General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

THE CONDITION OF A FINITE MARKOV CHAIN AND

PERTURBATION BOUNDS FOR THE LIMITING PROBABILITIES[*]

by

Carl D. Meyer, Jr.
Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27650

May 1, 1979

Abstract. Let T denote the transition matrix of an ergodic chain, C, and let
A = I - T. Let E be a perturbation matrix such that $\tilde{T} = T - E$ is also the
transition matrix of an ergodic chain, $\tilde{C}$. Let ω and $\tilde{\omega}$ denote the limiting
probability (row) vectors for C and $\tilde{C}$. The purpose of this paper is to exhibit
inequalities bounding the relative error $\dfrac{\|\omega - \tilde{\omega}\|}{\|\omega\|}$ by a very simple function
of E and A. Furthermore, the inequality will be shown to be the best one which
is possible. This bound can be significant in the numerical determination of
the limiting probabilities for an ergodic chain.

In addition to presenting a sharp bound for $\dfrac{\|\omega - \tilde{\omega}\|}{\|\omega\|}$, an explicit
expression for $\tilde{\omega}$ will be derived in which $\tilde{\omega}$ is given as a function of E, A, ω
and some other related terms.

THE CONDITION OF A FINITE MARKOV CHAIN AND

PERTURBATION BOUNDS FOR THE LIMITING PROBABILITIES

1. **Introduction.** Let T denote the transition matrix of an ergodic chain, C, and let $A = I - T$. (The terminology and notation will be that used in [5] and [6].) Let E be a perturbation matrix such that $\tilde{T} = T - E$ is also the transition matrix of an ergodic chain, $\tilde{C}$. Let ω and $\tilde{\omega}$ denote the limiting probability (row) vectors for C and $\tilde{C}$. The purpose of this paper is to exhibit inequalities bounding the relative error $\dfrac{\| \omega - \tilde{\omega} \|}{\| \omega \|}$ by a very simple function of E and A. Furthermore, the inequality will be shown to be the best one which is possible. This bound can be significant in the numerical determination of the limiting probabilities for an ergodic chain.

In addition to presenting a sharp bound for $\dfrac{\| \omega - \tilde{\omega} \|}{\| \omega \|}$, an explicit expression for $\tilde{\omega}$ will be derived in which $\tilde{\omega}$ is given as a function of E, A, ω and some other related terms.

The approach taken in this paper differs from the traditional methods of past authors in that group properties of the matrix A are used to produce the desired results where as previous results have relyed upon the so called "fundamental matrix" given in [5]. (See [9]) Examples will be given which show that the use of the group properties produce superior results to those which can be produced using the traditional theories.

2. Group Properties.

2. Group Properties. The fundamental fact on which the analysis of this paper is based is the following.

THEOREM. If $A = I - T$ where T is any row stochastic matrix, then A belongs to a multiplicative matrix group.

A proof of this is given in [2] and [6]. It also follows from well known results found in [4] and [8].

Since A belongs to some multiplicative group, G, A must possess an inverse in G. This matrix is called the group inverse of A and is denoted by $A^{\#}$. The identity in G is $P = AA^{\#}$, the projector whose range is R(A) and whose nullspace is N(A).

As is shown in [6] and [2], almost all of the important information concerning an ergodic chain is available in terms of the entries of $A^{\#}$. In particular, the limiting matrix, W, for a chain with transition matrix T is given by

$$(2.1) \qquad W = \lim_{n \to \infty} \frac{I + T + T^2 + \ldots + T^{n-1}}{n} = I - AA^{\#}. \quad \text{(See [6] or [2])}$$

As pointed out in [6], the computation of $A^{\#}$ is not unduly complicated. Indeed, computing $A^{\#}$ is less of a chore than calculating the "fundamental matrix." Further properties of $A^{\#}$ are available in [2].

3. **A Perturbation Formula for $(A + E)^{\#}$.** Suppose T and $\tilde{T}$ are transition matrices

for ergodic chains C and $\tilde{C}$, respectively, where $\tilde{T} = T - E$ so that $\tilde{A} = A + E$.

In order to analyze $\tilde{C}$, it suffices to analyze $\tilde{A}^{\#}$. The purpose of this section

is to provide an expression for $(A + E)^{\#}$ which will hold for all possible values

of E. Notice that E cannot be arbitrary. Since $\tilde{T}$ must be a stochastic matrix,

the elements, e_{ij}, of E are constrained so that $|e_{ij}| \leq 1$. There are, of course,

other additional restrictions.

