). \)

Proceeding as in the two-dimensional case, form an 
orthogonal matrix \( Q_{n+1} \) whose first column is \( x_1, \) the associated characteristic vector of characteristic root \( \lambda_1, \) and 
whose other columns are designated \( x_1, x_2, \ldots, x_{n+1}, \) so that

\[ Q_{n+1}^T T_{n+1} Q_{n+1} = (q_{ij})^T (t_{ij}) (q_{ij}) \]

where \( q_{ij} = (x_j)_1, (i, j = 1, 2, \ldots, n + 1). \) Letting \( t_i \) denote the row of \( T \) with \( t_{i1}, t_{i2}, \ldots, t_{i,n+1}, \) one finds
Let \( (y_j)_i = (t_1, x_j) \), \( i = 1, 2, \ldots, n+1 \), then
\[
Q_{n+1}^T T_{n+1} Q_{n+1} = (q_{ij})^T \begin{bmatrix}
(t_1, x_1) & (t_1, x_2) & \cdots & (t_1, x_{n+1}) \\
(t_2, x_1) & (t_2, x_2) & \cdots & (t_2, x_{n+1}) \\
\vdots & \vdots & \ddots & \vdots \\
(t_{n+1}, x_1) & (t_{n+1}, x_2) & \cdots & (t_{n+1}, x_{n+1})
\end{bmatrix}.
\]

Since \( T_{n+1} x_1 = \lambda_1 x_1 \),
\[
Q_{n+1}^T T_{n+1} Q_{n+1} = (q_{ij})^T \begin{bmatrix}
\lambda_1(x_1, 1) & (t_1, x_2) & \cdots & (t_1, x_{n+1}) \\
\lambda_1(x_1, 2) & (t_2, x_2) & \cdots & (t_2, x_{n+1}) \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_1(x_1, n+1) & (t_{n+1}, x_2) & \cdots & (t_{n+1}, x_{n+1})
\end{bmatrix}.
\]

Using the fact that \( Q_{n+1} \) is orthogonal and \( Q_{n+1}^T T_{n+1} Q_{n+1} \) is symmetric, one can verify that the elements of the first row and column are zero with the exception of the diagonal element, which will be \( \lambda_1 \). Also the \( n \times n \) matrix formed by deleting the first row and column of \( Q_{n+1}^T T_{n+1} Q_{n+1} \) is symmetric and can be denoted by \( S_n \). Therefore,
Since the characteristic equation of $Q_{n+1}^T T_{n+1} Q_{n+1}$ is
\[ \det (\lambda I - Q_{n+1}^T T_{n+1} Q_{n+1}) = 0 \]
and using Theorem 14, one finds the eigenvalues of $S_n$ are the remaining eigenvalues of $T_{n+1}$ which will be denoted by $\lambda_2, \lambda_3, \ldots, \lambda_{n+1}$.

Let $Q_n$ be an orthogonal matrix which reduces $S_n$ to diagonal form. Form the $(n + 1)$ dimensional matrix

\[
W_{n+1} = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 0 & Q_n & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & 0
\end{bmatrix}
\]

which is also orthogonal. It is readily verified that
\[
W_{n+1}^T (Q_{n+1}^T T_{n+1} Q_{n+1}) W_{n+1} = \text{diag} (\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{n+1}).
\]

Since $W^T (Q^T Q) W = (Q W)^T T (Q W)$, one sees that $(Q_{n+1} W_{n+1})$ is the required diagonalizing orthogonal matrix for $T_{n+1}$. Thus the following theorem has been proven (1, pp. 50-54).
Theorem 16: If $T$ is a real symmetric matrix, then $T$ may be transformed into diagonal form by an orthogonal transformation, i.e., there is an orthogonal matrix $Q$ such that $Q^T T Q = \text{diag} (\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n)$ where $\lambda_i$ is a characteristic root of $T$.

If one changes the matrices $T_i$ to Hermitian matrices and used the conjugate transpose of $Q_i$ instead of the transpose, i.e., $Q_i^T T_i Q_i$ instead of $Q_i^T T_i Q_i$, and parallels the procedure used in proving Theorem 16, one proves the following theorem. (1, p. 59).

Theorem 17: If $H$ is a Hermitian matrix, there exists a unitary matrix $U$ such that $H = U A U^*$. 

The Companion Matrix

In Theorem 13, it was shown that the characteristic equation of a given matrix was a polynomial of degree $n$ where $n$ was the order of the matrix. Now, suppose that $\phi(z) = z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n$ is a polynomial of degree $n$. Is there an $n \times n$ matrix whose characteristic polynomial is $\phi(z)$? If so, it is not unique since if $T$ is such a matrix, $P^{-1} T P$ is another for any non-singular matrix $P$. However, there does not exist one.

Theorem 18: Every polynomial of degree $n$ is the characteristic polynomial of an $n \times n$ matrix.
Proof: Consider the matrix

\[
T = \begin{bmatrix}
-a_1 & -a_2 & -a_3 & \cdots & -a_{n-1} & -a_n \\
1 & 0 & 0 & \cdots & 0 & 0 \\
0 & 1 & 0 & \cdots & 0 & 0 \\
\vdots & & & \ddots & & \vdots \\
0 & 0 & 0 & \cdots & 1 & 0 \\
\end{bmatrix}
\]

The characteristic polynomial of \( T \) is

\[
\det (\lambda I - T) = \begin{bmatrix}
a_1 + \lambda & a_2 & a_3 & \cdots & a_{n-1} & a_n \\
-1 & \lambda & 0 & \cdots & 0 & 0 \\
0 & -1 & \lambda & \cdots & 0 & 0 \\
\vdots & & & \ddots & & \vdots \\
0 & 0 & 0 & \cdots & -1 & \lambda \\
\end{bmatrix}
\]

By multiplying column one by \( \lambda \) and adding to column two, multiplying column two by \( \lambda \) and adding to column three, and continuing until column \( n-1 \) has been multiplied by \( \lambda \) and added to column \( n \), one is able to evaluate the \( \det (\lambda I - T) \) handily.

\[
\det (\lambda I - T) = \begin{bmatrix}
a_1 + \lambda & \sum_{i=0}^{2} a_i \lambda^{2-i} & \sum_{i=0}^{3} a_i \lambda^{3-i} & \cdots & \phi(\lambda) \\
-1 & 0 & 0 & \cdots & 0 \\
0 & -1 & 0 & \cdots & 0 \\
\vdots & & & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 \\
\end{bmatrix}
\]

where \( a_0 = 1 \). Expanding by row one,
\[
\det (\lambda I - T) = 0 + 0 + 0 + \ldots + (-1)^{n+1} \phi(\lambda) (-1)^{n-1} \\
= (-1)^{2n} \phi(\lambda) = \phi(\lambda).
\]

Letting \( \lambda = z \), \( \det (zI - T) = \phi(z) \) so that \( \phi(z) \) is the characteristic polynomial of \( T \).

**Definition 36:** The companion matrix of a polynomial \( \phi(z) \) is the matrix of the form of \( T \) in Theorem 18.

**Bordering Matrices**

**Definition 37:** The process of building an \((n+1) \times (n+1)\) matrix \( \tilde{T} \) from an \( n \times n \) matrix \( T \) is called bordering if

\[
\tilde{T} = \begin{bmatrix} T & u \\ v^T & \alpha \end{bmatrix}
\]

where each of \( u \) and \( v \) is a column vector and \( \alpha \) is a complex number (real if \( T \) is Hermitian).

**Theorem 19:** If \( T \) is Hermitian, then \( \tilde{T} \) is Hermitian if and only if \( u = v \).

**Proof:** If \( u = v \),

\[
t_{ij} = \overline{t_{ji}}, \quad (i, j = 1, 2, \ldots, n)
\]

since \( T \) is Hermitian.

\[
t_{n+1,j} = \overline{t_{j,n+1}}, \quad (j = 1, 2, \ldots, n)
\]

since \( \overline{u_j} = t_{n+1,j} = \overline{v_j} \) and \( \alpha = \overline{\alpha} \) since \( \alpha \) is real. Therefore \( \tilde{T} \) is Hermitian.

If \( \tilde{T} \) is Hermitian,

\[
\tilde{t}_{ij} = \overline{\tilde{t}_{ji}}, \quad (i, j = 1, 2, \ldots, n).
\]

Now, \( \overline{v^T} = (\overline{\tilde{t}_{n1}}, \overline{\tilde{t}_{n2}}, \ldots, \overline{\tilde{t}_{nn}})^T \) and \( u = (\tilde{t}_{1n}, \tilde{t}_{2n}, \ldots, \tilde{t}_{nn}) \).

Hence, \( u = v \).
Due to the result of Theorem 19, one sees that

$$\tilde{T} = \begin{bmatrix} T & u \\ u^T & \alpha \end{bmatrix}$$

if $T$ is Hermitian. It is of particular interest to discover what happens to the eigenvalues and eigenvectors of a matrix when it is bordered.

Let $y$ be an $n$ component vector, let $\beta$ be a complex number, and let $x = (y, \beta)$ be the $(n+1)$ component vector whose first $n$ components are the $n$ components of $y$ and whose $(n+1)$ component is $\beta$. Suppose $x$ is an eigenvector of $\tilde{T}$, then

$$\begin{bmatrix} T & u \\ u^T & \alpha \end{bmatrix} \begin{bmatrix} y \\ \beta \end{bmatrix} = \lambda \begin{bmatrix} y \\ \beta \end{bmatrix}$$

so that $Ty + \beta u = \lambda y$ and

$$(v)_1(y)_1 + (v)_2(y)_2 + \ldots + (v)_n(y)_n + \alpha \beta = \lambda \beta$$

or $(v, y) + \alpha \beta = \lambda \beta$. Suppose $T$ has diagonal form $T = P\Lambda P^{-1}$ where $P$ is the polar matrix of $T$. Let $y = Pw$ and one finds

$$TPw + \beta u = \lambda Pw$$

or

$$P\Lambda w + \beta u = \lambda Pw.$$ 

Multiplying the latter equation by $P^{-1}$,

$$\Lambda w + \beta P^{-1}u = \lambda w$$

or

$$w = \beta(\Lambda I - \Lambda)^{-1}P^{-1}u$$

which gives the eigenvector if the eigenvalue is known. Also
(v, Pw) = (\lambda - \alpha)\beta

or

(v, P(\lambda I - \Lambda)^{-1}P^{-1}u) = (\lambda - \alpha).

This is an algebraic equation from which the eigenvalues of T can be determined.

Consider the expression \( P(\lambda I - \Lambda)^{-1}P^{-1} \) where P is the polar matrix of \( T(x_j) \), \( (i, j = 1, 2, \ldots, n) \). The diagonal matrix \( (\lambda I - \Lambda)^{-1} \) is formed by subtracting each eigenvalue of T from \( \lambda \) and taking the inverse so that

\[
(\lambda I - \Lambda)^{-1} = \begin{bmatrix}
\frac{1}{\lambda - \lambda_1} & 0 & \cdots \\
0 & \frac{1}{\lambda - \lambda_2} & \cdots \\
\vdots & \vdots & \ddots & \ddots \\
0 & 0 & \cdots & \frac{1}{\lambda - \lambda_n}
\end{bmatrix}.
\]

Denoting the elements of \( P^{-1} \) by \( (x_j)^{-1} \), \( (i = 1, 2, \ldots, n) \),

\( P(\lambda I - \Lambda)^{-1}P^{-1} = (\xi_{km}) \) where

\[
\xi_{km} = \frac{\sum_{i=1}^{n} (x_i)_k (x_m)_i^{-1}}{\lambda - \lambda_m}
\]

for \( k, m = 1, 2, \ldots, n \). Substituting into the equation

\[
(v, P(\lambda I - \Lambda)^{-1}P^{-1}u) = \lambda - \alpha
\]

one finds that

\[
\sum_{k=1}^{n} (v)_k \left[ \sum_{m=1}^{n} \left( \frac{\sum_{i=1}^{n} (x_i)_k (x_k)_i^{-1}}{\lambda - \lambda_m} \right) (u)_m \right] = \lambda - \alpha.
\]
If $T$ is Hermitian, $P$ is a unitary matrix, $v = u$, and the above equation simplifies to

$$\sum_{k=1}^{n} \frac{|(u, x_k)|^2}{\lambda - \lambda_k} = \lambda - \alpha.$$ 

By plotting the left and right sides of this equation as a function of $\lambda$, it is easy to see that an eigenvalue of $\tilde{T}$ lies between each pair of eigenvalues of $T$. One eigenvalue lies to the right of all of them and one lies to the left of all of them. If $T$ has a multiple eigenvalue $\lambda$ repeated $p$ times, then $\tilde{T}$ has the eigenvalue repeated $p - 1$ times (3, p. 27).
CHAPTER BIBLIOGRAPHY


The escalator method (2; 1, pp. 265-272) is a method for determining the eigenvalues and eigenvectors of a matrix $T_{k+1}$, of order $k + 1$, by using the eigenvalues and eigenvectors of the matrices $T_k$ and $T_k^*$ where $T_k$ is the principal submatrix of order $k$ obtained from $T_{k+1}$ by deleting the $(k + 1)$th row and column. The matrix $T_k$ is bordered so as to obtain $T_{k+1}$. It is then possible to set up an equation to determine the eigenvalues of $T_{k+1}$ and to compute by simple formulas the components of the eigenvectors for $T_{k+1}$ and $T_{k+1}^*$. Application of the method is begun by finding the eigenvectors of a second order matrix.

The great value of the method is the existence of a powerful control which makes it possible for the computations to be verified at each step in terms of their own calculations and without loss of significance.

The method is based on the use of orthogonality properties for the eigenvectors of the matrix $T$ and its conjugate transpose $T^*$. Consider the matrix $T_k = (t_{ij})$, $(i, j = 1, 2, \ldots, k)$. The conjugate transpose $T_k^* = (t_{ji})$. Let $\lambda_{ki}$ and $\overline{\lambda}_{ki}$, $(i = 1, 2, \ldots, k)$, be the eigenvalues of $T_k$ and $T_k^*$ respectively.
Furthermore, let $x_{ki}$ and $x_{^T ki}$ be the eigenvectors corresponding to $\lambda_{ki}$ and $\overline{\lambda}_{ki}$, $(i=1, 2, \ldots, k)$, for $T_k$ and $T_k^*$. 

**Note:** The vector $(x_{ki})^T$ is not necessarily equivalent to $x_{^T ki}$. The $x_{^T ki}$ notation is used simply to denote the $i$th eigenvector of $T_k^* = T_k^T$. Now, $x_{ki} = ((x_{ki})_1, (x_{ki})_2, \ldots, (x_{ki})_k)$ and $x_{^T ki} = ((x_{^T ki})_1, (x_{^T ki})_2, \ldots, (x_{^T ki})_k)$.

The eigenvectors of the matrices $T_k$ and $T_k^*$ are rectified if the following condition is satisfied.

\[
\begin{bmatrix}
(x_{k1})_1 & (x_{k1})_2 & \cdots & (x_{k1})_k \\
(x_{k2})_1 & (x_{k2})_2 & \cdots & (x_{k2})_k \\
\vdots & \vdots & & \vdots \\
(x_{kk})_1 & (x_{kk})_2 & \cdots & (x_{kk})_k
\end{bmatrix}
\begin{bmatrix}
(x_{k1})_1 & (x_{k2})_1 & \cdots & (x_{kk})_1 \\
(x_{k1})_2 & (x_{k2})_2 & \cdots & (x_{kk})_2 \\
\vdots & \vdots & & \vdots \\
(x_{k1})_k & (x_{k2})_k & \cdots & (x_{kk})_k
\end{bmatrix}
= I_k,
\]

or

\[(2-1) \quad \sum_{m=1}^{k} (x_{^T ki})_m (x_{kj})_m = \delta_{ij}, \quad (i, j = 1, 2, \ldots, k).\]

Let

\[
Z_1^T = (x_{k1})_1 Z_1 + (x_{k2})_1 Z_2 + \cdots + (x_{kk})_1 Z_k \\
Z_2^T = (x_{k1})_2 Z_1 + (x_{k2})_2 Z_2 + \cdots + (x_{kk})_2 Z_k \\
\vdots \\
Z_k^T = (x_{k1})_k Z_1 + (x_{k2})_k Z_2 + \cdots + (x_{kk})_k Z_k
\]

or

\[(2-2) \quad Z_m^T = \sum_{j=1}^{k} (x_{kj})_m Z_j, \quad (m = 1, 2, \ldots, k),\]
where the \((x_{kj})_m\) are the components of the rectified eigenvectors of \(T_k\) and the \(Z_j\) are any quantities whatever. After multiplying each \(Z_j^T\) by \((x_{kl})_m\), \(m = 1, 2, \ldots, k\), adding, and using the properties of (2-1), one can immediately verify that

\[
(x_{kl})_1 Z_1^T + (x_{kl})_2 Z_2^T + \ldots + (x_{kl})_k Z_k^T = Z_1.
\]

Similarly, using \((x_{kj})_m\) as multiplier, \(j = 2, 3, \ldots, k\), one obtains

\[
(x_{kj})_1 Z_1^T + (x_{kj})_2 Z_2^T + \ldots + (x_{kj})_k Z_k^T = Z_2
\]

\[
\vdots
\]

\[
(x_{kk})_1 Z_1^T + (x_{kk})_2 Z_2^T + \ldots + (x_{kk})_k Z_k^T = Z_k,
\]

which can be condensed to

\[
(2-3) \quad Z_j = \sum_{m=1}^{k} (x_{kj})_m z_m, \quad (i = 1, 2, \ldots, k).
\]

Substituting for \(Z_j\) from (2-3) into (2-2), one finds that

\[
Z_i^T = \sum_{j=1}^{k} (x_{kj})_i \sum_{m=1}^{k} (x_{kj})_m z_m
\]

or

\[
(2-4) \quad Z_i^T = \sum_{m=1}^{k} \sum_{j=1}^{k} (x_{kj})_i (x_{kj})_m z_m^T.
\]

Hence it follows from (2-4) that

\[
(2-5) \quad \sum_{j=1}^{k} (x_{kj})_i (x_{kj})_m = \delta_{im}, \quad (i, m = 1, 2, \ldots, k)
\]

since when \(i = m\), \(\sum_{j=1}^{k} (x_{kj})_i (x_{kj})_m = 1\), and when \(i \neq m\),
\[ \sum_{j=1}^{\lambda} (x_{kj})_j \left( \overline{x_{kj}} \right)_m = 0 \] in order for (2-4) to be valid. Thus one can rectify the eigenvectors of a matrix by satisfying either equation (2-1) or (2-5).

