COLLAPSED MARKOV CHAINS AND THE CHAPMAN-KOLMOGOROV EQUATION

By Jack Hachigian

Indiana University

1. Introduction. Functions of a finite state Markov chain were considered by Burke and Rosenblatt in [2]. They obtained, as a result of these considerations, conditions under which the Chapman-Kolmogorov equation implies a process is Markovian. In [8] Rosenblatt considered functions of Markov chains in some generality but was not concerned with the Chapman-Kolmogorov equation and its implications. This paper extends some results obtained in [2] to a denumerable state space with the Chapman-Kolmogorov equation in mind. Also an example is given which shows limitations to this approach. The example is one more counter-example showing that the Chapman-Kolmogorov equation does not always imply a process is Markovian [5], [7], [10].

2. Collapsed Markov chains with any initial distribution. Let $X(t)$, $0 \leqq t < \infty$ be a Markov chain having a stationary transition probability matrix $P(t) = (p_{ij}(t); i, j, = 1, 2, \cdots)$, $P[X(t + \tau) = j \mid X(\tau) = i] = p_{ij}(t)$ with any initial distribution $w = (w_i > 0; i = 1, 2, \cdots)$. The $p_{ij}(t)$, $i, j, = 1, 2, \cdots$ are assumed to have the following properties

$$(1)\qquad \begin{aligned} 0 \leqq p_{ij}(t) \leqq 1, &\qquad \sum_j p_{ij}(t) = 1 \\ p_{ij}(t + \tau) &= \sum_{k=1} p_{ik}(t) p_{kj}(\tau) \end{aligned}$$

and $w = (w_i > 0)$ is such that $\sum_i w_i = 1$. Consider now a new process $Y(t) = f(X(t))$ (called herein the collapsed process), where f is a given function on the states $i = 1, 2, 3, \cdots$. The function f is a many-one function on the state space of $X(t)$ onto the state space of $Y(t)$. The states i of $X(t)$ on which f assumes the same value are collapsed into a single state of the $Y(t)$ process. We label the states of $Y(t)$ S_α, $\alpha = 1, 2, \cdots$, for convenience [2], [9].

Theorem 1. *Let $X(t)$, $0 \leqq t < \infty$ be a Markov chain having stationary transition mechanism $P(t) = (p_{ij}(t); i, j = 1, 2, \cdots)$ such that $\lim_{t\to 0} p_{ij}(t) = \delta_{ij}$ uniformly in i.* (Note that this is equivalent to requiring that $g_i < M < \infty$ for all i, where $g_i = \lim_{t\to 0} [1 - p_{ij}(t)/t]$, (see Doob, [3] p. 266).) *Then $Y(t) = f(X(t))$ is Markovian, whatever the initial distribution $w = (w_i > 0)$ for $X(t)$, if and only if its transition probabilities satisfy the Chapman-Kolmogorov equation.* The schema of proof follows that given by Burke and Rosenblatt in [2].

Proof. We need not consider the necessity. Assume then that $Y(t)$ satisfies the collapsed Chapman-Kolmogorov equations,

$$(2)\qquad Q_w^{(t)} Q_{wP(t)}^{(\tau)} = Q_w^{(t+\tau)}$$

Received March 19, 1962.

where $Q_w^{(t)} = (B'D_wB)^{-1}B'D_wP(t)B$, B' being the transpose of B. B is defined as $B = (b_{ij})$,

$$b_{ij} = \begin{cases} 1 & \text{for} \quad i \varepsilon S_j \\ 0 & \text{otherwise} \end{cases}$$

and $D = \text{diag}(w_i)$ [2].

We carry out the following differentiations formally; the required justifications are easily verified using the results in [1] and standard techniques. Differentiating (2) with respect to τ and evaluating at $\tau = 0$ we obtain:

$$Q_w^{(t)}(B'D_{wP(t)}B)^{-1}B'D_{wP(t)}GB = (B'D_wB)^{-1}B'D_wP(t)GB, \tag{3}$$

where G is the infinitesimal generator having elements g_i and g_{ij}. Differentiating (3) with respect to t at $t = 0$ we have:

$$B'D_wGB(B'D_wB)^{-1}B'D_wGB - (B'D_wB)^{-1}B'D_{wG}BB'D_wGB + B'D_{wG}GB = B'D_wG^2B. \tag{4}$$