If $j = [1,1,1,\ldots,1]^T$, then $Aj = 0$ and $(A + E)j = 0$ so that $Ej = 0$. If

ω and $\tilde{\omega}$ denote the limiting probability (row) vectors for C and $\tilde{C}$, respectively,

then (2.1) implies that

$$E(I - AA^{\#}) = E(j\omega) = 0$$

so that

$$(3.1) \qquad EAA^{\#} = E. \quad \text{(i.e., Row Sp(E)} \subseteq \text{Row Sp(A))}$$

Since $A_{n \times n}$ belongs to a matrix group, there exist nonsingular matrices P and

$C_{(n-1) \times (n-1)}$ such that

$$(3.2) \qquad A = P\begin{bmatrix} C & | & 0 \\ - & + & - \\ 0 & | & 0 \end{bmatrix} P^{-1}, \quad A^{\#} = P\begin{bmatrix} C^{-1} & | & 0 \\ - & + & - \\ 0 & | & 0 \end{bmatrix} P^{-1}, \quad \text{and} \quad I - AA^{\#} = P\begin{bmatrix} 0 & | & 0 \\ - & + & - \\ 0 & | & 1 \end{bmatrix} P^{-1}.$$

(These statements are evident, but the reader may wish to consult [2].) Write E

in the form

$$(3.3) \qquad E = P\begin{bmatrix} E_1 & | & E_3 \\ - & -+ & - \\ E_2 & | & E_4 \end{bmatrix} P^{-1}$$

where E_1 is $(n - 1) \times (n - 1)$. The fact that $EAA^\# = E$ implies that $E_3 = 0$ and $E_4 = 0$ so that

$$(3.4) \qquad \tilde{A} = A + E = P \begin{bmatrix} C + E_1 & \vdots & 0 \\ -\ -\ -\ -\ & \vdots & -\ - \\ E_2 & \vdots & 0 \end{bmatrix} P^{-1}.$$

Since $\tilde{C}$ is again an ergodic chain, it must be the case that the limiting matrix, $\tilde{W}$, must be a rank 1 matrix. By virtue of (2.1), it follows that $\text{rank}(I - \tilde{A}\tilde{A}^\#) = 1$. By using the formula

$$(3.5) \qquad \begin{bmatrix} X & \vdots & 0 \\ -\ -\ & \vdots & - \\ Y & \vdots & 0 \end{bmatrix}^\# = \begin{bmatrix} X^\# & \vdots & 0 \\ -\ -\ & \vdots & - \\ YX^{2\#} & \vdots & 0 \end{bmatrix} \qquad \text{(found in [2] or [7])}$$

it is easy to see from (3.4) that

$$I - \tilde{A}\tilde{A}^\# = P \begin{bmatrix} I - (C + E_1)(C + E_1)^\# & \vdots & 0 \\ -\ -\ -\ -\ -\ -\ -\ -\ -\ -\ -\ -\ & \vdots & - \\ -E_2(C + E_1)^\# & \vdots & 1 \end{bmatrix} P^{-1}.$$

The fact that $\text{rank}(I - \tilde{A}\tilde{A}^\#) = 1$ now implies that $I - (C + E_1)(C + E_1)^\# = 0$. That is, $C + E_1$ is a nonsingular matrix. Since $C + E_1 = (I + E_1 C^{-1})C$, it follows that $(I + E_1 C^{-1})$ is nonsingular so that

$$I + EA^\# = P \begin{bmatrix} I + E_1 C^{-1} & \vdots & 0 \\ -\ -\ -\ -\ -\ & \vdots & - \\ E_2 C^{-1} & \vdots & 1 \end{bmatrix} P^{-1}$$

is also nonsingular and

$$(3.6) \qquad (I + EA^\#)^{-1} = P \begin{bmatrix} (I + E_1 C^{-1})^{-1} & \vdots & 0 \\ -\ -\ -\ -\ -\ -\ -\ -\ -\ -\ & \vdots & - \\ -E_2 C^{-1}(I + E_1 C^{-1})^{-1} & \vdots & 1 \end{bmatrix} P^{-1}.$$

Now write the expression for $(A + E)^{\#}$. Using (3.4), (3.5), together with the fact that $(I + E_1 C^{-1})$ is nonsingular, yields

$$(3.7) \qquad (A + E)^{\#} = P \left[\begin{array}{c|c} C^{-1}(I + E_1 C^{-1})^{-1} & 0 \\ \hline E_2 C^{-1}(I + E_1 C^{-1})^{-1} C^{-1}(I + E_1 C^{-1})^{-1} & 0 \end{array} \right] P^{-1}.$$

From (3.2) and (3.6) it is easy to see that

$$A^{\#}(I + EA^{\#})^{-1} = P \left[\begin{array}{c|c} C^{-1}(I + E_1 C^{-1})^{-1} & 0 \\ \hline 0 & 0 \end{array} \right] P^{-1},$$

$$(I - AA^{\#})(I + EA^{\#})^{-1} = P \left[\begin{array}{c|c} 0 & 0 \\ \hline -E_2 C^{-1}(I + E_1 C^{-1})^{-1} & 1 \end{array} \right] P^{-1}$$

and

$$(I - AA^{\#})(I + EA^{\#})^{-1}A^{\#}(I + EA^{\#})^{-1} = P \left[\begin{array}{c|c} 0 & 0 \\ \hline -E_2 C^{-1}(I + E_1 C^{-1})^{-1} C^{-1}(I + E_1 C^{-1})^{-1} & 0 \end{array} \right] P^{-1}.$$

so that (3.7) becomes

$$(3.8) \qquad (A + E)^{\#} = A^{\#}(I + EA^{\#})^{-1} - (I - AA^{\#})(I + EA^{\#})^{-1}A^{\#}(I + EA^{\#})^{-1}.$$

By using the identity $(I + EA^{\#})^{-1} = I - EA^{\#}(I + EA^{\#})^{-1}$, together with (2.1), one arrives at the following result.