Let \( T_{k+1} \) be the \((k + 1)\)th order matrix obtained from \( T_k \) by bordering and let \( x_{k+1} \) be the eigenvector corresponding to the eigenvalue \( \lambda_{k+1} \). One has, letting \( p = k + 1 \),

\[
\lambda_p(x_p)_1 = t_{11}(x_p)_1 + t_{12}(x_p)_2 + \cdots + t_{1p}(x_p)_p \\
\lambda_p(x_p)_2 = t_{21}(x_p)_1 + t_{22}(x_p)_2 + \cdots + t_{2p}(x_p)_p \\
\vdots \\
\lambda_p(x_p)_p = t_{p1}(x_p)_1 + t_{p2}(x_p)_2 + \cdots + t_{pp}(x_p)_p.
\]

Similarly for \( T_p^* \),

\[
\overline{\lambda}_p(x_p^T)_1 = \overline{t}_{11}(x_p^T)_1 + \overline{t}_{21}(x_p^T)_2 + \cdots + \overline{t}_{1p}(x_p^T)_p \\
\overline{\lambda}_p(x_p^T)_2 = \overline{t}_{12}(x_p^T)_1 + \overline{t}_{22}(x_p^T)_2 + \cdots + \overline{t}_{2p}(x_p^T)_p \\
\vdots \\
\overline{\lambda}_p(x_p^T)_p = \overline{t}_{p1}(x_p^T)_1 + \overline{t}_{p2}(x_p^T)_2 + \cdots + \overline{t}_{pp}(x_p^T)_p.
\]

If one multiplies the first \( k \) equations of (2-6) by \( x_{k1} \), \( x_{k2} \), \ldots, \( x_{kk} \) respectively and adds, then

\[
(2-8) \quad \overline{\lambda}_{p_{j=1}^{k}} (x_{ki})_j (x_p)_j = \sum_{m=1}^{p} \left( \sum_{j=1}^{k} \overline{t}_{j m} (x_{ki})_j \right) (x_p)_m
\]

\[
= \sum_{m=1}^{p} \left( \sum_{j=1}^{k} \overline{t}_{j m} (x_{ki})_j \right) (x_p)_m.
\]
Since \( x_{ki}^T \) is an eigenvector of \( T_k^* \), (2-8) can be written as

\[
\lambda_p \sum_{j=1}^{k} (x_{ki}^T) j (x_p) j = \lambda_{ki} (x_{ki})_1 (x_p)_1 + \lambda_{ki} (x_{ki})_2 (x_p)_2 + \cdots
\]

\[
+ \lambda_{ki} (x_{ki})_k (x_p)_k + \left[ t_{1p} (x_{ki})_1 + t_{2p} (x_{ki})_2 + \cdots + t_{kp} (x_{ki})_k \right] (x_p)_p
\]

(2-9)

Consequently,

\[
(2-10) \quad (\lambda_{ki} - \lambda_p) \sum_{j=1}^{k} (x_{ki}^T) j (x_p) j = - \sum_{j=1}^{k} t_{jp} (x_{ki})_j (x_p)_p.
\]

Letting

\[
(2-11) \quad P_{pi}^T = \sum_{j=1}^{k} t_{jp} (x_{ki})_j,
\]

(2-10) can be written as

\[
(2-12) \quad (\lambda_{ki} - \lambda_p) \sum_{j=1}^{k} (x_{ki}^T) j (x_p) j = - P_{pi}^T (x_p)_p.
\]

If one multiplies the first \( k \) equations of (2-7) by \( (x_{ki})_1, (x_{ki})_2, \ldots, (x_{ki})_k \) respectively, and adds, then

\[
(2-13) \quad \lambda_{p} \sum_{j=1}^{k} (x_p^T) j (x_{ki})_j = \sum_{m=1}^{p} \left[ \sum_{j=1}^{k} t_{mj} (x_{ki})_j \right] (x_p)_m
\]

\[
= \sum_{m=1}^{p} \left[ \sum_{j=1}^{k} t_{mj} (x_{ki})_j \right] (x_p)_m.
\]

Since \( x_{ki} \) is an eigenvector of \( T_k \), (2-13) can be written as

\[
\lambda_{p} \sum_{j=1}^{k} (x_p^T) j (x_{ki})_j = \lambda_{ki} \sum_{j=1}^{k} (x_{ki})_j (x_p^T) j + \left[ \sum_{j=1}^{k} t_{pj} (x_{ki})_j \right] (x_p)_p
\]
or
\[(2-14) \quad \lambda_p \sum_{j=1}^{k} (x_p^T) j (x_{ki})_j = \lambda_{ki} \sum_{j=1}^{k} (x_{ki})_j (x_p^T) j + \left[ \sum_{j=1}^{k} t_{pj} (x_{ki})_j \right] (x_p^T)_p.\]

Consequently,
\[(2-15) \quad (\lambda_{ki} - \lambda_p) \sum_{j=1}^{k} (x_p^T) j (x_{ki})_j = - \left[ \sum_{j=1}^{k} t_{pj} (x_{ki})_j \right] (x_p^T)_p.\]

Letting
\[(2-16) \quad P_{pi} = \sum_{j=1}^{k} t_{pj} (x_{ki})_j,\]
equation (2-15) can be written as
\[(2-17) \quad (\lambda_{ki} - \lambda_p) \sum_{j=1}^{k} (x_p^T) j (x_{ki})_j = - P_{pi} (x_p^T)_p.\]

In view of the orthogonality properties of (2-1),
\[(2-18) \quad \sum_{i=1}^{k} P_{pi} \left[ \sum_{j=1}^{k} (x_p^T) j (x_{ki})_j \right] = P\]

where
\[(2-19) \quad P = \sum_{j=1}^{k} t_{pj} (x_p)_j = -(t_{pp} - \lambda_p) (x_p)_p.\]

Now (2-18) becomes
\[(2-20) \quad \sum_{i=1}^{k} P_{pi} \left[ \sum_{j=1}^{k} (x_{ki})_j (x_p^T) j \right] = -(t_{pp} - \lambda_p) (x_p)_p.\]

Similarly,
\[(2-21) \quad \sum_{i=1}^{k} P_{pi} \left[ \sum_{j=1}^{k} (x_{ki})_j (x_p^T) j \right] = -(t_{pp} - \lambda_p) (x_p^T)_p.\]
If one multiplies (2-20) by \( \prod_{t=1}^{k} (\lambda_{kt} - \lambda_p) \), one obtains

\[
(2-22) \quad \prod_{t=1}^{k} (\lambda_{kt} - \lambda_p) \sum_{i=1}^{k} P_{pi} \left[ \sum_{j=1}^{k} (x_{ki}^j) j(x_p)_j \right] = D
\]

where

\[
D = - \prod_{t=1}^{k} (\lambda_{kt} - \lambda_p)(t_{pp} - \lambda_p)(x_p)_p.
\]

Substituting appropriately from (2-12) into (2-22),

\[
-D = \prod_{t=2}^{k} (\lambda_{kt} - \lambda_p)P_{pl}x_{pl}^T(x_p)_p
\]

\[
\quad + (\lambda_{k1} - \lambda_p) \prod_{t=3}^{k} (\lambda_{kt} - \lambda_p)P_{p2}x_{p2}^T(x_p)_p + \ldots
\]

\[
\quad + \sum_{t=i+1}^{k} (\lambda_{kt} - \lambda_p) \prod_{t=i+1}^{k} (\lambda_{kt} - \lambda_p)P_{pi}x_{pi}^T(x_p)_p + \ldots
\]

\[
\quad + \prod_{t=1}^{k} (\lambda_{ki} - \lambda_p)P_{pk}x_{pk}^T(x_p)_p
\]

or

\[
(2-23) \quad \sum_{i=1}^{k-1} \prod_{t=1}^{i-1} (\lambda_{kt} - \lambda_p) \prod_{t=i+1}^{k} (\lambda_{kt} - \lambda_p)P_{pi}x_{pi}^T = -D.
\]

Equation (2-23) shall be called the escalator equation.

If \( \lambda_{ki} \neq \lambda_p \), then (2-23) can be written

\[
(2-24) \quad \sum_{i=1}^{k} \left( P_{pi}x_{pi}^T/(\lambda_{ki} - \lambda_p) \right) = (t_{pp} - \lambda_p)
\]

and one can determine the eigenvalues of \( T_k = T_{k+1} \) from (2-24) if the eigenvalues of \( T_k \) are distinct from those of \( T_k \). The case where \( \lambda_{ki} = \lambda_p \) shall be considered later in the chapter.
If one multiplies equation (2-12) by \((x_{k1})_j\), \((i = 1, 2, 3, \ldots, k)\), one obtains

\[
\begin{align*}
(x_{k1})_j & \sum_{j=1}^{k} (x_{k1})_j (x_p)_j = - \frac{P_{p1}^T (x_p)_j (x_{k1})_1}{\lambda_{k1} - \lambda_p} \\
(x_{k2})_j & \sum_{j=1}^{k} (x_{k2})_j (x_p)_j = - \frac{P_{p2}^T (x_p)_j (x_{k2})_1}{\lambda_{k2} - \lambda_p} \\
& \quad \vdots \\
(x_{kk})_j & \sum_{j=1}^{k} (x_{kk})_j (x_p)_j = - \frac{P_{pk}^T (x_p)_j (x_{kk})_1}{\lambda_{kk} - \lambda_p}
\end{align*}
\]

Adding the previous equations,

\[
\sum_{i=1}^{k} \sum_{j=1}^{k} (x_{ki})_j (x_p)_j = - \frac{P_{pi}^T (x_p)_j (x_{ki})_1}{\lambda_{ki} - \lambda_p}
\]

where \(\lambda_{ki} \neq \lambda_p\). Considering the orthogonality properties of (2-1), one finds

\[
\frac{(x_{p1})_1}{(x_p)_p} = - \sum_{i=1}^{k} \frac{P_{pi}^T (x_{ki})_1}{\lambda_{ki} - \lambda_p} \quad , \lambda_{ki} \neq \lambda_p.
\]

Similarly, if (2-12) is multiplied by \((x_{ki})_j\), \((i = 1, 2, \ldots, k)\), and the equations are added for \(j = 1, 2, \ldots, k\), the following is true considering the orthogonality properties of (2-1).

\[
\begin{align*}
\frac{(x_{p2})_2}{(x_p)_p} & = - \sum_{i=1}^{k} \frac{P_{pi}^T (x_{ki})_2}{\lambda_{ki} - \lambda_p} \\
& \quad \vdots \\
(2-25)
\end{align*}
\]
\[
\frac{(x_p)_k}{(x_p)_p} = -\sum_{i=1}^{k} \frac{P_{pi}^{T}(x_{ki})_k}{\lambda_{ki} - \lambda_p}
\]
if \(\lambda_{ki} \neq \lambda_p\). Analogously
\[
\frac{(x_p)_1^T}{(x_p)_p^T} = -\sum_{i=1}^{k} \frac{P_{pi}^{T}(x_{ki})_1}{\lambda_{ki} - \lambda_p}
\]
\[
\frac{(x_p)_2^T}{(x_p)_p^T} = -\sum_{i=1}^{k} \frac{P_{pi}^{T}(x_{ki})_2}{\lambda_{ki} - \lambda_p}
\]
\[
\vdots
\]
\[
\frac{(x_p)_k^T}{(x_p)_p^T} = -\sum_{i=1}^{k} \frac{P_{pi}^{T}(x_{ki})_k}{\lambda_{ki} - \lambda_p}
\]
if \(\lambda_{ki} \neq \lambda_p\). Thus by finding the eigenvalues \(\lambda_p\) from (2-24) one can determine the eigenvectors of \(T_{k+1}\) and \(T_{k+1}^*\), which correspond to \(\lambda_p\) and \(\overline{\lambda}_p\), accurate within a numerical factor. To continue the process one must rectify the eigenvectors in the sense of (2-1).

In order to keep the notation standard, it is convenient to replace \(p\) in equations (2-24), (2-25), and (2-26) with \(pr\) where \(p = k + 1\) and \(r = 1, 2, \ldots, p\). This notation allows one to distinguish between the eigenvalues and eigenvectors of \(T_p\) and \(T_p^*\).
Considering (2-1), (2-25), and (2-26), one can immediately verify that

\[
\sum_{j=1}^{\kappa} \frac{\mathbf{x}_p^T \mathbf{x}_p^T}{\mathbf{x}_p^T \mathbf{x}_p} = \frac{\mathbf{P}_p^T \mathbf{P}_p}{(\lambda_{ki} - \lambda_{pr})^2}
\]

\[(2-27)\]

\( (r = 1, 2, \ldots, p) \). Adding \( \frac{\mathbf{x}_p^T \mathbf{x}_p^T}{\mathbf{x}_p^T \mathbf{x}_p} \) to both sides of

\[(2-27)\], one sees that

\[
\sum_{j=1}^{\kappa} \frac{\mathbf{x}_p^T \mathbf{x}_p^T}{\mathbf{x}_p^T \mathbf{x}_p} = 1 + \sum_{i=1}^{\kappa} \frac{\mathbf{P}_p^T \mathbf{P}_p}{(\lambda_{ki} - \lambda_{pr})^2}
\]

\[(r = 1, 2, \ldots, p)\]. Considering the orthogonality properties of (2-1), the rectification conditions are satisfied for

\[
\frac{1}{\mathbf{x}_p^T \mathbf{x}_p} = 1 + \sum_{i=1}^{\kappa} \frac{\mathbf{P}_p^T \mathbf{P}_p}{(\lambda_{ki} - \lambda_{pr})^2}
\]

Let

\[
\lambda_{pr} = \lambda_{pr}^T = \mathbf{t}_{pp} + \mathbf{\lambda}_{pr} + \sum_{i=1}^{\kappa} \frac{\mathbf{P}_p^T \mathbf{P}_p}{\lambda_{ki} - \lambda_{pr}} = 0
\]

\[(r = 1, 2, \ldots, p),\) then

\[
f'(\lambda_{pr}) = \frac{df(\lambda_{pr})}{d(\lambda_{pr})} = 1 + \sum_{i=1}^{\kappa} \frac{\mathbf{P}_p^T \mathbf{P}_p}{(\lambda_{ki} - \lambda_{pr})^2}
\]

\[
= \frac{1}{\mathbf{x}_p^T \mathbf{x}_p}
\]
Without loss of generality, one can let \((x_{pr})_p = \overline{\pm}(x_{pr})_p\), choosing the sign so that \(1/(x_{pr})_p^2 = \pm f'(\overline{\lambda}_{pr})\) is positive. Therefore,

\[
(x_{pr})_p = \frac{1}{\sqrt{f'}(\overline{\lambda}_{pr})}
\]

\[
(x^T_{pr})_p = \frac{1}{\sqrt{f'}(\overline{\lambda}_{pr})}
\]

if \(f'(\overline{\lambda}_{pr}) > 0\), and

\[
(x_{pr})_p = \frac{1}{\sqrt{-f'}(\overline{\lambda}_{pr})}
\]

\[
(x^T_{pr})_p = \frac{1}{\sqrt{-f'}(\overline{\lambda}_{pr})}
\]

if \(f'(\overline{\lambda}_{pr}) < 0\).

The valuable control quantities can be determined using (2-1) and (2-11) for the \(t_{ip}\) and (2-1) and (2-16) for the \(t_{pi}\)

\[
\sum_{i=1}^{p} \lambda_{pi} = \sum_{i=1}^{p} t_{ii} = \text{Tr} (T_p)
\]

\[
\sum_{i=1}^{k} P_{pi}(x_{ki})_1 = t_{p1}
\]

\[
\sum_{i=1}^{k} P_{pi}(x_{ki})_2 = t_{p2}
\]

\[
\vdots
\]

\[
\sum_{i=1}^{k} P_{pi}(x_{ki})_k = t_{pk}
\]

\[
\sum_{i=1}^{p} P^T_{pi}(x_{ki})_1 = t_{1p}
\]

\[
\sum_{i=1}^{p} P^T_{pi}(x_{ki})_2 = t_{2p}
\]
\[ \sum_{i=1}^{k} P_{pi}^T (x_{ki})_i = t_{3p} \]
\[ \vdots \]
\[ \sum_{i=1}^{k} P_{pi}^T (x_{ki})_k = t_{kp}. \]

It is necessary here to consider what happens when one or more eigenvalues of \( T_p \) are not distinct from the eigenvalues of \( T_k \).

If \( \lambda_{pr} = \lambda_{ki} \) for some \( i \) and some \( r \) (\( i = 1, 2, \ldots, k; r = 1, 2, \ldots, p \)), say \( i = a \) and \( r = b \), then \( \lambda_{pb} = \lambda_{ka} \). Since \( x_{ka} \) is the eigenvector associated with \( \lambda_{ka} \), the vector \((x_{ka}, 0)\) will be an eigenvector of \( T_p \) associated with \( \lambda_{pb} \) if \( P_{pa} = 0 \). Similarly, \((x_{ka}, 0)\) will be an eigenvector of \( T_p^* \) associated with \( \lambda_{pb} \) if \( P_{pa}^T = 0 \).

If \( \lambda_{pb} = \lambda_{ka} \), then each of \( P_{pa}^T (x_{pb})_p \) and \( P_{pa} (x_{pb}^T)_p = 0 \) by (2-12) and (2-17). In this case, \( P_{pa}^T = 0 \) or \( (x_{pb})_p = 0 \) or both; and \( P_{pa} = 0 \) or \( (x_{pb}^T)_p = 0 \) or both. If \( P_{pa}^T = 0 \), it will be permissible for \( x_{pb}^T = (x_{ka}^T, 0) \). If \( P_{pa} = 0 \), it will be permissible for \( x_{pb} = (x_{ka}, 0) \).

