ELSEVIER

Statistics & Probability Letters 27 (1996) 313–318

STATISTICS & PROBABILITY LETTERS

On weak lumpability of a finite Markov chain

Nan-Fu Peng

Institute of Statistics, National Chiao Tung University, Hsin Chu, Taiwan

Received December 1993; revised August 1994

Abstract

An irreducible and homogeneous Markov chain with finite state space is considered. Under a mild condition Γ on the transition probability matrix, a necessary and sufficient condition for weak lumpability, and the characterization of the set of initial starting vectors which make it lumpable are obtained. A similar result is obtained for those transition probability matrices without the restriction of condition Γ.

Keywords: Markov chain; Weak lumpability

1. Introduction

Consider an irreducible Markov chain $X(n)$, $n = 0, 1, 2, \ldots$, with a finite state space $E = \{1, 2, \ldots, N\}$, transition probability matrix $P = (P_{ij})$, $i, j \in E$, and an initial probability (row) vector $\boldsymbol{v} = (v_i)$, $i \in E$. Let $A = \{A(1), A(2), \ldots, A(M)\}$ be a partition of E where without loss of generality we assume

$$A(1) = \{1, \ldots, n(1)\}$$
$$\vdots$$
$$A(m) = \{n(1) + \cdots + n(m-1) + 1, \ldots, n(1) + \ldots + n(m)\}$$
$$\vdots$$
$$A(M) = \{n(1) + \cdots + n(M-1) + 1, \ldots, N\}.$$

With the given process X and the partition A, we associate the aggregated process Y, defined by

$$Y(n) = m \text{ iff } X(n) = A(m), \quad \forall n \geqslant 0.$$

The usefulness of discussing $Y(n)$ was stated in Hachgian (1963), Kemeny and Snell (1976) and Rubino and Sericola (1989), etc. Unfortunately, $Y(n)$ is not necessarily Markov nor even homogeneous. Conditions under which $Y(n)$ is Markov for any initial probability vectors (this property is called strong lumpability), were studied in Burke and Rosenblatt (1958), Hachgian (1963) and Kemeny and Snell (1976). Hachgian (1963) also discussed lumpability of a Markov chain with a denumerable state space. Weak lumpability first appeared in (Kemeny and Snell, 1976) which proposed that it is possible that there exists a proper subset S_μ of the set of all

SSDI 0167-7152(95)00091-7

initial probability vectors S such that the process $Y(n)$ starting with any $\alpha \in S_\mu$ is Markov homogeneous while it is not when starting with $\alpha \notin S_\mu$. In this more general situation, $X(n)$ is said to be weakly lumpable with respect to the given partition A. Kemeny and Snell (1976) also provided a local necessary and sufficient condition and a useful sufficient condition to weak lumpability. Abdel-Moneim and Leysieffer (1982) gave another but incorrect necessary and sufficient condition as was shown later with a counterexample in Rubino and Sericola (1989). The same author (Rubino and Sericola, 1991) obtained a finite characterization of weak lumpability by means of an algorithm which computes the set S_μ. Under a mild condition, a necessary and sufficient condition to weak lumpability as well as S_μ are given in the paper. Discussion on that mild condition follows.

2. Preliminaries

Most of the following definitions are similar to those of Kemeny and Snell (1976) and Rubino and Sericola (1989). For a given $\alpha \in S$, the restriction of α to $A(k)$, denoted by $\alpha^{(k)}$, is the vector of S defined by: $\alpha^{(k)}(i) = \alpha(i)/Q$ if i belongs to $A(k)$, 0 otherwise, where $Q = \sum_{j \in A(k)} \alpha(j)$ for all α and k such that $Q \neq 0$. If $Q = 0$, $\alpha^{(k)}$ is not defined. For example, suppose $N = 5$, $M = 2$, $A(1) = \{1, 2\}$ and $A(2) = \{3, 4, 5\}$. For $\alpha = (\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{12}, \frac{3}{12})$, we have $\alpha^{(1)} = (\frac{1}{2}, \frac{1}{2}, 0, 0, 0)$ and $\alpha^{(2)} = (0, 0, \frac{1}{2}, \frac{1}{8}, \frac{3}{8})$. Let U_α be the $M \times N$ matrix with the ith row $\alpha^{(i)}$, and V be the $N \times M$ matrix with the jth column a vector with 1's in the components corresponding to states in $A(j)$ and 0's otherwise. In the above example,

$$U_\alpha = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{8} & \frac{3}{8} \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

The next theorem, a combination of those in Kemeny and Snell (1976), characterizes strong lumpability.