THEOREM 3.1. Let C be an ergodic chain with transition matrix T and limiting matrix W and let $\tilde{C}$ be an ergodic chain with transition matrix $\tilde{T} = T - E$. If $A = I - T$, then

$$(A + E)^{\#} = A^{\#} - A^{\#}EA^{\#}(I + EA^{\#})^{-1} - W(I + EA^{\#})^{-1}A^{\#}(I + EA^{\#})^{-1}.$$

It is clear that this theorem guarantees that for the situation under question,

$$\lim_{E \to 0} (A + E)^{\#} = A^{\#}$$

so that the following corollary is obtained

COROLLARY 3.1. For the situation of Theorem 3.1, the elements of $A^{\#}$ depend continuously on the elements of A.

This result can also be proven using the information in [2] or [3].

Now that an explicit representation for $(A + E)^{\#}$ is known, one can obtain almost all of the important information regarding $\tilde{C}$ through the results of [6]. However, the purpose here is to now concentrate on the problem of obtaining a perturbation formula and bounds for the limiting probabilities because it is these quantities which lie at the heart of any analysis of the chain.

4. A Perturbation Formula and Perturbation Bounds for the Limiting Probabilities.

If W and $\tilde{W}$ are the limiting matrices for ergodic chains C and $\tilde{C}$, respectively, then using Theorem 3.1 together with (2.1) yields an explicit expression for $\tilde{W}$ One has the following result.

THEOREM 4.1. If C and $\tilde{C}$ are ergodic chains with transition matrices T and $\tilde{T} = T - E$ and limiting matrices W and $\tilde{W}$, respectively, then

$$\tilde{W} = W(I + EA^{\#})^{-1} = W - WEA^{\#}(I + EA^{\#})^{-1}$$

where $A = I - T$.

In passing, it is pointed out that as a corollary one obtains $\lim_{E \to 0} \tilde{W} = W$, which is of course the well known result stating that the limiting probabilities are continuous functions of the elements of T. By making use of (3.1) and (2.1) another important corollary of Theorem 4.1 is obtained. It is the one which reveals the structure necessary in order for the limiting probabilities to remain invariant under a perturbation.

COROLLARY 4.1. For ergodic chains C and $\tilde{C}$, it is the case that $W = \tilde{W}$ if and only if $R(E) \subseteq R(A)$. (i.e., the limiting probabilities are unaltered if and only if the columns of E are linear combinations of columns of A.)

Consider now the problem of bounding the relative error term $\dfrac{\|\omega - \tilde{\omega}\|}{\|\omega\|}$ where ω and $\tilde{\omega}$ are the limiting probability vectors for C and $\tilde{C}$, respectively. Since every row of W is equal to ω and every row of $\tilde{W}$ is equal to $\tilde{\omega}$, Theorem 4.1 yields

(4.1) $\omega - \tilde{\omega} = \tilde{\omega} E A^{\#}$

and

(4.2) $\omega - \tilde{\omega} = \omega E A^{\#} (I + E A^{\#})^{-1}.$

For the vector 1-norm ($\| x \|_1 = \sum_j |x_j|$) the induced matrix norm is

$$\| A \|_1 = \max_{\| x \|_1 = 1} \| x A \|_1 = \max_i \sum_j |a_{ij}|$$ because one is dealing with row vectors

and left hand multiplication. A trivial observation is that the relative error in ω for the 1-norm is always bounded by 2. That is,

$$\frac{\| \omega - \tilde{\omega} \|_1}{\| \omega \|_1} = \| \omega - \tilde{\omega} \|_1 < 2$$

and $\| \omega - \tilde{\omega} \|_1$ can be made to be arbitrarily close to 2 with particular choices of ω and $\tilde{\omega}$. However, this does not take into account the relative size of $\| E \|_1$. The expression in (4.1) can provide a more useful bound in the case of the 1-norm. Using (4.1) to bound the relative error in ω provides an additional desirable feature. Namely, that the bound is obtainable without having to impose any additional hypothesis on the magnitude of the elements of E.

The above remarks are summarized in the following.

THEOREM 4.2. For ergodic chains C and $\tilde{C}$ with transition matrices T and $\tilde{T} = T - E$ and limiting probability vectors ω and $\tilde{\omega}$, the relative error in ω for the 1-norm is

$$\frac{\| \omega - \tilde{\omega} \|_1}{\| \omega \|_1} = \| \omega - \tilde{\omega} \|_1 \leq \| E A^{\#} \|_1 \leq \left\| \frac{E}{A} \right\|_1 \kappa_1(C)$$

where $A = I - T$ and $\kappa_1(C) = \| A \|_1 \| A^{\#} \|_1$.