If either \( P_{pa} \) or \( P_{pa}^T \) is zero, it is convenient to eliminate them from the escalator equation. In order to remove \( P_{pa} \) one should consider (2-20) in the following form:
(2-30) \[ \sum_{i=1}^{c-1} P_i \left[ \sum_{j=1}^{k} \left( x_{ki} \right) j (x_p) j \right] \]

\[ + \sum_{i=a+1}^{k} P_i \left[ \sum_{j=1}^{k} \left( x_{ki} \right) j (x_p) j \right] = -(t_{pp} - \lambda_p)(x_p)_p. \]

Using (2-23) and eliminating the \( P_{pa} \) term,

\[ (2-31) \sum_{i=1}^{c-1} \sum_{t=1}^{i-1} (\lambda_{kt} - \lambda_{pt}) \sum_{t=1}^{k} (\lambda_{kt} - \lambda_{pt}) P_i P_i^T \]

\[ + \sum_{i=a+1}^{k} \sum_{t=1}^{i-1} (\lambda_{kt} - \lambda_{pt}) \sum_{t=1}^{k} (\lambda_{kt} - \lambda_{pt}) P_i P_i^T = -D \]

where

\[ D = - \sum_{t=1}^{k} (\lambda_{kt} - \lambda_{pt})(t_{pp} - \lambda_{pp}). \]

One can see that \( P_{pa} \) has been eliminated from the escalator equation. In a similar manner, \( P_{pa} \) can be eliminated from (2-21) and (2-31) will result. Equation (2-31) will also be called the escalator equation.

Assuming that the remaining eigenvalues of \( T_k \) are distinct from the eigenvalues of \( T_p \), the escalator equation may be written

\[ (2-32) \sum_{i=1}^{c-1} \frac{P_i P_i^T}{(\lambda_{ki} - \lambda_{pi})} \]

\[ + \sum_{i=a+1}^{k} \frac{P_i P_i^T}{(\lambda_{ki} - \lambda_{pi})} = (t_{pp} - \lambda_{pp}). \]

If, however, \( P_{pi} \) or \( P_{pi}^T \) is zero for some other \( i \), then either one or both must be eliminated from the escalator equation in the same manner.
The eigenvectors corresponding to the eigenvalues of equation (2-32) are determined in a similar manner to (2-25) and (2-26) and they are, accurate within a numerical factor,

\[(2-33) \quad \frac{(x_{pr})_j}{(x_{pr})'_p} = - \sum_{i=1}^{k} \frac{p_{pi}(x_{ki})_j}{(\lambda_{ki} - \lambda_{pr})} \]

\[(2-34) \quad \frac{(x_{pr})^T_j}{(x_{pr})'_p} = - \sum_{i=1}^{k} \frac{p_{pi}(x_{ki})^T_j}{(\lambda_{ki} - \lambda_{pr})} \]

\[(j = 1, 2, \ldots, k; \ r = 1, 2, \ldots, b-1, b+1, \ldots, p).\]

Considering (2-1), (2-33), and (2-34) one can verify

\[(2-35) \quad \sum_{j=1}^{k} (x_{pr})^T_j (x_{pr})'_p \frac{(x_{pr})^T_j (x_{pr})'_p}{(x_{pr})'_p (x_{pr})'_p} = \sum_{i=1}^{k} \frac{p_{pi}(x_{ki})^T_j}{(\lambda_{ki} - \lambda_{pr})^2}.\]

Adding \(\frac{(x_{pr})^T_p (x_{pr})'_p}{(x_{pr})'_p (x_{pr})'_p}\) to both sides of (2-35) one sees that

\[\sum_{j=1}^{p} (x_{pr})^T_j (x_{pr})'_p = 1 + \sum_{i=1}^{k} \frac{p_{pi}(x_{ki})^T_j}{(\lambda_{ki} - \lambda_{pr})^2},\]

\[(r = 1, 2, \ldots, b-1, b+1, \ldots, p).\]

Considering the orthogonality properties of (2-1), the rectification conditions are satisfied for

\[\frac{1}{(x_{pr})'_p (x_{pr})'_p} = 1 + \sum_{i=1}^{k} \frac{p_{pi}(x_{ki})^T_j}{(\lambda_{ki} - \lambda_{pr})^2}.\]
Let

\[ f(\lambda_{pr}) = -t_{pp} + \lambda_{pr} + \sum_{i=1}^{k} \frac{P_i \sigma_i^2}{\lambda_{xi} - \lambda_{pr}} \]

then

\[ f'(\lambda_{pr}) = \frac{1}{T} \frac{T}{(x_{pr})_p (x_{pr})_p} \]

\( (r = 1, 2, \ldots, b-1, b+1, \ldots, p) \). Without loss of generality, one can let \( (x_{pr})_p = \pm (x_{pr}^T)_p \) where the sign is chosen so that \( 1/(x_{pr})^2 = \pm f'(\lambda_{pr}) \) is positive. Therefore,

\[ (x_{pr})_p = 1/\sqrt{f'(\lambda_{pr})} \]

\[ (x_{pr}^T)_p = (1/\sqrt{-f'(\lambda_{pr})}) \]

if \( f'(\lambda_{pr}) > 0 \), and

\[ (x_{pr})_p = 1/\sqrt{-f'(\lambda_{pr})} \]

\[ (x_{pr}^T)_p = (1/\sqrt{-f'(\lambda_{pr})}) \]

if \( f'(\lambda_{pr}) < 0 \), \( (r = 1, 2, \ldots, b-1, b+1, \ldots, p) \).

One now needs to determine \( x_{pb}^T \) if \( P_{pa} = 0 \) or \( x_{pb} \) if \( P_{pa} = 0 \). Using the last equation of (2-7), one can determine \( (x_{pb}^T)_p \) in terms of \( (x_{pb}^T)_j \) and using the last equation of (2-6) one can determine \( (x_{pb})_p \), i.e.,

\[ (2-37) \quad (x_{pb}^T)_p = - \left[ \sum_{i=1}^{k} t_{ip} (x_{pb})_j \right] / (t_{pp} - \lambda_{pb}) \]

and
\[(2-38) \quad (x_{pb})_p = - \left[ \frac{k}{i=1} t_{pi}(x_{pb})_i \right] / (t_{pp} - \lambda_{pb}).\]

In order to determine \((x^T_{pb})_j\) or \((x_{pb})_j\), \((j = 1, 2, \ldots, k)\), one must satisfy the rectification conditions of \((2-1)\) or \((2-5)\). Considering \((2-5)\),

\[(2-39) \quad \sum_{i=1}^{P} (x_{pi})_j (x_{pi}^T)_m = \delta_{jm},\]

\((j, m = 1, 2, \ldots, p)\), and one sees immediately that

\[\sum_{i=1}^{P} (x_{pi})_j (x_{pi}^T)_j = 1.\]

In this case, every element is known except \((x^T_{pb})_j\) or \((x_{pb})_j\), so that the element can be determined for \(j = 1, 2, \ldots, k\).

Using \((2-37)\) or \((2-38)\) appropriately, \((x^T_{pb})_p\) or \((x_{pb})_p\) can be computed. Thus the eigenvalues and eigenvectors of \(T_p\) and \(T^*_p\) can be determined if \(P_{pi} = 0\) for some \(i\), or \(P^T_{pi} = 0\) for some \(i\), i.e., if \(\lambda_{ki} = \lambda_{pr}\). If both \(P_{pi}\) and \(P^T_{pi}\) are zero for some \(i\), it is sufficient to say that \(x_{pb} = (x_{ka}, 0)\) and \(x^T_{pb} = (x^T_{ka}, 0)\) with the remaining eigenvectors being determined from \((2-33)\), \((2-34)\), and \((2-36)\).

If \(\lambda_{ki} = \lambda_{pr}\) for more than one \(i\) and \(r\), then \(P_{pi}\), if it is zero, or \(P^T_{pi}\), if it is zero, must be eliminated from the escalator equation and all other pertinent equations used in determining the eigenvectors. The same method used in eliminating \(P_{pi}\) and \(P^T_{pi}\) for one \(i\) is applicable. Hence it is left to the reader. The control quantities of \((2-29)\) are still applicable.
Example 1: Find the eigenvalues and eigenvectors of the 4 X 4 matrix $T_4$ where

$$T_4 = \begin{bmatrix}
i & -1 & 0 & 0 \\
0 & 1-i & i & 0 \\
0 & 0 & 2 & 0 \\
1 & -2i & 0 & 1+i
\end{bmatrix}$$

and

$$T_4^* = \begin{bmatrix}
-i & 0 & 0 & 1 \\
1-1 & 1+i & 0 & 2i \\
0 & -i & 2 & 0 \\
0 & 0 & 0 & 1-i
\end{bmatrix}.$$

Considering the 2 X 2 matrix $T_2$ formed by deleting the last two rows and columns of $T_4$, one finds that

$$T_2 = \begin{bmatrix}
i & -1 \\
0 & 1-i
\end{bmatrix}.$$

The characteristic equation of $T_2$ is $\det (T_2 - \lambda I)$ and

$$\det (T_2 - \lambda I) = \begin{vmatrix}
i-\lambda & -1 \\
0 & 1-i-\lambda
\end{vmatrix} = (i - \lambda)(1 - i - \lambda) = 0$$

so that $\lambda_{21} = i$ and $\lambda_{22} = 1 - i$ are the eigenvalues of $T_2$.

$$\text{Tr} (T_2) = 1 = \lambda_{21} + \lambda_{22} = 1$$

so that the first equation of (2-29) is satisfied.

To find the eigenvectors associated with $\lambda_{21}$ and $\lambda_{22}$ one must satisfy the equations $T_2 x_{21} = \lambda_{21} x_{21}$ and $T_2 x_{22} = \lambda_{22} x_{22}$.

From the first equation,
\[
\begin{bmatrix}
i & -1 \\
0 & 1-i
\end{bmatrix}
\begin{bmatrix}(x_{21})_1 \\
(x_{21})_2
\end{bmatrix} = \begin{bmatrix}i(x_{21})_1 \\
i(x_{21})_2
\end{bmatrix}
\]
\[
i(x_{21})_1 - (x_{21})_2 = i(x_{21})_1
\]
\[
(1 - i)(x_{21})_2 = i(x_{21})_2
\]
\[
(x_{21})_2 = 0
\]
\[
(x_{21})_1 = (x_{21})_1.
\]
From the second equation,
\[
\begin{bmatrix}
i & -1 \\
0 & 1-i
\end{bmatrix}
\begin{bmatrix}(x_{22})_1 \\
(x_{22})_2
\end{bmatrix} = \begin{bmatrix}(1 - i)(x_{22})_1 \\
i(x_{22})_2
\end{bmatrix}
\]
\[
i(x_{22})_1 - (x_{22})_2 = (1 - i)(x_{22})_1
\]
\[
(1 - i)(x_{22})_2 = (1 - i)(x_{22})_2
\]
\[
(x_{22})_2 = (x_{22})_2
\]
\[
(x_{22})_1 = (1/(2i - 1))(x_{22})_2
\]
\[
(x_{22})_1 = -((1 + 2i)/5)(x_{22})_2.
\]
Since the eigenvalues of \(T_2^*\) are the complex conjugates of those of \(T_2\), i.e., \(\overline{\lambda}_{21} = -i\) and \(\overline{\lambda}_{22} = 1 + i\), one sees that
\[
T_2^* T_2 x_{21} = \lambda_{21} x_{21}
\]
\[
\begin{bmatrix}
-i & 0 \\
-1 & 1+i
\end{bmatrix}
\begin{bmatrix}(x_{21})_1 \\
(x_{21})_2
\end{bmatrix} = \begin{bmatrix}-i(x_{21})_1 \\
-i(x_{21})_2
\end{bmatrix}
\]
\[
-i(x_{21})_1 = -i(x_{21})_1
\]
\[
-(x_{21})_1 + (1 + i)(x_{21})_2 = -i(x_{21})_2
\]
\[ (x_{21}^T)_1 = (x_{21}^T)_1 \]
\[ (x_{21}^T)_2 = ((1 - 2i)/5)(x_{21}^T)_1 \]
\[ T_{22}^T = \lambda_{22} x_{22}^T \]
\[ \begin{bmatrix} -i & 0 \\ -1 & 1+i \end{bmatrix} \begin{bmatrix} x_{22}^T \end{bmatrix} = \begin{bmatrix} (1 + i)(x_{22}^T)_1 \\ (1 + i)(x_{22}^T)_2 \end{bmatrix} \]
\[-i(x_{22}^T)_1 = (1 + i)(x_{22}^T)_2 \]
\[-(x_{22}^T)_1 + (1 + i)(x_{22}^T)_2 = (1 + i)(x_{22}^T)_2 \]
\[ (x_{22}^T)_1 = 0 \]
\[ (x_{22}^T)_2 = (x_{22}^T)_2. \]

Now to rectify \( x_{21}, x_{22}, x_{21}^T, \) and \( x_{22}^T, \) one must satisfy (2-1)

\[
\left[ \begin{array}{c}
(x_{21}^T)_1 \\
(x_{22}^T)_2
\end{array} \right] = \left[ \begin{array}{c}
(x_{21}^T)_1 - \frac{(1+2i)}{5}(x_{22}^T)_2 \\
0
\end{array} \right]
\]
\[
\begin{bmatrix}
(x_{21}^T)_1 & (1+2i)/5(x_{21}^T)_1 \\
0 & (x_{22}^T)_2
\end{bmatrix}
\begin{bmatrix}
(x_{21}^T)_1 - (1+2i)/5(x_{22}^T)_2 \\
0 & (x_{22}^T)_2
\end{bmatrix}
= \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]
\[
(x_{21}^T)_1(x_{21}^T)_1 + \frac{(1 + 2i)}{5}(x_{21}^T)_1(0) = 1
\]
\[-\frac{(1 + 2i)/5}{21} (x_{21}^T)_1(x_{22}^T)_2 + \frac{(1 + 2i)/5}{21} (x_{21}^T)_1(x_{22}^T)_2 = 0
\]
\[
(x_{22}^T)_2(x_{22}^T)_2 = 1
\]
\[
(x_{21}^T)_1(x_{21}^T)_1 = 1
\]
\[(x_{21}^T)_{1}(x_{22})_2 = c\]

where \(c\) is arbitrary, say 1 and
\[(x_{22}^T)_{2}(x_{22})_2 = 1.\]

Let \((x_{21})_{1} = 1\), then \((x_{21}^T)_{1} = 1\), \((x_{22})_2 = 1\) and \((x_{22}^T)_{2} = 1.\)

Hence \(x_{21} = (1, 0), x_{22} = (-(1 + 2i)/5, 1), x_{22}^T = (0, 1)\), and
\(x_{21}^* = (1, (1 - 2i)/5)\) are the rectified eigenvalues of \(T_2\) and \(T_2^*\) respectively.

From (2-11) and (2-16),
\[
P_{31} = t_{31}(x_{21})_1 + t_{32}(x_{22})_2 = (0)(1) + (0)(0) = 0
\]
\[
P_{32} = t_{31}(x_{21})_1 + t_{32}(x_{22})_2 = (0)(-1+2i)/5 + (0)(1) = 0
\]
\[
P_{31} = t_{13}(x_{21}^T)_{1} + t_{23}(x_{22}^T)_{2} = (0)(1) + (i)((1+2i)/5)
\]
\[
= (-2 + i)/5
\]
\[
P_{32} = t_{13}(x_{22}^T)_{1} + t_{23}(x_{22}^T)_{2} = (0)(0) + (i)(1) = i.
\]

For control, using the equations of (2-29),
\[
P_{31} = t_{31}(x_{21})_1 + P_{31} = 0 = (0)(1) + (0)(0)
\]
\[
P_{31} = t_{32}(x_{22})_2 = t_{32} = 0 = (0)(1+2i)/5 + (0)(1)
\]
\[
P_{31} = t_{13} = 0 = ((-2+1)/5)(1) + (i)(1-2i)/5
\]
\[
P_{31} = t_{23} = i = ((-2+i)/5)(0) + (i)(1).
\]

The escalator equation, determined from (2-31), is
\[0 = (\lambda_{21} - \lambda_{31})(\lambda_{22} - \lambda_{32})(t_{33} - \lambda_{33})\]
\[ 0 = (i - \lambda_{31})(1 - i - \lambda_{32})(2 - \lambda_{33}) \]

so that the eigenvalues of \( T_3 \) are \( \lambda_{31} = i, \lambda_{32} = 1 - i, \) and \( \lambda_{33} = 2. \) Using the first equation of (2-29) as a check,

\[ \text{Tr} (T_3) = 3 = i + 1 - i + 2 = \lambda_{31} + \lambda_{32} + \lambda_{33}. \]

Remembering that \((x_{ka}, 0)\) is an eigenvector of \( T_p \) associated with \( \lambda_{pb} \) if \( P_{pa} = 0, \) one can verify that \( x_{31} = (x_{21}, 0) \) and \( x_{32} = (x_{22}, 0), \) i.e.,

\[ x_{31} = (1, 0, 0) \]
\[ x_{32} = (-(1 + 2i)/5, 1, 0). \]

From (2-33) and (2-34) with \( P_{31} = 0 \) and \( P_{32} = 0, \)

\[ \frac{(x_{33})_1}{(x_{33})_3} = - \frac{((-2 + i)/5)(1)}{(i - 2)} - \frac{(i)(-1/5)(1 + 2i)}{1 - 1 - 2} = - \frac{1 - 3i}{10}, \]

\[ \frac{(x_{33})_2}{(x_{33})_3} = - \frac{((-2 + i)/5)(0)}{(i - 2)} - \frac{i(1)}{1-1-2} = \frac{i}{1 + i} = \frac{1 + i}{2}, \]

\[ \frac{(x_{33})_1}{(x_{33})_3} = 0, \]

\[ \frac{(x_{33})_2}{(x_{33})_3} = 0. \]

Using equation (2-36), one sees that \((x_{33})_3 \) and \((x_{33})_3^T = 1. \)

Hence \( x_{33} = (-1 + 3i)/10, (1+1)/2, 1 \) and \( x_{33}^T = (0, 0, 1). \)
From equation (2-39),

\[
\begin{align*}
(x_{31}^T)_1(x_{31})_1 + (x_{32}^T)_1(x_{32})_1 + (x_{33}^T)_1(x_{33})_1 &= 1 \\
(x_{31}^T)_2(x_{31})_1 + (x_{32}^T)_2(x_{32})_2 + (x_{33}^T)_2(x_{33})_2 &= 0 \\
(x_{31}^T)_1 + ((-1 + 2i)/5)(x_{32}^T)_1 &= 1 \\
(x_{32}^T)_1 &= 0
\end{align*}
\]

so that \((x_{32}^T)_1 = 0\) and \((x_{31}^T)_1 = 1\). Also,

\[
\begin{align*}
(x_{31}^T)_1(x_{31})_2 + (x_{32}^T)_1(x_{32})_2 + (x_{33}^T)_1(x_{33})_2 &= 0 \\
(x_{31}^T)_2 + (x_{32}^T)_2(x_{32})_2 + (x_{33}^T)_2(x_{33})_2 &= 1 \\
(x_{31}^T)_2 + ((-1 + 2i)/5)(x_{32}^T)_2 &= 0 \\
(x_{32}^T)_2 &= 1
\end{align*}
\]

so that \((x_{32}^T)_2 = 1\) and \((x_{31}^T)_2 = (1 - 2i)/5\). Using (2-37),

\[
\begin{align*}
(x_{31}^T)_3 &= \left[-\left((0)(x_{31}^T)_1 + (-1)(x_{31}^T)_2\right)/(2 + i)\right]/(2 + i) = 1/5 \\
(x_{32}^T)_3 &= \left[-\left((0)(x_{32}^T)_1 + (-1)(x_{32}^T)_2\right)/(2 - 1 - i)\right]/(2 - 1 - i) = i/(1 - i) = (-1 + i)/2.
\end{align*}
\]

Hence,

\[
\begin{align*}
x_{31}^T &= (1, (1 - 2i)/5, 1/5) \\
x_{32}^T &= (0, 1, (-1 + i)/2) \\
x_{33}^T &= (0, 0, 1)
\end{align*}
\]
$x_{31} = (1, 0, 0)$

$x_{32} = -(1 + 2i)/5, 1, 0)$

$x_{33} = -(1 + 3i)/10, (1 + i)/2, 1)$.

