AUGMENTED TRUNCATIONS OF INFINITE STOCHASTIC MATRICES

DIANA GIBSON and E. SENETA
Department of Mathematical Statistics
University of Sydney
N.S.W. 2006 Australia

Summary

We consider the problem of approximating the stationary distribution of a positive-recurrent Markov chain with infinite transition matrix P, by stationary distributions computed from $(n \times n)$ stochastic matrices formed by augmenting the entries of the $(n \times n)$ northwest corner truncations of P, as $n \to \infty$. A generalization to quasi-stationary distributions is also considered.

KEY WORDS AND PHRASES: STATIONARY DISTRIBUTION, TRUNCATION, AUGMENTATION, LAST-EXIT PROBABILITIES, UPPER-HESSENBERG, LOWER-HESSENBERG, QUASI-STATIONARY DISTRIBUTION.

26th May, 1986.

1. INTRODUCTION

We are concerned throughout with approximating the stationary distribution $\underset{\sim}{\pi}$ of an infinite positive-recurrent Markov chain on the positive integers $\mathbb{N}$, with transition matrix P, through the finite northwest corner truncations of P. Let ${}_{(n)}P$ denote the truncation of size n. It is aesthetically pleasing to try to approximate the stationary distribution $\underset{\sim}{\pi} = \{\pi_i\}$ by a sequence of stationary distributions $\{{}_{(n)}\underset{\sim}{\pi}\}_{n=1}^{\infty}$. We consider ${}_{(n)}\underset{\sim}{\pi}$ obtained from an $n \times n$ stochastic matrix ${}_{(n)}\tilde{P}$ where ${}_{(n)}\tilde{P} \geqslant {}_{(n)}P$ elementwise, and ask for what kinds of P and what sequences $\{{}_{(n)}\tilde{P}\}_{n=1}^{\infty}$ is it true that ${}_{(n)}\underset{\sim}{\pi} \to \underset{\sim}{\pi}$. (By convergence of probability vectors we mean convergence in ℓ_1 which is equivalent to elementwise (see Wolf (1975), Lemma 1)).

In this paper, we prove that for a Markov matrix P or an upper-Hessenberg P any sequence $\{{}_{(n)}\tilde{P}\}_{n=1}^{\infty}$ will do; that certain methods of constructing ${}_{(n)}\tilde{P}$ work for <u>all</u> P; and that for lower-Hessenberg P we must be somewhat careful in generalizing these. The motivating papers in the investigation of this problem are Seneta (1980) and Wolf (1980) Section 5, although earlier papers by both authors play a role.

We also consider in Section 5 the problem of approximating the quasi-stationary distribution (Seneta (1981), Ch. 7) of an infinite non-negative irreducible R-positive matrix T whose left R-invariant vector is summable, and the special case when T is stochastic and positive-recurrent.

Returning to our basic context of positive-recurrent P, with $(n \times n)$ northwest corner truncation ${}_{(n)}P$, and $(n \times n)$ stochastic ${}_{(n)}\tilde{P}$ where ${}_{(n)}\tilde{P} \geqslant {}_{(n)}P$, let $\ell_{ij}^{(k)}$, ${}_{(n)}\ell_{ij}^{(k)}$, ${}_{(n)}\tilde{\ell}_{ij}^{(k)}$ denote the last-exit probabilities from state i to state j, and

$L_{ij}(z)$, ${}_{(n)}L_{ij}(z)$, ${}_{(n)}\tilde{L}_{ij}(z)$ the corresponding generating functions, $|z| \leqslant 1$. (See Seneta, 1981, Chapters 5 and 6 for amplification on these and the following introductory remarks.)

Note that

$$(1.1) \qquad \pi_j/\pi_i = L_{ij}(1)$$

and similarly if C_n is any essential class of indices (states) of ${}_{(n)}\tilde{P}$

$$(1.2) \qquad {}_{(n)}\pi_j/{}_{(n)}\pi_i = {}_{(n)}\tilde{L}_{ij}(1) \qquad i,j \in C_n$$

where ${}_{(n)}\underset{\sim}{\pi} = \{{}_{(n)}\pi_i\}$ is the corresponding stationary distribution of ${}_{(n)}\tilde{P}$. Finally recall that as $n \to \infty$

$$(1.3) \qquad {}_{(n)}L_{ij}(1) \uparrow L_{ij}(1).$$

In consequence of these relations, last exit generating functions will play a centrol role in our discussion.

---

2. GENERAL AUGMENTATION FOR SPECIAL MATRICES

Lemma 2.1. Let ${}_{(n)}\tilde{P}$ be any $(n \times n)$ stochastic matrix with ${}_{(n)}\tilde{P} \geqslant {}_{(n)}P$, and suppose for all sufficiently large n, ${}_{(n)}\tilde{P}$ has a unique essential class C_n, which contains, for all such n, a fixed pair of indices i,j. Let ${}_{(n)}\underset{\sim}{\pi}$ then denote the unique stationary distribution of ${}_{(n)}\tilde{P}$. Then as $n \to \infty$

$${}_{(n)}\pi_j/{}_{(n)}\pi_i \to \pi_j/\pi_i$$

Proof: Since ${}_{(n)}P \leqslant {}_{(n)}\tilde{P}$

$${}_{(n)}L_{ij}(1) \leqslant {}_{(n)}\tilde{L}_{ij}(1) \qquad \text{if} \quad n \geqslant \max(i,j).$$

Now for n so large that $i,j \in C_n$, from (1.2)

$$ {}_{(n)}\tilde{L}_{ij}(1) = {}_{(n)}\pi_j / {}_{(n)}\pi_i = 1/{}_{(n)}\tilde{L}_{ji}(1) $$

