Introduction to Markov Chains

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

Subsection 1

Q-Matrices and Their Exponentials

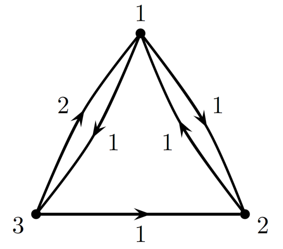
Q-Matrices

- Let I be a countable set.
- A Q-**matrix** on I is a matrix $Q = (q_{ij} : i, j \in I)$ satisfying the following conditions:
  (i) $0 \leq -q_{ii} < \infty$, for all i;
  (ii) $q_{ij} \geq 0$, for all $i \neq j$;
  (iii) $\sum_{j \in I} q_{ij} = 0$, for all i.
- Thus in each row of Q we can choose the off-diagonal entries to be any nonnegative real numbers, subject only to the constraint that the off-diagonal row sum is finite:

$$q_i = \sum_{j \neq i} q_{ij} < \infty.$$

- The diagonal entry q_{ii} is then $-q_i$, making the total row sum zero.

Diagrams and Q-Matrices



- A convenient way to present the data for a continuous-time Markov chain is by means of a diagram.
- Each diagram then corresponds to a unique Q-matrix, in this case

$$Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}.$$

- Thus each off-diagonal entry q_{ij} gives the value we attach to the (i, j) arrow on the diagram, which we shall interpret later as the **rate of going from i to j**.
- The numbers q_i are not shown, but can be worked out from the other information given.
- We shall later interpret q_i as the **rate of leaving i**.

Interpolating Into a Discrete Sequence

- We may think of the discrete parameter space $\{0, 1, 2, \ldots\}$ as embedded in the continuous parameter space $[0, \infty)$.
- For $p \in (0, \infty)$ a natural way to interpolate the discrete sequence $(p^n : n = 0, 1, 2, \ldots)$ is by the function

$$(e^{tq} : t \geq 0), \quad q = \log p$$

- Consider a *finite set* I and a matrix $P = (p_{ij} : i, j \in I)$.
- We ask for a natural way to fill in the gaps in the discrete sequence

$$(P^n : n = 0, 1, 2, \ldots).$$

Interpolating Into a Discrete Sequence (Cont'd)

- Consider any matrix $Q = (q_{ij} : i, j \in I)$.
- The series
$$\sum_{k=0}^{\infty} \frac{Q^k}{k!}$$
converges componentwise.
- We denote its limit by e^Q.
- If two matrices Q_1 and Q_2 commute, then
$$e^{Q_1+Q_2} = e^{Q_1} e^{Q_2}.$$
- Suppose that we can find a matrix Q with $e^Q = P$.
- Then
$$e^{nQ} = (e^Q)^n = P^n.$$
- So $(e^{tQ} : t \geq 0)$ fills in the gaps in the discrete sequence.

Properties of e^{tQ}

Theorem

Let Q be a matrix on a finite set I. Set $P(t) = e^{tQ}$. Then $(P(t) : t \geq 0)$ has the following properties:

(i) $P(s+t) = P(s)P(t)$, for all s, t (**semigroup property**);

(ii) $(P(t) : t \geq 0)$ is the unique solution to the **forward equation**

$$\frac{d}{dt}P(t) = P(t)Q, \quad P(0) = I;$$

(iii) $(P(t) : t \geq 0)$ is the unique solution to the **backward equation**

$$\frac{d}{dt}P(t) = QP(t), \quad P(0) = I;$$

(iv) For $k = 0, 1, 2, \ldots$, we have $(\frac{d}{dt})^k \mid_{t=0} P(t) = Q^k$.

Properties of e^{tQ}

- For any $s, t \in \mathbb{R}$, sQ and tQ commute.
  So $e^{sQ}e^{tQ} = e^{(s+t)Q}$ proving the semigroup property.
  The matrix-valued power series
  $$P(t) = \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!}$$
  has infinite radius of convergence.
  So each component is differentiable and the derivative is given by term-by-term differentiation.
  We differentiate term-by-term
  $$P'(t) = \sum_{k=1}^{\infty} \frac{t^{k-1}Q^k}{(k-1)!} = P(t)Q = QP(t).$$
  Hence $P(t)$ satisfies the forward and backward equations.
  Moreover by repeated term-by-term differentiation we obtain (iv).

