QUASISTATIONARITY IN MARKOVIAN MODELS

by

Andrew Hart, The University of Queensland
Phil Pollett, The University of Queensland
Pauline Schrijner, The University of Durham

MARKOV CHAINS WITH POSITIVE DRIFT

Let $(X_t, t \geq 0)$ be a continuous time Markov chain over a denumerable state space S. For simplicity take $S = \{0, 1, \dots\}$.

Let $Q = (q_{ij}, i, j \in S)$ be the matrix of transition rates (q-matrix), assumed to be *stable* and *conservative*, so that $q_{ij}\,(\geq 0)$, for $j \neq i$, represents the transition rate from state i to state j and $q_{ii} = -q_i$, where $q_i = \sum_{j \neq i} q_{ij}\,(< \infty)$ represents the transition rate out of state i.

Let $P(t) = (p_{ij}(t), i, j \in S)$, where

$$p_{ij}(t) = \Pr(X_t = j | X_0 = i),$$

and suppose that ${p_{ij}}'(0+) = q_{ij}$. It will not be necessary to assume that the transition function P is determined uniquely by Q.

We assume only that P is *irreducible*, that is, $p_{ij}(t) > 0$ for all $i, j \in S$, and *transient*, that is,

$$\int_0^\infty p_{ij}(t)\,dt < \infty, \qquad i, j \in S.$$

LIMITING CONDITIONAL DISTRIBUTIONS

We shall be concerned with the existence of a *limiting conditional distribution* (LCD) for the process:

$$\lim_{t\to\infty} \Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0).$$

The conditional probability can be evaluated as follows:

$$\Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0) \\ = \frac{p_{ij}(t) b_j}{\sum_{k \in S} p_{ik}(t) b_k},$$

where

$$b_i = \Pr(X_t = 0 \text{ for some } t > 0 | X_0 = i).$$

Note that $b_0 = 1$ and (by irreducibility and transience) $0 < b_i < 1$, $i \in S$.

LCDs for ABSORBING CHAINS

Compare the above definition with the widely studied LCDs for *absorbing chains*:

Append to S an absorbing state (or, more generally, an *absorbing set*) θ such that $p_{i\theta}(t) > 0$ for some (and then all) $i \in S$ and $t > 0$.

In this context we study the limit (as $t \to \infty$) of

$$\Pr(X_t = j | X_0 = i, X_t \in S) = \frac{p_{ij}(t)}{\sum_{k \in S} p_{ik}(t)}, \quad i, j \in S,$$

or, more generally (when a_j, the probability of absorption starting in j, is *less than* 1),

$$\Pr(X_t = j | X_0 = i, X_t \in S, X_{t+s} = \theta \text{ for some } s) = \frac{p_{ij}(t) a_j}{\sum_{k \in S} p_{ik}(t) a_k}, \quad i, j \in S.$$

HITTING PROBABILITIES

Let $b_i(t) = \Pr(X_s = 0 \text{ for some } s > t | X_0 = i)$, so that $b_i(0) = b_i$. This can be evaluated as

$$b_i(t) = \frac{\int_t^\infty p_{i0}(u)du}{\int_0^\infty p_{00}(u)du}, \qquad i \in S, t \geq 0.$$

Claim. $b_i(t) = \sum_{j \in S} p_{ij}(t) b_j$, $i \in S$, $t \geq 0$.

It follows that $b = (b_i, i \in S)$ is a *subinvariant vector for* P; in particular,

$$\sum_{j \in S} p_{ij}(t) b_j = b_i - \frac{1}{u_0} \int_0^t p_{i0}(u)du \ (< b_i \text{ for } t > 0),$$

where $u_0 = \int_0^\infty p_{00}(u)du \ (< \infty)$.

Claim. b is a *subinvariant vector for* Q; in fact, b is 'almost invariant':

$$\sum_{j \in S} q_{ij} b_j = \begin{cases} 0, & \text{if } i \geq 1, \\ -\frac{1}{u_0}, & \text{if } i = 0. \end{cases}$$

CONSTRUCTION OF THE DUAL

Define $\bar{P}(t) = (\bar{p}_{ij}(t), i, j \in S)$ and $\bar{Q} = (\bar{q}_{ij}, i, j \in S)$ by

$$\bar{p}_{ij}(t) = \frac{p_{ij}(t) b_j}{b_i}, \qquad i, j \in S,$$

and

$$\bar{q}_{ij} = \frac{q_{ij} b_j}{b_i}, \qquad i, j \in S.$$

Clearly $\bar{P}$ is a (standard) transition function and $\bar{Q}$ is a q-matrix such that ${\bar{p}_{ij}}'(0+) = \bar{q}_{ij}$. However, P is *dishonest* with

$$\sum_{j \in S} \bar{p}_{ij}(t) = \frac{b_i(t)}{b_i} \ (< 1 \text{ for } t > 0),$$

and $\bar{Q}$ is non-conservative with

$$\sum_{j \in S} \bar{q}_{ij} = \begin{cases} 0, & \text{if } i \geq 1, \\ -\frac{1}{u_0}, & \text{if } i = 0. \end{cases}$$

(Recall that $b_0 = 1$.)