So

$$ {}_{(n)}L_{ij}(1) \leqslant {}_{(n)}\pi_j / {}_{(n)}\pi_i \leqslant 1/{}_{(n)}L_{ji}(1) \tag{2.1} $$

But $\lim_{n\to\infty} {}_{(n)}L_{ij}(1) = L_{ij}(1) = \pi_j/\pi_i = 1/L_{ji}(1) = \lim_{n\to\infty} 1/{}_{(n)}L_{ji}(1)$

using (1.1) and (1.3)

$\therefore$ $\lim_{n\to\infty} {}_{(n)}\pi_j / {}_{(n)}\pi_i$ exists and equals π_j/π_i. □

This result leaves open the general question of convergence of ${}_{(n)}\underset{\sim}{\pi}$ to $\underset{\sim}{\pi}$ for a positive-recurrent P, which as we shall see from Section 3, does not necessarily hold under the conditions of the lemma. However, it does hold if the infinite matrix P has special structure.

Definition 2.1 A stochastic matrix $P = \{p_{ij}\}$ is said to be a *Markov matrix* if the elements of at least one column are bounded away from zero i.e. there exists a j_0 and an $\varepsilon > 0$ such that $p_{ij_0} > \varepsilon$, all i. □

Such a matrix has single essential class, which is positive-recurrent, aperiodic, and contains j_0.

Theorem 2.1 Let P be a Markov matrix and for each $n \in \mathbb{N}$, let ${}_{(n)}\tilde{P}$ be an $(n \times n)$ stochastic matrix satisfying ${}_{(n)}\tilde{P} \geqslant {}_{(n)}P$. Then for all n sufficiently large ${}_{(n)}\tilde{P}$ has a unique stationary distribution ${}_{(n)}\underset{\sim}{\pi}$ and ${}_{(n)}\underset{\sim}{\pi} \to \underset{\sim}{\pi}$ as $n \to \infty$.

Proof: ${}_{(n)}\tilde{P}$ is a Markov matrix for all n sufficiently large to take in the column (say j_0-th) uniformly bounded from 0 in P. The rest of the proof is precisely as in Seneta (1980), §2 or Seneta (1981), Theorem 7.3, where the unnecessary assumption is made that the column j_0 is augmented in ${}_{(n)}P$ to form ${}_{(n)}\tilde{P}$. □

<u>Definition 2.2</u> A stochastic matrix $P = \{p_{ij}\}$ is said to be upper-Hessenberg if $p_{ij} = 0$ if $i > j + 1$. □

Since any Markov chain governed by such a P in passing from a state i to a state j, where $i > j$, must pass through every intermediate state, it follows that

$$\ell_{ij}^{(k)} = {}_{(n)}\ell_{ij}^{(k)}, \; n \geq i > j, \text{ and } k \in \mathbb{N}.$$

Therefore for such i,j

$$L_{ij}(z) = {}_{(n)}L_{ij}(z), \; |z| \leq 1. \tag{2.2}$$

In the sequel the blanket assumption that P is positive-recurrent is to be understood.

<u>Theorem 2.2</u> Suppose P is upper-Hessenberg and for each $n \in \mathbb{N}$ let ${}_{(n)}\tilde{P}$ be an $n \times n$ stochastic matrix satisfying ${}_{(n)}\tilde{P} \geq {}_{(n)}P$. Then ${}_{(n)}\tilde{P}$ has unique stationary distribution ${}_{(n)}\underset{\sim}{\pi}$ and ${}_{(n)}\underset{\sim}{\pi} \to \underset{\sim}{\pi}$ as $n \to \infty$.

<u>Proof</u>: Since P is irreducible, all entries on its subdiagonal are positive, i.e. $p_{i+1,i} > 0$, $\forall\, i \in \mathbb{N}$. Hence $j \to 1$ with respect to ${}_{(n)}\tilde{P}$ for all $j \in \{2,\ldots,n\}$. So ${}_{(n)}\tilde{P}$ has just one essential class, C_n, say, and $1 \in C_n$.

Note that $\bigcup_{n=1}^{\infty} C_n = \mathbb{N}$ since any index j communicates with 1 for large enough n.

Take $i = 1$ in (2.1) and sum over j to obtain

$$\sum_{j \in C_n} {}_{(n)}L_{1j}(1) \leq \sum_{j \in C_n} {}_{(n)}\pi_j / {}_{(n)}\pi_1 = 1/{}_{(n)}\pi_1 \leq \sum_{j \in C_n} 1/{}_{(n)}L_{j1}(1).$$

But ${}_{(n)}L_{1j}(1) \leq L_{1j}(1)$ any $j \leq n$, so by dominated convergence and (1.3)

$$\lim_{n\to\infty} \sum_{j \in C_n} {}_{(n)}L_{1j}(1) = \sum_{j=1}^{\infty} L_{1j}(1)$$

$$= 1/\pi_1 \qquad \text{by (1.1).}$$

Also, from (2.2),

$$\sum_{j\varepsilon C_n} 1/{}_{(n)}L_{j1}(1) = 1/{}_{(n)}L_{11}(1) + \sum_{\substack{j\varepsilon C_n \\ j\neq 1}} 1/L_{j1}(1)$$
$$\to \sum_{j=1}^{\infty} 1/L_{j1}(1) \qquad \text{by (1.3)}$$
$$= 1/\pi_1 \qquad \text{by (1.1) again.}$$