Properties of e^{tQ} (Cont'd)

- It remains to show that $P(t)$ is the only solution of the forward and backward equations.

  If $M(t)$ satisfies the forward equation, then

$$\begin{aligned}\frac{d}{dt}(M(t)e^{-tQ}) &= (\frac{d}{dt}M(t))e^{-tQ} + M(t)(\frac{d}{dt}e^{-tQ}) \\ &= M(t)Qe^{-tQ} + M(t)(-Q)e^{-tQ} \\ &= 0.\end{aligned}$$

  So $M(t)e^{-tQ}$ is constant.

  Thus, $M(t) = P(t)$.

  A similar argument proves uniqueness for the backward equation.

Focusing on Q-Matrices: Notation

- The preceding theorem was about matrix exponentials in general.
- We look at what happens to Q-matrices.
- Recall that a matrix $P = (p_{ij} : i, j \in I)$ is stochastic if it satisfies:
  (i) $0 \leq p_{ij} < \infty$, for all i, j;
  (ii) $\sum_{j \in I} p_{ij} = 1$ for all i.
- We recall the conventions that, in the limit $t \to 0$, the statement:
  - $f(t) = O(t)$ means that, for some $C < \infty$, $\frac{f(t)}{t} \leq C$, for all sufficiently small t;
  - $f(t) = o(t)$ means $\frac{f(t)}{t} \to 0$ as $t \to 0$.

Characterization of Q-Matrices

Theorem

A matrix Q on a finite set I is a Q-matrix if and only if $P(t) = e^{tQ}$ is a stochastic matrix for all $t \geq 0$.

- As $t \searrow 0$ we have
$$P(t) = I + tQ + O(t^2).$$
So $q_{ij} \geq 0$ for $i \neq j$ if and only if $p_{ij}(t) \geq 0$, for all i, j and $t \geq 0$ sufficiently small.
But $P(t) = P(\frac{t}{n})^n$ for all n.
So $q_{ij} \geq 0$ for $i \neq j$ if and only if $p_{ij}(t) \geq 0$ for all i, j and all $t \geq 0$.

Characterization of Q-Matrices (Cont'd)

- If Q has zero row sums then so does Q^n for every n:

$$\sum_{k\in I} q_{ik}^{(n)} = \sum_{k\in I}\sum_{j\in I} q_{ij}^{(n-1)} q_{jk} = \sum_{j\in I} q_{ij}^{(n-1)} \sum_{k\in I} q_{jk} = 0.$$

So

$$\sum_{j\in I} p_{ij}(t) = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{j\in I} q_{ij}^{(n)} = 1.$$

Conversely, suppose, for all $t \geq 0$,

$$\sum_{j\in I} p_{ij}(t) = 1.$$

Then

$$\sum_{j\in I} q_{ij} = \frac{d}{dt}\Big|_{t=0} \sum_{j\in I} p_{ij}(t) = 0.$$

Interpolating in Processes

- Now, if P is a stochastic matrix of the form e^Q for some Q-matrix, we can do some sort of filling-in of gaps at the level of processes.
- Fix some large integer m.
- Let $(X_n^m)_{n\geq 0}$ be discrete-time Markov$(\lambda, e^{Q/m})$.
- We define a process indexed by $\{\frac{n}{m} : n = 0, 1, 2, \ldots\}$ by
$$X_{n/m} = X_n^m.$$
- Then $(X_n : n = 0, 1, 2, \ldots)$ is discrete-time Markov$(\lambda, (e^{Q/m})^m)$.
- Moreover,
$$(e^{Q/m})^m = e^Q = P.$$
- Thus we can find discrete-time Markov chains with arbitrarily fine grids $\{\frac{n}{m} : n = 0, 1, 2, \ldots\}$ as time-parameter sets which give rise to Markov(λ, P) when sampled at integer times.
- It should not then be too surprising that there is, as we will see, a continuous-time process $(X_t)_{t\geq 0}$ which also has this property.