CONSTRUCTION OF THE DUAL

Append to S an absorbing state θ and extend the definition of $\bar{P}$ and $\bar{Q}$ to $\bar{S} = S \cup \{\theta\}$ as follows:

$$\bar{p}_{\theta j}(t) = \delta_{\theta j}, \quad j \in \bar{S},$$
$$\bar{p}_{i\theta}(t) = 1 - \frac{b_i(t)}{b_i}, \quad i \in S,$$

and

$$\bar{q}_{\theta j} = 0, \quad j \in \bar{S},$$
$$\bar{q}_{i\theta} = \frac{\delta_{i0}}{u_0}, \quad i \in S.$$

In this way $\bar{Q}$ will be conservative and $\bar{P}$ will be honest.

$\bar{P}$ is a $\bar{Q}$-transition function since, for $i \in S$,

$$\begin{aligned}
\lim_{t \downarrow 0} \frac{\bar{p}_{i\theta}(t)}{t} &= \lim_{t \downarrow 0} \frac{1}{t}\left(1 - \frac{b_i(t)}{b_i}\right) \\
&= \frac{1}{b_i u_0} \lim_{t \downarrow 0} \frac{1}{t} \int_0^t p_{i0}(u)\,du \\
&= \frac{1}{b_i u_0} \lim_{t \downarrow 0} p_{i0}(t) = \frac{\delta_{i0}}{b_i u_0} = \bar{q}_{i\theta}.
\end{aligned}$$

INTERPRETATION OF THE DUAL

Let L be the last exit time of state 0, that is,

$$L(\omega) = \sup\{t \geq 0 : X_t(\omega) = 0\},$$

and define $(\bar{X}_t, t \geq 0)$ by

$$\bar{X}_t(\omega) = \begin{cases} X_t(\omega), & \text{if } t < L(\omega), \\ \theta, & \text{if } t \geq L(\omega); \end{cases}$$

$\bar{X}_t$ is X_t *killed* at time L. Then*, since $b_i = \Pr(L > 0 | X_0 = i)$, $\bar{X}_t$ is Markovian with transition function $\bar{P}$, and, if $X_0 = i$, its initial distribution is given by

$$\Pr(\bar{X}_0 = j) = \begin{cases} b_i, & \text{if } j = i, \\ 1 - b_i, & \text{if } j = \theta, \\ 0, & \text{otherwise.} \end{cases}$$

Notice that $\bar{X}_t$ is absorbed with probability 1:

$$\bar{a}_i := \lim_{t\to\infty} \bar{p}_{i\theta}(t) = \lim_{t\to\infty} \left(1 - \frac{b_i(t)}{b_i}\right) = 1, \quad i \in S.$$

*Meyer, P.A., Smithe, R.T., and Walsh, J.B. (1971) Birth and death of Markov processes, *Proc. 6th Berkeley Symposium*, III, pp. 295–305, University of California Press.

EXISTENCE OF LCDs

By the definition of $\bar{P}$, we see immediately that, for every $i, j \in S$,

$$\begin{aligned}&\Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0)\\ &\quad= \frac{p_{ij}(t) b_j}{\sum_{k \in S} p_{ik}(t) b_k} = \frac{\bar{p}_{ij}(t)}{\sum_{k \in S} \bar{p}_{ik}(t)}.\end{aligned}$$

> *Thus, the existence or otherwise of an LCD can be established via the dual transition function using results for absorbing chains.*

Here is one such result for absorbing chains:

Theorem. Suppose S is λ-positive recurrent and let $\bar{m} = (\bar{m}_j, j \in S)$ be the (essentially unique) λ-invariant measure for $\bar{P}$. Then, if $\sum_{j \in S} \bar{m}_j < \infty$,

$$\lim_{t \to \infty} \frac{\bar{p}_{ij}(t)}{\sum_{k \in S} \bar{p}_{ik}(t)} = \frac{\bar{m}_j}{\sum_{k \in S} \bar{m}_k}, \quad i, j \in S.$$

λ-CLASSIFICATION

$\bar{P}$ and P have the *same* λ-classification.