One can verify the properties of (2-1) with little difficulty as a means of checking.

Equations (2-11) and (2-16) yield

\[ P_{41} = t_{41}(x_{31})_1 + t_{42}(x_{31})_2 + t_{43}(x_{31})_3 \]
\[ = (1)(1) + (-2i)(0) + (0)(0) \]
\[ = 1 \]

\[ P_{42} = t_{41}(x_{32})_1 + t_{42}(x_{32})_2 + t_{43}(x_{32})_3 \]
\[ = (1)(-(1 + 2i)/5) + (-2i)(1) + (0)(0) \]
\[ = -(1 + 12i)/5 \]

\[ P_{43} = t_{41}(x_{33})_1 + t_{42}(x_{33})_2 + t_{43}(x_{33})_3 \]
\[ = (1)(-(1 + 3i)/10) + (-2i)(1 + i)/2) + (0)(1) \]
\[ = (9 - 13i)/10 \]

\[ P_{41}^T = t_{14}(\overline{x_{31}})_1 + t_{24}(\overline{x_{31}})_2 + t_{34}(\overline{x_{31}})_3 \]
\[ = (0)(1) + (0)(1 + 2i)/5 + (0)(1/5) \]
\[ = 0 \]

\[ P_{42}^T = t_{14}(\overline{x_{32}})_1 + t_{24}(\overline{x_{32}})_2 + t_{34}(\overline{x_{32}})_3 \]
\[ = 0 \]

\[ P_{43}^T = t_{14}(\overline{x_{33}})_1 + t_{24}(\overline{x_{33}})_2 + t_{34}(\overline{x_{33}})_3 \]
\[ = 0. \]
Using equation (2-29) as a check, one can verify that
\[ P_{41}(x_{31})_1 + P_{42}(x_{32})_1 + P_{43}(x_{33})_1 = 1 = t_{41} \]
\[ P_{41}(x_{31})_2 + P_{42}(x_{32})_2 + P_{43}(x_{33})_2 = -2i = t_{42} \]
\[ P_{41}(x_{31})_3 + P_{42}(x_{32})_3 + P_{43}(x_{33})_3 = 0 = t_{43}. \]

Similarly,
\[ \sum_{i=1}^{3} P_{4i}(x_{3i})_{j} = t_{j4} = 0 \]

since \( t_{j4} = 0 \) and \( P_{4i} = 0 \) for \( j = 1, 2, 3. \)

From (2-31), the escalator equation of \( T_4 \) is
\[ 0 = (\lambda_{31} - \lambda_{41})(\lambda_{32} - \lambda_{42})(\lambda_{33} - \lambda_{43})t_{44} - \lambda_{44} \]
\[ = (i - \lambda_{41})(1 - i - \lambda_{42})(2 - \lambda_{43})(1 + i - \lambda_{44}) \]
so that the eigenvalues of \( T_4 \) are \( \lambda_{41} = i, \lambda_{42} = 1 - i, \lambda_{43} = 2, \) and \( \lambda_{44} = 1 + i. \)

\[ \text{Tr} (T_4) = 4 + i = i + 1 - i + 2 + 1 + i \]
\[ = 4 + i \]
\[ = \sum_{i=1}^{4} \lambda_{4i}. \]

Since \( (x^T_{4a}, 0) \) is an eigenvector of \( T_4 \) associated with \( \lambda_{4b} \) if \( P_{4a} = 0, \) it is immediately obvious that \( x^T_{41} = (x^T_{31}, 0), \)
\( x^T_{42} = (x^T_{32}, 0), \) and \( x^T_{43} = (x^T_{33}, 0), \) i.e.,
\[ x^T_{41} = (1, (1 - 2i)/5, 1/5, 0) \]
\[ x_{4,2}^T = (0, 1, (-1 + i)/2, 0) \]
\[ x_{4,3}^T = (0, 0, 1, 0). \]

With \( p_{4,4}^T = 0 \), \( (i = 1, 2, 3) \), equation (2-33) gives \( (x_{44})_1 = 0 \), \( (x_{44})_2 = 0 \), and \( (x_{44})_3 = 0 \). From equation (2-34),

\[
\begin{pmatrix}
\frac{\langle x_{4,4}^T \rangle_1}{\langle x_{4,4}^T \rangle_4}
\end{pmatrix}
= -\frac{(-1+12i)/5)(0)}{1-i-1} - \frac{((9-13i)/10)(0)}{2-1-i}
= \frac{(1)(1)}{1-i-1} - \frac{((-1+12i)/5)(1)}{1-i-1} - \frac{((9-13i)/10)(1)}{2-1-i}
= -(2 - i)/2
\]

\[
\begin{pmatrix}
\frac{\langle x_{4,4}^T \rangle_2}{\langle x_{4,4}^T \rangle_4}
\end{pmatrix}
= -\frac{(-1+12i)/5)(0)}{1-i-1} - \frac{((-1+12i)/5)(1)}{1-i-1} - \frac{((9-13i)/10)(1)}{2-1-i}
= \frac{(1)(1/5)}{1} - \frac{((-1+12i)/5)(-1+1/2)}{2} - \frac{((9-13i)/10)(1)}{1-i}
= (-1 + 3i)/4
\]

and using (2-36), \( (x_{44})_4 = 1 \) and \( (x_{44}^T)_4 = 1 \). Hence,

\[ x_{44}^T = (1, -(2 + i)/2, -(1 + 3i)/4, 1) \]
\[ x_{44} = (0, 0, 0, 1). \]

In order to determine the remaining eigenvectors of \( T_4 \) one must satisfy equation (2-39), i.e.,

\[
\sum_{i=1}^{4} (x_{4i})_j (x_{4i}^T)_m = \delta_{jm},
\]

\((j, m = 1, 2, 3)\). Substituting appropriately into (2-39),
\[
\begin{align*}
(x_{41})_1(1) & = 1 \\
(x_{41})_1(1 + 2i)/5 + (x_{42})_1(1) & = 0 \\
(x_{41})_1(1/5) + (x_{42})_1(-1 - i)/2 + (x_{43})_1(1) & = 0
\end{align*}
\]

Also
\[
\begin{align*}
(x_{41})_2(1) & = 0 \\
(x_{41})_2(1 + 2i)/5 + (x_{42})_2(1) & = 1 \\
(x_{41})_2(1/5) + (x_{42})_2(-1 - i)/2 + (x_{43})_2(1) & = 0
\end{align*}
\]

and
\[
\begin{align*}
(x_{41})_3(1) & = 0 \\
(x_{41})_3(1 + 2i)/5 + (x_{42})_3(1) & = 0 \\
(x_{41})_3(1/5) + (x_{42})_3(-1 - i)/2 + (x_{43})_3(1) & = 1
\end{align*}
\]

\[
\begin{align*}
(x_{41})_3 & = 0 \\
(x_{42})_3 & = 0 \\
(x_{43})_3 & = 1.
\end{align*}
\]
Now to determine the fourth component using (2-38),

\[(x_{41})_4 = -\frac{[(1)(1) + (-2i)(0) + (0)(0)]}{(1+i-1-i)} = -1\]

\[(x_{42})_4 = -\frac{[(1)(-(1+2i)/5) + (-2i)(1) + (0)(0)]}{(1+i-1+i)}\]

\[= \frac{(12 - i)/10}{(1+i-1+i)}\]

\[(x_{43})_4 = -\frac{[(1)(-(1+3i)/10) + (-2i)(1+i)/2 + (0)(1)]}{(1+i-2)}\]

\[= \frac{(22 - 4i)/20}{(1+i-2)}\]

Therefore,

\[x_{41} = (1, 0, 0, -1)\]

\[x_{42} = \frac{(-1 + 2i)/5, 1, 0, (12 - i)/10}{(1+i-1+i)}\]

\[x_{43} = \frac{(-1 + 3i)/10, (1 + i)/2, 1, (22 - 4i)/20}{(1+i-2)}\]

\[x_{44} = (0, 0, 0, 1)\]

\[x_{41}^T = (1, (1 - 2i)/5, 1/5, 0)\]

\[x_{42}^T = (0, 1, (-1 + i)/2, 0)\]

\[x_{43}^T = (0, 0, 1, 0)\]

\[x_{44}^T = (1, -(2 + i)/2, -(1 + 3i)/4, 1)\].

Hence the eigenvalues and eigenvectors of \(T_4\) have been found.

Although the escalator method is voluminous when done by hand, it can be adapted to a computer without a great deal of difficulty. The difficulty would arise in the complex arithmetic. If, however, the matrix considered were real, then very little difficulty should be encountered, since the entire
formulation would be simplified. One will note that the equations are greatly simplified if the matrix is real and symmetric, since each element with a T would be equivalent to the element without the T (1, p. 268).

The form of the escalator equation for a real matrix allows one to use Newton's approximation method for finding the roots of a polynomial. It is, if employed, the only approximation in the escalator method.

**Example 2:** Find the eigenvalues and eigenvectors of $T_3$ where

$$T_3 = \begin{bmatrix} 4 & -7 & 3 \\ 1 & 2 & 5 \\ -1 & 2 & -1 \end{bmatrix}$$

and

$$T_3^* = \begin{bmatrix} 4 & 1 & -1 \\ -7 & 2 & 2 \\ 3 & 5 & -1 \end{bmatrix}.$$  

Considering the 2 X 2 matrix $T_2$ formed by deleting the last row and column of $T_3$, one finds that

$$T_2 = \begin{bmatrix} 4 & -7 \\ 1 & 2 \end{bmatrix},$$

$$\det (T_2 - \lambda I) = \begin{vmatrix} 4-\lambda & -7 \\ 1 & 2-\lambda \end{vmatrix} = (4 - \lambda)(2 - \lambda) + 7 = 0$$

so that $\lambda_{21} = 3 + \sqrt{5}i$ and $\lambda_{22} = 3 - \sqrt{5}i$ are the eigenvalues
of $T_2$. Using (2-29) as a check,

$\text{Tr } (T_2) = 6 = \lambda_{21} + \lambda_{22} = 6.$

$T_{2x}^{x_2} = \lambda_{21}^{x_2} \lambda_{21}^{x_2}$

$\begin{bmatrix} 4 & -7 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{21} \end{bmatrix} = \begin{bmatrix} (3 + \sqrt{3}i) x_{21} \\ (3 + \sqrt{3}i) x_{21} \end{bmatrix}$

$4(x_{21})_1 - 7(x_{21})_2 = (3 + \sqrt{3}i)(x_{21})_1$

$(x_{21})_1 + 2(x_{21})_2 = (3 + \sqrt{3}i)(x_{21})_2$

$(x_{21})_1 = (1 + \sqrt{3}i)(x_{21})_2$

$T_{2x}^{x_2} = \lambda_{22}^{x_2} \lambda_{22}^{x_2}$

$\begin{bmatrix} 4 & -7 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_{22} \\ x_{22} \end{bmatrix} = \begin{bmatrix} (3 - \sqrt{3}i) x_{22} \\ (3 - \sqrt{3}i) x_{22} \end{bmatrix}$

$4(x_{22})_1 - 7(x_{22})_2 = (3 - \sqrt{3}i)(x_{22})_1$

$(x_{22})_1 + 2(x_{22})_2 = (3 - \sqrt{3}i)(x_{22})_1$

$(x_{22})_1 = (1 - \sqrt{3}i)(x_{22})_2$.

Since $T_2$ and $T_2^*$ are complex conjugates, the eigenvalues of $T_2^*$

are the complex conjugates of those of $T_2$, i.e.,

$\lambda_{21} = 3 - \sqrt{3}i$

$\lambda_{22} = 3 + \sqrt{3}i$.

$T_{2x}^{x_2} = T_{2x}^{x_2} = \bar{\lambda}_{21}^{x_2} \lambda_{21}^{x_2}$
\[
\begin{bmatrix}
4 & 1 \\
-7 & 2
\end{bmatrix}
\begin{bmatrix}
x_{21}^T \\
x_{22}^T
\end{bmatrix}_1
= \begin{bmatrix}
(3 - \sqrt{5}i)(x_{21}^T)_1 \\
(3 - \sqrt{5}i)(x_{21}^T)_2
\end{bmatrix}
\]

\[
\begin{bmatrix}
4 & 1 \\
-7 & 2
\end{bmatrix}
\begin{bmatrix}
x_{21}^T \\
x_{22}^T
\end{bmatrix}_2
= \begin{bmatrix}
(3 - \sqrt{5}i)(x_{21}^T)_1 \\
(3 - \sqrt{5}i)(x_{21}^T)_2
\end{bmatrix}
\]

\[
4(x_{21}^T)_1 + (x_{21}^T)_2 = (3 - \sqrt{5}i)(x_{21}^T)_1
\]

\[
-7(x_{21}^T)_1 + 2(x_{21}^T)_2 = (3 - \sqrt{5}i)(x_{21}^T)_2
\]

\[
(x_{21}^T)_1 = \frac{-1 - \sqrt{5}i}{7}(x_{21}^T)_2
\]

\[
(x_{21}^T)_2 = \frac{1}{\sqrt{2}}x_{22}^T
\]

\[
\begin{bmatrix}
4 & 1 \\
-7 & 2
\end{bmatrix}
\begin{bmatrix}
x_{22}^T \\
x_{22}^T
\end{bmatrix}_1
= \begin{bmatrix}
(3 + \sqrt{5}i)(x_{22}^T)_1 \\
(3 + \sqrt{5}i)(x_{22}^T)_2
\end{bmatrix}
\]

\[
\begin{bmatrix}
4 & 1 \\
-7 & 2
\end{bmatrix}
\begin{bmatrix}
x_{22}^T \\
x_{22}^T
\end{bmatrix}_2
= \begin{bmatrix}
(3 + \sqrt{5}i)(x_{22}^T)_1 \\
(3 + \sqrt{5}i)(x_{22}^T)_2
\end{bmatrix}
\]

\[
4(x_{22}^T)_1 + (x_{22}^T)_2 = (3 + \sqrt{5}i)(x_{22}^T)_1
\]

\[
-7(x_{22}^T)_1 + 2(x_{22}^T)_2 = (3 + \sqrt{5}i)(x_{22}^T)_2
\]

\[
(x_{22}^T)_1 = \frac{-1 + \sqrt{5}i}{7}(x_{22}^T)_2
\]

Rectifying $x_{21}$, $x_{22}$, $x_{21}^T$, and $x_{22}^T$, one finds

\[
(x_{21}^T)_2(x_{21})_2 = \frac{6 + \sqrt{5}i}{12}
\]

\[
(x_{22}^T)_2(x_{22})_2 = \frac{6 - \sqrt{5}i}{12}.
\]

Letting $(x_{21})_2 = \frac{6 + \sqrt{5}i}{12}$ and $(x_{22})_2 = \frac{6 - \sqrt{5}i}{12}$,

one sees that $(x_{21})_2 = 1$ and $(x_{22})_2 = 1$ so that the remaining components of the eigenvectors can be determined, i.e.,

\[
(x_{21})_1 = (1 + \sqrt{5}i)(6 + \sqrt{5}i)/12 = (7\sqrt{5}i)/12
\]
\[(x_{22}^T)_1 = (1 - \sqrt{6} i)(6 - \sqrt{6} i)/12 = (-7\sqrt{6} i)/12\]

\[(x_{21}^T)_1 = -(1 - \sqrt{6} i)/7\]

\[(x_{22}^T)_1 = -(1 + \sqrt{6} i)/7\].

Hence,

\[x_{21} = (7\sqrt{6} i/12, (6 + \sqrt{6} i)/12)\]

\[x_{22} = (-7\sqrt{6} i/12, (6 - \sqrt{6} i)/12)\]

\[x_{21}^T = (-(1 - \sqrt{6} i)/7, 1)\]

\[x_{22}^T = (-(1 + \sqrt{6} i)/7, 1)\],

and these are the rectified eigenvectors of \(T_2\) and \(T_*\).