The 1-norm may not be the most desirable choice of norms. It seems that the ∞-norm is a more natural choice of norm when investigating the sensitivity of the limiting probabilities to perturbations in the transition probabilities.

It is worth completing the statement on norms by noting that for any two probability vectors, ω and $\tilde{\omega}$, the following relations always hold for an n-state chain.

$$\| \omega \|_1 = 1 \text{ and } \| \omega - \tilde{\omega} \|_1 \leq 2.$$

$$\frac{1}{\sqrt{n}} \leq \| \omega \|_2 \leq 1 \text{ and } \| \omega - \tilde{\omega} \|_2 \leq \sqrt{2}, \text{ so that } \frac{\| \omega - \tilde{\omega} \|_2}{\| \omega \|_2} \leq \sqrt{2n}.$$

$$\frac{1}{n} \leq \| \omega \|_\infty \leq 1 \text{ and } \| \omega - \tilde{\omega} \|_\infty \leq 1, \text{ so that } \frac{\| \omega - \tilde{\omega} \|_\infty}{\| \omega \|_\infty} \leq n.$$

Consider now an arbitrary vector norm and a compatible matrix norm such that $\| I \| = 1$. Take the norm of both sides of (4.2) to obtain

$$\frac{\| \omega - \tilde{\omega} \|}{\| \omega \|} < \| EA^{\#} \| \ \| (I + EA^{\#})^{-1} \|.$$

If $\| EA^{\#} \| < 1$, then

$$\| (I + EA^{\#})^{-1} \| \leq \frac{1}{1 - \| EA^{\#} \|},$$

and the inequality takes a familiar form which is given below.

THEOREM 4.3. Let C and $\tilde{C}$ be ergodic chains with transition matrices T and $\tilde{T}$, respectively, where $\tilde{T} = T - E$. Let $d = \omega - \tilde{\omega}$ where ω and $\tilde{\omega}$ are the limiting probability vectors for C and $\tilde{C}$, respectively, and let $A = I - T$. If $\| EA^{\#} \| < 1$, then

$$\frac{\| d \|}{\| \omega \|} \leq \frac{\| EA^{\#} \|}{1 - \| EA^{\#} \|}.$$

If $\| E \| \ \| A^\# \| \le 1$, then

$$(4.3) \qquad \left\| \frac{d}{\omega} \right\| \le \frac{\left\| \frac{E}{A} \right\| \kappa(C)}{1 - \left\| \frac{E}{A} \right\| \kappa(C)}$$

where $\kappa(C) = \| A \| \ \| A^\# \|$. Moreover, there are nontrivial cases where equality is actually attained in each of the above.

Note that (3.1) guarantees that $\left\| \frac{E}{A} \right\| \le \| EA^\# \|$, which is less than 1, by hypothesis.

The term $\left\| \frac{d}{\omega} \right\|$ is the relative error in ω while $\left\| \frac{E}{A} \right\|$ is the relative error in A. This inequality is exactly of the same form as the familiar inequality obtained when analyzing a perturbed nonsingular linear system of equations. The only difference is the term $\kappa(C)$. The fact that the analysis of any ergodic chain revolves about the limiting probabilities, together with the appearance of $\kappa(C)$ in Theorems 4.2 and 4.3, motivates one to make the following definition.

DEFINITION. Let C be an ergodic chain whose transition matrix is T, and let $A = I - T$. The <u>condition of the chain</u> C is defined to be the number $\kappa(C) = \| A \| \ \| A^\# \|$.

Clearly, if the condition of the chain is relatively small, then the limiting probabilities will be relatively insensitive to small changes in the transition probabilities. If the condition of the chain is relatively large,

then the limiting probabilities may or may not be sensitive. Although the bound in (4.3) can sometimes be pessimistic, it is important to point out that there are nontrivial cases where equality is actually attained. Examples are given in following sections.

As a final observation, note that since $AA^{\#} = I - W$ where $W_{n \times n}$ is the limiting matrix, one has $\| AA^{\#} \|_1 = 2 - 2\min \omega_i$, $\| AA^{\#} \|_2 \geq 1$, and $\| AA^{\#} \|_\infty = 1 + (n - 2)\max \omega_i$ so that

$$\kappa_1(C) \geq 2 - 2\min \omega_i \geq 2 - \frac{2}{n},$$

$$\kappa_2(C) \geq 1,$$

$$\kappa_\infty(C) \geq 1 + (n - 2)\max \omega_i \geq 2 - \frac{2}{n}.$$

A special case which is of frequent interest is that in which the perturbation affects only a single state. That is, only the probabilities for leaving (or entering) the i-th state are perturbed. The question is "how does this effect ω_i and perhaps the rest of ω?"