Thus

(2.3) $$1/{}_{(n)}\pi_1 \to 1/\pi_1 \quad \text{as} \quad n \to \infty .$$

Since $j\varepsilon C_n$ for large enough n, we can use (2.3) together with Lemma 2.1 to show

$${}_{(n)}\pi_j = \frac{{}_{(n)}\pi_j/{}_{(n)}\pi_1}{1/{}_{(n)}\pi_1} \to \frac{\pi_j/\pi_1}{1/\pi_1} = \pi_j$$

$$\text{as} \quad n \to \infty . \qquad \square$$

A version of this theorem (where ${}_{(n)}\tilde{P}$ is formed from ${}_{(n)}P$ by augmenting only the *last* column of ${}_{(n)}P$ but leading to a stronger conclusion) was proved by Golub and Seneta (1974)(see Seneta (1981), Lemma 7.3). Indeed, then ${}_{(n)}\tilde{\ell}_{nj}^{(k)} = {}_{(n)}\ell_{nj}^{(k)} = \ell_{nj}^{(k)}$, $1 \leqslant j < n$, so ${}_{(n)}\pi_j/{}_{(n)}\pi_n = \pi_j/\pi_n$, $1 \leqslant j \leqslant n$. (Similar notions apply to the generalized renewal matrix treated in the same sources.)

---

3. LINEAR AUGMENTATION FOR GENERAL MATRICES

Consider the method, which we shall call linear augmentation, of constructing a stochastic $(n \times n)$ matrix ${}_{(n)}\tilde{P} \geqslant {}_{(n)}P$ suggested in Seneta (1980):

(3.1) $${}_{(n)}\tilde{P} = {}_{(n)}P + ({}_{(n)}I - {}_{(n)}P)\, {}_{(n)}\underset{\sim}{1}\, {}_{(n)}\underset{\sim}{\alpha}'$$

where ${}_{(n)}\underset{\sim}{\alpha}$ is a probability n-vector, and ${}_{(n)}\underset{\sim}{1}$ is an n-vector of 1's.

Seneta (1980); (1981) Section 7.2, showed that ${}_{(n)}\tilde{P}$ thus formed has unique essential class, and correspondingly unique stationary distribution given by

(3.2) $$\frac{{}_{(n)}\underset{\sim}{\alpha}'({}_{(n)}I - {}_{(n)}P)^{-1}}{{}_{(n)}\underset{\sim}{\alpha}'({}_{(n)}I - {}_{(n)}P)^{-1}\, {}_{(n)}\underset{\sim}{1}} = {}_{(n)}\pi' .$$

Let ${}_{(n)}\underset{\sim}{f}_i$ be the n-vector with unity in the i-th position, zeros elsewhere, $1 \leqslant i \leqslant n$.

<u>Theorem 3.1</u> For fixed $i \geqslant 1$, and $n \geqslant i$, let ${}_{(n)}\tilde{P}$ be formed from ${}_{(n)}P$ by linear augmentation using ${}_{(n)}\underset{\sim}{\alpha} = {}_{(n)}\underset{\sim}{f}_i$ (i.e. by increasing the elements of its i-th column only), and let ${}_{(n)}\underset{\sim}{\pi}$ be the unique stationary distribution of ${}_{(n)}\tilde{P}$. Then as $n \to \infty$ ${}_{(n)}\underset{\sim}{\pi} \to \underset{\sim}{\pi}$.

<u>Proof</u>: For $n \geqslant i$, let C_n denote the unique essential class of indices corresponding to ${}_{(n)}\tilde{P}$. Then $i \in C_n$, otherwise C_n would be an essential class of the infinite matrix P, contradicting its assumed irreducibility.

Let $j \in \mathbb{N}$ $(j \neq i)$. For sufficiently large n, again by irreducibility of P, $j \in C_n$. Then by (1.2)

$$(3.3) \qquad {}_{(n)}\pi_j / {}_{(n)}\pi_i = {}_{(n)}\tilde{L}_{ij}(1).$$

But, since ${}_{(n)}\tilde{P}$ differs from ${}_{(n)}P$ only in the i-th column ${}_{(n)}\tilde{\ell}^{(k)}_{ij} = {}_{(n)}\ell^{(k)}_{ij}$ whence:

$${}_{(n)}\tilde{L}_{ij}(1) = {}_{(n)}L_{ij}(1).$$

By (1.3) and (3.3), as $n \to \infty$

$$(3.4) \qquad {}_{(n)}\pi_j / {}_{(n)}\pi_i \uparrow \pi_j / \pi_i$$

(c.f. Lemma 2.1). For $j \notin C_n$, ${}_{(n)}\pi_j = 0$. Hence

$$(3.5) \quad 1/{}_{(n)}\pi_i = \sum_{j=1}^{n} {}_{(n)}\pi_j / {}_{(n)}\pi_i = \sum_{j=1}^{\infty} {}_{(n)}\pi_j / {}_{(n)}\pi_i \uparrow \sum_{j=1}^{\infty} \pi_j / \pi_i = 1/\pi_i$$

by dominated convergence. Hence for fixed $j \in \mathbb{N}$, by (3.4) and (3.5)

$${}_{(n)}\pi_j = ({}_{(n)}\pi_j / {}_{(n)}\pi_i) / (1/{}_{(n)}\pi_i) \to \pi_j$$

as $n \to \infty$. □

Equation (3.4) was proved by Seneta (1967)(1968) in a different guise; see Seneta (1980). Theorem 3.1 was proved by Wolf (1975), Satz 3, essentially by the above augment; see also Wolf (1980), Section 5, and Allen, Anderssen and Seneta (1977). The theorem is included here for completeness, central importance, and focus on the role of last-exit probabilities. The result can be extended as follows: we omit the proof for brevity.