From (2-11) and (2-16),

\[p_{31} = (-1)(7\sqrt{6} i/12) + (2)(6 + \sqrt{6} i)/12 = (12 - 5\sqrt{6} i)/12\]

\[p_{32} = (-1)(-7\sqrt{6} i/12) + (2)(6 - \sqrt{6} i)/12 = (12 + 5\sqrt{6} i)/12\]

\[p_{21}^T = (3)(-(1 + \sqrt{6} i)/7) + (5)(1) = (32 - 3\sqrt{6} i)/7\]

\[p_{32}^T = (3)(-(1 - \sqrt{6} i)/7) + (5)(1) = (32 + 3\sqrt{6} i)/7\].

The escalator equation, determined from (2-24), is

\[
\frac{p_{31}^T p_{31}}{\lambda_{21} - \lambda_3} + \frac{p_{32}^T p_{32}}{\lambda_{22} - \lambda_3} = (\nu_{33} - \lambda_3)
\]

\[
\begin{bmatrix}
12 - 5\sqrt{6} i \\
12 \\
3 + \sqrt{3} i - \lambda_3
\end{bmatrix}
\begin{bmatrix}
12 - 5\sqrt{6} i \\
12 \\
3 + \sqrt{3} i - \lambda_3
\end{bmatrix}
\begin{bmatrix}
12 + 3\sqrt{6} i \\
12 \\
3 - \sqrt{3} i - \lambda_3
\end{bmatrix}
\begin{bmatrix}
12 + 3\sqrt{6} i \\
12 \\
3 - \sqrt{3} i - \lambda_3
\end{bmatrix}
= -1 - \lambda_3
\]
\(-7 - 7\lambda_3 = -\lambda_3^3 + 5\lambda_3^2 - 9\lambda_3 - 15\)
\[\lambda_3^3 - 5\lambda_3^2 + 2\lambda_3 + 8 = 0\]

\((\lambda_3 - 4)(\lambda_3 - 2)(\lambda_3 + 1) = 0\)

so that the eigenvalues of \(T_3\) are \(\lambda_{31} = 4\), \(\lambda_{32} = 2\), and \(\lambda_{33} = -1\). The \(\text{Tr}(T_3) = 4 + 2 - 1 = 5 = \lambda_{31} + \lambda_{32} + \lambda_{33}\).

From (2-25),
\[
\begin{bmatrix}
\frac{x_{31}}{x_{31}} \\
\frac{x_{32}}{x_{32}} \\
\frac{x_{33}}{x_{33}}
\end{bmatrix} = -\begin{bmatrix}
\frac{32 - 3\sqrt{3}}{7} & 7\sqrt{3} & 7 \\
12 & 12 & 12 \\
3 + \sqrt{3} & 1 - \sqrt{3}
\end{bmatrix}
- \begin{bmatrix}
\frac{32 + 3\sqrt{3}}{7} & -7\sqrt{3} & -7 \\
12 & 12 & 12 \\
3 - \sqrt{3} & 1 + \sqrt{3}
\end{bmatrix} = -\frac{29}{7}
\]
\[
\begin{bmatrix}
\frac{x_{31}}{x_{31}} \\
\frac{x_{32}}{x_{32}} \\
\frac{x_{33}}{x_{33}}
\end{bmatrix} = -\begin{bmatrix}
\frac{32 - 3\sqrt{3}}{7} & 6 + \sqrt{3} & 6 - \sqrt{3} \\
12 & 12 & 12 \\
-1 + \sqrt{3} & 1 - \sqrt{3}
\end{bmatrix}
- \begin{bmatrix}
\frac{32 + 3\sqrt{3}}{7} & 6 - \sqrt{3} & 6 + \sqrt{3} \\
12 & 12 & 12 \\
-1 - \sqrt{3} & 1 + \sqrt{3}
\end{bmatrix} = \frac{3}{7}
\]

\[
\begin{bmatrix}
\frac{x_{31}}{x_{31}} \\
\frac{x_{32}}{x_{32}} \\
\frac{x_{33}}{x_{33}}
\end{bmatrix} = -\begin{bmatrix}
\frac{32 - 5\sqrt{3}}{12} & 1 - \sqrt{3} & 1 + \sqrt{3} \\
12 & 12 & 12 \\
-1 + \sqrt{3} & 1 - \sqrt{3}
\end{bmatrix}
- \begin{bmatrix}
\frac{32 + 5\sqrt{3}}{12} & 1 + \sqrt{3} & 1 - \sqrt{3} \\
12 & 12 & 12 \\
-1 - \sqrt{3} & 1 + \sqrt{3}
\end{bmatrix} = 0
\]
\[
\begin{bmatrix}
\frac{x_{31}}{x_{31}} \\
\frac{x_{32}}{x_{32}} \\
\frac{x_{33}}{x_{33}}
\end{bmatrix} = -\begin{bmatrix}
\frac{12 - 5\sqrt{3}}{12} & 12 & 12 \\
12 & 12 & 12 \\
-1 + \sqrt{3} & -1 - \sqrt{3}
\end{bmatrix}
- \begin{bmatrix}
\frac{12 + 5\sqrt{3}}{12} & 12 & 12 \\
12 & 12 & 12 \\
-1 - \sqrt{3} & -1 + \sqrt{3}
\end{bmatrix} = 1.
\]

\[z'^{(\lambda_{31})} = 1 + \frac{117 - 98\sqrt{3}i}{42(-1 + \sqrt{3}i)^2} + \frac{117 + 98\sqrt{3}i}{42(-1 - \sqrt{3}i)^2}\]
\[= 1 + \frac{(1/42)(441 + 784\sqrt{3}i + 441 - 784\sqrt{3}i)/49}{49}\]
\[= 1 + \frac{21/49}{49}\]
\[= \frac{70}{49}\]
\[= \frac{10}{7}.
\]
Using equation (2-26), one finds that \((x_3)_3 = \sqrt{7}/10\) and \((x^{T}_{31})_3 = \sqrt{7}/10\). Hence
\[x_{31} = \left( -(29/7)\sqrt{7}/10, (3/7)\sqrt{7}/10, \sqrt{7}/10 \right)\]
\[x^T_{31} = (0, \sqrt{7}/10, \sqrt{7}/10).\]

Similarly for \(\lambda_{32}\) and \(\lambda_{33}\)
\[
\frac{(x_{32})_1}{(x_{32})_3} = -5
\]
\[
\frac{(x_{32})_2}{(x_{32})_3} = -1
\]
\[
\frac{(x_{32}^T)_1}{(x_{32}^T)_3} = \frac{2}{7}
\]
\[
\frac{(x_{32}^T)_2}{(x_{32}^T)_3} = \frac{3}{7}
\]
\[2! \lambda_{32} = -\frac{6}{7}
\]
\[\lambda_{32}_3 = \sqrt{7}/6\]
\[\lambda_{32}^T_3 = -\sqrt{7}/6\]

so that
\[x_{32} = \left(-5\sqrt{7}/\sqrt{6}, -\sqrt{7}/6, \sqrt{7}/6 \right)\]
\[x^T_{32} = \left(-2\sqrt{7}/7\sqrt{6}, -3\sqrt{7}/7\sqrt{6}, -\sqrt{7}/6 \right)\]

and
\[
\frac{(x_{33})_1}{(x_{33})_3} = -2
\]
so that

\[
\begin{align*}
\left( \frac{x_{33}}{x_{33}} \right)_2 &= -1 \\
\left( \frac{x_{33}}{x_{33}} \right)_3 &= -\frac{3}{2} \\
\left( \frac{x_{33}}{x_{33}} \right)_1 &= \frac{5}{2} \\
\left( \frac{x_{33}}{x_{33}} \right)_2 &= \frac{22}{\sqrt{15}} \\
\left( x_{33} \right)_3 &= \sqrt{\frac{22}{15}} \\
\left( x_{33}^T \right)_3 &= \sqrt{\frac{22}{15}}
\end{align*}
\]

Thus the eigenvalues and eigenvectors of \( T_3 \) have been determined. As a check, one can verify that these eigenvalues and eigenvectors satisfy the control equations.

**Example 3:** Find the eigenvalues and eigenvectors of the matrix \( T_4 \) where

\[ T_4 = \begin{bmatrix}
4 & 7 & 7 & 0 \\
7 & -2 & 0 & 0 \\
7 & 0 & 10 & 0 \\
0 & 0 & 0 & -2
\end{bmatrix}. \]
Note: Since $T_4$ is symmetric, the escalator method is simplified for $T_4 = T_4^T$ so that any term with a $T$ will be the same as the term without the $T$.

$$T_2 = \begin{bmatrix} 4 & 7 \\ 7 & -2 \end{bmatrix}$$

$$\det (T_2 - \lambda I) = \begin{vmatrix} 4-\lambda & 7 \\ 7 & -2-\lambda \end{vmatrix} = (4 - \lambda)(-2 - \lambda) - 49 = 0$$

so that $\lambda_{21} = 1 + \sqrt{58}$ and $\lambda_{22} = 1 - \sqrt{58}$ are the eigenvalues of $T_2$. From (2-29) one can immediately verify that

$$\text{Tr}(T_2) = 2 = \lambda_{21} + \lambda_{22} = 2.$$

$$T_2x_{21} = \lambda_{21}x_{21}$$

$$\begin{bmatrix} 4 & 7 \\ 7 & -2 \end{bmatrix} \begin{bmatrix} x_{211} \\ x_{212} \end{bmatrix} = \begin{bmatrix} (1 + \sqrt{58})x_{211} \\ (1 + \sqrt{58})x_{212} \end{bmatrix}$$

$$4(x_{211})_1 + 7(x_{212})_1 = (1 + \sqrt{58})(x_{211})_1$$

$$7(x_{211})_1 - 2(x_{212})_1 = (1 + \sqrt{58})(x_{212})_1$$

$$(x_{211})_1 = (3 + \sqrt{58})/7)(x_{211})_2 = (x_{211})_1$$

$$T_2x_{22} = \lambda_{22}x_{22}$$

$$\begin{bmatrix} 4 & 7 \\ 7 & -2 \end{bmatrix} \begin{bmatrix} x_{221} \\ x_{222} \end{bmatrix} = \begin{bmatrix} (1 - \sqrt{58})x_{221} \\ (1 - \sqrt{58})x_{222} \end{bmatrix}$$

$$4(x_{221})_1 + 7(x_{222})_1 = (1 - \sqrt{58})(x_{221})_1$$

$$7(x_{221})_1 - 2(x_{222})_1 = (1 - \sqrt{58})(x_{222})_1$$

$$(x_{221})_1 = (3 - \sqrt{58})/7)(x_{222})_2 = (x_{221})_1$$
Remembering that \((x_{21})_2 = (x_{21}^n)_2\) and \((x_{22})_2 = (x_{22}^n)_2\), one finds by rectifying the eigenvectors of \(T_2\) and \(T^*_2\) that
\[
\begin{align*}
(3 + \sqrt{58})/7^2 + 1 &= 1 \\
-(x_{21})_2(x_{22})_2 + (x_{21})_2(x_{22})_2 &= 0 \\
-(x_{22})_2(x_{21})_2 + (x_{21})_2(x_{22})_2 &= 0 \\
(3 - \sqrt{58})/7^2 + 1 &= 1 \\
\end{align*}
\]
\[
\begin{align*}
(x_{21})_2 &= \sqrt{\frac{58 - 3\sqrt{58}}{2(58)}} = (x_{21}^T)_2 \\
(x_{21})_1 &= \frac{3 + \sqrt{58}}{7} \sqrt{\frac{58 - 3\sqrt{58}}{2(58)}} = (x_{21}^T)_1 \\
(x_{22})_2 &= \sqrt{\frac{58 + 3\sqrt{58}}{2(58)}} = (x_{22}^T)_2 \\
(x_{22})_1 &= \frac{3 - \sqrt{58}}{7} \sqrt{\frac{58 + 3\sqrt{58}}{2(58)}} = (x_{22}^T)_1. \\
\end{align*}
\]
From (2-16),
\[
\begin{align*}
P_{31} &= (3 + \sqrt{58})\sqrt{\frac{58 - 3\sqrt{58}}{2(58)}} = P_{31}^T \\
P_{32} &= (3 - \sqrt{58})\sqrt{\frac{58 + 3\sqrt{58}}{2(58)}} = P_{32}^T. \\
\end{align*}
\]
The escalator equation of \(T_3\) is
\[
\begin{align*}
\frac{P_{31}^2}{\lambda_{21} - \lambda_3} + \frac{P_{32}^2}{\lambda_{22} - \lambda_3} &= \tau_{33} - \lambda_3 \\
\frac{19(58 + 3\sqrt{58})}{2(58)} + \frac{19(58 - 3\sqrt{58})}{2(58)} &= 10 - \lambda_3 \\
\end{align*}
\]
\[
\begin{align*}
&\sqrt{58}, \begin{bmatrix} -232 & -116\lambda_3 \end{bmatrix} = (10 - \lambda_3)(\lambda_3^2 - 2\lambda_3 - 57) \\
-58 + 49\lambda_3 &= -\lambda_3^3 + 12\lambda_3^2 + 37\lambda_3 - 570
\end{align*}
\]
\[ \lambda_3^3 - 12\lambda_3^2 - 86\lambda_3 + 472 = 0 \]

\[(\lambda_3 - 4)(\lambda_3 - 4 - \sqrt{134})(\lambda_3 - 4 + \sqrt{134}) = 0.\]

Hence \(\lambda_{31} = 4\), \(\lambda_{32} = 4 + \sqrt{134}\), and \(\lambda_{33} = 4 - \sqrt{134}\) are the eigenvalues of \(T_3\).

Using (2-25), the eigenvectors for \(T_3\) can be determined.

\[
\begin{align*}
\frac{x_{31}}{(x_{31})_3} &= \frac{(3 + \sqrt{58}) \sqrt{58 - 3\sqrt{58}} + 3 + \sqrt{58} \sqrt{58 - 3\sqrt{58}}}{1 + \sqrt{58} - 4} \\
&\quad - \frac{(3 - \sqrt{58}) \sqrt{58 + 3\sqrt{58}} - 3 - \sqrt{58} \sqrt{58 + 3\sqrt{58}}}{1 - \sqrt{58} - 4} \\
&= - \frac{7}{2(58)} \left[ \frac{58 + 3\sqrt{58}}{-3 + \sqrt{58}} + \frac{58 - 3\sqrt{58}}{-3 - \sqrt{58}} \right] \\
&= - \frac{6}{7}.
\end{align*}
\]

\[
\begin{align*}
\frac{x_{32}}{(x_{32})_3} &= \frac{(3 + \sqrt{58}) \sqrt{58 - 3\sqrt{58}} \sqrt{58 - 3\sqrt{58}}}{1 + \sqrt{58} - 4} \\
&\quad - \frac{(3 - \sqrt{58}) \sqrt{58 + 3\sqrt{58}} \sqrt{58 + 3\sqrt{58}}}{1 - \sqrt{58} - 4} \\
&= -1.
\end{align*}
\]

\[
\begin{align*}
\pi'(\lambda_3) &= 1 + \frac{(3 + \sqrt{58})^2 \left( \frac{58 - 3\sqrt{58}}{2(58)} \right)^2}{(1 + \sqrt{58} - 4)^2} + \frac{(3 - \sqrt{58})^2 \left( \frac{58 + 3\sqrt{58}}{2(58)} \right)^2}{(1 - \sqrt{58} - 4)^2} \\
&= 1 + 85/49 = 134/49 \\
\Rightarrow (x_{31})_3 &= 1/\sqrt{134}/49 = 7\sqrt{134}/134 = (x_{31})_3 \\
\Rightarrow (x_{31})_1 &= -(6/7)(7\sqrt{134}/134) = -3\sqrt{134}/67
\end{align*}
\]
\[(x_{12})_2 = -7\sqrt{134}/134.\]

Therefore,

\[x_{12} = (-3\sqrt{134}/57, -7\sqrt{134}/134, 7\sqrt{134}/134) = x_{31}^T.\]

Also

\[
\frac{(x_{12})_1}{(x_{12})_3} = -\frac{(3 + \sqrt{58})\left(\frac{58 - 3\sqrt{58}}{2(58)}\right) + \frac{58 - 3\sqrt{58}}{2(58)}}{1 + \sqrt{58} - 4 - \sqrt{134}}
\]

\[
\frac{(x_{12})_2}{(x_{12})_3} = -\frac{(3 - \sqrt{58})\left(\frac{58 + 3\sqrt{58}}{2(58)}\right) - \frac{58 + 3\sqrt{58}}{2(58)}}{1 + \sqrt{58} - 4 - \sqrt{134}}
\]

\[
\frac{(x_{32})_1}{(x_{32})_3} = -\frac{(3 + \sqrt{58})\left(\frac{58 - 3\sqrt{58}}{2(58)}\right) + \frac{58 - 3\sqrt{58}}{2(58)}}{1 + \sqrt{58} - 4 - \sqrt{134}}
\]

\[
\frac{(x_{32})_2}{(x_{32})_3} = -\frac{(3 - \sqrt{58})\left(\frac{58 + 3\sqrt{58}}{2(58)}\right) - \frac{58 + 3\sqrt{58}}{2(58)}}{1 + \sqrt{58} - 4 - \sqrt{134}}
\]

\[
x^T(\Lambda_{32}) = 1 + \frac{(3 + \sqrt{58})^2\left(\frac{58 - 3\sqrt{58}}{2(58)}\right)^2}{(1 + \sqrt{58} - 4 - \sqrt{134})^2} + \frac{(3 + \sqrt{58})^2\left(\frac{58 + 3\sqrt{58}}{2(58)}\right)^2}{(1 + \sqrt{58} - 4 - \sqrt{134})^2}
\]

\[
= 1 + \frac{993571 - 76792\sqrt{134}}{117649}
\]

\[
= 268(65 - 6\sqrt{134})/2401.
\]

\[
(x_{32})_3 = 1/\sqrt{268(65 - 6\sqrt{134})/2401}
\]

\[
= 49/(2\sqrt{87(65 - 6\sqrt{134})})
\]

\[
= \sqrt{87(65 + 6\sqrt{134})}/134
\]

\[
= (67 + 3\sqrt{134})/134.
\]
\((x_{32})_1 = 7\sqrt{134}/134\)

\((x_{32})_2 = (67 - 3\sqrt{134})/134\).