Theorem 2.1 (Kemeny and Snell, 1976). *For a given homogeneous, irreducible finite Markov chain $X(n)$ with probability transition matrix P, the following conditions are equivalent.*
(i) *$X(n)$ is strongly lumpable with respect to a partition A.*
(ii) *The rows of PV in each $A(i)$ are the same and the (unique) probability transition matrix of the lumped process $Y(n)$ is $\hat{P} = U_\alpha PV$ where α is the vector with each entry being $1/N$.*
(iii) *$VU_\alpha PV = PV$ for some $\alpha \in S$ or, equivalently, for all $\alpha \in S$.*

It is not difficult to see that $\hat{P} = U_\alpha PV$ for all α such that U_α is defined.

Results on weak lumpability were obtained in Abdel-Moneim and Leysieffer (1982); Kemeny and Snell (1976) and Rubino and Sericola (1989, 1991). First, more notations are needed. For any $\alpha \in S$, let

$$\begin{aligned} \alpha_1 &= \alpha^{(i)} \\ \alpha_2 &= (\alpha_1 P)^{(j)} \\ &\vdots \\ \alpha_m &= (\alpha_{m-1} P)^{(s)}. \end{aligned}$$

Denote by Z_s the totality of vectors α_m obtained by considering all finite sequences $A(i)$, $A(j)$, ..., $A(s)$, ending in $A(s)$. Define also $P_\beta(X(1) \in G)$ be the probability of $X(1)$ belonging to G with starting vector β. Next theorem is the origin of the idea used in Abdel-Moneim and Leysieffer (1982) and Rubino and Sericola (1989, 1991).

Theorem 2.2 (Kemeny and Snell, 1976). *A given finite homogeneous irreducible Markov chain $X(n)$ is weakly lumpable with respect to a partition A iff for each pair of t and s in E, $P_{\beta}(X(1) \in A(t))$ is the same $\forall \beta \in Z_s$.*

Theorem 2.3 (Kemeny and Snell, 1976). *If $X(n)$ is weakly lumpable with respect to A, then $\pi \in S_\mu$ where π is the fixed vector of P. The unique transition probability matrix of Y is $\hat{P} = U_\pi PV$.*

Similarly, $\hat{P} = U_\pi PV = U_\alpha PV$ for all $\alpha \in S_\mu$ such that U_α is defined. A useful sufficient condition to weak lumpability is given next.

Theorem 2.4 (Kemeny and Snell, 1976). *$U_\pi PVU_\pi = U_\pi P$ implies weak lumpability of $X(n)$.*

See that $U_\pi PVU_\pi = U_\pi P$ if and only if $(\pi^{(i)}P)^{(j)} = \pi^{(j)}\ \forall i,j = 1, \ldots, M$. The following theorem describe the form of S_μ.

Theorem 2.5 (Rubino and Sericola, 1989). *S_μ is a convex closed set.*

Theorems 2.3 and 2.5 yield improvement on results of Rubino and Scricola (1989). Let us define the set

$$S_\pi = \left\{\lambda_1 \pi^{(1)} + \cdots + \lambda_M \pi^{(M)}:\ \sum_{i=1}^{M} \lambda_i = 1, \lambda_i \geqslant 0, \forall i\right\}.$$

If $S_\mu \neq \phi$, then $S_\pi \subseteq S_\mu$ and $\dim(S_\pi) = M$, so the number of steps needed to compute S_μ is necessarily less than or equal to $N - M$, instead of N.

3. Main results

It was shown in Burke and Rosenblatt (1958) that under the condition of reversibility of $X(n)$, the aggregated process $Y(n)$ satisfying the Chapman–Kolmogorov equations are equivalent to strong lumpability of $X(n)$. Now we want to show that under the following condition, the Chapman–Kolmogorov equations are equivalent to weak lumpability of $X(n)$. A characterization of it is also obtained.