In this case, the i-th row of T, denoted by t_i, is perturbed so as to produce $\bar{t}_i$, the i-th row of $\tilde{T}$. If $u_i = [0,0,\ldots,0,1,0,\ldots,0]^T$ is the i-th unit vector, then E is the rank 1 matrix $E = u(t_i - \bar{t}_i)$. Equation (4.1) degenerates to

$$d = \omega - \tilde{\omega} = \omega_i e_i A^{\#}(I + EA^{\#})^{-1}$$

where $e_i = t_i - \bar{t}_i$. Since $E = u_i e_i$, one can write

$$(I + EA^{\#})^{-1} = (I + u_i e_i A^{\#})^{-1} = I - \frac{u_i e_i A^{\#}}{1 + e_i A^{\#} u_i}$$

so that (4.2) reduces to the following.

COROLLARY 4.2. If C is an ergodic chain and the transition probabilities for leaving the i-th state are perturbed so as to form an ergodic chain $\tilde{C}$, then

$$(4.4) \qquad \omega - \tilde{\omega} = \omega_i \left[\frac{e_i A^{\#}}{1 + e_i A^{\#} u_i} \right]$$

where ω, $\tilde{\omega}$, e_i, $A^{\#}$, and u_i are as described earlier.

In particular,

$$\frac{\omega_i - \tilde{\omega}_i}{\omega_i} = \frac{\sigma}{1 + \sigma}$$

where $\sigma = e_i A^{\#} u_i$ and

$$\frac{\omega_k - \tilde{\omega}_k}{\omega_k} = \frac{\omega_i}{\omega_k} \left[\frac{e_i A^{\#} u_k}{1 + e_i A^{\#} u_i} \right] .$$

If $\| e_i A^{\#} \| < 1$, then

$$\left| \frac{\omega_i - \tilde{\omega}_i}{\omega_i} \right| \leq \frac{\| e_i A^{\#} \|}{1 - \| e_i A^{\#} \|} \quad \text{and} \quad \left| \frac{\omega_k - \tilde{\omega}_k}{\omega_k} \right| \leq \frac{\omega_i}{\omega_k} \left[\frac{\| e_i A^{\#} \|}{1 - \| e_i A^{\#} \|} \right] .$$

If $\| e_i \| \, \| A^{\#} \| < 1$, then

$$\left| \frac{\omega_i - \tilde{\omega}_i}{\omega_i} \right| \leq \frac{\frac{\| e_i \|}{\| A \|} \kappa(C)}{1 - \frac{\| e_i \|}{\| A \|} \kappa(C)} \quad \text{and} \quad \left| \frac{\omega_k - \tilde{\omega}_k}{\omega_k} \right| \leq \frac{\omega_i}{\omega_k} \left[\frac{\frac{\| e_i \|}{\| A \|} \kappa(C)}{1 - \frac{\| e_i \|}{\| A \|} \kappa(C)} \right] .$$

5. Examples. Below, a general example is constructed to show equality in 4.3 can be attained for the ∞-norm as well as the 2-norm. Note that the fact that row vectors, rather than column vectors, are involved means that the ∞-matrix norm is given by $\| A \|_\infty = \max\limits_{\| x \|_\infty = 1} \| xA \|_\infty = \max\limits_{j} \sum\limits_{i} |a_{ij}|$.

Consider the regular chain whose transition matrix is the symmetric circulant

$$
T = \frac{1}{4n}
\begin{bmatrix}
3 & 1 & 3 & 1 & \cdot & \cdot & \cdot & 3 & 1 \\
1 & 3 & 1 & 3 & \cdot & \cdot & \cdot & 1 & 3 \\
3 & 1 & 3 & 1 & \cdot & \cdot & \cdot & 3 & 1 \\
\cdot & & & & & & & \\
\cdot & & & & & & & \\
\cdot & & & & & & & \\
1 & 3 & 1 & 3 & \cdot & \cdot & \cdot & 1 & 3
\end{bmatrix}_{2n \times 2n}
$$

Since T is symmetric, the limiting probability vector is

$$
\omega = \frac{1}{2n} [1,1,\ldots,1].
$$

It is easy to check that $A^{\#}$ is given by the symmetric circulant

$$
A^{\#} = \frac{1}{n}
\begin{bmatrix}
n & -1 & 0 & -1 & 0 & \cdot & \cdot & \cdot & -1 & 0 & -1 \\
-1 & n & -1 & 0 & -1 & \cdot & \cdot & \cdot & 0 & -1 & 0 \\
0 & -1 & n & -1 & 0 & \cdot & \cdot & \cdot & -1 & 0 & -1 \\
\cdot & & & & & & & \\
\cdot & & & & & & & \\
\cdot & & & & & & & \\
-1 & 0 & -1 & 0 & -1 & \cdot & \cdot & \cdot & -1 & 0 & n
\end{bmatrix}
$$

by verifying that $AA^{\#}A = A$, $A^{\#}AA^{\#} = A^{\#}$, and $AA^{\#} = A^{\#}A$. (These three conditions suffice to define $A^{\#}$. See [2] or [6].) Note that $\| A^{\#} \|_\infty = 2$.

