SIMILAR MARKOV CHAINS

by

Phil Pollett

The University of Queensland

MAIN REFERENCES

Convergence of Markov transition probabilities and their spectral properties

1. Vere-Jones, D. Geometric ergodicity in denumerable Markov chains. Quart. J. Math. Oxford Ser. 2 13 (1962) 7–28.

2. Vere-Jones, D. On the spectra of some linear operators associated with queueing systems. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 2 (1963) 12–21.

3. Vere-Jones, D. Ergodic properties of nonnegative matrices. I. Pacific J. Math. 22 (1967) 361–386.

4. Vere-Jones, D. Ergodic properties of nonnegative matrices. II. Pacific J. Math. 26 (1968) 601–620.

Classification of transient Markov chains and quasi-stationary distributions

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

6. Vere-Jones, D. Some limit theorems for evanescent processes. Austral. J. Statist. 11 (1969) 67–78.

Related work

7. Vere-Jones, D.; Kendall, David G. A commutativity problem in the theory of Markov chains. Teor. Veroyatnost. i Primenen. 4 (1959) 97–100.

8. Vere-Jones, D. A rate of convergence problem in the theory of queues. Teor. Verojatnost. i Primenen. 9 (1964) 104–112.

9. Vere-Jones, D. Note on a theorem of Kingman and a theorem of Chung. Ann. Math. Statist. 37 (1966) 1844–1846.

10. Heathcote, C. R.; Seneta, E.; Vere-Jones, D. A refinement of two theorems in the theory of branching processes. Teor. Verojatnost. i Primenen. 12 (1967) 341–346.

11. Rubin, H.; Vere-Jones, D. Domains of attraction for the subcritical Galton-Watson branching process. J. Appl. Probability 5 (1968) 216–219.

12. Seneta, E.; Vere-Jones, D. On the asymptotic behaviour of subcritical branching processes with continuous state space. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 10 (1968) 212–225.

13. Fahady, K. S.; Quine, M. P.; Vere-Jones, D. Heavy traffic approximations for the Galton-Watson process. Advances in Appl. Probability 3 (1971) 282–300.

14. Pollett, P. K.; Vere-Jones, D. A note on evanescent processes. Austral. J. Statist. 34 (1992), no. 3, 531–536.

Important early work on quasi-stationary distributions

Yaglom, A.M. Certain limit theorems of the theory of branching processes (Russian) Doklady Akad. Nauk SSSR (N.S.) 56 (1947) 795–798.

Bartlett, M.S. Deterministic and stochastic models for recurrent epidemics. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. IV, pp. 81–109. University of California Press, Berkeley and Los Angeles, 1956.

Bartlett, M.S. Stochastic population models in ecology and epidemiology. Methuen's Monographs on Applied Probability and Statistics Methuen & Co., Ltd., London; John Wiley & Sons, Inc., New York, 1960.

Darroch, J. N.; Seneta, E. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. J. Appl. Probability 2 (1965) 88–100.

Darroch, J. N.; Seneta, E. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. J. Appl. Probability 4 (1967) 192–196.

Important early work on quasi-stationary distributions

Mandl, Petr Sur le comportement asymptotique des probabilités dans les ensembles des états d'une chaîne de Markov homogène (Russian) Časopis Pěst. Mat. 84 (1959) 140–149.

Mandl, Petr On the asymptotic behaviour of probabilities within groups of states of a homogeneous Markov process (Czech) Časopis Pěst. Mat. 85 (1960) 448–456.

Ewens, W.J. The diffusion equation and a pseudo-distribution in genetics. J. Roy. Statist. Soc., Ser B 25 (1963) 405–412.

Kingman, J.F.C. The exponential decay of Markov transition probabilities. Proc. London Math. Soc. 13 (1963) 337–358.

Ewens, W.J. The pseudo-transient distribution and its uses in genetics. J. Appl. Probab. 1 (1964) 141–156.

Seneta, E. Quasi-stationary distributions and time-reversion in genetics. (With discussion) J. Roy. Statist. Soc. Ser. B 28 (1966) 253–277.

Seneta, E. Quasi-stationary behaviour in the random walk with continuous time. Austral. J. Statist. 8 (1966) 92–98.

DISCRETE-TIME CHAINS

Setting: $\{X_n, n = 0, 1, \dots\}$, a time-homogeneous Markov chain taking values in a countable set S with transition probabilities

$$p_{ij}^{(n)} = \text{Pr}(X_{m+n} = j | X_m = i), \quad i, j \in S.$$

Let C be any irreducible and (for simplicity) aperiodic class.

DVJ1: For $i \in C$, $\{p_{ii}^{(n)}\}^{1/n} \to \rho$ as $n \to \infty$. The limit ρ does not depend on i and it satisfies $0 < \rho \leq 1$. Moreover, $p_{ii}^{(n)} \leq \rho^n$ and indeed, for $i, j \in C$, $p_{ij}^{(n)} \leq M_{ij}\rho^n$, where $M_{ij} < \infty$.

