ANALYTICAL AND COMPUTATIONAL METHODS FOR MODELLING THE LONG-TERM BEHAVIOUR OF EVANESCENT RANDOM PROCESSES

by

Phil Pollett

Department of Mathematics
The University of Queensland

AN AUTO-CATALYTIC REACTION

Consider the following reaction scheme:

$$A \stackrel{X}{\to} B,$$

where X is a catalyst. Suppose that there are two stages, namely

$$A + X \stackrel{k_1}{\to} 2X \quad \text{and} \quad 2X \stackrel{k_2}{\to} B.$$

Let $X(t) =$ number of X molecules at time t.

Suppose that the concentration of A is held constant; let a be the number of molecules of A. The state space is $S = \{0, 1, 2, \ldots\}$ and the transition rates, $Q = (q_{xy},\ x, y \in S)$, are given by

$$q_{xy} = \begin{cases} k_1 a x & \text{if } y = x + 1, \\ k_2 \binom{x}{2} & \text{if } y = x - 2, \\ 0 & \text{otherwise.} \end{cases}$$

AN EPIDEMIC MODEL
(Ridler-Rowe)

The state at time t:

$X(t)$ - number of susceptibles

$Y(t)$ - number of infectives

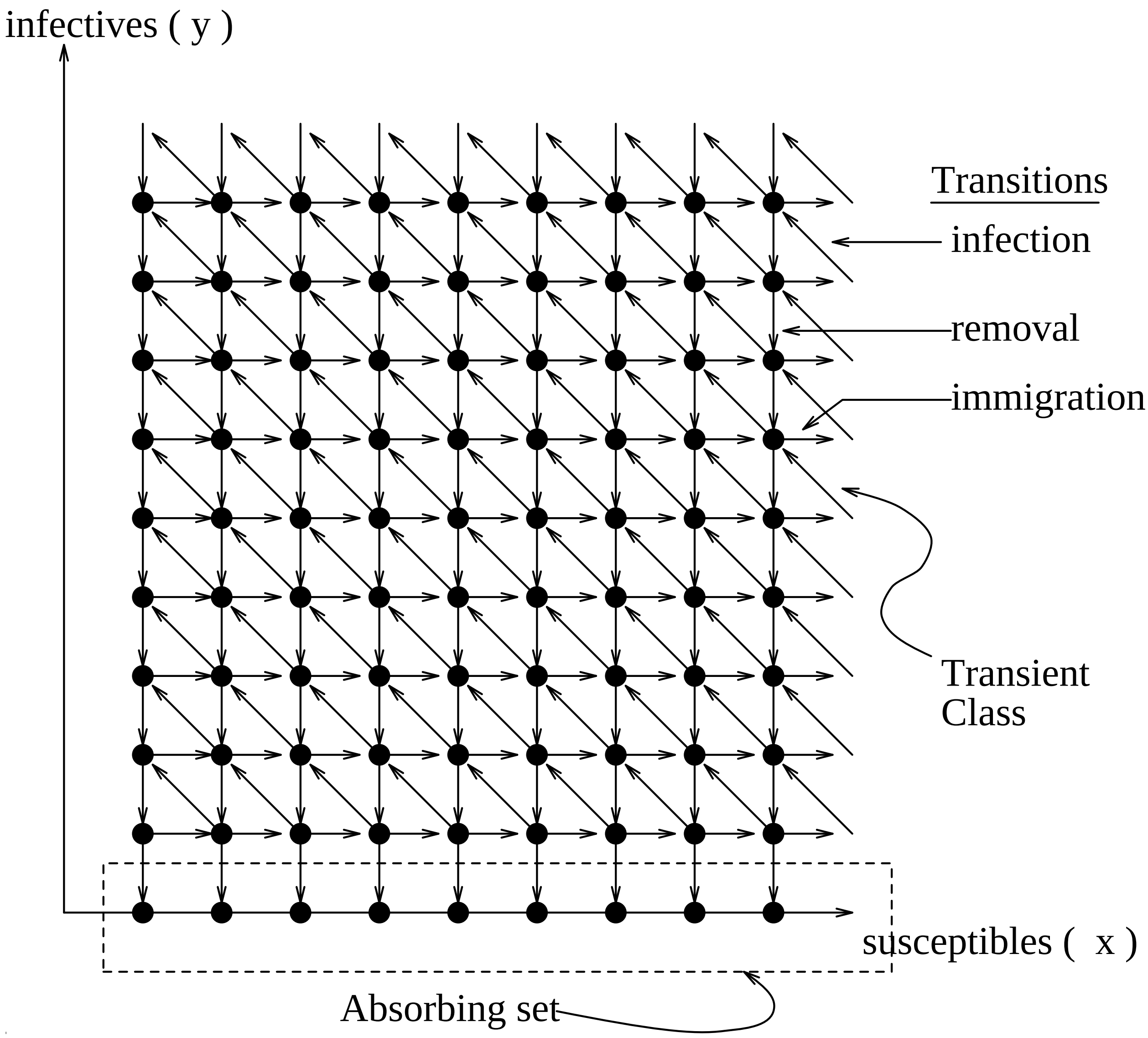
State space:

