ELEMENTARY PROOF OF THE WIELANDT-HOFFMAN THEOREM
AND OF ITS GENERALIZATION

BY

J. H. WILKINSON

TECHNICAL REPORT NO. CS 150

JANUARY 1970

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

Reproduced by the CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va 22151
Abstract: An elementary proof is given of the Wielandt-Hoffman Theorem for normal matrices and of a generalization of this theorem. The proof makes no direct appeal to results from linear-programming theory.
1. Introduction

In [2] Wielandt and Hoffman proved a theorem on the eigenvalues of normal matrices which is of considerable importance in the error analysis of eigenvalue algorithms based on the use of unitary transformations [4,5]. Their proof was very elegant and was based on the use of linear programming techniques. In [5] Wilkinson gave an elementary proof in the case when the matrices are Hermitian, which was based on an earlier proof due to Givens [1]. This proof did not extend easily to the general case. Here we give an elementary proof for the general case which applies immediately to a generalization of the Wielandt-Hoffman theorem due to Kahan [3]. Not surprisingly the proof involves techniques which are familiar in the area of linear programming but no direct appeal is made to results from that field.

2. The Basic Theorem

The proof depends on a theorem which is not directly concerned with normal matrices. Before stating this theorem we give two definitions.

DEFINITION 1. The set of n elements \( a_{i_1,i_1}, a_{i_2,i_2}, \ldots, a_{i_n,i_n} \) of an nxn matrix \( A \) is called a diagonal of \( A \) if \( i_1, i_2, \ldots, i_n \) is a permutation of the integers 1,2,\ldots,n. If \( i_j = j \) (\( j = 1, \ldots, n \)) then we have the principal diagonal.

DEFINITION 2. A matrix \( X \) is called a doubly stochastic matrix if \( x_{ij} \geq 0 \) and \( \sum_{i=1}^{n} x_{ij} = \sum_{j=1}^{n} x_{ij} = 1 \) (\( j = 1, \ldots, n \)) i.e., all row and column runs are unity.
THEOREM 1. If $P$ is a real matrix such that the sum of the elements on the principal diagonal is not greater than the sum of the elements on any other diagonal, and $X$ is any doubly stochastic matrix, then $S(X) = \sum \sum p_{ij} x_{ij}$ is a minimum when $X = I$.

PROOF. The minimum is attained, possibly for many different $X$. Let us choose $X$ to be a minimizing doubly stochastic matrix having the maximum number of zero off-diagonal elements. We shall show that all its off-diagonals must be zero. For suppose that this is not true. Let $x_{i_1, i_2}$ be a non-zero off-diagonal. Then $x_{i_2, i_2} < 1$ and hence there is a non-zero element $x_{i_2, i_3}$ (say) in row $i_2$. If $i_3 \neq i_2$ then similarly there is a non-zero element $x_{i_k, i_k}$ in row $i_k$. Continue in this way until we reach an $x_{i_{m-1}, i_m}$ for which $i_m$ equals some earlier $i_k$. Let $x$ be the smallest of the positive elements $x_{i_{k+1}, i_{k+1}}, x_{i_{k+2}, i_{k+2}}, \ldots, x_{i_{m-1}, i_m}$.

Construct a matrix $Y$ such that

$$y_{i_s, i_s} = x_{i_s, i_s} + x, \quad s = k, k+1, \ldots, m-1 \quad (2.1)$$

$$y_{i_s, i_{s+1}} = x_{i_s, i_{s+1}} - x, \quad s = k, k+1, \ldots, m-1 \quad (2.2)$$

$$y_{ij} = x_{ij} \quad \text{otherwise.} \quad (2.3)$$

Then $Y$ is clearly a doubly stochastic matrix and
The expression in brackets cannot be positive since otherwise by replacing the elements \( p_{i_s,i_s} \) in the principal diagonal by the elements \( p_{i_s,i_{s+1}} \) we could obtain a smaller diagonal sum. Hence

\[
\sum_{i,j} p_{ij}y_{ij} - \sum_{i,j} p_{ij}x_{ij} = X \left[ \sum_{s=k}^{m-1} p_{i_s,i_s} - \sum_{s=k}^{m-1} p_{i_s,i_{s+1}} \right].
\] (2.4)

But \( Y \) is clearly a doubly stochastic matrix and it has at least one more off-diagonal zero than \( X \), contradicting the hypothesis. Hence all off-diagonal elements of \( X \) must be zero, i.e., \( X = I \).