Let $w_i = u_ih$ for $i \varepsilon S_\alpha$ and let $h \to 0$. The first term on the left-hand side and the term on the right-hand side of equality (4) both go to zero. The elementwise expression of the remainder is:

$$-\sum_{i \not\varepsilon S_\alpha} w_i g_{iS_\alpha} \cdot u_{S_\alpha}^{-1} \cdot \sum_{i \varepsilon S_\alpha} u_i g_{iS_\beta} + \sum_{i \not\varepsilon S_\alpha} w_i \sum_{k \varepsilon S_\alpha} g_{ik} g_{kS_\beta} = 0 \tag{5}$$

where $g_{iS_\alpha} = \sum_{j \varepsilon S_\alpha} g_{ij}$. This is valid, if and only if

$$g_{iS_\alpha} \cdot u_{S_\alpha}^{-1} \sum_{i \varepsilon S_\alpha} u_i g_{iS_\beta} = \sum_{k \varepsilon S_\alpha} g_{ik} g_{kS_\beta} \tag{6}$$

for all $i \not\varepsilon S_\alpha$. The "if" portion of this remark is obvious; the "only if" portion follows from the fact that both terms of (5) converge and (5) holds for all w_i.

Since (6) holds for all u_i we have for all $j \varepsilon S_\alpha$ and $i \not\varepsilon S_\alpha$

$$g_{iS_\alpha} g_{jS_\beta} = \sum_{k \varepsilon S_\alpha} g_{ik} g_{kS_\beta}. \tag{7}$$

Two cases must be considered

$$\text{(i) } g_{iS_\alpha} = 0 \quad \text{for all} \quad i \not\varepsilon S_\alpha$$

or

$$\text{(ii) } g_{iS_\alpha} \neq 0 \quad \text{for some} \quad i \not\varepsilon S_\alpha.$$

In the first case it is easily shown that $g_{iS}^{(v)} = 0$, $v = 0, 1, 2, \cdots$, and for all $i \not\varepsilon S_\alpha$, and hence $p_{iS_\alpha}(r) = 0$ for all $i \not\varepsilon S_\alpha$. In case (ii) we see that $g_{jS_\beta} = K_{S_\alpha, S_\beta}$ for all $j \varepsilon S_\alpha$. Again one can show that $g_{jS}^{(v)} = K_{S_\alpha, S_\beta}$ for all $j \varepsilon S_\alpha$ and $\beta = 1, 2, \cdots$, and we conclude in this case that $p_{iS_\beta}(t) = C_{S_\alpha, S_\beta}(t)$ for all $i \varepsilon S_\alpha$, $\beta = 1, 2, \cdots$.

These conditions i.e., (1°) $p_{iS_\alpha}(t) \equiv 0$ for all $i \not\varepsilon S_\alpha$ or (2°.) $p_{iS_\beta}(t) = C_{S_\alpha, S_\beta}(t)$ for every $i \varepsilon S_\alpha$ and all $\beta = 1, 2, 3, \cdots$, are sufficient to show that $Y(t)$ is Markovian. The proof of this remark is immediate; this concludes the proof of the theorem.

3. Example. In this section we show by counter example that one cannot relax the condition "$\cdots$ whatever the initial distribution of $X(t)$." in Theorem 1. To construct the example we need the following result. The ideas are based on Feller [5], Rosenblatt [10] and Levy [7].

THEOREM 2. *Let X_m, $m = 0, 1, 2, \cdots$ be a stationary, discrete parameter, denumerable state Markov chain with transition matrix $P = (p_{ij})$ and initial distribution vector $p = (p_i)$. Let $N(t)$ be a continuous parameter, denumerable state Markov chain, stochastically independent of X_m, and with a stationary transition mechanism $Q(t) = (q_{ij}(t))$ where*

$$q_{ij}(t) = \begin{cases} f(t, j - i) & i \leqq j \\ 0 & \end{cases}$$

otherwise $q = (q_i)$ is the initial vector for $N(t)$. Then $X(t) = X_{N(t)}$ is a continuous parameter Markov chain.

The proof is merely a verification of the Markov property in the form $P[X(t_1) = i_1, \cdots X(t_n) = i_n] = P[X(t_1) = i_1]P[X(t_2) = i_2 \mid X(t_1) = i_1] \cdots P[X(t_n) = i_n \mid X(t_{n-1}) = i_{n-1}]$ and is omitted.