Also,

\[
\begin{align*}
\frac{(x_{33})_1}{(x_{33})_3} &= -\frac{1}{7}(3 + \sqrt{58})^2 \left(\frac{58 - 3\sqrt{58}}{2(58)}\right)^2 - \frac{1}{7}(3 - \sqrt{58})^2 \left(\frac{58 + 3\sqrt{58}}{2(58)}\right)^2 \\
&= -(6 + \sqrt{134})/7.
\end{align*}
\]

\[
\begin{align*}
\frac{(x_{33})_2}{(x_{33})_3} &= -\frac{1}{7}(3 + \sqrt{58})^2 \left(\frac{58 - 3\sqrt{58}}{2(58)}\right)^2 - \frac{1}{7}(3 - \sqrt{58})^2 \left(\frac{58 + 3\sqrt{58}}{2(58)}\right)^2 \\
&= \frac{(65 + 6\sqrt{134})}{49}.
\end{align*}
\]

\[
\begin{align*}
\pi'((\gamma_{33})_1) &= 1 + \frac{(3 + \sqrt{58})^2 \left(\frac{58 - 3\sqrt{58}}{2(58)}\right)^2}{(-3 + \sqrt{58} + \sqrt{134})^2} + \frac{(3 - \sqrt{58})^2 \left(\frac{58 + 3\sqrt{58}}{2(58)}\right)^2}{(-3 + \sqrt{58} + \sqrt{134})^2} \\
&= 1 + \frac{998571 + 73792\sqrt{134}}{117649} \\
&= 268(85 + 6\sqrt{134})/2401.
\end{align*}
\]

\[
(x_{33})_3 = 1/\sqrt{268(85 + 6\sqrt{134})}/2401
\]

\[
= 49/(2\sqrt{67(85 + 6\sqrt{134})})
\]

\[
= \sqrt{67(85 - 6\sqrt{134})}/134
\]

\[
= (67 - 3\sqrt{134})/134.
\]

\((x_{33})_1 = -7\sqrt{134}/134\)

\((x_{33})_2 = (67 + 3\sqrt{134})/134\).

From (2-16),

\[
P_{41} = \pi_{41}(x_{31})_1 + \pi_{42}(x_{31})_2 + \pi_{43}(x_{31})_3 = 0 = P^T_{41}
\]

\[
P_{42} = \pi_{41}(x_{32})_1 + \pi_{42}(x_{32})_2 + \pi_{43}(x_{32})_3 = 0 = P^T_{42}
\]
\[ P_{43} = t_{41}(x_{33})_1 + t_{42}(x_{33})_2 + t_{43}(x_{33})_3 = 0 = P^T_{43} \]

since \( t_{4i} = 0 \) for \( i = 1, 2, 3 \). The control equations of (2-29) are satisfied since \( P_{41} = 0 \) and \( t_{4i} = 0 \) for \( i = 1, 2, 3 \).

From (2-31), the escalator equation of \( T_i \) is
\[ 0 = (4 - \lambda_{41})(4 + \sqrt{134} - \lambda_{42})(4 - \sqrt{134} - \lambda_{43})(-2 - \lambda_{44}) \]
so that the eigenvalues of \( T_i \) are \( \lambda_{41} = 4, \lambda_{42} = 4 + \sqrt{134}, \lambda_{43} = 4 - \sqrt{134}, \) and \( \lambda_{44} = -2 \).

Since \( P_{pi} = P^T_{pi} = 0, i = 1, 2, 3 \), it is sufficient for
\[ x_{41} = (x_{31}, 0) = x^T_{41} \]
\[ x_{42} = (x_{32}, 0) = x^T_{42} \]
\[ x_{43} = (x_{33}, 0) = x^T_{43} \]

One now needs to determine \( x_{44} \) and \( x^T_{44} \) so that \( x_{41}, x_{42}, x_{43}, x_{44}, x^T_{41}, x^T_{42}, x^T_{43} \), and \( x^T_{44} \) are rectified. In order to do this it is sufficient to satisfy (2-39). One can see without much difficulty that it will be sufficient for \( x_{44} = (0, 0, 0, 1) \) and \( x^T_{44} = (0, 0, 0, 1) \) in order that (2-39) be satisfied.

Hence,
\[ x_{41} = \left( -3\sqrt{134}/57, -7\sqrt{134}/134, 7\sqrt{134}/134, 0 \right) = x^T_{41} \]
\[ x_{42} = \left( 7\sqrt{134}/134, (67 - 3\sqrt{134})/134, (67 + 3\sqrt{134})/134, 0 \right) = x^T_{42} \]
\[ x_{43} = \left( -7\sqrt{134}/134, (67 + 3\sqrt{134})/134, (67 - 3\sqrt{134})/134, 0 \right) = x^T_{43} \]
\[ x_{44} = (0, 0, 0, 1) = x^T_{44}. \]
CHAPTER BIBLIOGRAPHY


CHAPTER III

THE METHOD OF ORTHOGONALIZATION

OF SUCCESSIVE ITERATIONS

The method of orthogonalization of successive iterations (1, pp. 277-286) is aimed toward finding a linear combination (for a sequence of iterations of an arbitrary real vector for a real diagonalizable matrix \( T \) of order \( n \)) which is equal to zero. In this method, the orthogonalization process shall be applied to achieve this goal.

Starting with a real non-zero vector \( \mathbf{x}_1 \), construct its iteration \( \mathbf{T}\mathbf{x}_1 \) and orthogonalize it with \( \mathbf{x}_1 \). This is done by constructing a vector \( \mathbf{x}_2 = \mathbf{T}\mathbf{x}_1 + s_{11}\mathbf{x}_1 \) such that \( (\mathbf{x}_1, \mathbf{x}_2) = 0 \).

Now,

\[
(x_1, x_2) = 0 = (x_1, T x_1 + s_{11} x_1)
= (x_1, T x_1) + (x_1, s_{11} x_1)
= (x_1, T x_1) + s_{11}(x_1, x_1)
\]

so that

\[
s_{11} = -\frac{(x_1, T x_1)}{(x_1, x_1)} = -\frac{(T x_1, x_1)}{(x_1, x_1)}.
\]

Note: If \((x_1, x_1) = 1\), then the above process is equivalent to the Gram-Schmidt process.
Furthermore, construct the vector $T x_2$ and orthogonalize it with the vector $x_1$ and $x_2$. As a result, one gets the vector $x_3 = T x_2 + \varepsilon_{12} x_1 + \varepsilon_{22} x_2$ with $(x_1, x_2) = 0$ and $(x_2, x_3) = 0$. Hence

$$(x_1, x_3) = 0 = (x_1, T x_2 + \varepsilon_{12} x_1 + \varepsilon_{22} x_2)$$

$$= (x_1, T x_2) + (x_1, \varepsilon_{12} x_1) + (x_1, \varepsilon_{22} x_2)$$

$$= (x_1, T x_2) + \varepsilon_{12}(x_1, x_1) + \varepsilon_{22}(x_1, x_2)$$

$$(x_2, x_3) = 0 = (x_2, T x_2 + \varepsilon_{12} x_1 + \varepsilon_{22} x_2)$$

$$= (x_2, T x_2) + (x_2, \varepsilon_{12} x_1) + (x_2, \varepsilon_{22} x_2)$$

$$= (x_2, T x_2) + \varepsilon_{12}(x_2, x_1) + \varepsilon_{22}(x_2, x_2).$$

Consider the equations

$$(x_1, T x_2) + \varepsilon_{12}(x_1, x_1) + \varepsilon_{22}(x_1, x_2) = 0$$

$$(x_2, T x_2) + \varepsilon_{12}(x_2, x_1) + \varepsilon_{22}(x_2, x_2) = 0.$$
The process may be continued in a natural way by the formula

$$x_{i+1} = T x_i + s_{i1} x_1 + s_{i2} x_2 + \cdots + s_{in} x_n,$$

$i \geq 1$, and $s_{ki}$ can be determined in a similar manner as above and it will be

$$s_{ki} = - \frac{(T x_i, x_k)}{(x_k, x_k)} , \ k = 1, 2, \ldots, i.$$

The process continues until the null vector is obtained. The choice of the vector $x_1$ is arbitrary except that it is required that the null vector not be obtained until $x_{n+1}$. Thus one finds that

$$x_2 = T x_1 + s_{11} x_1,$$
$$x_3 = T x_2 + s_{12} x_1 + s_{22} x_2,$$
$$x_4 = T x_3 + s_{13} x_1 + s_{23} x_2 + s_{33} x_3,$$

$$\ddots$$

$$x_{n+1} = T x_n + s_{1n} x_1 + s_{2n} x_2 + \cdots + s_{nn} x_n = 0.$$

Consider (5-3) in the following form with $0_n$ as the null vector with $n$ components.

$$T x_1 + s_{11} x_1 - x_2 = 0_n,$$
$$T x_2 + s_{12} x_2 + s_{22} x_2 - x_3 = 0_n,$$
$$T x_3 + s_{13} x_1 + s_{23} x_2 + s_{33} x_3 - x_4 = 0_n,$$

$$\ddots$$

$$T x_{n-2} + s_{1,n-2} x_1 + \cdots + s_{n-2,n-2} x_{n-2} - x_{n-1} = 0_n.$$
\[
\begin{align*}
&\sum_{i=1}^{n-1} x_i^n + g_{1,n-1} x_1 + \cdots + g_{n-1,n-1} x_{n-1} - x_n = 0 \\
&\sum_{i=1}^{n} x_i^n + g_{1,n} x_1 + g_{2,n} x_2 + g_{3,n} x_3 + \cdots + g_{n,n} x_n = 0 \\
\end{align*}
\]

or
\[(3-4) \quad T x_i + g_{i1} x_1 + g_{i2} x_2 + g_{i3} x_3 + \cdots + g_{in} x_n = 0_n \]

where \( i = 1, 2, \ldots, n, \) and
\[
\begin{align*}
&g_{ki} = -\frac{(T x_i, x_k)}{(x_k, x_k)} , \quad k = 1, 2, \ldots, i; \quad k \leq n. \\
&(3-5) \quad g_{ki} = \begin{cases} 
-1 & k = i + 1; \ k \leq n. \\
0 & k = i + 2, i + 3, \ldots.
\end{cases}
\end{align*}
\]

Let \( X \) be the \( n \times n \) matrix formed by using \( x_i \) as its columns, \( i = 1, 2, \ldots, n \). Let \( G \) be the \( n \times n \) matrix with \( g_{ij} \), \( i, j = 1, 2, \ldots, n \), as its elements defined by (3-5). Now considering (3-4), one can write the matrix equation
\[TX + XG = 0.\]

Also, \( TX = -XG \). Since \( X \) is nonsingular,
\[(3-6) \quad T = -XGX^{-1}.\]

From (3-5) one sees that
\[
\begin{bmatrix}
g_{11} & g_{12} & g_{13} & \cdots & g_{1,n-1} & g_{1n} \\
-1 & g_{22} & g_{23} & \cdots & g_{2,n-1} & g_{2n} \\
0 & -1 & g_{33} & \cdots & g_{3,n-1} & g_{3n} \\
0 & 0 & -1 & \cdots & g_{4,n-1} & g_{4n} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & -1 & g_{nn}
\end{bmatrix}
\]

(3-7)
Consider the matrices

\[ (\mathbf{3}) = (\mathbf{s}_{11}), \quad - \begin{bmatrix} s_{11} & s_{12} \\ -1 & s_{22} \end{bmatrix}, \quad - \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ -1 & s_{22} & s_{23} \\ 0 & -1 & s_{33} \end{bmatrix}, \ldots, - \mathbf{G}. \]

If one defines \( \phi_i(\lambda) \) as the characteristic polynomial of the \( i \)th order matrix of \((\mathbf{3})\), then one can verify that

\[ \phi_1(\lambda) = \lambda + s_{11} \]
\[ \phi_2(\lambda) = (\lambda + s_{11})(\lambda + s_{22}) + s_{12} \]
\[ \phi_3(\lambda) = [(\lambda + s_{11})(\lambda + s_{22}) + s_{12}](\lambda + s_{33}) + s_{23}(\lambda + s_{11}) + s_{13} \]
\[ \phi_4(\lambda) = \left[ [(\lambda + s_{11})(\lambda + s_{22}) + s_{12}](\lambda + s_{33}) + s_{23}(\lambda + s_{11}) + s_{13} \right] \left[ (\lambda + s_{11})(\lambda + s_{22}) + s_{12} \right] + s_{34}[(\lambda + s_{11})(\lambda + s_{22}) + s_{12}] + s_{24}(\lambda + s_{11}) + s_{14} \]
\[ \vdots \]

and in general, by expanding the characteristic determinant of a matrix of order \( i \) by the last row, one can determine the recursion relation

\[ (\mathbf{3}-9) \quad \phi_i(\lambda) = (\lambda + s_{11})\phi_{i-1}(\lambda) + s_{i-1,1}\phi_{i-2}(\lambda) + \]
\[ + s_{i-2,1}\phi_{i-3}(\lambda) + \ldots + s_{11}\phi_{0}(\lambda) \]

where \( \phi_0(\lambda) = 1 \) and \( i = 1, 2, \ldots, n \). This is indicated by \( \phi_2(\lambda) \) above. One will note that \( \phi_n(\lambda) \) is the characteristic equation of \( T \) since \( T = -\mathbf{XGX}^{-1} \).
As soon as the \( g_{ij} \) have been computed, the characteristic polynomial can be written and from it the eigenvalues can be determined. Using the eigenvalue \( \lambda_i \), one can determine the corresponding eigenvector, since \( T = -XG^{-1} \) and the eigenvalue of \( G \) corresponding to \( -\lambda_i \) is \( y_i \) so that

\[
T = -XG^{-1}
\]

\[
Tx_i = -XG^{-1}y_i
\]

\[
= -Xy_i
\]

\[
= -x(-\lambda_i y_i)
\]

\[
= \lambda_i x_i.
\]

Now let the eigenvector of \( T \) corresponding to \( \lambda_i \) be \( z_i \), one sees that

\[
Tz_i = \lambda_i z_i
\]

so that

\[
z_i = Xy_i
\]

since \( Xy_i \) is the eigenvector of \( T \) corresponding to \( \lambda_i \). Hence, as soon as \( y_i \) has been determined, \( z_i \) is readily determined.

One must consider the following equation to determine \( y_i \).

\[
Gy_i = -\lambda_i y_i
\]
The above equation yields the following:

\[
\begin{align*}
\varepsilon_{11}(y^1_1) + \varepsilon_{12}(y^1_2) + \ldots + \varepsilon_{1n}(y^1_n) &= -\lambda_1(y^1_1) \\
(-1)(y^1_1) + \varepsilon_{22}(y^1_2) + \ldots + \varepsilon_{2n}(y^1_n) &= -\lambda_1(y^1_2) \\
\vdots \\
(-1)(y^1_{n-1}) + \varepsilon_{nn}(y^1_n) &= -\lambda_1(y^1_n).
\end{align*}
\]

(3-11)

One sees that starting with the last equation of (3-11) and substituting an arbitrarily chosen value for \((y^1_n, y^1_{n-1})\) can be determined. Using these values, one can compute \((y^1_j)\) for \(j = n-2, n-3, \ldots, 1\). The first equation of (3-11) is of help in determining \((y^1_j)\) since the last \(n-1\) equations contain all components of the eigenvectors \(y^1\). It can, however, be used as a means of control.

The method of orthogonalization of successive iterations is rather lengthy in this case. However, it is considerably simplified if \(T\) is symmetric, for \(G\) will be tridiagonal.

**Example 1:** Let \(T\) be the \(3 \times 3\) matrix

\[
T = \begin{bmatrix}
4 & -7 & 3 \\
1 & 2 & 5 \\
-1 & 2 & -1
\end{bmatrix}.
\]

Let \(x_1 = (0, 1, 0)\).

\[
Tx_1 = \begin{bmatrix}
4 & -7 & 3 \\
1 & 2 & 5 \\
-1 & 2 & -1
\end{bmatrix} \begin{bmatrix}
0 \\
1 \\
0
\end{bmatrix} = \begin{bmatrix}
-7 \\
2 \\
2
\end{bmatrix}.
\]
\[ S_{11} = - \frac{(Tx_1, x_1)}{(x_1, x_1)} = - \frac{2}{1} = -2. \]

\[ x_2 = Tx_1 + S_{11}x_1 = \begin{bmatrix} -7 \\ 2 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -7 \\ 0 \\ 2 \end{bmatrix}. \]

\[ T_{x_2} = \begin{bmatrix} 4 & -7 & 3 \\ 1 & 2 & 5 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} -7 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -22 \\ 3 \\ 5 \end{bmatrix}. \]

\[ S_{12} = - \frac{(Tx_2, x_1)}{(x_1, x_1)} = - \frac{3}{1} = -3. \]

\[ S_{22} = - \frac{(Tx_2, x_2)}{(x_2, x_2)} = - \frac{164}{53}. \]

\[ x_3 = Tx_2 + S_{12}x_1 + S_{22}x_2 = \begin{bmatrix} -22 \\ 3 \\ 5 \end{bmatrix} - 3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{164}{53} \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -18/53 \\ 0 \\ -63/53 \end{bmatrix}. \]

\[ T_{x_3} = \begin{bmatrix} 4 & -7 & 3 \\ 1 & 2 & 5 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} -18/53 \\ 0 \\ -63/53 \end{bmatrix} = \begin{bmatrix} -261/53 \\ -333/53 \\ 81/53 \end{bmatrix}. \]

\[ S_{13} = - \frac{(Tx_3, x_1)}{(x_1, x_1)} = - \frac{-353/53}{1} = \frac{333}{53}. \]

\[ S_{23} = - \frac{(Tx_3, x_2)}{(x_2, x_2)} = - \frac{(1989/53)}{53} = - \frac{1989}{53^2}. \]

\[ S_{33} = - \frac{(Tx_3, x_3)}{(x_3, x_3)} = - \frac{-405/53^2}{4293/53^2} = \frac{405}{4293} = \frac{5}{53}. \]
\[
\begin{align*}
\chi_1 &= x_3 + 333x_1 + 233x_2 + 53x_3 \\
&= \begin{bmatrix}
-261/53 & 333 & 0 & -18/53 \\
-333/53 & 1 & -1989/53 & 5 \\
81/53 & 0 & 2 & 0 \\
\end{bmatrix}
\begin{bmatrix}
x_3 \\
x_1 \\
x_2 \\
x_3 \\
\end{bmatrix}
\end{align*}
\]

Using equation (3-9), one finds that
\[
\phi_0(\lambda) = 1
\]
\[
\phi_1(\lambda) = (\lambda - 2)
\]
\[
\phi_2(\lambda) = (\lambda - 164/53)(\lambda - 2) - 3
\]
\[
\phi_3(\lambda) = (\lambda + 5/53) [ (\lambda - 164/53)(\lambda - 2) - 3 ] - (1989/53^2)(\lambda - 2) + 333/53 = 0.
\]
Now
\[
\phi_2(\lambda) = (1/53^2) [ (53\lambda + 5)(53\lambda^2 - 270\lambda + 169) - 1989\lambda + 21627 ] = 0
\]
\[
= (1/53^2) [ 53^2\lambda^3 - 14045\lambda^2 + 56187 + 22472 ] = 0
\]
\[
= \lambda^3 - 5\lambda^2 + 2\lambda + 8 = 0
\]
\[
= (\lambda - 4)(\lambda - 2)(\lambda + 1) = 0.
\]

Hence, \( \lambda_1 = 4, \lambda_2 = 2, \) and \( \lambda_3 = -1. \) From (3-7)
\[
C = \begin{bmatrix}
-2 & -3 & 333/53 \\
-1 & -164/53 & -1989/53^2 \\
0 & -1 & 5/53 \\
\end{bmatrix}.
\]
Let $y_1$ be the eigenvector of $G$ corresponding to $-\lambda_1$.