An exactly analogous theorem holds when the principal diagonal has the maximum sum.

3. The Wielandt-Hoffman Theorem

**Theorem 2.** If \( A \) and \( B \) are normal matrices and \( C = A - B \), and if \( a_1 \) and \( b_1 \) are the eigenvalues of \( A \) and \( B \) arranged so that \( \sum_{i=1}^{n} |a_i - b_i|^2 \) is a minimum for all possible orderings, then

\[
\sum_{i=1}^{n} |a_i - b_i|^2 \leq \|C\|^2_F. \quad (\|C\|^2_F = \text{the Frobenius norm of } C) \quad (3.1)
\]

**Proof.** Since \( A \) and \( B \) are normal there exist unitary \( Q_1 \) and \( Q_2 \) such that

\[
A = Q_1 \text{diag}(a_1)Q_1^H, \quad B = Q_2 \text{diag}(b_1)Q_2^H. \quad (3.2)
\]
(Note then we are free to prescribe the ordering of the $a_i$ and $b_i$ and we choose the ordering which gives $\sum |a_i - b_i|^2$ a minimum value. Hence

$$ A - B = Q_1 \text{diag}(a_i)q_1^H - Q_2 \text{diag}(b_i)q_2^H = C $$

(3.3)

giving

$$ \text{diag}(a_i)q_1^H q_2 - q_2^H q_2 \text{diag}(b_i) = q_1^H C Q_2 $$

(3.4)

Writing $Q = Q_1^H Q_2$, a unitary matrix, we have

$$ \|\text{diag}(a_i)Q - Q \text{diag}(b_i)\|_F^2 = \|C\|_F^2 $$

(3.5)

since the Frobenius norm is unitarily invariant. Hence

$$ \sum \sum |a_i - b_j|^2 |q_{ij}|^2 = \|C\|_F^2 $$

(3.6)

Now the matrix $P$ with $p_{ij} = |a_i - b_j|^2$ is real and from the ordering of the $a_i$ and $b_i$ its principal diagonal is minimal. Further, since $Q$ is unitary, the matrix $Z$ with $z_{ij} = |q_{ij}|^2$ is a doubly stochastic matrix. Hence by Theorem 1 and equation (3.6)

$$ \sum_{i=1}^{n} |a_i - b_i|^2 \leq \sum \sum |a_i - b_j|^2 |q_{ij}|^2 = \|C\|_F^2 $$

(3.7)

and the result is proved.

When $A$ and $B$ are Hermitian, the $a_i$ and $b_i$ are real, and it is easy to prove that the orderings $a_1 \geq a_2 \geq \ldots \geq a_n$, $b_1 \geq b_2 \geq \ldots \geq b_n$ give the minimal value. In fact, returning to Theorem 1 in the case when $p_{ij} = (a_i - b_j)^2$ with $a_i$ and $b_i$ real and monotonically ordered, the proof is much simpler. For if $X$ has a non-zero off diagonal element
in row 1 or column 1 it must have at least one such in both. Suppose $x_{1r}$ and $x_{sr}$ are non-zero and $x$ is the smaller. If we increase $x_{11}$ and $x_{sr}$ by $x$ and diminish $x_{1r}$ and $x_{sl}$ by $x$ the sum is changed by

$$x[(a_1-b_1)^2+(a_s-b_r)^2-(a_1-b_r)^2-(a_s-b_1)^2] = x(a_1-a_s)(b_r-b_1) < 0 \quad (4.3)$$

Hence continuing in this way the minimizing $X$ has no non-zero off-diagonal elements in row 1 or column 1, and continuing again the minimizing $X$ is $I$. (Notice we do not even have to show that for this $P$, the principal diagonal is minimal; this emerges from the proof.)

4. Generalization of the Wielandt-Hoffman Theorem

A generalization of the Wielandt-Hoffman Theorem which is of practical importance is the following.