(If C is recurrent, $\sum_n p_{ii}^{(n)} = \infty$ implies $\rho = 1$. When C is transient, we can have $\rho = 1$, or, $\rho < 1$, which is called *geometric ergodicity*.)

DVJ2: For any real $r > 0$, the series $\sum_n p_{ij}^{(n)} r^n$, $i, j \in C$, converge or diverge together; in particular, they have the same radius of convergence R, and $R = 1/\rho$. And, all or none of the sequences $\{p_{ij}^{(n)} r^n\}$ tend to zero.

TRANSIENT CHAINS

The key to unlocking this "quasi-stationarity" is to examine the behaviour of the transition probabilities *at* the radius of convergence R.

Suppose that C is transient class which is geometrically ergodic ($\rho < 1$, $R > 1$). Although $p_{ij}^{(n)} \to 0$, it might be true that $p_{ij}^{(n)} R^n \to m_{ij}$, where $m_{ij} > 0$. How does this help?

For $i, j \in C$,

$$\begin{aligned} &\Pr(X_n = j | X_n \in C, X_0 = i) \\ &\qquad = \frac{\Pr(X_n = j | X_0 = i)}{\Pr(X_n \in C | X_0 = i)} = \frac{p_{ij}^{(n)}}{\sum_{k \in C} p_{ik}^{(n)}} \\ &\qquad\qquad = \frac{p_{ij}^{(n)} R^n}{\sum_{k \in C} p_{ik}^{(n)} R^n} \to \frac{m_{ij}}{\sum_{k \in C} m_{ik}}, \end{aligned}$$

provided that we can justify taking limit under summation.

DVJ3: C is said to be R-*transient* or R-*recurrent* according as $\sum_n p_{ij}^{(n)} R^n$ converges or diverges. If C is R-recurrent, then it is said to be R-*positive* or R-*null* according to whether the limit of $p_{ij}^{(n)} R^n$ is positive or zero.

DVJ4: If C is R-recurrent, then, for $i \in C$, the inequalities

$$\sum_{i \in C} m_i p_{ij}^{(n)} \leq m_j \rho^n \qquad \sum_{i \in C} p_{ji}^{(n)} x_i \leq x_j \rho^n$$

have unique positive solutions $\{m_j\}$ and $\{x_j\}$ and indeed they are eigenvectors:

$$\sum_{i \in C} m_i p_{ij}^{(n)} = m_j \rho^n \qquad \sum_{i \in C} p_{ji}^{(n)} x_i = x_j \rho^n.$$

C is then R-positive recurrent *if and only if* $\sum_{k \in C} m_k x_k < \infty$, in which case

$$p_{ij}^{(n)} R^n \to \frac{x_i m_j}{\sum_{k \in C} x_k m_k},$$

and, if $\sum_k m_k < \infty$, then

$$\lim_{n \to \infty} \sum_{k \in C} p_{ik}^{(n)} R^n = \sum_{k \in C} \lim_{n \to \infty} p_{ik}^{(n)} R^n = x_i \sum_{k \in C} m_k \, .$$

AND FINALLY

S-DVJ: If C is R-positive recurrent and the left-eigenvector satisfies $\sum_k m_k < \infty$, then the *limiting conditional* (or *quasi-stationary*) distribution exists: as $n \to \infty$,

$$\Pr(X_n = j | X_n \in C, X_0 = i) \to \frac{m_j}{\sum_{k \in C} m_k}.$$

... AND MUCH MORE

Other kinds of QSD, more general and more precise statements, continuous-time chains, general state spaces, numerical methods and in particular truncation methods, MCMC, countless applications of QSDs: chemical kinetics, population biology, ecology, epidemiology, reliability, telecommunications. A full bibliography is maintained at my web site:

`http://www.maths.uq.edu.au/~pkp/research.html`

SIMILAR MARKOV CHAINS

New setting: $(X_t, t \geq 0)$, a time-homogeneous Markov chain in *continuous time* taking values in a countable set S, with transition function $P = (p_{ij}(t))$, where

$$p_{ij}(t) = \text{Pr}(X_{s+t} = j | X_s = i), \quad i, j \in S.$$

Assuming that $p_{ij}(0+) = \delta_{ij}$ (*standard*), the transitions rates are defined by $q_{ij} = p'_{ij}(0+)$. Set $q_i = -q_{ii}$ and assume $q_i < \infty$ (*stable*).

Definition: Two such chains X and $\tilde{X}$ are said to be *similar* if their transition functions, P and $\tilde{P}$, satisfy $\tilde{p}_{ij}(t) = c_{ij} p_{ij}(t)$, $i, j \in S$, $t > 0$, for some collection of positive constants c_{ij}, $i, j \in S$.