<u>Theorem 3.2</u> Let $\underset{\sim}{\alpha} = \{\alpha_i\}_1^\infty$ be a probability vector with $\sum_{j=1}^{N} \alpha_j = 1$ for some fixed finite N, and let ${}_{(n)}\underset{\sim}{\alpha}$ consist of the first n entries of $\underset{\sim}{\alpha}$, $n \geq N$. Let ${}_{(n)}\tilde{P}$ be formed by linear augmentation of ${}_{(n)}P$ using ${}_{(n)}\underset{\sim}{\alpha}$, $n \geq N$. Then ${}_{(n)}\underset{\sim}{\pi} \to \underset{\sim}{\pi}$ as $n \to \infty$, where ${}_{(n)}\pi$ is the unique stationary distribution of ${}_{(n)}\tilde{P}$. □

That arbitrary linear augmentation is not always successful, and the need to restrict the manner of growth of the probability vector ${}_{(n)}\underset{\sim}{\alpha}$ as $n \to \infty$, is demonstrated by the following example, where ${}_{(n)}\underset{\sim}{\alpha} = {}_{(n)}\underset{\sim}{f}_n$ (so augmentation of ${}_{(n)}P$ to form ${}_{(n)}\tilde{P}$ occurs only in the *last* column).

EXAMPLE: Consider a stochastic renewal matrix

$$(3.6)\quad P = \begin{bmatrix} q_1 & p_1 & 0 & 0 & \cdot & \cdot & \cdot \\ q_2 & 0 & p_2 & 0 & \cdot & \cdot & \cdot \\ q_3 & 0 & 0 & p_3 & & & \\ \cdot & & & & \cdot & & \\ \cdot & & & & & \cdot & \\ \cdot & & & & & & \cdot \end{bmatrix}$$

where $0 < p_i < 1 \quad \forall\, i \in \mathbb{N}$. P is clearly irreducible. Define $a_0 = 1$ and $a_j = \prod_{i=1}^{j} p_i \quad j \in \mathbb{N}$.

It's easy to see that P positive-recurrent is equivalent to $\sum_{j=0}^{\infty} a_j < \infty$ (e.g. Seneta (1981), Section 5.6).

In this case, the stationary equations yield $\pi_1 = 1/\sum_{j=0}^{\infty} a_j$,

$\pi_j = a_{j-1}\,\pi_1 \quad j \in \mathbb{N}$.

Fix $N \geqslant 3$ and define

$$p_j = \begin{cases} (j/j+1)^2 & \text{if } j = 1 \text{ or } j \equiv 2,3,\ldots,N-1 \pmod N \\ 1-(1/j^2) & \text{if } j \equiv 0 \pmod N \\ \dfrac{(j-1)^4}{((j-1)^2-1)} \cdot \dfrac{1}{(j+1)^2} & \text{if } j \equiv 1 \pmod N \text{ but } j \neq 1. \end{cases}$$

Then for $j \geqslant 1$

$$a_j = \begin{cases} (1/j+1)^2 & \text{if } j \not\equiv 0 \pmod N \\ (j^2-1)/j^4 & \text{if } j \equiv 0 \pmod N \end{cases}$$

so P is positive recurrent.

Notice that ${}_{(n)}\tilde{P}$ is irreducible for all $n \in \mathbb{N}$, and that the conditions of Lemma 2.1 are satisfied.

The stationary equations ${}_{(n)}\underset{\sim}{\pi}' = {}_{(n)}\underset{\sim}{\pi}'\,{}_{(n)}\tilde{P}$ give ${}_{(n)}\pi_1 = 1/(\sum_{j=0}^{n-2} a_j + \frac{a_{n-1}}{q_n})$, since

$$ {}_{(n)}\pi_j = \begin{cases} {}_{(n)}\pi_1\; a_{j-1} & j = 1,\ldots,n-1 \\ {}_{(n)}\pi_1\; \dfrac{a_{j-1}}{q_j} & j = n \end{cases}$$

But for $n \equiv 0 \pmod N$, $\frac{a_{n-1}}{q_n} = 1$

Hence ${}_{(n)}\pi_1 \not\to \pi_1$ as $n \to \infty$. □

---

4. LOWER-HESSENBERG P.

Definition 4.1 A stochastic matrix $P = \{p_{ij}\}$ is said to be lower-Hessenberg if $p_{ij} = 0$, $j > i + 1$. □

Such matrices satisfy a property dual to (2.2). Specially, if $f_{ij}^{(k)}$, ${}_{(n)}f_{ij}^{(k)}$ denote the first-passage probabilities from state i to state j, then $f_{ij}^{(k)} = {}_{(n)}f_{ij}^{(k)}$, $i < j \leqslant n$, whence

$$(4.1) \qquad F_{ij}(z) = {}_{(n)}F_{ij}(z), \qquad |z| \leqslant 1,\ i < j \leqslant n.$$

We should also note the properties dual to (1.1) and (1.3) for positive-recurrent P (Seneta (1981), Chapter 5): as $n \to \infty$

$$(4.2) \qquad {}_{(n)}F_{ij}(1) \uparrow F_{ij}(1) = 1.$$

Although there is an obvious duality between upper - and lower - Hessenberg P, property (2.2) of the former is far more pertinent to our problem than property (4.1) of the latter, because it links the *left* Perron-Frobenius structure of the truncations with that of the infinite matrix. The Example of Section 3 shows how difficulties may arise with positive-recurrent lower-Hessenberg P, in contrast to Theorem 2.2 for such upper-Hessenberg P.