Assume that a discrete parameter Markov chain $X_m = (Y_{m+1}, Y_m)$ $m = 0, 1, 2, \cdots$ is given where the random variables Y_m assume values $i = 0, 1, 2, \cdots, r - 1$ $(r < \infty)$. The transition probabilities are given by:

$$\begin{aligned} P[Y_{m+2} = u_2 \mid Y_{m+1} = u_1, Y_m = u_0] &= (1/r)[1 - \cos(2\pi/r)(2u_2 - u_1 - u_0)] \\ &= P[X_{m+1} = (u_2, u_1) \mid X_m = (u_1, u_0)] \end{aligned}$$

and initial distribution $P[Y_0 = u_0, Y_1 = u_1] = 1/r^2 = P[X_0 = (u_1, u_0)]$ where $u_0, u_1, u_2 = 0, 1, 2, \cdots, r - 1$. This example was constructed by Rosenblatt [10] and he has shown that X_m is stationary and persistent in [9]. Moreover Rosenblatt has shown that Y_m as a function of X_m is not Markovian and yet the one-step transition probabilities

$$P[Y(\tau) = u_\tau \mid Y(\sigma) = u_\sigma] = 1/r \qquad 1 \leqq \sigma < \tau, \quad \sigma, \tau = 0, 1, 2, \cdots$$

satisfy the Chapman-Kolmogorov equation.

Choose $N(t)$, $0 \leqq t < \infty$ to be a Poisson process, stochastically independent of X_m, with mean $\lambda = 1$. Consider the chain defined by $X_{N(t)} = X(N(t)) = [Y(N(t) + 1), Y(N(t))]$. Clearly $X_{N(t)}$ satisfies Theorem 2 by its very definition and hence must be Markovian.

$X_{N(t)} = (Y_{N(t)+1}, Y_{N(t)})$ defines the functional relation between $X_{N(t)}$ and $Y_{N(t)}$. We restrict our attention to $Y_{N(t)} = Y(t)$ and show that it is not Markovian; we will then show that the transition probabilities of $Y(t)$ satisfy the Chapman-Kolmogorov equation.

To show $Y(t)$ is not Markovian we show that

$$\begin{aligned} P[Y(\tau) = u_n \mid Y(t) = u_m, Y(s) = u_k] &\neq P[Y(\tau) = u_n \mid Y(t) = u_m], \\ & 0 \leqq s < t < \tau < \infty. \end{aligned} \tag{8}$$

Consider then

$$P[Y(\tau) = u_n, Y(t) = u_m, Y(s) = u_k]$$

$$(9) \qquad = \sum_{k,m,n} \frac{e^{-s}s^k}{k!} \frac{e^{-(t-s)}(t-s)^{m-k}}{(m-k)!} \frac{e^{-(\tau-t)}(\tau-t)^{n-m}}{(n-m)!} \cdot P[Y_k = u_k, Y_m = u_m, Y_n = u_n].$$

The computation of $P[Y_k = u_n, Y_m = u_m, Y_n = u_n]$ gives rise to seven distinct cases summarized here for brevity; k is taken equal to zero by stationarity of the Y_k process.

I: $n = m = 0$; $P[Y_0 = u_0, Y_m = u_m, Y_n = u_n] = \delta_{u_0 u_m}\delta_{u_m u_n}/r$
II: $m = 0, n \geqq 1$; $P[Y_0 = u_0, Y_m = u_m, Y_n = u_n] = \delta_{u_0 u_m}/r^2$
III: $m \geqq 1, m = n$; $P[Y_0 = u_0, Y_m = u_m, Y_n = u_n] = \delta_{u_m u_n}/r^2$
IV: $m = 1, n = 2$;
$P[Y_0 = u_0, Y_m = u_m, Y_n = u_n] = (1/r^3)[1 - \cos(2\pi/r)(2u_2 - u_1 - u_0)]$
V: $m = 2, n = 3$;
$P[Y_0 = u_0, Y_m = u_m, Y_n = u_n] = (1/r^3)[1 + \frac{1}{2}\cos(2\pi/r)(-2u_3 + 3u_2 - u_0)]$
VI: $m \geqq 1, n \geqq m + 2$; $P[Y_0 = u_0, Y_m = u_m, Y_n = u_n] = 1/r^3$
VII: $m \geqq 3, n = m + 1$; $P[Y_0 = u_0, Y_m = u_m, Y_n = u_n] = 1/r^3$

The exact expression for (9) is now

$$\begin{aligned}
P[Y(s) = u_0, Y(t) &= u_m, Y(\tau) = u_n] \\
&= (1/r)\delta_{u_0 u_m}\delta_{u_m u_n}e^{-(\tau-s)} + (1/r^2)\delta_{u_0 u_m}e^{-(\tau-s)}[1 - e^{-(\tau-t)}] \\
&+ (1/r^2)\delta_{u_m u_n}e^{-(\tau-s)}[1 - e^{-(t-s)}] \\
(10) \qquad &+ (1/r^3)[1 - \cos(2\pi/r)(2u_2 - u_1 - u_0)]e^{-(\tau-s)}(t-s)(\tau-t) \\
&+ (1/r^3)[1 + \tfrac{1}{2}\cos(2\pi/r)(-2u_3 + 3u_2 - u_0)]e^{-(\tau-s)}[(t-s)^2/2!](\tau-t) \\
&+ (1/r^3)e^{-(\tau-s)}[1 - e^{-(t-s)}][1 - e^{-(\tau-t)} - (\tau-t)e^{-(\tau-t)}] \\
&+ (1/r^3)e^{-(\tau-s)}(\tau-t)[1 - e^{-(t-s)} - (t-s)e^{-(t-s)} - [(t-s)^2/2!]e^{-(t-s)}].
\end{aligned}$$