$$\lambda_i = n - \sum_{k=1}^{n} \nu_i^{2k-1} \quad , \quad i = 1,2,3,\ldots,2n$$

where $\{\nu_i | i = 1,2,\ldots,2n\}$ is the set of the 2n-th roots of unity. (See [1])
Since $nA^{\#}$ is symmetric, $\| nA^{\#} \|_2 = \sqrt{\max |\lambda_i|}$ and one can see that $\max |\lambda_i| = \lambda_{n+1} = 2n$ so that

$$\| A^{\#} \|_2 = \frac{1}{n} \| nA^{\#} \|_2 = \frac{2}{\sqrt{2n}} \quad .$$

From (5.1), $\| d \|_2 = \dfrac{\epsilon \sqrt{2n}}{n(1 - 2\epsilon)}$. Now $\| \omega \|_2 = \dfrac{\sqrt{2n}}{2n}$ and $\| E \|_2 = \epsilon \sqrt{2n}$ so that

$$\frac{\| d \|_2}{\| \omega \|_2} = \frac{2\epsilon}{1 - 2\epsilon} = \frac{\| EA^{\#} \|_2}{1 - \| EA^{\#} \|_2} = \frac{\| E \|_2 \| A^{\#} \|_2}{1 - \| E \|_2 \| A^{\#} \|_2} = \frac{\dfrac{\| E \|_2}{\| A \|_2} \kappa_2(C)}{1 - \dfrac{\| E \|_2}{\| A \|_2} \kappa_2(C)}$$

6. Why Not Treat This Strictly as an Eigenvector Problem?

In principle, the problem is an eigenvector problem. That is, one is analyzing the normalized left hand eigenvector associated with the eigenvalue $\lambda_T = 1$ for a row stochastic matrix T, or equivalently, the normalized left hand eigenvector associated with the eigenvalue $\lambda_A = 0$ for $A = I - T$. The known facts concerning the eigenvectors of a perturbed matrix are therefore, to some extent, relevant. However, there are some peculiar aspects which the general eigenvector theory does not capitalize upon. For example, the stochastic nature of the problem sets it apart. The fact that the relevant eigenvalue, $\lambda_A = 0$, (as well as its multiplicity) is unaltered by the perturbation is certainly special. The perturbation E is constrained to be one of a special kind, namely one which preserves the ergodic nature of the chain.

Moreover, the problem at hand is not concerned with the sensitivity of the entire eigensystem of a stochastic matrix. Only a very special eigenvalue and eigenvector with peculiar properties are involved. One should therefore not be too surprised to find some sort of special behavior exhibited which is not present in the general theory.

In general, if x is an eigenvector for B such that $(B - \lambda_1 I)x = 0$ and there exists another eigenvalue λ_2, of B which is close to λ_1, then one expects x to be sensitive to perturbations in B. (See [10]) However, this can produce some wrong impressions when applied to the special case at hand. The following example illustrates how applying this general theory can be somewhat misleading. Let C_1 and C_2 be two ergodic chains whose transition matrices are given by

$$(6.1) \qquad T_1 = \begin{bmatrix} .99995 & .00005 & 0 \\ 0 & .99995 & .00005 \\ .00005 & 0 & .00005 \end{bmatrix}$$

and

$$T_2 = \begin{bmatrix} .99995 & .00005 \\ .00005 & .99995 \end{bmatrix}.$$

The eigenvalues for $A_1 = I - T_1$ are $\lambda_1 = 0$, $\lambda_2 \approx .000100002$ and $\lambda_3 \approx .999949998$ while the limiting probability vector is $\omega_1 \approx (.4999875, .4999875, .000025)$. The eigenvalues for $A_2 = I - T_2$ are $\mu_1 = 0$ and $\mu_2 = .0001$ and the limiting probability vector is $\omega_2 = (.5, .5)$. In each case the matrices have another eigenvalue very close to the eigenvalue 0. The general perturbed eigenvector theory therefore suggests that the eigenvectors associated with the eigenvalue 0 should be "sensitive" to perturbations in the elements of each of the matrices A_1 and A_2.

However, if one allows the term _sensitive_ to mean that small relative errors in the A matrix can produce large relative errors in the limiting vector ω, then the sensitivity of the limiting probabilities may or may not be greatly influenced by the distance between the eigenvalue 0 and the other eigenvalues of A.