$$S = \{(x, y) : x, y = 0, 1, 2, \ldots\}$$

Transition rates $Q = (q_{ij},\ i, j \in S)$:

if $i = (x, y)$, then

$$q_{ij} = \begin{cases} \beta xy & \text{if } j = (x-1, y+1), \\ \gamma y & \text{if } j = (x, y-1), \\ \alpha & \text{if } j = (x+1, y), \\ 0 & \text{otherwise.} \end{cases}$$



Clearly,

$$C = \{(x, y) : x = 0, 1, \ldots ; y = 1, 2, \ldots\}$$

is an irreducible transient class.

Ridler-Rowe (1967): Q is regular (nonexplosive) and absorption occurs with probability 1.

QUASISTATIONARY DISTRIBUTIONS

Notation. Let $X(t)$ be the state of the system at time t and suppose that $X(t) \in S$, where S, the state space, is a discrete set. Let $Q = (q_{ij},\ i,j \in S)$ be the matrix of transition rates (q-matrix), 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.

Assumptions. Take 0 to be the sole absorbing state, that is, $q_0 = 0$, and suppose that $S = \{0\} \cup C$, where $C = \{1, 2, \ldots\}$ is an irreducible class. In order that there be a positive probability of ever reaching 0, given that the process starts in C, suppose that $q_{i0} > 0$ for at least one $i \in C$.

Let $p_{ij}(t) = \Pr(X(t) = j | X(0) = i)$, $t \geq 0$.

Definition. Let $\pi = (\pi_j,\ j \in C)$ be a probability distribution over C and let

$$p_j(t) = \sum_{i \in C} \pi_i p_{ij}(t), \quad j \in S,\ t \geq 0.$$

Then, π is a *quasistationary distribution* if

$$\frac{p_j(t)}{\sum_{i \in C} p_i(t)} = \pi_j, \quad j \in C,\ t \geq 0.$$

That is, if the process has π as its initial distribution, then π is a quasistationary distribution if the state probabilities at time t, conditional on the process being in C at t, are the same for all t.

If π is a quasistationary distribution then, under certain conditions,

$$\lim_{t \to \infty} \Pr(X(t) = j | X(0) = i, X(t) \in C) = \pi_j,$$

no matter what the value of $i (\in C)$.