If, however, as suggested by the Example and Theorem 3.1 we require that the sequence $\{{}_{(n)}\tilde{P}\}_{n=1}^{\infty}$ be constructed by linear augmentation (3.1) using a sequence $\{{}_{(n)}\underset{\sim}{\alpha}\}_{n=1}^{\infty}$ which is more "stable" than the sequence $\{{}_{(n)}\underset{\sim}{f}_n\}$, then the desired convergence of the corresponding stationary distributions obtains for lower-Hessenberg P.

We need additional notation. Define ${}_{(n)}B = ({}_{(n)}b_{ij})_{i,j \in \mathbb{N}}$ by

$${}_{(n)}b_{ij} = \begin{cases} (({}_{(n)}I - {}_{(n)}P)^{-1})_{ij}, & i,j \in \{1,\dots,n\} \\ 0, & \text{otherwise.} \end{cases}$$

(According to Seneta (1980), $0 \leqslant {}_{(n)}b_{ij} < \infty$).

For any element $\underset{\sim}{\beta} = (\beta_i)_{i \in \mathbb{N}}$ of ℓ_1 define

$${}_{(n)}B(\underset{\sim}{\beta}) = \underset{\sim}{\beta}' \, {}_{(n)}B.$$

Then ${}_{(n)}B \in L(\ell_1)$ and

$$\| {}_{(n)}B \| = \sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} {}_{(n)}b_{ij}.$$

<u>Theorem 4.1</u> Suppose that P is lower-Hessenberg and let $\{{}_{(n)}\underset{\sim}{\alpha}\}_{n=1}^{\infty}$ be a tight sequence of probability vectors with $\sum_{j=1}^{n} {}_{(n)}\alpha_j = 1 \;\forall\; n \in \mathbb{N}$. If ${}_{(n)}\underset{\sim}{\pi}$ is defined by (3.1) for each $n \in \mathbb{N}$, then ${}_{(n)}\underset{\sim}{\pi} \to \underset{\sim}{\pi}$ as $n \to \infty$.

In other words $\dfrac{{}_{(n)}B({}_{(n)}\underset{\sim}{\alpha})}{\|{}_{(n)}B({}_{(n)}\underset{\sim}{\alpha})\|} \to \underset{\sim}{\pi}$ as $n \to \infty$.

We first prove an auxiliary lemma.

<u>Lemma 4.1</u> Assume P is lower-Hessenberg. Let $i_0 \in \mathbb{N}$ and $\delta > 0$. Then there exists a constant $C = C(i_0,\delta) \geq 1$ such that for any probability vector $\underset{\sim}{\beta}$ with $\beta_{i_0} > \delta$

$$\frac{\|{}_{(n)}B\|}{\|{}_{(n)}B(\underset{\sim}{\beta})\|} \leq C \qquad \forall\; n \geq i_0.$$

<u>Proof</u>: For $n \geq i_0$, $\|{}_{(n)}B\| = \max_{1 \leq i \leq n} \{\sum_{j=1}^{n} {}_{(n)}b_{ij}\}$

and $$\|{}_{(n)}B(\underset{\sim}{\beta})\| = \sum_{j=1}^{n} \sum_{i=1}^{n} \beta_i \; {}_{(n)}b_{ij}$$
$$\geq \delta \sum_{j=1}^{n} {}_{(n)}b_{i_0 j} ,$$

so it is sufficient to find $C \geq 1$ such that
$$a_{n,i} = \frac{\sum_{j=1}^{n} {}_{(n)}b_{ij}}{\sum_{j=1}^{n} {}_{(n)}b_{i_0 j}} \leq C \qquad \forall\; n \geq i_0,\; 1 \leq i \leq n .$$

We use the notation of Seneta (1968); (1981), Chapter 6 in defining

$${}_{(n)}C_{ij}(1) = \text{cofactor of } (i,j)\text{th element of } ({}_{(n)}I - {}_{(n)}P)$$

and

$${}_{(n)}\Delta(1) = \det({}_{(n)}I - {}_{(n)}P)$$

Seneta (1981), Chapter 6 showed

(4.3) $${}_{(n)}L_{ij}(1) = \frac{{}_{(n)}C_{ji}(1)}{{}_{(n)}C_{ii}(1)} \quad \text{if} \quad j \neq i$$

A dual argument gives

$$(4.4) \qquad {}_{(n)}F_{ij}(1) = \frac{{}_{(n)}C_{ji}(1)}{{}_{(n)}C_{jj}(1)}, \qquad j \neq i .$$

Hence

$$\begin{aligned} {}_{(n)}b_{ij} &= {}_{(n)}C_{ji}(1)/{}_{(n)}\Delta(1) \\ &= \begin{cases} {}_{(n)}L_{ij}(1)\ {}_{(n)}C_{ii}(1)/{}_{(n)}\Delta(1) & i \neq j \\ {}_{(n)}C_{ii}(1)/{}_{(n)}\Delta(1) & i = j \end{cases} \end{aligned}$$

$\therefore$

$$a_{n,i} = \left\{ \frac{1 + \sum_{\substack{j=1 \\ j\neq i}}^{n} {}_{(n)}L_{ij}(1)}{1 + \sum_{\substack{j=1 \\ j\neq i_0}}^{n} {}_{(n)}L_{i_0 j}(1)} \right\} \left\{ \frac{{}_{(n)}C_{ii}(1)}{{}_{(n)}C_{i_0 i_0}(1)} \right\} .$$

$$\begin{aligned} \text{The first factor} &\leqslant 1 + \sum_{\substack{j=1 \\ j\neq i}}^{n} {}_{(n)}L_{ij}(1) \\ &\leqslant 1 + \sum_{\substack{j=1 \\ j\neq i}}^{\infty} L_{ij}(1) && \text{by (1.3)} \\ &= 1/\pi_i && \text{by (1.1)}. \end{aligned}$$

$$\begin{aligned} \text{The second factor} &= \frac{{}_{(n)}C_{ii_0}(1)}{{}_{(n)}C_{i_0 i_0}(1)} \cdot \frac{{}_{(n)}C_{ii}(1)}{{}_{(n)}C_{ii_0}(1)} \\ &= \frac{{}_{(n)}L_{i_0 i}(1)}{{}_{(n)}F_{i_0 i}(1)} && \text{by (4.3) and (4.4)} \\ &\to \pi_i/\pi_{i_0} && \text{by (1.3) and (4.2)} \\ &\text{as } n \to \infty . \end{aligned}$$

