Adv. Appl. Prob. **20,** 228–230 (1988)
Printed in N. Ireland

LETTERS TO THE EDITOR

PERTURBATION OF THE STATIONARY DISTRIBUTION MEASURED BY ERGODICITY COEFFICIENTS

E. SENETA*, *University of Sydney*

Abstract

It is shown that an easily calculated ergodicity coefficient of a stochastic matrix P with a unique stationary distribution $\boldsymbol{\pi}^{\mathrm{T}}$, may be used to assess sensitivity of $\boldsymbol{\pi}^{\mathrm{T}}$ to perturbation of P.

MARKOV CHAINS; FINITE STOCHASTIC MATRIX; STABILITY

1. Introduction

Suppose $P = \{p_{ij}\}$ is an $(n \times n)$ stochastic matrix containing a single irreducible set of states, so that there is a unique stationary distribution vector $\boldsymbol{\pi}^{\mathrm{T}}$ ($\boldsymbol{\pi}^{\mathrm{T}}P = \boldsymbol{\pi}^{\mathrm{T}}, \boldsymbol{\pi}^{\mathrm{T}}\mathbf{1} = 1$). Let $\bar{P}$ be any other $(n \times n)$ stochastic matrix with this structure (the irreducible sets need not coincide), and $\bar{\boldsymbol{\pi}}^{\mathrm{T}}$ its unique stationary distribution. In an important early paper on the practical problem of the effect on $\boldsymbol{\pi}$ of the perturbation $E = \bar{P} - P$ to P, Schweitzer ([2], Theorem 2) showed that

$$\bar{\boldsymbol{\pi}}^{\mathrm{T}} = \boldsymbol{\pi}^{\mathrm{T}}(I - EZ)^{-1} \tag{1}$$

where $Z = (I - P + \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}})^{-1}$ is the 'fundamental matrix' of the Markov chain described by P. Thus

$$\bar{\boldsymbol{\pi}}^{\mathrm{T}} - \boldsymbol{\pi}^{\mathrm{T}} = \bar{\boldsymbol{\pi}}^{\mathrm{T}}EZ. \tag{2}$$

More recently (e.g. [1]), attention has focused on finding easily-evaluated scalar measures of the discrepancy (2). We shall show, using the ergodicity coefficient

$$\tau_1(P) = \tfrac{1}{2} \max_{i,j} \sum_{s=1}^{n} |p_{is} - p_{js}| = 1 - \min_{i,j} \sum_{s=1}^{n} \min(p_{is}, p_{js})$$

(thus $0 \leqq \tau_1(P) \leqq 1$ always), *that if* $\tau_1(P) < 1$ (i.e. *the matrix P is 'scrambling'*) *then*

$$\frac{\|\bar{\boldsymbol{\pi}}^{\mathrm{T}} - \boldsymbol{\pi}^{\mathrm{T}}\|_1 / \|\boldsymbol{\pi}^{\mathrm{T}}\|_1}{\|E\|_1 / \|P\|_1} \left(= \frac{\|\bar{\boldsymbol{\pi}}^{\mathrm{T}} - \boldsymbol{\pi}^{\mathrm{T}}\|_1}{\|E\|_1} \right) \leqq (1 - \tau_1(P))^{-1}. \tag{3}$$

Here $\|\cdot\|_p$, $1 \leqq p \leqq \infty$ is the l_p norm on the space of real row vectors R^n. For the background to ergodicity coefficients and scrambling matrices, see [4], Sections 4.3–4.4. The result (3), in providing a bound on the relative effect on $\boldsymbol{\pi}^{\mathrm{T}}$ of the perturbation E, extends to a large class of matrices the general direction of [1], and incorporates into a more general framework the sharpening [5] of Funderlic and Meyer's [1] result. This sharpening was obtained directly, without the powerful notion of ergodicity coefficient. Clearly, if $\tau_1(P)$ is not close to 1, then $\boldsymbol{\pi}$ is relatively insensitive to perturbation in P.

Received 23 July 1987.
* Postal address: Department of Mathematical Statistics, University of Sydney, NSW 2006 Australia.

2. Use of ergodicity coefficients

The coefficient of ergodicity, using the l_p norm, is defined for any stochastic P by

$$\tau_p(P) = \sup_{\substack{\|\boldsymbol{\delta}\|_p=1 \\ \boldsymbol{\delta}^{\mathrm{T}}\mathbf{1}=0}} \|\boldsymbol{\delta}^{\mathrm{T}}P\|_p$$

and as is well known satisfies $\tau_p(P^k) \leqq (\tau_p(P))^k$, integer $k \geqq 0$. Now if the single irreducible set of P is aperiodic (and this is implied by $\tau_1(P) < 1$)

$$Z = (I - Q)^{-1} = \sum_{k=0}^{\infty} Q^k$$

where $Q = P - \mathbf{1}\boldsymbol{\pi}^{\mathrm{T}}$, since Q has as its spectral radius the non-unit eigenvalue of P of largest modulus, whence $Q^k \to 0$ as $k \to \infty$. Thus, putting $\boldsymbol{\alpha}^{\mathrm{T}} = \bar{\boldsymbol{\pi}}^{\mathrm{T}}E$ from (2)

$$\|\bar{\boldsymbol{\pi}}^{\mathrm{T}} - \boldsymbol{\pi}^{\mathrm{T}}\|_p = \left\|\sum_{k=0}^{\infty} \boldsymbol{\alpha}^{\mathrm{T}}Q^k\right\|_p \leqq \sum_{k=0}^{\infty} \|\boldsymbol{\alpha}^{\mathrm{T}}Q^k\|_p \leqq \infty.$$