A CHARACTERIZATION

If π is a quasistationary distribution, then clearly

$$\sum_{i \in C} \pi_i p_{ij}(t) = g(t)\pi_j, \quad j \in C, \ t > 0,$$

where $g(t) = \sum_{i \in C} p_i(t)$. It is easy to show that g satisfies:

$$g(s+t) = g(s)g(t), \quad s, t \geq 0,$$

and $0 < g(t) < 1$. Thus, $g(t) = e^{-\mu t}$, for some $\mu > 0$. Since the converse is also true, we have the following result:

Proposition. A probability distribution, $\pi = (\pi_j, \ j \in C)$, over C is a quasistationary distribution if and only if, for some $\mu > 0$, π is a μ-invariant measure, that is

$$\sum_{i \in C} \pi_i p_{ij}(t) = e^{-\mu t}\pi_j, \quad j \in C, \ t \geq 0. \qquad (1)$$

HOW DO WE DETERMINE π from Q?

Rewrite (1) as

$$\sum_{i \in C:\, i \neq j} \pi_i p_{ij}(t) = \Big((1 - p_{jj}(t)) - (1 - e^{-\mu t})\Big)\, \pi_j$$

and remember that q_{ij} is the right-hand derivative of $p_{ij}(\cdot)$ near 0. So, on dividing by t and letting $t \downarrow 0$, we get (formally)

$$\sum_{i \in C:\, i \neq j} \pi_i q_{ij} = (q_j - \mu)\pi_j, \qquad j \in C,$$

or, equivalently,

$$\sum_{i \in C} \pi_i q_{ij} = -\mu \pi_j, \qquad j \in C. \tag{2}$$

Accordingly, we shall say that π is a μ-invariant measure for Q whenever (2) holds.

Proposition. If π is a quasistationary distribution then, for some $\mu > 0$, π is a μ-invariant measure for Q.

IS THE CONVERSE TRUE?

Suppose that, for some $\mu > 0$, π is a μ-invariant measure for Q, that is

$$\sum_{i \in C} \pi_i q_{ij} = -\mu \pi_j, \quad j \in C.$$

Is π a quasistationary distribution?

Sum this equation over $j \in C$: we get (formally)

$$\sum_{i \in C} \pi_i q_{i0} = -\sum_{i \in C} \pi_i \sum_{j \in C} q_{ij} = -\sum_{j \in C} \sum_{i \in C} \pi_i q_{ij}$$

$$= \mu \sum_{j \in C} \pi_j = \mu.$$

Theorem. Let $\pi = (\pi_j,\ j \in C)$ be a probability distribution over C and suppose that π is a μ-invariant measure for Q. Then, $\mu \geq \sum_{j \in C} \pi_j q_{j0}$ with equality if and only if π is a quasistationary distribution.

AN EXAMPLE

The birth-death-catastrophe process. Let $S = \{0, 1, 2, \ldots\}$ and suppose that

$$\begin{aligned} q_{i,i+1} &= a\rho i, & i &\geq 0, \\ q_{i,i} &= -\rho i, & i &\geq 0, \\ q_{i,i-k} &= \rho i b_k, & i \geq 2,\ k &= 1, 2 \ldots i-1, \\ q_{i,0} &= \rho i \textstyle\sum_{k=i}^{\infty} b_k, & i &\geq 1, \end{aligned}$$

where $\rho > 0$, $a > 0$, $b_i > 0$ for at least one value of $i \geq 1$ and $a + \sum_{i=1}^{\infty} b_i = 1$. Thus, jumps occur at a constant "per capita" rate ρ and, at a jump time, a birth occurs with probability a, or otherwise a catastrophe occurs, the size of which is determined by the probabilities b_i, $i \geq 1$.

Clearly, 0 is an absorbing state and $C = \{1, 2, \ldots\}$ is an irreducible class.

So, does the process admit a quasistationary distribution?