**Theorem 3.** If $X$ is an $n \times r$ matrix with orthonormal columns, $A$ is an $n \times n$ normal matrix, $B$ is an $r \times r$ normal matrix and $R$ an $n \times r$ matrix is defined by

$$AX - XB = R \quad (4.1)$$

if the eigenvalues $a_i$ ($i = 1, \ldots, n$) of $A$ and $b_i$ ($i = 1, \ldots, r$) of $B$ are ordered so that $\sum_{i=1}^{r} |a_i - b_i|^2$ is a minimum, then

$$\sum_{i=1}^{r} |a_i - b_i|^2 \leq \|R\|_F^2 \quad (4.2)$$

A weaker result with $\|R\|_F^2$ replaced by $2^{1/2} \|R\|_F^2$ was given by Wilkinson in [5] and the result itself by Kahan [3].
Notice we are interested only in the selection and ordering of the relevant $r$ of the $a_i$ to be associated with the $b_j$. Writing

$$A = Q_1 \text{diag}(a_1)q_1^H, \quad B = Q_2 \text{diag}(b_1)q_2^H$$

(4.3)

with the prescribed ordering of the $a_i$ and $b_j$, we have

$$\|\text{diag}(a_1)Q - Q \text{diag}(b_1)\|_F^2 = \|Q_1^H R Q_2\|_F^2 = \|R\|_F^2$$

(4.4)

where $Q$ is an $n \times r$ matrix with orthonormal columns. Hence

$$\sum_{i=1}^r \sum_{j=1}^n |a_{ij} - b_{ij}|^2 |q_{ij}|^2 = \|R\|_F^2$$

(4.5)

Let $Y = [Q \mid Z]$ be an $n \times n$ unitary matrix given by the completion of $Q$; then if

$$P_{ij} = |a_{ij} - b_{ij}|^2 \quad (j \leq r), \quad P_{ij} = 0 \quad (j > r)$$

(4.6)

$$\sum_{i=1}^n \sum_{j=1}^n P_{ij} |y_{ij}|^2 = \sum_{j=1}^r \sum_{i=1}^n |a_{ij} - b_{ij}|^2 |q_{ij}|^2$$

(4.7)

and from the definition of the ordering of the $a_i$ and $b_j$, the diagonal of $P$ is minimal. Hence by Theorem 1 and Equation (4.5)

$$\sum_{i=1}^n P_{ii} = \sum_{i=1}^r |a_{ii} - b_{ii}|^2 \leq \sum_{j=1}^r \sum_{i=1}^n |a_{ij} - b_{ij}| |q_{ij}|^2 = \|R\|_F^2$$

(4.8)

This theorem is of practical value when $r$ orthonormal approximate eigenvectors $x_1, \ldots, x_r$ are known corresponding to alleged eigenvalues $\mu_1, \ldots, \mu_r$. If
\[ Ax_i - \mu_i x_i = r_i \quad (i = 1, \ldots, r) \quad (4.9) \]

Then

\[ AX - X \text{diag}(\mu_1) = R \quad (4.10) \]

with an obvious notation, and \( \text{diag}(\mu_1) \) is the matrix \( B \) of Theorem 3. This then states that there exist \( r \) eigenvalues \( a_1, \ldots, a_r \) of \( A \) such that

\[ \sum_{i=1}^{r} (a_i - \mu_i)^2 = \| R \|_F^2 \quad (4.11) \]

Notice that the \( \mu_i \) can include multiple or pathologically chic eigenvalues. The result is well known when \( r = 1 \) and the Wielandt-Hoffman theorem corresponds to the case \( r = n \). We observe that by using less than \( r \) of the alleged eigenvectors we can obtain results of the type (4.11) corresponding to any \( s \) (\( \leq r \)) of the approximate eigenvalues.
References


ELEMENTARY PROOF OF THE WIELANDT-HOFFMAN THEOREM AND OF ITS GENERALIZATION

**Abstract**

An elementary proof is given of the Wielandt-Hoffman Theorem for normal matrices and of a generalization of this theorem. The proof makes no direct appeal to results from linear programming theory.
<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROLE</td>
<td>WT</td>
<td>ROLE</td>
</tr>
<tr>
<td>Eigenvalue problems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion and exclusion theorems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>