To compute the left-hand side of (8) we evaluate $P[Y(s) = u_0, Y(t) = u_m]$ and divide it into (10):

$$(11) \qquad \begin{aligned} P[Y(s) = u_0, Y(t) = u_m] &= (1/r)e^{-t} + (1/r^2)[1 - e^{-t}] \quad &\text{if} \quad u_m = u_0 \\ &= (1/r^2)[1 - e^{-t}] \quad &\text{if} \quad u_m \neq u_0. \end{aligned}$$

The right-hand side of (8) can be computed from

$$\begin{aligned} P[Y(\tau) = u_n, Y(t) = u_m] &= (e^{-\tau}/r^2) + (1/r^2)[1 - e^{-\tau}] \quad &\text{if} \quad u_m = u_n \\ &= (1/r^2)[1 - e^{-\tau}] \quad &\text{if} \quad u_m \neq u_n \end{aligned}$$

and $P[Y(t) = u_m] = 1/r$, i.e.

$$(12) \qquad \begin{aligned} P[Y(\tau) = u_n \mid Y(t) = u_m] &= (1/r)(1 - e^{-\tau}) \quad &\text{if} \quad u_m \neq u_n \\ &= e^{-\tau} + (1/r)(1 - e^{-\tau}) \quad &\text{if} \quad u_m = u_n. \end{aligned}$$

A comparison of (12) and the ratio of (10) and (11) verifies the validity of (8); we conclude $Y(t)$ is not Markovian. However the transition mechanism of $Y(t)$ satisfies the Chapman-Kolmogorov equation.

Let $P[Y(\tau + t) = \lambda \mid Y(\tau) = v] = p_{v\lambda}(t)$, then the Chapman-Kolmogorov equation states

$$\sum_{\mu=0}^{r-1} p_{v\lambda}(s)p_{\lambda\mu}(t) = p_{v\mu}(s + t). \tag{13}$$

Consider the case when $v = \mu$, then

$$p_{vv}(s + t) = e^{-(s+t)} + (1/r^2)(1 - e^{-(s+t)}). \tag{13}$$

On the other hand for $v = \mu$,

$$\sum_{\lambda=0}^{r-1} p_{v\lambda}(s)p_{\lambda\mu}(t) = p_{vv}(t)p_{vv}(s) + \sum_{\lambda\neq\mu} (1/r^2)(1 - e^{-t})(1 - e^{-s})$$
$$= e^{-(s+t)} + (1/r)(1 - e^{-(s+t)}),$$

hence (13) is satisfied for the case $v = \mu$. A similar computation shows (13) to be satisfied for the case $v \neq \mu$.

This then is an example of a Markov chain with a specific initial distribution which is collapsed by a given function, where the transition probabilities of the collapsed chain satisfy the Chapman-Kolmogorov equation but the collapsed chain is not Markovian.

I would like to thank the referee for his careful reading of this paper and pointing out appropriate corrections.

REFERENCES

[1] Austin, D. G. (1956). Some differentiation properties of Markov transition probability functions. *Proc. Amer. Math. Soc.* **7**, No. 5 751–761.

[2] Burke, C. J. and Rosenblatt, M. (1958). A Markovian function of a Markov chain. *Ann. Math. Statist.* **29** 1112–1122.

[3] Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[4] Feller, W. (1940). On the integro-differential equations of purely discontinuous Markov processes. *Trans. Amer. Math. Soc.* **48** 488–515.

[5] Feller, W. (1959). Non-Markovian processes with the semi-group property. *Ann. Math. Statist.* **30** 1252–1253.

[6] Hachigian, J. and Rosenblatt, M. Functions of Reversible Markov Processes that are Markovian. To be published in the *J. Math. Mech.* (Nov. 1962).

[7] Lévy, P. (1949). Exemples de processus pseudo-Markovien. *C. R. Acad. Sci. Paris.* **228** 2004–2006.

[8] Rosenblatt, M. (1959). Functions of a Markov process that are Markovian. *J. Math. Mech.* **8** No. 4 585–596.

[9] Rosenblatt, M. (1962). *Random Processes*. Oxford Press.

[10] Rosenblatt, M. (1960). On aggregation problems for Markov chains. *Information and Decision Processes*, McGraw-Hill, New York 87–92.