Condition Γ. The probability transition matrix P of a finite homogeneous irreducible Markov chain satisfies condition Γ with respect to a Partition A if the columns of $V, PV, P^2V, \ldots$ generate R^N.

Theorem 3.1. *Under condition Γ, the following are equivalent:*

(i) *X is weakly lumpable.*

(ii) *Y satisfies Chapman–Kolmogorov equations when the initial distribution is π.*

(iii) *$U_\pi PVU_\pi = U_\pi P$, i.e. $(\pi^{(i)}P)^{(j)} = \pi^{(j)}\ \forall i,j = 1, \ldots, M$.*

(iv) *X is weakly lumpable and*

$$S_\mu = \left\{\lambda_1 \pi^{(1)} + \cdots + \lambda_M \pi^{(M)}:\ \sum_{i=1}^{M} \lambda_i = 1, \lambda_i \geqslant 0, \forall i\right\}.$$

Proof. Trivially, (i) implies (ii). Since $\hat{P}^k = U_\pi P^k V$, it is also not difficult to see that (ii) can be reworded in an equivalent form:

$$U_\pi P(VU_\pi - I)P^k V = 0, \quad \forall k = 1, 2, \ldots, \tag{3.1}$$

where $\boldsymbol{\pi}$ is the fixed vector of P. Since $U_\pi V = I_{M \times M}$, we also have

$$U_\pi P(VU_\pi - I)V = 0. \tag{3.2}$$

(3.1), (3.2) and condition Γ give the result of (iii), and (iii) implies (i) by Theorem 2.4.

We now show that (ii) and (iv) are equivalent. Since $\hat{P}^k = U_\pi P^k V = U_\alpha P^k V$ for $\boldsymbol{\alpha} \in S_\mu$ such that U_α is defined, (3.1) and (3.2) can be rewritten as

$$\begin{aligned} &U_\alpha P(VU_\pi - I)P^k V = 0, \quad k = 0, 1, 2, \ldots, \\ &U_\alpha P(VU_\alpha - I)P^k V = 0. \end{aligned} \tag{3.3}$$

Condition Γ combined with (3.3) immediately yields the following:

$$U_\alpha PVU_\pi = U_\alpha P = U_\alpha PVU_\alpha,$$

which implies for all i, j

$$\boldsymbol{\pi}^{(j)} = (\boldsymbol{\alpha}^{(i)} P)^{(j)} = \boldsymbol{\alpha}^{(j)}. \tag{3.4}$$

Theorem 2.5 together with (3.4) concludes the desired result. □

Checking condition Γ is not a serious matter if we use the following procedure.

1. $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_M$, the columns of V, are linearly independent of each other. Set $\gamma = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_M\}$.
2. Check whether $P\boldsymbol{v}_1$ is linearly independent of the vectors of γ. Add $P\boldsymbol{v}_1$ to γ if the answer is yes, discard $P\boldsymbol{v}_1$ if it is not.
3. Do procedure 2 for $P\boldsymbol{v}_2, \ldots, P\boldsymbol{v}_M$.
4. Do procedure 2 and 3 for P^j, $j = 2, 3, \ldots$, but only to those $\boldsymbol{v}_i$'s such that $P^{j-1}\boldsymbol{v}_i$'s were added to γ.
5. Stop when γ contains N vectors, which means Γ is satisfied, or stop at the smallest j such that all $P^j\boldsymbol{v}$'s are linearly dependent to the vectors in γ that contains vectors fewer than N, which means Γ is failed.

Remark. (1) If a stochastic matrix P is strongly lumpable, then $VUPV = PV$ by (iii) of Theorem 2.1. Multiply P^{k-1} to both sides of the above equality, we have $VUP^kV = P^kV$. Thus, the columns of P^k's are linear combinations of the columns of V. Hence, P does not satisfy condition Γ.

(2) Moreover, a stochastic matrix P with at least two identical rows in the same block $A(m)$ does not satisfy condition Γ.

(3) By the nature of P, the last column $P\boldsymbol{v}_M$ is not added to γ. Thus, the last step in procdure 3 should be actually dropped.