Now, since $\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{1} = 0$, it follows that $\boldsymbol{\alpha}^{\mathrm{T}}Q^k = \boldsymbol{\alpha}^{\mathrm{T}}P^k$, so that

$$\|\bar{\boldsymbol{\pi}}^{\mathrm{T}} - \boldsymbol{\pi}^{\mathrm{T}}\|_p \leqq \sum_{k=0}^{\infty} \tau_p(P^k)\, \|\boldsymbol{\alpha}^{\mathrm{T}}\|_p$$

$$\leqq \|\boldsymbol{\alpha}^{\mathrm{T}}\|_p \sum_{k=0}^{\infty} (\tau_p(P))^k,$$

i.e.

(4) $$\|\bar{\boldsymbol{\pi}}^{\mathrm{T}} - \boldsymbol{\pi}^{\mathrm{T}}\|_p \leqq \|\bar{\boldsymbol{\pi}}^{\mathrm{T}}\|_p\, \|E\|_p\, (1 - \tau_p(P))^{-1} \leqq \|E\|_p\, (1 - \tau_p(P))^{-1}$$

for $p \geqq 1$ if $\tau_p(P) < 1$. The result (3) with $\tau_1(P)$ now follows, since $\|\boldsymbol{\pi}^{\mathrm{T}}\|_1 = 1 = \|P\|_1$. Of course, result (4) is also useful for any τ_p with $p > 1$ providing the single irreducible set of P is aperiodic and $\tau_p(P) < 1$, if $\tau_p(P)$ can be easily evaluated in terms of the elements of P. This last is the case in particular with $\tau_\infty(P)$ [6].

3. Applications

1. The transition matrices in [1]: (i) p. 10 and (ii) p. 12 (this is an 8×8 matrix arising in radiochemistry):

(i) $\tau_1(P) = \frac{1}{2} = \tau_\infty(P)$ (ii) $\tau_1(P) = 0{\cdot}912$, $\tau_\infty(P) = 1{\cdot}140$.

Thus quick calculations on P show that $\boldsymbol{\pi}^{\mathrm{T}}$ is relatively insensitive to perturbation of P for either matrix (i) or (ii).

2. The transition matrix in [7], Example 2:

$$\tau_1(P) = 0{\cdot}375, \qquad \tau_\infty(P) = 0{\cdot}1875$$

where $\tau_\infty(P)$, optimal in the sense of being the spectral radius of P, does better than $\tau_1(P)$.

4. A related result

Recently, Schweitzer [3] has derived bounds for $\boldsymbol{\pi} - \boldsymbol{x}$ for any distribution $\boldsymbol{x}$ and shown they have certain desirable properties if $\|r(\boldsymbol{x}^{\mathrm{T}}P)\|_p < \|r(\boldsymbol{x}^{\mathrm{T}})\|_p$ whenever $\|r(\boldsymbol{x}^{\mathrm{T}})\|_p > 0$, where $r(\boldsymbol{x}^{\mathrm{T}}) = \boldsymbol{x}^{\mathrm{T}}P - \boldsymbol{x}^{\mathrm{T}}$. We point out that these desirable properties clearly hold if $p = 1$ and $\tau_1(P) < 1$, for then $r(\boldsymbol{x}^{\mathrm{T}}P) = \boldsymbol{x}^{\mathrm{T}}(P - I)P = r(\boldsymbol{x}^{\mathrm{T}})P$, and $r(\boldsymbol{x}^{\mathrm{T}})\mathbf{1} = 0$, so $\|r(\boldsymbol{x}^{\mathrm{T}}P)\|_1 \leqq \tau_1(P)\, \|r(\boldsymbol{x}^{\mathrm{T}})\|_1$. Again, other $\tau_p(P)$ may be useful.

References

[1] Funderlic, R. E. and Meyer, C. D. (1986) Sensitivity of the stationary distribution vector for an ergodic Markov chain. *Linear Algebra Appl.* **76,** 1–17.

[2] Schweitzer, P. J. (1968) Perturbation theory and finite Markov chains. *J. Appl. Prob.* **5,** 401–413.

[3] Schweitzer, P. J. (1986) Posterior bounds on the equilibrium distribution of a finite Markov chain. *Commun. Statist.-Stochastic Models* **2,** 323–338.

[4] Seneta, E. (1981) *Non-Negative Matrices and Markov Chains,* 2nd edn. Springer-Verlag, New York.

[5] Seneta, E. (1987) Sensitivity to perturbation of the stationary distribution: some refinements (submitted).

[6] Tan, C. P. (1982) A functional form for a particular coefficient of ergodicity. *J. Appl. Prob.* **19,** 858–863.

[7] Tan, C. P. (1983) Coefficients of ergodicity with respect to vector norms. *J. Appl. Prob.* **20,** 277–287.