It follows that for $i \leqslant i_0$ (i.e. finitely many i) there exists a constant $C' \geqslant 1$ such that

$$\frac{{}_{(n)}C_{ii}(1)}{{}_{(n)}C_{i_0 i_0}(1)} \leqslant C' \frac{\pi_i}{\pi_{i_0}} \qquad \forall \ n \geqslant i_0 .$$

For $i > i_0$ and $n \geqslant i$,

$$\frac{{}_{(n)}C_{ii}(1)}{{}_{(n)}C_{i_0 i_0}(1)} = {}_{(n)}L_{i_0 i}(1)/F_{i_0 i}(1) \qquad \text{by (4.1)}$$
$$= {}_{(n)}L_{i_0 i}(1) \qquad \text{by (4.2)}$$
$$\leqslant L_{i_0 i}(1) \qquad \text{by (1.3)}$$
$$= \frac{\pi_i}{\pi_{i_0}} \qquad \text{by (1.1)}$$

So $$\frac{{}_{(n)}C_{ii}(1)}{{}_{(n)}C_{i_0 i_0}(1)} \leqslant C' \frac{\pi_i}{\pi_{i_0}} \qquad \forall \; n \geqslant i_0 \text{ and } 1 \leqslant i \leqslant n.$$

Take $C = \dfrac{C'}{\delta \pi_{i_0}}$ and the lemma is proved. □

<u>Proof of Theorem 4.1</u>

<u>Step 1</u>

Consider a subsequence $\{{}_{(n_k)}\underset{\sim}{\alpha}\}_{k=1}^{\infty}$ of $\{{}_{(n)}\underset{\sim}{\alpha}\}_{n=1}^{\infty}$ such that ${}_{(n_k)}\underset{\sim}{\alpha}$ converges to a probability vector, $\underset{\sim}{\alpha}$ say, as $k \to \infty$.

We will prove

(4.5) $$\frac{{}_{(n_k)}B({}_{(n_k)}\underset{\sim}{\alpha})}{\| {}_{(n_k)}B({}_{(n_k)}\underset{\sim}{\alpha}) \|} \to \underset{\sim}{\pi} \quad \text{as} \quad n \to \infty .$$

Let $\varepsilon > 0$ be arbitrary. Since $\underset{\sim}{\alpha}$ is a probability vector, for some $i_0 \in \mathbb{N}$, $\alpha_{i_0} > 0$.

Let $\delta = \alpha_{i_0}/2$.

Then because ${}_{(n_k)}\underset{\sim}{\alpha} \to \underset{\sim}{\alpha}$ we can find K_1 so large that ${}_{(n_k)}\alpha_{i_0} > \delta \quad \forall \quad k \geqslant K_1$.

Now use Lemma 4.1 to find $C \geqslant 1$ such that

$$\frac{\| {}_{(n_k)}B \|}{\| {}_{(n_k)}B({}_{(n_k)}\underset{\sim}{\alpha}) \|} \leqslant C \qquad \forall \quad k \geqslant K_1 \quad \text{such that} \quad n_k \geqslant i_0 .$$

The sequence $\{{}_{(n_k)}\underset{\sim}{\alpha}\}_{k=1}^{\infty}$ is tight, being convergent to a probability vector, so there exists N so large that $N \geqslant i_0$ and

(4.6) $$\sum_{j=N+1}^{\infty} {}_{(n_k)}\alpha_j \leqslant \frac{\varepsilon}{6C^2} \qquad \forall \ k \in \mathbb{N}.$$

Define: ${}_{(0)}\underset{\sim}{\alpha}$ by $${}_{(0)}\alpha_j = \begin{cases} \alpha_j & j \leqslant N \\ 0 & j > N \end{cases}$$

Next, since ${}_{(n_k)}\underset{\sim}{\alpha} \to \underset{\sim}{\alpha}$, there exists K_2 so large that

(4.7) $$\sum_{j=1}^{N} |{}_{(n_k)}\alpha_j - \alpha_j| \leqslant \frac{\varepsilon}{6C^2} \qquad \forall \ k \geqslant K_2$$

so when $k \geqslant K_2$

$$\| {}_{(n_k)}\underset{\sim}{\alpha} - {}_{(0)}\underset{\sim}{\alpha} \| = \sum_{j=1}^{N} |{}_{(n_k)}\alpha_j - \alpha_j| + \sum_{j=N+1}^{\infty} {}_{(n_k)}\alpha_j$$

$$\leqslant \frac{\varepsilon}{6C^2} + \frac{\varepsilon}{6C^2} \quad \text{by (4.6) and (4.7)}$$

i.e.