Immediate consequences of the definition: Since both chains are standard, $c_{ii} = 1$ and the transition rates must satisfy $\tilde{q}_{ij} = c_{ij} q_{ij}$, in particular, $\tilde{q}_i = q_i$. They share the same irreducible classes and the same classification of states.

Birth-death chains: Lenin et al.* proved that for birth-death chains the "similarity constants" must *factorize* as $c_{ij} = \alpha_i \beta_j$. (Note that $c_{ij} = \beta_j/\beta_i$, since $c_{ii} = 1$.)

Is this true more generally?

Definition: Let C be a subset of S. Two chains are said to be *strongly similar* over C if $\tilde{p}_{ij}(t) = p_{ij}(t)\beta_j/\beta_i$, $i, j \in C$, $t > 0$, for some collection of positive constants β_j, $j \in C$.

Proposition: If C is recurrent, then $c_{ij} = 1$.

(**Proof:** $\tilde{f}_{ij} = c_{ij} f_{ij}$.)

*Lenin, R., Parthasarathy, P., Scheinhardt, W. and van Doorn, E. (2000) Families of birth-death processes with similar time-dependent behaviour. *J. Appl. Probab.* **37,** 835–849.

EXTENSION OF DJV THEORY

Kingman: If C is irreducible, then, for each $i, j \in C$, $-t^{-1} \log p_{ij}(t) \to \lambda$ (≥ 0), $p_{ij}(t) \leq M_{ij} e^{-\lambda t}$, for some $M_{ij} < \infty$, et cetera.

Definition: C is said to be λ-*transient* or λ-*recurrent* according as $\int_0^\infty p_{ij}(t) e^{\lambda t}\, dt$ converges or diverges. If C is λ-recurrent, then it is said to be λ-*positive* or λ-*null* according to whether the limit of $p_{ij}(t)e^{\lambda t}$ is positive or zero.

Theorem: If C is λ-recurrent, then, for $i \in C$, the inequalities

$$\sum_{i \in C} m_i p_{ij}(t) \leq e^{-\lambda t} m_j \qquad \sum_{i \in C} p_{ji}(t) x_i \leq e^{-\lambda t} x_j$$

have unique positive solutions $\{m_j\}$ and $\{x_j\}$ and indeed they are eigenvectors:

$$\sum_{i \in C} m_i p_{ij}(t) = e^{-\lambda t} m_j \qquad \sum_{i \in C} p_{ji}(t) x_i = e^{-\lambda t} x_j.$$

C is then λ-positive recurrent *if and only if* $\sum_{k \in C} m_k x_k < \infty$, in which case

$$p_{ij}(t) e^{\lambda t} \to \frac{x_i m_j}{\sum_{k \in C} x_k m_k}.$$

Suppose that P and $\tilde{P}$ are similar. They share the same λ and the same "λ-classification".

Theorem: If C is a λ-positive recurrent class, then P and $\tilde{P}$ are strongly similar over C. We may take $\beta_j = \tilde{m}_j / m_j$, where $\{m_j\}$ and $\{\tilde{m}_j\}$ are the essentially unique λ-invariant measures (left eigenvectors) on C for P and for $\tilde{P}$, respectively.

Proof: Let $t \to \infty$ in $\tilde{p}_{ij}(t)e^{\lambda t} = c_{ij} p_{ij}(t) e^{\lambda t}$. We get, in an obvious notation,

$$c_{ij} = E \frac{\tilde{x}_i}{x_i} \frac{\tilde{m}_j}{m_j}, \qquad E = \sum_{i \in C} m_i x_i \Bigg/ \sum_{i \in C} \tilde{m}_i \tilde{x}_i,$$

and, since $c_{ii} = 1$, we have $E \tilde{x}_i \tilde{m}_i = x_i m_i$.

Again: Are similar chains always strongly similar?

In the λ-null recurrent case, it may still be possible to deduce the desired factorization, for, although $e^{\lambda t}p_{ij}(t) \to 0$, it may be possible to find a $\kappa > 0$ such that $t^{\kappa}e^{\lambda t}p_{ij}(t)$ tends to a strictly positive limit. (Similar chains will have the same κ.)

Lemma: Assume that C is λ-null recurrent and suppose that there is a $\kappa > 0$, which does not depend on i and j, such that $t^{\kappa}e^{\lambda t}p_{ij}(t)$ tends to a strictly positive limit π_{ij} for all $i, j \in C$. Then, there is a positive constant d such that $\pi_{ij} = dx_i m_i$, $i, j \in C$, where $\{m_j\}$ and $\{x_j\}$ are, respectively, the essentially unique λ-invariant measure and vector (left- and right-eigenvectors) on C for P.

Remark: Even in the λ-transient case it might still be possible to find a $\kappa > 0$ such that $t^{\kappa}e^{\lambda t}p_{ij}(t)$ tends to a positive limit, and for the conclusions to the lemma to hold good. (Note that, by the usual irreducibility arguments, κ will be the same for all i and j in any given class.)