Some amends can be done for those P beyond the restriction of Γ. Without loss of generality, we assume the first $q = n(1) + \cdots + n(Q)$ elements of each columns of V, PV, P^2V, $\ldots$ generate R^q where $Q \leqslant M$, $q \leqslant N$.

Theorem 3.2. *Under the assumption above, X being weakly lumpable implies* $(\pi^{(i)}P)^{(j)} = \pi^{(j)}$ $i = 1, \ldots, M$ *and* $j = 1, \ldots, Q$. *Furthermore,* S_μ *is of the form*

$$S_\mu = \left\{ \lambda_1 \pi^{(1)} + \cdots + \lambda_Q \pi^{(Q)} + \lambda_{Q+1}\alpha_{Q+1} + \cdots + \lambda_M \alpha_M : \sum_{i=1}^{M} \lambda_i = 1, \lambda_i \geqslant 0, \forall i \right\}$$

with unknown α_i*'s having non-zero entries only in the ith block* $A(i)$.

Proof. Since X is weakly lumpable, we have, as in the proof of Theorem 3.1, that for all $\alpha, \beta \in S_\mu$

$$U_\alpha P(VU_\beta - I)P^K V = 0, \quad k = 0, 1, \ldots$$

Let

$$U_\pi P = \begin{pmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & & \\ a_{M1} & \cdots & a_{MM} \end{pmatrix},$$

where a_{ij} is a $1 \times n(j)$ vector.

Also let $a_i^{(j)} = (0 \cdots 0\; a_{ij}/c_{ij}\; 0 \cdots 0)$ with a_{ij} in the jth block and c_{ij} being the sum of the elements of a_{ij}, and let

$$U_x = \begin{pmatrix} \pi^{(1)} \\ \vdots \\ \pi^{(Q)} \\ a_i^{(Q+1)} \\ \vdots \\ a_i^{(M)} \end{pmatrix}.$$

Trivially, any such x belongs to S_μ.

A short computation shows the ith row of $U_\pi P(VU_x - I)$ is of the form,

$$d_1(a_i^{(1)} - \pi^{(1)}) + \cdots + d_Q(a_i^{(Q)} - \pi^{(Q)}),$$

where those d's are some constants.

Hence the assumption leads to the first part of the theorem

$$a_i^{(j)} = \pi^{(j)}, \quad j = 1, \ldots, Q.$$

The second part can be proved similarly by taking

$$U_x = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{(Q)} \\ a_i^{(Q+1)} \\ \vdots \\ a_i^{(M)} \end{pmatrix}. \qquad \square$$

Therefore the dimension needed in the calculation method of Rubino and Sericola (1991) can be reduced. Theorem 3.2 also implicitly implies that $(\alpha P)^{(j)} = \pi^{(j)}$ for all $\alpha \in S_\mu$ and $j = 1, 2, \ldots, Q$. Thus we are able to check continuously on the necessary condition $((\pi^{(i)}P)^j P)^{(k)} = \pi^{(k)}$ for $i = 1, \ldots, M$, $j = Q + 1, \ldots, M$ and $k = 1, \ldots, Q$.

Acknowledgements

I am very grateful to the referee for helpful comments and corrections.

References

Abdel-Moneim, A.M. and F.W. Leysieffer (1982), Weak lumpability in finite Markov chains, *J. Appl. Prob.* **19**, 685–691.
Burke, C.J. and M. Rosenblatt (1958), A Markovian function of a Markov chain, *Ann. Math. Statist.* **29**, 1112–1122.
Hachgian, J. (1963), Collapsed Markov chains and the Chapman–Kolmogorov equation, *Ann. Math. Statist.* **34**, 233–237.
Kemeny, J.G. and J.L. Snell (1976), *Finite Markov chains.* 2nd edn. (Springer, New York).
Rubino, G. and B. Sericola (1989), On weak lumpability in Markov chains, *J. Appl. Probab.* **26**, 446–457.
Rubino, G. and B. Sericola (1991), A finite characterization of weak lumpable Markov processes, *Stoch. Proc. Appl.* **38**, 195–204.