(4.8) $$\| {}_{(n_k)}\underset{\sim}{\alpha} - {}_{(0)}\underset{\sim}{\alpha} \| \leqslant \varepsilon/3C^2 \quad \text{for } k \geqslant K_2.$$

Now consider

$$\left\| \frac{(n_k)^{B({}_{(n_k)}\underset{\sim}{\alpha})}}{\|(n_k)^{B({}_{(n_k)}\underset{\sim}{\alpha})}\|} - \underset{\sim}{\pi} \right\| \quad \text{for } k \geqslant \max(K_1, K_2, K_3)$$

where K_3 is so large that $n_{K_3} \geqslant i_0$.

By the Δ- inequality, this is $\leqslant t_1 + t_2 + t_3$

where $$t_1 = \left\| \frac{(n_k)^{B({}_{(n_k)}\underset{\sim}{\alpha})}}{\|(n_k)^{B({}_{(n_k)}\underset{\sim}{\alpha})}\|} - \frac{(n_k)^{B({}_{(n_k)}\underset{\sim}{\alpha})}}{\|(n_k)^{B({}_{(0)}\underset{\sim}{\alpha})}\|} \right\|$$

$$t_2 = \left\| \frac{(n_k)^{B({}_{(n_k)}\underset{\sim}{\alpha})}}{\|(n_k)^{B({}_{(0)}\underset{\sim}{\alpha})}\|} - \frac{(n_k)^{B({}_{(0)}\underset{\sim}{\alpha})}}{\|(n_k)^{B({}_{(0)}\underset{\sim}{\alpha})}\|} \right\|$$

and $$t_3 = \left\| \frac{(n_k)^{B({}_{(0)}\underset{\sim}{\alpha})}}{\|(n_k)^{B({}_{(0)}\underset{\sim}{\alpha})}\|} - \underset{\sim}{\pi} \right\|$$

But $t_1 = \left\| \left(\frac{\| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha}) \| - \| (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \|}{\| (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \| \; \| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha}) \|} \right) (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \right\|$

$$\leqslant \frac{\left| \| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha}) \| - \| (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \| \right|}{\| (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \|} \cdot \frac{\| (n_k)^{B} \|}{\| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha}) \|} \cdot \| \underset{\sim}{(n_k)}^{\alpha} \|$$

$$\leqslant \frac{\| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha} - \underset{\sim}{(n_k)}^{\alpha}) \|}{\| (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \|} \cdot \frac{\| (n_k)^{B} \|}{\| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha}) \|} \cdot 1$$

$$\leqslant \frac{\| (n_k) B \|}{\| (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \|} \cdot \frac{\| (n_k)^{B} \|}{\| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha}) \|} \cdot \| \underset{\sim}{(0)}^{\alpha} - \underset{\sim}{(n_k)}^{\alpha} \|$$

$$\leqslant C \cdot C \cdot \frac{\varepsilon}{3C^2} = \varepsilon/3 \quad \text{for} \quad k \geqslant \max(K_1, K_2, K_3) \quad \text{by (4.8)}$$

and construction of C.

Similarly

$$t_2 \leqslant \frac{\| (n_k)^{B} \|}{\| (n_k)^{B}(\underset{\sim}{(0)}^{\alpha}) \|} \cdot \| \underset{\sim}{(n_k)}^{\alpha} - \underset{\sim}{(0)}^{\alpha} \|$$

$$\leqslant C \cdot \frac{\varepsilon}{3C^2} \quad \text{for} \quad k \geqslant \max(K_1, K_2, K_3)$$

$$\leqslant \varepsilon/3.$$

Finally by Theorem 3.2, since $\underset{\sim}{(0)}^{\alpha}$ has only finitely many non-zero components, there exists K_4 so large that $t_3 \leqslant \varepsilon/3$ for $k \geqslant K_4$.

$\therefore$ If $k \geqslant \max(K_1, K_2, K_3, K_4)$

$$\left\| \frac{(n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha})}{\| (n_k)^{B}(\underset{\sim}{(n_k)}^{\alpha}) \|} - \underset{\sim}{\pi} \right\| \leqslant \varepsilon$$

Since $\varepsilon > 0$ was arbitrary, (4.5) is established.

Step 2

It remains to prove that

$$a := \limsup_{n\to\infty} \left\| \frac{{}_{(n)}B({}_{(n)}\underset{\sim}{\alpha})}{\| {}_{(n)}B({}_{(n)}\underset{\sim}{\alpha}) \|} - \underset{\sim}{\pi} \right\| = 0.$$

Let $\{n_k\}_{k=1}^{\infty}$ be a subsequence of $\mathbb{N}$ such that

$$a = \lim_{k\to\infty} \left\| \frac{{}_{(n_k)}B({}_{(n_k)}\underset{\sim}{\alpha})}{\| {}_{(n_k)}B({}_{(n_k)}\underset{\sim}{\alpha}) \|} - \underset{\sim}{\pi} \right\| .$$

Since $\{{}_{(n)}\underset{\sim}{\alpha}\}_{n=1}^{\infty}$ is tight, we can find a subsequence $\{n_{k_\ell}\}_{\ell=1}^{\infty}$ of $\{n_k\}_{k=1}^{\infty}$ such that ${}_{(n_{k_\ell})}\underset{\sim}{\alpha}$ converges to a probability vector as $\ell \to \infty$.

By (4.5)

$$a = \lim_{\ell\to\infty} \left\| \frac{{}_{(n_{k_\ell})}B({}_{(n_{k_\ell})}\underset{\sim}{\alpha})}{\| {}_{(n_{k_\ell})}B({}_{(n_{k_\ell})}\underset{\sim}{\alpha}) \|} - \underset{\sim}{\pi} \right\| = 0. \qquad \square$$

---

5. APPROXIMATING THE QUASI-STATIONARY DISTRIBUTION

We consider only those infinite non-negative R-recurrent T each of whose truncations are irreducible, and denote the Perron-Frobenius eigenvalue of ${}_{(n)}T$ by $1/{}_{(n)}R$ and corresponding left probability eigenvector by ${}_{(n)}\underset{\sim}{\tau}$. We denote by $\underset{\sim}{\tau}$ a left R-invariant vector (unique to constant multiples).