Transition Probabilities

- We will see that a continuous-time Markov chain $(X_t)_{t\geq 0}$ with Q-matrix Q satisfies

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n),$$

  for all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq \cdots \leq t_{n+1}$ and all states $i_0, \ldots, i_{n+1}$, where $p_{ij}(t)$ is the (i,j) entry in e^{tQ}.
- In particular, the **transition probability** from i to j in time t is given by

$$\mathbb{P}_i(X_t = j) := \mathbb{P}(X_t = j | X_0 = i) = p_{ij}(t).$$

Example

- We calculate $p_{11}(t)$ for the continuous-time Markov chain with Q-matrix $Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$.

  We begin by writing down the characteristic equation for Q.

$$\det(x - Q) = 0$$
$$\det \begin{pmatrix} x+2 & -1 & -1 \\ -1 & x+1 & 0 \\ -2 & -1 & x+3 \end{pmatrix} = 0$$
$$(x+2)(x+1)(x+3) - 1 - 2(x+1) - (x+3) = 0$$
$$x^3 + 6x^2 + 11x + 6 - 1 - 2x - 2 - x - 3 = 0$$
$$x^3 + 6x^2 + 8 = 0$$
$$x(x+2)(x+4) = 0$$
$$x = 0, x = -2, x = -4.$$

  Thus, Q has distinct eigenvalues $0, -2, -4$.

Example (Cont'd)

Claim: $p_{11}(t)$ has the form

$$p_{11}(t) = a + be^{-2t} + ce^{-4t},$$

for some constants a, b and c.

We could diagonalize Q by an invertible matrix U:

$$Q = U \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -4 \end{pmatrix} U^{-1}.$$

Example (Cont'd)

- Then

$$\begin{aligned} e^{tQ} &= \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!} \\ &= U \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} 0^k & 0 & 0 \\ 0 & (-2t)^k & 0 \\ 0 & 0 & (-4t)^k \end{pmatrix} U^{-1} \\ &= U \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-2t} & 0 \\ 0 & 0 & e^{-4t} \end{pmatrix} U^{-1}. \end{aligned}$$

So $p_{11}(t)$ must be of the form

$$p_{11}(t) = a + be^{-2t} + ce^{-4t}.$$

Example (Cont'd)

- To determine the constants we use

$$\begin{aligned} 1 = p_{11}(0) &= a + b + c, \\ -2 = q_{11} = p'_{11}(0) &= -2b - 4c, \\ 7 = q_{11}^{(2)} = p''_{11}(0) &= 4b + 16c. \end{aligned}$$

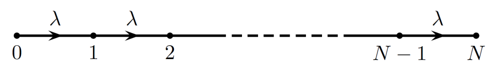
So we get

$$\left\{ \begin{array}{rcl} a+b+c &=& 1 \\ -2b-4c &=& -2 \\ 4b+16c &=& 7 \end{array} \right. \Rightarrow \left\{ \begin{array}{rcl} a+b+c &=& 1 \\ b+2c &=& 1 \\ 8c &=& 3 \end{array} \right. \Rightarrow \left\{ \begin{array}{rcl} a &=& \frac{3}{8} \\ b &=& \frac{1}{4} \\ c &=& \frac{3}{8} \end{array} \right.$$

So $p_{11}(t) = \frac{3}{8} + \frac{1}{4}e^{-2t} + \frac{3}{8}e^{-4t}$.

Example

- We calculate $p_{ij}(t)$ for the continuous time Markov chain with diagram



The Q-matrix is $Q = \begin{pmatrix} -\lambda & \lambda & & & & \\ & -\lambda & \lambda & & & \\ & & \ddots & \ddots & & \\ & & & -\lambda & \lambda & \\ & & & & -\lambda & \lambda \\ & & & & & 0 \end{pmatrix}$, where entries off the diagonal and super-diagonal are all zero.

The exponential of an upper-triangular matrix is upper-triangular.

So $p_{ij}(t) = 0$, for $i > j$.