Using (3-11) with $(y_1)_3 = 1$ and $\lambda_1 = 4, \lambda_2 = 2, \lambda_3 = -1$, one finds:

$$-(y_1)_2 = -(5/53 + 4) = -217/53$$
$$-(y_1)_1 = -(154/53 + 4)(217/53) + 1989/53^2$$
$$= -8427/53^2 = 159/53 = 3.$$

Therefore, $y_1 = (3, 217/53, 1)$. Using the first equation of (3-11) as a means of checking,

$$2(y_1)_1 - 3(y_1)_2 + (333/53)(y_1)_3 = -4(y_1)_1$$
$$2(3) - 3(217/53) + 333/53 = 0.$$

Also

$$-(y_2)_2 = -(5/53 + 2) = -111/53$$
$$-(y_2)_1 = -(164/53 + 2)(111/53) + 1989/53^2$$
$$= 8427/53^2 = 3.$$

Hence, $y_2 = (-2, 111/53, 1)$. Checking,

$$(-2 + 2)(y_2)_1 - 3(y_2)_2 + (333/53)(y_2)_3 = -333/53 + 333/53 = 0.$$

Now

$$-(y_3)_2 = -(5/53 - 1) = 48/53$$
$$-(y_3)_1 = -(164/53 - 1)(-48/53) + 1989/53^2$$
$$= -8427/53^2 = -3.$$

Thus $y_3 = (3, -48/53, 1)$. Checking,

$$(-2 - 1)(3) - 3(-48/53) + (333/53)(1) = 0.$$

From equation (3-10),

$$z_1 = xy_1.$$
\[
\begin{bmatrix}
0 & -7 & -18/53 & 3 \\
1 & 0 & 0 & 217/53 \\
0 & 2 & -63/53 & 1
\end{bmatrix}
\begin{bmatrix}
-29 \\
3 \\
7
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & -7 & -18/53 & -3 \\
1 & 0 & 0 & 111/53 \\
0 & 2 & -63/53 & 1
\end{bmatrix}
\begin{bmatrix}
-15 \\
-3 \\
3
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & -7 & -18/53 & 3 \\
1 & 0 & 0 & -48/53 \\
0 & 2 & -63/53 & 1
\end{bmatrix}
\begin{bmatrix}
6 \\
3 \\
-3
\end{bmatrix}
\]

**Example 2:** Let \( T \) be the 4 \( \times \) 4 matrix

\[
T = \begin{bmatrix}
4 & 7 & 7 & 0 \\
7 & -2 & 0 & 0 \\
7 & 0 & 10 & 0 \\
0 & 0 & 0 & -2
\end{bmatrix}
\]

Let \( x_1 = (0, 1, -1, 1) \).

\[
x_1 = \begin{bmatrix}
4 & 7 & 7 & 0 \\
7 & -2 & 0 & 0 \\
7 & 0 & 10 & 0 \\
0 & 0 & 0 & -2
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
-1 \\
1
\end{bmatrix}
\begin{bmatrix}
0 \\
-2 \\
-10 \\
-2
\end{bmatrix}
\begin{bmatrix}
0 \\
1 \\
5 \\
1
\end{bmatrix}
\]

\[
-x_1 = \frac{\langle x_1, x_1 \rangle}{\langle x_1, x_1 \rangle} = -6/3 = -2
\]
\[ x_2 = -2 \begin{bmatrix} 0 \\ 1 \\ 5 \\ 1 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = -4 \begin{bmatrix} 0 \\ -4 \\ -8 \\ -4 \\ 1 \end{bmatrix} = -4 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \]

\[ T_x_2 = -4 \begin{bmatrix} 4 & 7 & 7 & 0 & 0 \\ 7 & -2 & 0 & 0 & 1 \\ 7 & 0 & 10 & 0 & 2 \\ 0 & 0 & 0 & -2 & 1 \end{bmatrix} = -4 \begin{bmatrix} 21 \\ -2 \\ 20 \\ -2 \end{bmatrix} \]

\[ \varepsilon_{12} = -96/3 = -32 \]

\[ \varepsilon_{22} = -36(16)/96 = -6 \]

\[ x_3 = -4 \begin{bmatrix} 21 \\ -2 \\ 20 \\ -2 \end{bmatrix} - 32 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix} + 24 \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} = -84 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \]

\[ T_x_3 = \begin{bmatrix} 4 & 7 & 7 & 0 \\ 7 & -2 & 0 & 0 \\ 7 & 0 & 10 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} -84 \\ 0 \\ -588 \\ 0 \end{bmatrix} = -336 \begin{bmatrix} -336 \\ -588 \\ 0 \end{bmatrix} \]

\[ \varepsilon_{13} = -0/3 = 0 \]

\[ \varepsilon_{23} = -7056/96 = -147/2 \]

\[ \varepsilon_{33} = -336/84 = -4 \]
\[
x_4 = \begin{bmatrix}
-336 \\
-588 \\
-588 \\
0
\end{bmatrix} + 0 + 294 \begin{bmatrix}
0 \\
1 \\
2 \\
1
\end{bmatrix} - 4 \begin{bmatrix}
0 \\
0 \\
0 \\
0
\end{bmatrix} = \begin{bmatrix}
-294 \\
0 \\
0 \\
294
\end{bmatrix}
\]

\[
T_x_4 = \begin{bmatrix}
4 & 7 & 7 & 0 \\
7 & -2 & 0 & 0 \\
7 & 0 & 10 & 0 \\
0 & 0 & 0 & -2
\end{bmatrix} \begin{bmatrix}
0 \\
0 \\
-294 \\
294
\end{bmatrix} = \begin{bmatrix}
-7(294) \\
2(294) \\
0 \\
0
\end{bmatrix}
\]

\[
\varepsilon_{14} = - \frac{0}{3} = 0
\]

\[
\varepsilon_{24} = - \frac{0}{96} = 0
\]

\[
\varepsilon_{34} = - \frac{-7(294)(-84)}{84^2} = - \frac{294}{12} = -\frac{49}{2}
\]

\[
\varepsilon_{44} = \frac{(2)(-294)(294) + (-2)(294)(294)}{2(294)(294)} = 2
\]

\[
x_5 = \begin{bmatrix}
-7(294) \\
2(294) \\
0 \\
-2(294)
\end{bmatrix} + 0 + 0 - (49/2) \begin{bmatrix}
0 \\
0 \\
0 \\
294
\end{bmatrix} + 2 \begin{bmatrix}
0 \\
0 \\
0 \\
294
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
0 \\
294
\end{bmatrix}
\]

\[
G = \begin{bmatrix}
-2 & -32 & 0 & 0 \\
-1 & -6 & -147/2 & 0 \\
0 & -1 & -4 & -49/2 \\
0 & 0 & -1 & 2
\end{bmatrix}.
\]

From equation (3-9),

\[
\phi_0(\lambda) = 1
\]

\[
\phi_1(\lambda) = \lambda - 2
\]

\[
\phi_2(\lambda) = (\lambda - 2)(\lambda - 6) - 32
\]
\( \Phi_3(\lambda) = [(\lambda - 2)(\lambda - 6) - 32](\lambda - 4) - (147/2)(\lambda - 2) + 0 \)
\( \Phi_4(\lambda) = [(\lambda - 2)(\lambda - 6) - 32](\lambda - 4) - (147/2)(\lambda - 2)^{1/2}(\lambda + 2) \)
- \( (49/2)[(\lambda - 2)(\lambda - 6) - 32] + 0 + 0. \)

Now
\( \Phi_4(\lambda) = \lambda^4 - 10\lambda^3 - 110\lambda^2 + 300\lambda + 944 = 0 \)
\( = (\lambda - 4)(\lambda + 2)(\lambda - 4 - \sqrt{134})(\lambda - 4 + \sqrt{134}) \)
\( = 0. \)

Hence, \( \lambda_1 = 4, \lambda_2 = -2, \lambda_3 = 4 + \sqrt{134}, \) and \( \lambda_4 = 4 - \sqrt{134}. \)

Let \( y_1 \) be the eigenvector of \( G \) corresponding to \( -\lambda_1. \)

Using (3-11) with \( (y_1)_4 = 1 \) and \( \lambda_1 = 4, \lambda_2 = -2, \lambda_3 = 4 + \sqrt{134}, \)
\( \lambda_4 = 4 - \sqrt{134}, \) one finds
\(- (y_1)_3 + 2(y_1)_4 = -4(y_1)_4 \)
\( (y_1)_3 = 6(1) = 6 \)
\(- (y_1)_2 - 4(y_1)_3 - (49/2)(y_1)_4 = -4(y_1)_3 \)
\( (y_1)_2 = -(49/2)(1) = -49/2 \)
\(- 6(y_1)_1 - 6(y_1)_2 - (147/2)(y_1)_3 + 0 = -4(y_1)_2 \)
\( (y_1)_1 = -2(-49/2) - (147/2)(6) \)
\( (y_1)_1 = -392. \)

Therefore, \( y_1 = (-392, -49/2, 6, 1). \) Using the first equation of (3-11) as a means of checking,
\(-2)(-392) - 32(-49/2) + 0 + 0 + 4(-392) = 0. \)

Similarly,
\( y_2 = (196, -49/2, 0, 1) \)
\[ y_3 = (144 + 24\sqrt{134}, 219/2 + 6\sqrt{134}, 6 + \sqrt{134}, 1) \]
\[ y_4 = (144 - 24\sqrt{134}, 219/2 - 6\sqrt{134}, 6 - \sqrt{134}, 1). \]

From equation (3-10),

\[
z_1 = \begin{bmatrix} 0 & 0 & -84 & 0 \\ 1 & -4 & 0 & -294 \\ -1 & -8 & 0 & 0 \\ 1 & -4 & 0 & 294 \end{bmatrix} \begin{bmatrix} -392 \\ -49/2 \\ 6 \\ 1 \end{bmatrix} = \begin{bmatrix} -504 \\ -588 \\ 588 \\ 0 \end{bmatrix}
\]

\[
z_2 = \begin{bmatrix} 0 & 0 & -84 & 0 \\ 1 & -4 & 0 & -294 \\ -1 & -8 & 0 & 0 \\ 1 & -4 & 0 & 294 \end{bmatrix} \begin{bmatrix} 196 \\ -49/2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 588 \end{bmatrix}
\]

\[
z_3 = \begin{bmatrix} 0 & 0 & -84 & 0 \\ 1 & -4 & 0 & -294 \\ -1 & -8 & 0 & 0 \\ 1 & -4 & 0 & 294 \end{bmatrix} \begin{bmatrix} 12(12 + 2\sqrt{134}) \\ 219/2 + 6\sqrt{134} \\ 6 + \sqrt{134} \\ 1 \end{bmatrix} = \begin{bmatrix} -84(6 + \sqrt{134}) \\ -588 \\ -12(85 + 6\sqrt{134}) \\ 0 \end{bmatrix}
\]

\[
z_4 = \begin{bmatrix} 0 & 0 & -84 & 0 \\ 1 & -4 & 0 & -294 \\ -1 & -8 & 0 & 0 \\ 1 & -4 & 0 & 294 \end{bmatrix} \begin{bmatrix} 12(12 - 2\sqrt{134}) \\ 219/2 - 6\sqrt{134} \\ 6 - \sqrt{134} \\ 1 \end{bmatrix} = \begin{bmatrix} -84(6 - \sqrt{134}) \\ -588 \\ -12(85 - 6\sqrt{134}) \\ 0 \end{bmatrix}
\]

Normalizing \( z_1, z_2, z_3, \) and \( z_4, \) one finds

\[ ||z_1|| = \sqrt{945504} = \sqrt{67(2)(84)^2} = 84\sqrt{134} \]
\[ ||z_2|| = 588 \]
\[ ||z_3|| = 24\sqrt{67(85 + 6\sqrt{134})} = 24(67 + 3\sqrt{134}) \]
\[ ||z_4|| = 24\sqrt{67(85 - 6\sqrt{134})} = 24(67 - 3\sqrt{134}) \]
\[ z_1 = \left( -\frac{3\sqrt{134}}{67}, -\frac{7\sqrt{134}}{134}, \frac{7\sqrt{134}}{134}, 0 \right) \]
\[ z_2 = (0, 0, 0, 1) \]
\[ z_3 = \left( -\frac{7\sqrt{134}}{134}, -\frac{(67 - 3\sqrt{134})}{134}, -\frac{(67 + 3\sqrt{134})}{134}, 0 \right) \]
\[ z_4 = \left( -\frac{7\sqrt{134}}{134}, -\frac{(67 + 3\sqrt{134})}{134}, -\frac{(67 - 3\sqrt{134})}{134}, 0 \right). \]
CHAPTER BIBLIOGRAPHY

CHAPTER IV

TRANSFORMATION OF SYMMETRIC MATRICES TO TRIDIAGONAL FORM BY MEANS OF ROTATION

Many methods have been developed to compute the eigenvalues and eigenvectors of a symmetric matrix. In this chapter the symmetric matrix shall be tridiagonalized by a series of rotations on the matrix and from this tridiagonal matrix the eigenvalues and eigenvectors shall be determined.

A rotation means a transformation of coordinates with the elementary matrix of rotation

\[
R(i,j) = \begin{bmatrix}
1 & 0 & \cdots & \cdots & 0 \\
0 & 1 & \cdots & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & 0 \\
0 & \cdots & \cdots & 0 & 1 \\
\end{bmatrix}
\]

for \( c^2 + s^2 = 1 \) and \( j > i > 1 \).
The matrix \( R(i,j) \) may also be characterized as follows:

\[
(R(i,j))_{kl} = \begin{cases} 
1, & \text{if } k = 1 \neq i \text{ or } k = 1 \neq j \\
ll, & \text{if } k = 1 = i \text{ or } k = 1 = j \\
0, & \text{if } k \neq 1 \neq i \text{ or } k \neq 1 \neq j \\
s, & \text{if } k = j \text{ and } l = i \\
s, & \text{if } k = i \text{ and } l = j 
\end{cases}
\]

where \((R(i,j))_{kl}\) is the element of \( R(i,j) \) appearing in the \( k \)th row and the \( l \)th column.

A rotation may be interpreted geometrically as a change in the basis vectors \( e_i \) and \( e_j \) by a certain angle, carried out in the plane spanned by the vectors \( e_i \) and \( e_j \) (1, p. 280). Since the columns of \( R(i,j) \) are mutually orthogonal normal vectors, the matrix \( R(i,j) \) is orthogonal.

Let \( T = (t_{ij}) \) be a real symmetric matrix. Let \( A(i,j) = TR(i,j) \) and \( B(i,j) = R(i,j)^T A(i,j) \); then one can verify that

\[
A(i,j) = \begin{bmatrix}
t_{11} & \cdots & ct_{1i} + st_{1j} & t_{1,i+1} & \cdots & -st_{1i} + ct_{1j} & t_{1,j+1} & \cdots & t_{1n} \\
t_{21} & \cdots & ct_{2i} + st_{2j} & t_{2,i+1} & \cdots & -st_{2i} + ct_{2j} & t_{2,j+1} & \cdots & t_{2n} \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
t_{j1} & \cdots & ct_{ji} + st_{jj} & t_{j,i+1} & \cdots & -st_{ji} + ct_{jj} & t_{j,j+1} & \cdots & t_{jn} \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
t_{ni} & \cdots & ct_{ni} + st_{nj} & t_{n,i+1} & \cdots & -st_{ni} + ct_{nj} & t_{n,j+1} & \cdots & t_{nn}
\end{bmatrix}
\]
One sees immediately that the elements \( a_{kl} \) of \( A(i,j) \) are the same as \( t_{kl} \) with the exception of column \( i \) and \( j \) where

\[
a_{ki} = ct_{ki} + st_{kj}
\]

\[
a_{kj} = -st_{ki} + ct_{kj}
\]

and \( k = 1, 2, \ldots, n \).