In this section the notation L_{ij}, ${}_{(n)}L_{ij}$, F_{ij} ${}_{(n)}F_{ij}$ will be used with reference to the matrix T and its truncations.

Note the following generalizations of (1.1) to (1.3)

(5.1) $$L_{ij}(R) = \tau_j/\tau_i$$

(5.2) $${}_{(n)}L_{ij}({}_{(n)}R) = {}_{(n)}\tau_j/{}_{(n)}\tau_i$$

(5.3) $${}_{(n)}L_{ij}(R) \uparrow L_{ij}(R) \quad \text{as} \quad n \uparrow \infty .$$

Also

(5.4) $\quad {}_{(n)}R \downarrow R$ as $n \uparrow \infty$. (See Seneta (1981) Ch.6.)

Using (5.2) and (5.4), we obtain

(5.5) $\quad {}_{(n)}L_{ij}(R) \leq {}_{(n)}\tau_j / {}_{(n)}\tau_i \leq 1/{}_{(n)}L_{ji}(R)$.

Applying (5.1) and (5.3) we deduce in analogy to Lemma 2.1

(5.6) $\quad {}_{(n)}\tau_j / {}_{(n)}\tau_i \to \tau_j/\tau_i$ as $n \to \infty$.

Suppose now that T is irreducible R-positive and its left R-invariant vector $\underset{\sim}{\tau}$ can be, and is, probability-normed, so that $\underset{\sim}{\tau}'\underset{\sim}{1} = 1$. It is then called the quasi-stationary distribution of T (Seneta and Vere-Jones (1966)). Suppose $T = \{t_{ij}\}$ is also upper-Hessenberg i.e. $t_{ij} = 0$, $i > j + 1$. Then in place of (2.2) we have $L_{ij}(z) = {}_{(n)}L_{ij}(z)$, $|z| \leq R$, $n \geq i > j$, and we may argue from (5.5) as in the proof of Theorem 2.2 to deduce that ${}_{(n)}\underset{\sim}{\tau} \to \underset{\sim}{\tau}$ as $n \to \infty$.

Results on convergence of finite quasi-stationary distributions ${}_{(n)}\underset{\sim}{\tau}$ to $\underset{\sim}{\tau}$ are also of interest when T = P is stochastic and positive recurrent, thereby providing another means of approximating its unique stationary distribution vector $\underset{\sim}{\pi}$ (Seneta, 1981, Section 7.3). Notice that, at least formally, this method is a particular instance of linear augmentation (3.1) of ${}_{(n)}P$, for if we take ${}_{(n)}\underset{\sim}{\alpha} = {}_{(n)}\underset{\sim}{\tau}$, it is readily seen from (3.2) that ${}_{(n)}\underset{\sim}{\pi} = {}_{(n)}\underset{\sim}{\tau}$. This has been noted by Seneta (1984); and Keilson and Ramaswamy (1984) who consider a continuous-time context. The random walk analogue of their birth-death process results is covered by the above extension of Theorem 2.2.

We are able to establish the following new result, whose somewhat lengthy proof we omit.

<u>Theorem 5.1</u> Let T be a stochastic positive-recurrent renewal matrix of form (3.6). Then ${}_{(n)}\underset{\sim}{\tau} \to \underset{\sim}{\pi}$ where $\underset{\sim}{\pi}$ is its stationary distribution. □

REFERENCES

Allen, B., Anderssen, R.S., and Seneta, E. (1977) Computation of stationary measures for infinite Markov chains. *TIMS Studies in the Management Sciences*, Vol.7. *Algorithmic Methods in Probability* (M.F. Neuts, Ed.) North-Holland, Amsterdam, pp.13-23.

Golub, G.H. and Seneta, E. (1974) Computation of the stationary distribution of an infinite stochastic matrix of special form. *Bull. Austral. Math. Soc.*, **10**, 255-261.

Keilson, J. and Ramaswamy, R. (1984) Convergence of quasi-stationary distributions in birth-death processes. *Stochastic Processes and their Applications*, **18**, 301-312.

Pollak, M. and Siegmund, D. (1986) Convergence of quasi-stationary distributions for stochastically monotone Markov processes. *J. Appl. Prob.* (to appear).

Seneta, E. (1967) Finite approximations to infinite non-negative matrices. *Proc. Cambridge Philos. Soc.*, **63**, 983-992; Part II: Refinements and applications *ibid*, **64** (1968) 465-470.

Seneta, E. (1980) Computing the stationary distribution for infinite Markov chains. *Linear Algebra and its Applications*, **34**, 259-267.

Seneta, E. (1981) *Non-Negative Matrices and Markov Chains* (2nd Edn.), Springer, New York.

Seneta, E. (1984) Iterative aggregation. *Economics Letters*, **14**, 357-361.

Seneta, E. and Vere-Jones, D. (1966) On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Prob.*, **3**, 403-434.

Wolf, D. (1975) Approximation homogener Markoff-Ketten mit abzählbarem Zustandraum durch solche mit endlichem Zustandraum. In: *Proceedings in Operations Research*, **5**, Physica-Verlag, Wurzburg, pp. 137-146.

Wolf, D. (1980) Approximation of the invariant probability measure of an infinite stochastic matrix. *Adv. Appl. Prob.*, **12**, 710-726.