Example (Cont'd)

- In components the forward equation $P'(t) = P(t)Q$ reads

$$\begin{array}{lll} p'_{ii}(t) = -\lambda p_{ii}(t), & p_{ii}(0) = 1, & \text{for } i < N, \\ p'_{ij}(t) = -\lambda p_{ij}(t) + \lambda p_{i,j-1}(t), & p_{ij}(0) = 0, & \text{for } i < j < N, \\ p'_{iN}(t) = \lambda p_{iN-1}(t), & p_{iN}(0) = 0, & \text{for } i < N. \end{array}$$

  We can solve these equations.
  - $p_{ii}(t) = e^{-\lambda t}$, for $i < N$;
  - For $i < j < N$, $(e^{\lambda t} p_{ij}(t))' = e^{\lambda t} p_{i,j-1}(t)$.
    So, by induction,

$$p_{ij}(t) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}.$$

  If $i = 0$, these are the Poisson probabilities of parameter λt.
  So, starting from 0, the distribution of the Markov chain at time t is the same as the distribution of $\min\{Y_t, N\}$, where Y_t is a Poisson random variable of parameter λt.

Subsection 2

Continuous Time Random Processes

Continuous Time Random Processes

- Let I be a countable set.
- A **continuous time random process**

$$(X_t)_{t\geq 0} = (X_t : 0 \leq t < \infty)$$

  with values in I is a family of random variables $X_t : \Omega \to I$.
- A continuous time random process is **right continuous** if, for all $\omega \in \Omega$ and $t \geq 0$, there exists $\varepsilon > 0$, such that

$$X_s(\omega) = X_t(\omega), \text{ for all } t \leq s \leq t + \varepsilon.$$

- We restrict our attention to right continuous processes.

Finite-Dimensional Distributions

- By a standard result of measure theory, the probability of any event depending on a right continuous process can be determined from its **finite dimensional distributions**, i.e., from the probabilities

$$\mathbb{P}(X_{t_0} = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n),$$

for $n \geq 0, 0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$ and $i_0, \ldots, i_n \in I$.

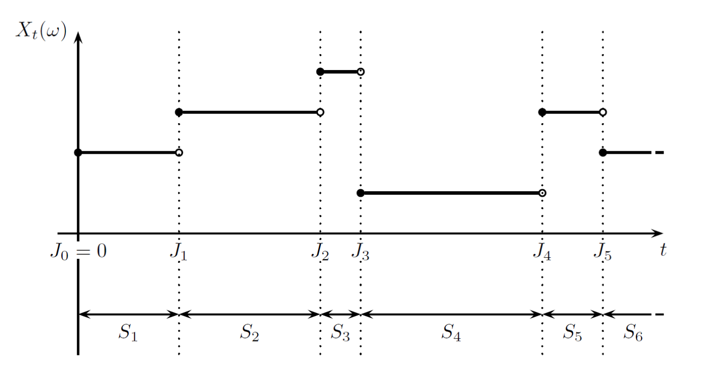
Example:

$$\begin{aligned}\mathbb{P}(X_t = i \text{ for some } t \in [0, \infty)) \\ = 1 - \lim_{n \to \infty} \sum_{j_1, \ldots, j_n \neq i} \mathbb{P}(X_{q_1} = j_i, \ldots, X_{q_n} = j_n),\end{aligned}$$

where $q_1, q_2, \ldots$ is an enumeration of the rationals.

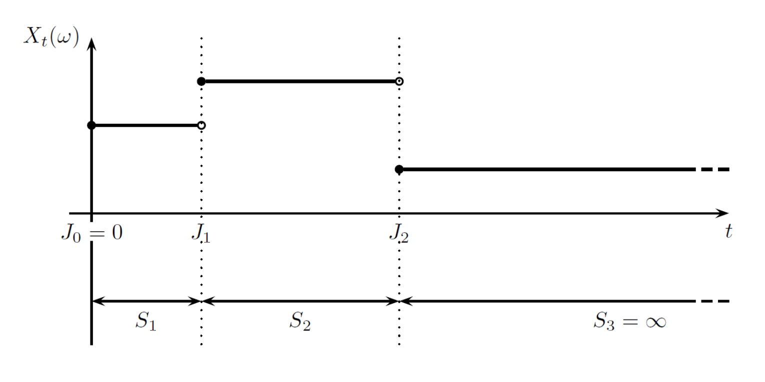
Right Continuous Process Type I

- The path makes infinitely many jumps, but only finitely many in any interval $[0, t]$:



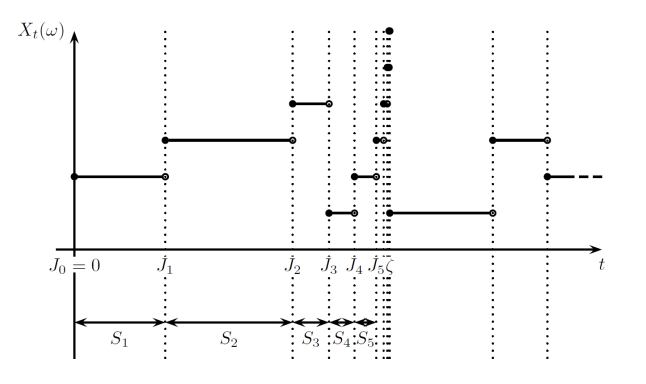
Right Continuous Process Type II

- The path makes finitely many jumps and then becomes stuck in some state forever.



Right Continuous Process Type III

- The process makes infinitely many jumps in a finite interval.



- In this case, after the explosion time ζ the process starts up again.
- It may explode again, maybe infinitely often, or it may not.

Jump Times and Holding Times

- We call $J_0, J_1, \ldots$ the **jump times** of $(X_t)_{t\geq 0}$.
- They are obtained from $(X_t)_{t\geq 0}$ by

$$J_0 = 0, \quad J_{n+1} = \inf\{t \geq J_n : X_t \neq X_{J_n}\}, \ n = 0, 1, \ldots,$$

  where $\inf \emptyset = \infty$.
- We call $S_1, S_2, \ldots$ the **holding times**.
- They are given, for $n = 1, 2, \ldots$, by

$$S_n = \begin{cases} J_n - J_{n-1}, & \text{if } J_{n-1} < \infty \\ \infty, & \text{otherwise.} \end{cases}$$

- Note that right continuity forces $S_n > 0$, for all n.
- If $J_{n+1} = \infty$, for some n, we define $X_\infty = X_{J_n}$, the final value, otherwise X_∞ is undefined.

Explosion Time and Jump Process

- The (**first**) **explosion time** ζ is defined by

$$\zeta = \sup_n J_n = \sum_{n=1}^{\infty} S_n.$$

- The discrete time process $(Y_n)_{n \geq 0}$ given by $Y_n = X_{J_n}$ is called the **jump process** of $(X_t)_{t \geq 0}$, or the **jump chain** if it is a discrete time Markov chain.
- This is the sequence of values taken by $(X_t)_{t \geq 0}$ up to explosion.

Minimal Processes

- We shall not consider what happens to a process after explosion.
- So it is convenient to:
  - Adjoin to I a new state, ∞ say;
  - Require that
$$X_t = \infty, \text{ if } t \geq \zeta.$$
- Any process satisfying this requirement is called **minimal**.
- The terminology "minimal" does not refer to the state of the process but to the interval of time over which the process is active.

The Process in terms of Holding Times and Jump Process

- Note that a minimal process may be reconstructed from its holding times and jump process.
- We, thus, obtain another "countable" specification of the probabilistic behavior of $(X_t)_{t\geq 0}$ by specifying the joint distribution of $S_1, S_2, \ldots$ and $(Y_n)_{n\geq 0}$.

  Example: The probability that $X_t = i$ is given by

$$\mathbb{P}(X_t = i) = \sum_{n=0}^{\infty} \mathbb{P}(Y_n = i \text{ and } J_n \leq t < J_{n+1}).$$

  Moreover,

$$\mathbb{P}(X_t = i \text{ for some } t \in [0, \infty)) = \mathbb{P}(Y_n = i \text{ for some } n \geq 0).$$

Subsection 3

Some Properties of the Exponential Distribution

Exponential Distributions

- A random variable $T : \Omega \to [0, \infty]$ has **exponential distribution of parameter** λ, $0 \leq \lambda < \infty$, if

$$\mathbb{P}(T > t) = e^{-\lambda t}, \text{ for all } t \geq 0.$$

- We write $T \sim E(\lambda)$ for short.
- If $\lambda > 0$, then T has density function

$$f_T(t) = \lambda e^{-\lambda t} 1_{t \geq 0}.$$

- The mean of T is given by

$$\mathbb{E}(T) = \int_0^\infty \mathbb{P}(T > t) dt = \frac{1}{\lambda}.$$

Memoryless Property

Theorem (Memoryless Property)