The *decay parameter*:

$$\bar{\lambda} = \lim_{t\to\infty} -\frac{1}{t}\log \bar{p}_{ij}(t) = \lim_{t\to\infty} -\frac{1}{t}\log p_{ij}(t) = \lambda.$$

Note that λ (≥ 0) is the same for each $i, j \in S$ and $p_{ij}(t) = O(e^{-\lambda t})$, $\bar{p}_{ij}(t) = O(e^{-\lambda t})$.

Geometric ergodicity: $\bar{\lambda} > 0 \Leftrightarrow \lambda > 0$.

λ-transience/recurrence: for $i, j \in S$,

$$\int_0^\infty e^{\lambda t}\bar{p}_{ij}(t)dt \begin{cases} < \infty \\ = \infty \end{cases} \Leftrightarrow \int_0^\infty e^{\lambda t}p_{ij}(t)dt \begin{cases} < \infty \\ = \infty \end{cases}$$

λ-positive/null recurrence: for $i, j \in S$,

$$\lim_{t\to\infty} e^{\lambda t}\bar{p}_{ij}(t) \begin{cases} > 0 \\ = 0 \end{cases} \Leftrightarrow \lim_{t\to\infty} e^{\lambda t}p_{ij}(t) \begin{cases} > 0 \\ = 0 \end{cases}$$

λ-INVARIANT MEASURES FOR P

Definition. A collection of positive numbers $m = (m_j, j \in S)$ is called a λ-subinvariant measure for P if, for all $t \geq 0$,

$$\sum_{i \in S} m_i p_{ij}(t) \leq e^{-\lambda t} m_j, \quad j \in S,$$

and λ-invariant if, for all $t \geq 0$,

$$\sum_{i \in S} m_i p_{ij}(t) = e^{-\lambda t} m_j, \quad j \in S.$$

Now, for any given collection $m = (m_j, j \in S)$, define $\bar{m} = (\bar{m}_j, j \in S)$ by $\bar{m}_j = m_j b_j$.

Claim. m is λ-(sub)invariant for P if and only if $\bar{m}$ is λ-(sub)invariant for $\bar{P}$.

Claim. If m is a λ-subinvariant measure for P, then $\sum_{i \in S} m_i b_i < \infty$ if $\lambda > 0$. If m is λ-invariant for P, then $\sum_{i \in S} m_i b_i < \infty$ *only if* $\lambda > 0$.

Thus, the usual premise, $\sum_{i \in S} \bar{m}_i < \infty$, that $\bar{m}$ be a finite λ-invariant measure is always satisfied in the present context of irreducible transient chains.

λ-INVARIANT MEASURES FOR Q

Definition. A collection of positive numbers $m = (m_j, j \in S)$ is called a λ-invariant measure for Q if, for all $t \geq 0$,

$$\sum_{i \in S} m_i q_{ij} = -\lambda m_j, \qquad j \in S.$$

Again we have that m is λ-invariant for Q if and only if $\bar{m}$ is λ-invariant for $\bar{Q}$.

> *Now suppose that Q is regular, so that P, now being the unique Q-transition function, is honest.*

We can then apply to $\bar{Q}$ the following result for absorbing chains (with a transient class S and absorbing state θ):

Theorem. If $\bar{m} = (\bar{m}_j, j \in S)$ is a *finite* λ-invariant measure for $\bar{Q}$, then it is λ-invariant for $\bar{P}$ if and only if

$$\lambda \sum_{i \in S} \bar{m}_i = \sum_{i \in S} \bar{m}_i \bar{q}_{i\theta}.$$

LCDs WHEN Q IS REGULAR

Theorem. Let $\lambda > 0$ and suppose that $m = (m_j, j \in S)$ is a λ-invariant measure for Q. Then, $\sum_{j\in S} m_j b_j < \infty$, and m is λ-invariant for P if (and only if)

$$\sum_{j\in S} m_j b_j = \frac{m_0}{\lambda u_0},$$

whence if S is λ-positive the LCD exists:

$$\lim_{t\to\infty} \Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0)$$
$$= \frac{m_j b_j}{\sum_{k\in S} m_k b_k}, \quad i, j \in S.$$

Other examples:

1. Kesten's bounded jump conditions*.
2. Asymptotic remoteness†.
3. Results for special processes: birth-death processes, branching processes, catastrophe processes.

*Kesten, H. (1995) A ratio limit theorem for sub Markov chains on $\{1,2,3\ldots\}$ with bounded jumps, *Adv. Appl. Probab.* **27**, 652–691.

†Ferrari, P.A., Kesten, H., Martínez, S. and Picco, P. (1995) Existence of quasi-stationary distributions. A renewal dynamic approach, *Ann. Probab.* **23**, 501–521.