For the two chains, C_1 and C_2, of the above example, one finds that

$$A_1^{\#} \approx \begin{bmatrix} 5000 & -4999.75 & -.25 \\ -5000.25 & 5000 & .25 \\ 4999.5 & -5000.25 & .75 \end{bmatrix}$$

and

$$A_2^{\#} = \begin{bmatrix} 5000 & -5000 \\ -5000 & 5000 \end{bmatrix}$$

so that $\kappa_\infty(C_1) \sim 15,000$ while $\kappa_\infty(C_2) = 1$. Theorem 4.3 guarantees that the chain C_2 is well conditioned while C_1 is more badly conditioned. Indeed, if C_1 is perturbed so as to produce $\tilde{C}_1$ with $\tilde{T}_1 = T_1 - E$ where

$$E = \begin{bmatrix} .001 & -.001 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

one finds that

$$\tilde{\omega}_1 \sim (.045452376, .954499897, .000047727)$$

so that a relative error of 10^{-3} in A_1 (using the ∞-norm) produces a relative error of about .91 in ω_1. In contrast, Theorem 4.3 guarantees that a relative error of 10^{-3} in A_2 can produce a relative error of <u>at most</u> $1/999 \sim 10^{-3}$ in ω_2.

The conclusion is that one cannot always use the distance between $\lambda = 0$ and the nearest nonzero eigenvalue of A as a measure of how sensitive the limiting probabilities are to perturbations.

7. Why Not Treat This Strictly as a System of Linear Equations?

If $T_{n \times n}$ is the transition matrix of an ergodic chain, it follows that $A = I - T$ has rank $n - 1$ and any subset of $n - 1$ columns of A is linearly independent. The problem of finding the limiting probability vector is simply that of solving the system

$$\omega A = 0, \quad \Sigma \omega_i = 1.$$

Clearly, this is equivalent to one $n \times n$ nonsingular system of the form $\omega M = b$ where M is obtained from A by replacing one column (say the k-th one) by the column $j = [1,1,\ldots,1]^T$ and b is the k-th unit vector.

Since M is nonsingular and b is not subject to perturbation, the standard result (which is the analogue of Theorem 4.3) holds. That is, if a perturbation of the transition probabilities causes M to go to $\bar{M} = M + F$ where $\| F \| \, \| M^{-1} \| < 1$, then

$$(7.1) \qquad \frac{\| \omega - \bar{\omega} \|}{\| \omega \|} \leq \frac{\frac{\| F \|}{\| M \|} \text{Cond}(M)}{1 - \frac{\| F \|}{\| M \|} \text{Cond}(M)} \qquad \text{where Cond}(M) = \| M \| \, \| M^{-1} \|.$$

and $\| I \| = 1$. (See [10].) This suggests that Cond(M) might also be used as a measure of the condition of the chain.

However, converting the singular matrix A into the nonsingular matrix M can drastically alter the condition of the problem. That is, although A is singular, it can be well conditioned in the sense that $\| A \| \, \| A^{\#} \|$ is small whereas the modified matrix M is nonsingular but $\| M \| \, \| M^{-1} \|$ can be very large.

For example, consider the chain C whose transition matrix is

$$T_{n \times n} = \begin{bmatrix} 1 - \epsilon & \dfrac{\epsilon}{n-1} & \dfrac{\epsilon}{n-1} & \cdots & \dfrac{\epsilon}{n-1} \\[2ex] \dfrac{\epsilon}{n-1} & 1 - \epsilon & \dfrac{\epsilon}{n-1} & \cdots & \dfrac{\epsilon}{n-1} \\ \vdots & & & & \\ \dfrac{\epsilon}{n-1} & \dfrac{\epsilon}{n-1} & \dfrac{\epsilon}{n-1} & \cdots & 1 - \epsilon \end{bmatrix}, \quad 0 < \epsilon < 1,$$

so that

$$A_{n \times n} = \dfrac{\epsilon}{n-1} \begin{bmatrix} n-1 & -1 & -1 & \cdots & -1 \\ -1 & n-1 & -1 & \cdots & -1 \\ \vdots & & & & \\ -1 & -1 & -1 & \cdots & n-1 \end{bmatrix}.$$

Clearly, A is positive semidefinite so that

$$\kappa_2(C) = \| A \|_2 \| A^\# \|_2 = \frac{\max \lambda_i}{\min_{\lambda_i \neq 0} \lambda_i} \quad \text{where } \lambda_i \text{ denotes eigenvalue.}$$

It is easy to verify that the eigenvalues of A are given by

$\{0, \dfrac{\epsilon n}{n-1}, \dfrac{\epsilon n}{n-1}, \dfrac{\epsilon n}{n-1}, \cdots \dfrac{\epsilon n}{n-1}\}$ so that $\kappa_2(C) = 1$, regardless of what value is

assigned to ϵ and what the size of n is. Now replace any column (say the k-th one)

of A by $j = [1,1,\ldots,1]^T$ so as to form the matrix M. The matrix $M^T M$ then has

the form

$$M^T M = n\left(\frac{\varepsilon}{n-1}\right)^2 \begin{bmatrix} n-1 & -1 & -1 \cdots & 0 & -1 & \cdots & -1 \\ -1 & n-1 & -1 \cdots & 0 & -1 & \cdots & -1 \\ \vdots & & & \vdots & & & \\ 0 & 0 & 0 \cdots & \left(\frac{n-1}{\varepsilon}\right)^2 & 0 & \cdots & 0 \\ -1 & -1 & -1 \cdots & 0 & n-1 & \cdots & -1 \\ \vdots & & & \vdots & & & \\ -1 & -1 & -1 \cdots & 0 & -1 & \cdots & n-1 \end{bmatrix}$$