Since \( B(i,j) = R(i,j)^T A(i,j) \), the elements \( b_{kl} \) of \( B(i,j) \) are the same as \( a_{kl} \) with the exception of the rows \( i \) and \( j \) where

\[
b_{il} = ca_{il} + sa_{jl}
\]

\[
b_{jl} = -sa_{il} + ca_{jl}
\]

and \( l = 1, 2, \ldots, n \). Now, \( A(i,j) = TR(i,j) \) so that

\[
B(i,j) = R(i,j)^T TR(i,j).
\]

Since \( T \) is symmetric, \( B(i,j) \) is symmetric for

\[
B(i,j)^T = (R(i,j)^T TR(i,j))^T
\]

\[
= (TR(i,j))^T (R(i,j)^T)^T
\]

\[
= R(i,j)^T TR(i,j)
\]

\[
= B(i,j).
\]

Therefore, \( b_{kl} = b_{lk} \). Hence as soon as matrix \( A(i,j) \) has been computed, matrix \( B(i,j) \) can be computed by finding only the elements \( b_{ii}, b_{ij}, b_{ji} \), and \( b_{jj}, (b_{ij} = b_{ji}) \), for the remaining elements of row \( i \) and row \( j \) of \( B(i,j) \) correspond to the elements of column \( i \) and column \( j \) of matrix \( A(i,j) \) and all other elements of \( B(i,j) \) are the same as the corresponding elements of \( A(i,j) \).
Since the ultimate goal is to rotate T to a tridiagonal matrix, the element $b_{i-1,j} = 0 = b_{j,i-1}$ must be true for $i = 2, \ldots, n-1$; $j = i+1, \ldots, n$. This implies that the rotation matrix must be used with $i = 2, \ldots, n-1$; $j = i+1, \ldots, n$, in which case the matrix $T$ will be replaced by $B(i,j)$ when $i \geq 2$ and $j \geq 3$ where $B(i,j)$ is the matrix obtained in (4-4) from the preceding step, i.e.,

$$T$$

$$B(2,3) = R(2,3)^T R(2,3)$$

$$B(2,4) = R(2,4)^T B(2,3) R(2,4)$$

$$\vdots$$

$$B(2,n) = R(2,n)^T B(2,n-1) R(2,n)$$

(4-6)

$$B(3,4) = R(3,4)^T B(2,3) R(3,4)$$

$$B(3,5) = R(3,5)^T B(3,4) R(3,5)$$

$$\vdots$$

$$B(3,n) = R(3,n)^T B(3,n-1) R(3,n)$$

$$\vdots$$

$$B(n-1,n) = R(n-1,n)^T B(n-1,n-1) R(n-1,n)$$

so that

$$B(n-1,n) = R(n-1,n)^T \cdots R(2,4)^T R(2,3)^T R(2,3) R(2,4) \cdots R(n-1,n)$$

One now needs to determine $c$ and $s$ so that $b_{i-1,j} = 0$.

Now

$$b_{i-1,j} = a_{i-1,j} = -st_{i-1,i} + ct_{i-1,j} = 0$$

$$-st_{i-1,i} = -ct_{i-1,j}$$
\[ s = c \frac{t_{i-1,i}}{t_{i-1,i}} \]
\[ c^2 + s^2 = 1 \]
\[ c^2 + c^2 \left[ \frac{t_{i-1,i}}{t_{i-1,i}} \right]^2 = 1 \]
\[ c^2 = \frac{1}{1 + \left[ \frac{t_{i-1,i}}{t_{i-1,i}} \right]^2} = \frac{t_{i-1,i}^2}{t_{i-1,i}^2 + t_{i-1,j}^2} \]
\[ c = \pm \sqrt{\frac{t_{i-1,i}^2}{t_{i-1,i}^2 + t_{i-1,j}^2}} \]
\[ s = \pm \frac{t_{i-1,i}}{t_{i-1,i}^2 + t_{i-1,j}^2} \cdot \frac{t_{i-1,i}}{t_{i-1,i}^2 + t_{i-1,j}^2} \]

Remembering that for each rotation \( R(i,j) \) the element \( b_{i-1,j} \) is zero for \( i = 2, 3, \ldots, n-1; j = i+1, \ldots, n \), then the elements of the first row beginning with the third element are annihilated by \( R(2,3), \ldots, R(2,n) \); the elements of the second row beginning with the fourth element are annihilated by \( R(3,4), \ldots, R(3,n) \); etc. It is clear that if an element is annihilated by a rotation, it will remain zero throughout the entire process. Since \( B(i,j) \) is symmetric, a tridiagonal matrix is obtained by the \( R(n-1,n) \) rotation.
Hence

\[
B(n-1,n) = S = \begin{bmatrix}
  b_{11} & b_{12} & 0 & \ldots & 0 & 0 \\
  b_{21} & b_{22} & b_{23} & \ldots & 0 & 0 \\
  0 & b_{32} & b_{33} & \ldots & 0 & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
  0 & 0 & 0 & b_{n-1,n-1} & b_{n-1,n} \\
  0 & 0 & 0 & b_{n,n-1} & b_{n,n}
\end{bmatrix}
\]

As in Chapter III, one needs to consider the matrices

\[
(4-10) \quad (s_{11}) = \begin{bmatrix}
s_{11} & s_{12} \\
s_{21} & s_{22}
\end{bmatrix}, \quad (s_{11} + s_{12} \lambda) = \begin{bmatrix}
s_{11} & s_{12} & 0 \\
0 & s_{21} & s_{22} \\
0 & 0 & s_{32} & s_{33}
\end{bmatrix}, \ldots, S,
\]

where \( s_{ij} = b_{ij} \) and define \( \phi_1(\lambda) \) as the characteristic polynomial so that

\[
\phi_1(\lambda) = \lambda - s_{11}
\]

\[
\phi_2(\lambda) = (\lambda - s_{11})(\lambda - s_{22}) - s_{12}s_{21}
\]

\[
= (\lambda - s_{11})(\lambda - s_{22}) - s_{12}^2
\]

\[
(4-11) \quad \phi_3(\lambda) = [(\lambda - s_{11})(\lambda - s_{22}) - s_{12}^2](\lambda - s_{33}) - s_{23}^2(\lambda - s_{11})
\]

\[
\vdots
\]

\[
\phi_i(\lambda) = (\lambda - s_{11})\phi_{i-1}(\lambda) - s_{i-1,i}^2\phi_{i-2}(\lambda)
\]

where \( \phi_0(\lambda) = 1 \), and where \( \phi_n(\lambda) = (-1)^n \phi(\lambda) \), the characteristic polynomial of \( S \). Now one needs to determine the latent roots of \( S \) from the polynomial \( \phi_n(\lambda) \).
The eigenvectors for the matrix $S$ can be computed just as in Chapter III by solving the corresponding triangular system

$$(s_{11} - \lambda_1)(y_1)_1 + s_{12}(y_1)_2 = 0$$
$$s_{12}(y_1)_1 + (s_{22} - \lambda_1)(y_1)_2 + s_{23}(y_1)_3 = 0$$
$$(4-12) \quad s_{23}(y_1)_2 + (s_{33} - \lambda_1)(y_1)_3 + s_{34}(y_1)_4 = 0$$
$$\vdots$$
$$s_{n-1,n}(y_1)_{n-1} + (s_{nn} - \lambda_1)(y_1)_n = 0$$

for the components $(y_1)_1$, $(y_1)_2$, ..., $(y_1)_n$ of the eigenvector $y_1$ of $S$ corresponding to $\lambda_1$. It is convenient here to choose the first component rather than the last, as in Chapter III, and then to compute the second, third, etc.

To determine the eigenvectors of $T$, one must consider $S$ in the form

$$S = (R(2,3)\ldots R(n-1,n))^T T(R(2,3)\ldots R(n-1,n)).$$

Now

$$Sy_1 = \lambda_1 y_1$$
$$(R(2,3)\ldots R(n-1,n))^T T(R(2,3)\ldots R(n-1,n))y_1 = \lambda_1 y_1$$
$$T(R(2,3)\ldots R(n-1,n))y_1 = \lambda_1 (R(2,3)\ldots R(n-1,n))y_1$$

and let $z_1 = (R(2,3)\ldots R(n-1,n))y_1$, one finds that

$$Tz_1 = \lambda_1 z_1$$

so that $z_1$ is the eigenvector of $T$ associated with $\lambda_1$. Hence as soon as $y_1$ is determined, $z_1$ can be determined by a series of multiplications of the rotation matrices $R(j,k)$. For each
separate multiplication, only two components of the preceding vector will be changed—the jth and kth. This can be formulized as follows:

\[(z'_i)_j = c(z_i)_j - s(z_i)_k\]
\[(z'_i)_k = s(z_i)_j + c(z_i)_k\]

where \((z'_i)_j\) and \((z'_i)_k\) are the components obtained after a multiplication \(R(j,k)\) and where \((z_i)_j\) and \((z_i)_k\) are the components of the preceding vector and \(c\) and \(s\) are the values used in the rotation matrix \(R(j,k)\).

**Example 1:** Find the eigenvalues and eigenvectors of \(T\) where

\[
T = \begin{bmatrix}
4 & 7 & 7 & 0 \\
7 & -2 & 0 & 0 \\
7 & 0 & 10 & 0 \\
0 & 0 & 0 & -2
\end{bmatrix}.
\]

From (4-7)

\[
c = \frac{t_{i-1,i}^2}{t_{i-1,i}^2 + t_{i-1,j}^2}.
\]

Letting \(i = 2\) and \(j = 3\),

\[
c = \sqrt{49/(49 + 49)} = \sqrt{1/2}.
\]

From (4-8)

\[
s = \frac{t_{i-1,i}^2}{t_{i-1,i}^2 + t_{i-1,j}^2} \times \frac{t_{i-1,i}^2}{t_{i-1,i}^2 + t_{i-1,j}^2}
\]

and with \(i = 2, j = 3, s = \sqrt{1/2}.
\]
Now from (4-1),

\[ R(2,3) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1/2 & -1/2 & 0 \\
0 & 1/2 & 1/2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}. \]

Using (4-3),

\[
\begin{align*}
a_{12} &= ct_{12} + st_{13} = \frac{\sqrt{1/2}(7) + \sqrt{1/2}(7)}{14} = 14 \frac{\sqrt{1/2}}{2} \\
a_{22} &= ct_{22} + st_{23} = \frac{\sqrt{1/2}(-2) + \sqrt{1/2}(0)}{2} = -2 \frac{\sqrt{1/2}}{2} \\
a_{32} &= ct_{32} + st_{33} = \frac{\sqrt{1/2}(0) + \sqrt{1/2}(10)}{10} = 10 \frac{\sqrt{1/2}}{2} \\
a_{42} &= ct_{42} + st_{43} = \frac{\sqrt{1/2}(0) + \sqrt{1/2}(0)}{0} = 0 \\
a_{13} &= -st_{12} + ct_{13}
\end{align*}
\]

so that

\[
\begin{align*}
a_{13} &= -\frac{\sqrt{1/2}(7) + \sqrt{1/2}(7)}{14} = 0 \\
a_{23} &= -\frac{\sqrt{1/2}(-2) + \sqrt{1/2}(0)}{2} = 2 \frac{\sqrt{1/2}}{2} \\
a_{33} &= -\frac{\sqrt{1/2}(0) + \sqrt{1/2}(10)}{10} = 10 \frac{\sqrt{1/2}}{2} \\
a_{43} &= -\frac{\sqrt{1/2}(0) + \sqrt{1/2}(0)}{0} = 0.
\end{align*}
\]

Hence,

\[ A(2,3) = \begin{bmatrix}
4 & 14 & 1/2 & 0 & 0 \\
7 & -2 & 1/2 & 2 & 1/2 & 0 \\
7 & 10 & 1/2 & 10 & 1/2 & 0 \\
0 & 0 & 0 & 0 & -2
\end{bmatrix}. \]

From (4-4),

\[ b_{21} = ca_{21} + sa_{31} \]

so that
\[ b_{21} = \sqrt{1/2}(7) + \sqrt{1/2}(7) = 14\sqrt{1/2} \]
\[ b_{22} = \sqrt{1/2}(-2\sqrt{1/2}) + \sqrt{1/2}(10\sqrt{1/2}) = 4 \]
\[ b_{23} = \sqrt{1/2}(2\sqrt{1/2}) + \sqrt{1/2}(10\sqrt{1/2}) = 6 \]
\[ b_{24} = \sqrt{1/2}(0) + \sqrt{1/2}(0) = 0 \]

and
\[ b_{31} = -sa_{21} + ca_{31} \]

so that
\[ b_{31} = -\sqrt{1/2}(7) + \sqrt{1/2}(7) = 0 \]
\[ b_{32} = -\sqrt{1/2}(-2\sqrt{1/2}) + \sqrt{1/2}(10\sqrt{1/2}) = 6 \]
\[ b_{33} = -\sqrt{1/2}(2\sqrt{1/2}) + \sqrt{1/2}(10\sqrt{1/2}) = 4 \]
\[ b_{34} = -\sqrt{1/2}(0) + \sqrt{1/2}(0) = 0. \]

Therefore,
\[
B(2,3) = \begin{bmatrix}
4 & 14\sqrt{1/2} & 0 & 0 \\
14\sqrt{1/2} & 4 & 6 & 0 \\
0 & 6 & 4 & 0 \\
0 & 0 & 0 & -2
\end{bmatrix}.
\]

Since \( B(2,3) \) is a tridiagonal symmetric matrix, the rotations \( R(2,4) \) and \( R(3,4) \) are unnecessary so that \( B(2,3) = S \). One now proceeds to determine \( \phi_n(\lambda) \) from (4-11).

\[ \phi_1(\lambda) = \lambda - 4 \]
\[ \phi_2(\lambda) = (\lambda - 4)(\lambda - 4) - 98 \]
\[ \phi_3(\lambda) = [(\lambda - 4)(\lambda - 4) - 98](\lambda - 4) - 36(\lambda - 4) \]
Thus
\[ \phi_4(\lambda) = 0 = \lambda^4 - 10\lambda^3 - 126\lambda^2 + 300\lambda + 944 \]
\[(\lambda - 4)(\lambda + 2)(\lambda - 4 - \sqrt{134})(\lambda - 4 + \sqrt{134}) = 0 \]
so that \( \lambda_1 = 4, \lambda_2 = -2, \lambda_3 = 4 + \sqrt{134}, \) and \( \lambda_4 = 4 - \sqrt{134}. \)

Using (4-12), one can determine the eigenvectors corresponding to \( \lambda_1. \) Letting \((y_1)_1 = -1,\) then
\[ (4 - 4)(-1) + 14\sqrt{1/2}(y_1)_2 = 0 \]
\[ (y_1)_2 = 0 \]
\[ 14\sqrt{1/2}(-1) + (4 - 4)(0) + 6(y_1)_3 = 0 \]
\[ (y_1)_3 = (7/6)\sqrt{2} \]
\[ 0(-(7/6)\sqrt{2}) + (-2 - 4)(y_1)_4 = 0 \]
\[ (y_1)_4 = 0. \]
Letting \((y_2)_1 = 0,\) then
\[ (4 + 2)(0) + 14\sqrt{1/2}(y_2)_2 = 0 \]
\[ (y_2)_2 = 0 \]
\[ (14\sqrt{1/2})(0) + (4 + 2)(0) + 6(y_2)_3 = 0 \]
\[ (y_2)_3 = 0 \]
\[ (0)(0) + (-2 + 2)(y_2)_4 = 0 \]
\[ (y_2)_4 \text{ is arbitrary, say 1.} \]
Letting \((y_3)_1 = -1\), then
\[
(4 - 4 - \sqrt{134})(-1) + 14\sqrt{1/2}(y_3)_2 = 0
\]
\[
(y_3)_2 = -\sqrt{134}/14\sqrt{1/2} = -\sqrt{67}/7
\]
\[
14\sqrt{1/2}(-1) + (4 - 4 - \sqrt{134})(-\sqrt{67}/7) + 6(y_3)_3 = 0
\]
\[
(y_3)_3 = -3\sqrt{2}/7
\]
\[
(0)(y_3)_3 + (-2 - 4 - \sqrt{134})(y_3)_4 = 0
\]
\[
(y_3)_4 = 0.
\]
Letting \((y_4)_1 = 1\), then
\[
(4 - 4 + \sqrt{134})(1) + 14\sqrt{1/2}(y_4)_2 = 0
\]
\[
(y_4)_2 = \sqrt{67}/7
\]
\[
14\sqrt{1/2}(1) + (4 - 4 + \sqrt{134})(-\sqrt{67}/7) + 6(y_4)_3 = 0
\]
\[
(y_4)_3 = 3\sqrt{2}/7
\]
\[
(0)(3\sqrt{2}/7) + (-2 - 4 + \sqrt{134})(y_4)_4 = 0
\]
\[
(y_4)_4 = 0.
\]
Now
\[
z_1 = R(2,3)y_1 = \frac{1}{5}
\]
\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \sqrt{1/2} & -\sqrt{1/2} & 0 \\
0 & \sqrt{1/2} & \sqrt{1/2} & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
-1 \\
0 \\
7\sqrt{2}/6 \\
0
\end{bmatrix}
= \frac{1}{5}
\begin{bmatrix}
-1 \\
7/6 \\
7/6 \\
0
\end{bmatrix}
\]

Similarly, \(z_1 = R(2,3)y_1\) so that \(z_2 = (0, 0, 0, 1)\), \(z_3 = (-1, (6 - \sqrt{134})/14, -(6 + \sqrt{134})/14, 0)\) and \(z_4 = (1, -(6 + \sqrt{134})/14, (6 - \sqrt{134})/14, 0)\).
If one normalizes the vectors \( z_1, z_2, z_3, \) and \( z_4, \) then

\[
\|z_1\| = \sqrt{1 + 49/36 + 49/36} = \sqrt{134/36}
\]

\[
z_1 = \left( -3\sqrt{134/67}, -7\sqrt{134/134}, 7\sqrt{134/134}, 0 \right)
\]

\[
\|z_2\| = 1
\]

\[
z_2 = \left( 0, 0, 0, 1 \right)
\]

\[
\|z_3\| = \sqrt{134/7}
\]

\[
z_3 = \left[ \frac{7\sqrt{134}}{134}, \frac{67 - 3\sqrt{134}}{134}, \frac{67 + 3\sqrt{134}}{134}, 0 \right]
\]

\[
\|z_4\| = \sqrt{134/7}
\]

\[
z_4 = \left[ \frac{7\sqrt{134}}{134}, \frac{67 + 3\sqrt{134}}{134}, \frac{67 - 3\sqrt{134}}{134}, 0 \right]
\]
CHAPTER BIBLIOGRAPHY

BIBLIOGRAPHY

Books


Articles