Substitute the transition rates in (2):

$$-(\rho-\mu)\pi_1+\sum_{k=2}^{\infty}k\rho b_{k-1}\pi_k=0,$$

and, for $j\geq 2$,

$$(j-1)\rho a\pi_{j-1}-(j\rho-\mu)\pi_j+\sum_{k=j+1}^{\infty}k\rho b_{k-j}\pi_k=0.$$

If we try a solution of the form $\pi_j = t^j$, the first equation tells us that $\mu=-\rho(f'(t)-1)$, where

$$f(s)=a+\sum_{i\in C}b_i s^{i+1}, \qquad |s|\leq 1,$$

and, on substituting both of *these* quantities in the second equation, we find that $f(t)=t$. This latter equation has a unique solution, σ, on $[0,1]$. Thus, by setting $t=\sigma$ we can obtain a positive solution, $\pi=(\pi_j,\ j\in C)$, to (2), such that $\sum_{j\in C}\pi_j=1$ whenever $\sigma<1$. Further, it is easy to show that $\sum_{i\in C}\pi_i q_{i0}=\mu$.

The condition $\sigma < 1$ is satisfied only in the subcritical case, that is, when the drift, D, given by

$$D = a - \sum_{i \in C} i b_i = 1 - \sum_{i \in C} (i+1) b_i,$$

is strictly negative; incidentally, this also guarantees that absorption occurs with probability 1.

Proposition. The subcritical birth-death-catastrophe process has a geometric quasistationary distribution, $\pi = (\pi_j,\ j \in C)$, given by

$$\pi_j = (1-\sigma)\sigma^{j-1}, \qquad j \in C,$$

where σ is the unique solution to $f(t) = t$ on the interval $[0, 1]$.

COMPUTATIONAL METHODS

Finite S : Mandl (1960) showed that the restriction of Q to C has eigenvalues with negative real parts and the one with maximal real part (called $-\mu$ above) is real and has multiplicity 1, and, the corresponding left eigenvector, $m = (m_j,\ j \in C)$, has positive entries; this is, of course, a μ-invariant measure for Q (unique up to constant multiples). Since S is finite, the quasistationary distribution, $\pi = (\pi_i,\ i \in C)$, exists and is given by

$$\pi_j = m_j / \sum_{k \in C} m_k, \quad j \in C.$$

Infinite S : Truncate the restriction of Q to an $n \times n$ matrix, $Q^{(n)}$, and construct a sequence, $\{m^{(n)}\}$, of eigenvectors and hope that this converges to a μ-invariant measure, m, for Q, et cetera.

HOW SHOULD WE EVALUATE m?

Consider the Ridler-Rowe epidemic model.

First truncate C to

$$C_N = \{(x, y) : x = 0, \ldots, N-1; y = 1, \ldots, N\}$$

and restrict Q to C_N. Use the transformation $i = x + N(y-1)$ to convert the restricted q-matrix into an $n \times n$ matrix, $Q = (q_{ij},\ i, j = 0, 1, \ldots, n-1)$, where $n = N^2$.

Evaluation of the eigenvectors of Q is not completely trivial. For example, if (as well shall assume) $N = 100$, that is $n = 10^4$, Q needs 400 Mbytes of storage.

However, for N large, this matrix is large and sparse. Does this help?

THE ARNOLDI METHOD

We need to solve $Ax = \lambda x$, where $A = Q^T$.

Using an initial estimate of x, the basic Arnoldi method produces an $m \times m$ (upper-Hessenberg) matrix H_m and an $n \times m$ matrix V_m with

$$V_m^T A V_m = H_m,$$

and such that if z_m is an eigenvector of H_m, then, for m large, $V_m z_m$ is close to an eigenvector of A.

Solve for z_m using standard (dense-matrix) methods.

AN ITERATIVE ARNOLDI METHOD

Take m small (we found that $m = 20$ worked best). Then, using an initial estimate, v_1, of the eigenvector x, apply the basic Arnoldi method (to obtain H_m and V_m) and set $\widehat{\lambda}$ to be the dominant eigenvalue of H_m if this is real, or set $\widehat{\lambda}$ equal to zero otherwise.

Now solve

$$(H_m - \widehat{\lambda} I)u_1 = z$$

with z chosen at random and repeat the procedure with a new initial estimate, given by

$$v_1 = V_m u_1 / \|V_m u_1\|_2.$$

Continue until the residual, $\|Av_1 - \widehat{\lambda} v_1\|_2$, is sufficiently small.