BIRTH-DEATH PROCESSES

Consider an irreducible birth-death process on $S = \{0, 1, \dots\}$. Its q-matrix is given by

$$q_{ij} = \begin{cases} \lambda_i, & \text{if } j = i+1, \\ -(\lambda_i + \mu_i), & \text{if } j = i, \\ \mu_i, & \text{if } j = i-1, \\ 0, & \text{otherwise,} \end{cases}$$

where $\lambda_i > 0$ $(i \geq 0)$, $\mu_i > 0$ $(i \geq 1)$ and $\mu_0 = 0$.

Define series A and C by

$$A = \sum_{i=0}^{\infty} \frac{1}{\lambda_i \pi_i} \quad \text{and} \quad C = \sum_{i=0}^{\infty} \frac{1}{\lambda_i \pi_i} \sum_{j=0}^{i} \pi_j \,,$$

where $\pi_0 = 1$ and, for $i \geq 1$,

$$\pi_i = \prod_{k=1}^{i} \frac{\lambda_{k-1}}{\mu_k} \,.$$

Assume that $C = \infty$ (Q regular) and that $A < \infty$ (S transient). Then, the hitting probabilities are given by

$$b_i = \frac{1}{A} \sum_{k=i}^{\infty} \frac{1}{\lambda_k \pi_k} \,.$$

Define polynomials $(\phi_i(\cdot),\ i \in S)$, where $\phi_i : \mathrm{R} \to \mathrm{R}$, by $\phi_0(x) = 0$, $\lambda_0\phi_1(x) = \lambda_0 + \mu_0 - x$ and, for $i \geq 1$, $\lambda_i\phi_{i+1}(x) = (\lambda_i + \mu_i - x)\phi_i(x) - \mu_i\phi_{i-1}(x)$, and let $m_i = \pi_i\phi_i(\lambda)$, $i \in S$, where λ is the decay parameter of S. Then, m is the essentially unique λ-invariant measure.

The dual is an absorbing birth-death process; we get (in an obvious notation):

$$\begin{gathered}
\bar{\lambda}_i = \lambda_i b_{i+1}/b_i,\ i \geq 0,\ \lambda_\theta = 0, \\
\bar{\mu}_i = \mu_i b_{i-1}/b_i,\ i \geq 1,\ \bar{\mu}_0\ (= \bar{q}_{0\theta})\ = \lambda_0(1 - b_1), \\
\bar{\pi}_i = b_i^2\pi_i,\ i \geq 0,\ \bar{\pi}_\theta = 1, \\
\bar{\phi}_i(x) = \phi_i(x)/b_i,\ i \geq 0,\ \phi_\theta(x) = 0, \\
\bar{m}_i = \bar{\pi}_i\bar{\phi}_i(\lambda) = b_i\pi_i\phi_i(\lambda),\ i \geq 0.
\end{gathered}$$

A direct application of van Doorn's* result for absorbing birth-death processes shows that the LCD exists if and only if $\lambda > 0$:

$$\begin{aligned}
&\lim_{t\to\infty} \Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0) \\
&\qquad = \frac{b_i\pi_i\phi_i(\lambda)}{\sum_{k\in S} b_k\pi_k\phi_k(\lambda)}, \quad i, j \in S.
\end{aligned}$$

*Van Doorn, E.A (1991) Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes, *Adv. Appl. Probab.* **23**, 683–700.

AN EXAMPLE

Let $f : [0, \infty) \to [0, \infty)$, $V > 0$ and $0 < \alpha < 1$, and define $\lambda_0 = q_0$,

$$\lambda_n = \alpha q_n, \;\; \mu_n = (1-\alpha) q_n, \quad n \geq 1,$$

where $q_n = V f(n/V)$.

This defines a density-dependent birth-death process with constant jump probabilities: think of V as "area" and n/V as "population density". The idea is that, as V gets large, the density process $x_t = X_t/V$ becomes "more deterministic":

$$\frac{dx}{dt} = \alpha f(x) - (1-\alpha) f(x).$$

Clearly the process is irreducible if $f(n) > 0$ for all $n \geq 0$, and then transient if (and only if) $\alpha > 1/2$, in which case

$$b_i = \left(\frac{1-\alpha}{\alpha}\right)^i, \quad i \geq 0.$$

Two choices for f:

$$f(x) = \frac{(\pi - x)^2}{(\pi + x)^3}, \quad f(x) = \frac{(\pi - x)^2 (2\pi - x)^2}{(\pi + x)^5}.$$

Each has $f(x) \to 0$ and $x \to \infty$.