It is not difficult to see that the eigenvalues of $M^T M$ are given by

$\{n, n\left(\frac{\varepsilon}{n-1}\right)^2, \left(\frac{n\varepsilon}{n-1}\right)^2, \ldots, \left(\frac{n\varepsilon}{n-1}\right)^2\}$ so that for large n or small ε,

$$\text{Cond}_2(M) = \frac{\text{max singular value}}{\text{min singular value}} = \frac{n-1}{\varepsilon}.$$

Thus $\text{Cond}_2(M)$ can be made arbitrarily large by either taking ε small or n large. Note also that $\text{Cond}_2(M)$ is independent of which column is selected to contain the 1's.

It clearly would be a mistake to use $\text{Cond}(M)$ as any sort of guide to the sensitivity the limiting probabilities might exhibit to perturbations in the transition probabilities. Aside from the theoretical hazards which the matrix M can produce, it is obvious that M could also present numerical difficulties if it were used in any sort of computational scheme.

The bound produced by using M and (7.1) is almost always inferior to the bound obtained from Theorem 4.3. As an example, consider again the three state chain C, whose transition matrix is given by (6.1). Suppose this chain is perturbed so that the transition matrix becomes $\tilde{T} = T - E$ where

$$E = \begin{bmatrix} -2.5 \times 10^{-5} & 2.5 \times 10^{-5} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then $\bar{A} = A + E$. M is obtained from $\bar{A}$ by replacing some column of A by j.
Assume that in M as well as in $\bar{M}$, the column which is j is taken to be the
second column. Then

$$F = \bar{M} - M = \begin{bmatrix} -2.5 \times 10^{-5} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is the perturbation in M in (7.1). Using the ∞-norm, one finds that
$\text{Cond}_\infty(M) \sim 60,000$ whereas $\kappa_\infty(C) \sim 15,000$. The bound for the relative error in
ω which (7.1) provides is approximately 1 whereas the bound produced by (4.3) is
about .6. In this case, the actual relative error (with the ∞-norm) is about
.33334.

This example exhibits only a single case where (4.3) is superior to (7.1).
However, experience has shown this to be typical. For each value of $n = 3$, 5,
10, 20, and 30, twenty n-state ergodic chains were randomly generated. A
random perturbation (which satisfied the hypothesis of Theorem 4.3 and (7.1))
was introduced and the bounds given by (4.3) and (7.1) were computed using the
∞-norm. For $n = 3$, (4.3) gave a better bound than (7.1) in 13 out of the 20
trials. For $n = 5$, (4.3) gave a better bound in 18 out of 20 trials. For each
of the cases $n = 10$, $n = 20$, and $n = 30$, (4.3) was found to be superior in 20
out of 20 trials. Moreover, for each of the 100 chains generated, $\kappa_\infty(C)$ was
never significantly greater than 5 whereas $\text{Cond}_\infty(M_{n \times n})$ was always in the
neighborhood of n^2.

Since the goal was not to use M in any sort of computational scheme, but rather to determine the degree to which characteristics of M (e.g., Cond(M)) relfect the relative sensitivity of the limiting probabilities, no attempt was made to scale M. This, of course, could be done and should be done if M is specifically given and is to be used in computations. However, when M is not specifically given, no theoretical advantage as far as producing a general analytical bound on the relative error can be realized.

REFERENCES

1. Browne, Edward T., Introduction to the Theory of Determinants and Matrices, The University of North Carolina Press, Chapel Hill, N.C., 1958.

2. Campbell, S. L., & Meyer, Carl D., Generalized Inverses of Linear Transformations, Pitman Pub. Co., (Surveys and Reference Works in Mathematics), 1979, London.

3. Campbell S. L., & Meyer, Carl D., Continuity Properties of The Drazin Pseudoinverse, Linear Algebra And Its Applications, 10(1975), pp. 77-83.

4. Gantmacher, F. R., Matrix Theory, Vol. 2., Chelsea Pub. Co., New York, 1960.

5. Kemeny, J. G. & Snell, J. L., Finite Markov Chains, D Van Nostrand Co., New York, 1960.

6. Meyer. Carl. D., The Role Of The Group Generalized Inverse In The Theory Of Finite Markov Chains, SIAM Rev., 17(1975), pp. 443-464.

7. Meyer, Carl D., & Rose, N. J., The Index And The Drazin Inverse Of Block Triangular Matrices, SIAM J. Appl. Math., 33(1977), pp. 1-7.

8. Mirsky, L., An Introduction to Linear Algebra, Oxford At The Clarendon Press, London, 1963.

9. Schweitzer, Paul J., Perturbation Theory and Finite Markov Chains, J. Appl. Prob., 5(1968), pp. 401-413

10. Wilkinson, J. H., The Algebraic Eigenvalue Problem, Oxford University Press, London, 1965.