A random variable $T : \Omega \to (0, \infty]$ has an exponential distribution if and only if it has the following memoryless property:

$$\mathbb{P}(T > s + t | T > s) = \mathbb{P}(T > t), \text{ for all } s, t \geq 0.$$

- Suppose $T \sim E(\lambda)$.

  Then

$$\begin{aligned} \mathbb{P}(T > s + t | T > s) &= \frac{\mathbb{P}(T>s+t)}{\mathbb{P}(T>s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} \\ &= \mathbb{P}(T > t). \end{aligned}$$

Memoryless Property (Converse)

- Suppose T has the memoryless property whenever $\mathbb{P}(T > s) > 0$. Then $g(t) = \mathbb{P}(T > t)$ satisfies

  $$g(s + t) = g(s)g(t), \text{ for all } s, t \geq 0.$$

  We assumed $T > 0$ so that $g(\frac{1}{n}) > 0$, for some n.

  Then, by induction

  $$g(1) = g\left(\frac{1}{n} + \cdots + \frac{1}{n}\right) = g\left(\frac{1}{n}\right)^n > 0.$$

  So $g(1) = e^{-\lambda}$, for some $0 \leq \lambda < \infty$.

Memoryless Property (Converse Cont'd)

- By the same argument, for integers $p, q \geq 1$,

$$g\left(\frac{p}{q}\right) = g\left(\frac{1}{q}\right)^p = g(1)^{p/q}.$$

So $g(r) = e^{-\lambda r}$, for all rationals $r > 0$.

For real $t > 0$, choose rationals $r, s > 0$ with $r \leq t \leq s$.

Since g is decreasing,

$$e^{-\lambda r} = g(r) \geq g(t) \geq g(s) = e^{-\lambda s}.$$

But we can choose r and s arbitrarily close to t.

This forces $g(t) = e^{-\lambda t}$.

So $T \sim E(\lambda)$.

Sum of Independent Exponential Random Variables

Theorem

Let $S_1, S_2, \ldots$ be a sequence of independent random variables with $S_n \sim E(\lambda_n)$ and $0 < \lambda_n < \infty$, for all n.

(i) If $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty$, then $\mathbb{P}(\sum_{n=1}^{\infty} S_n < \infty) = 1$.

(ii) If $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$, then $\mathbb{P}(\sum_{n=1}^{\infty} S_n = \infty) = 1$.

(i) Suppose $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty$.
By Monotone Convergence,

$$\mathbb{E}\left(\sum_{n=1}^{\infty} S_n\right) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty.$$

So $\mathbb{P}(\sum_{n=1}^{\infty} S_n < \infty) = 1$.

Sum of Independent Exponential Random Variables (ii)

(ii) Suppose instead that

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty.$$

Then

$$\prod_{n=1}^{\infty} \left(1 + \frac{1}{\lambda_n}\right) = \infty.$$

By Monotone Convergence and independence

$$\begin{aligned} \mathbb{E}(\exp\{-\textstyle\sum_{n=1}^{\infty} S_n\}) &= \textstyle\prod_{n=1}^{\infty} \mathbb{E}(\exp\{-S_n\}) \\ &= \textstyle\prod_{n=1}^{\infty}(1 + \frac{1}{\lambda_n})^{-1} \\ &= 0. \end{aligned}$$

So $\mathbb{P}(\sum_{n=1}^{\infty} S_n = \infty) = 1$.

Infimum of Independent Exponential Random Variables

Theorem

Let I be countable and let T_k, $k \in I$, be independent random variables with $T_k \sim E(q_k)$ and $0 < q := \sum_{k \in I} q_k < \infty$. Set $T = \inf_k T_k$. Then this infimum is attained at a unique random value K of k, with probability 1. Moreover, T and K are independent, with $T \sim E(q)$ and $\mathbb{P}(K = k) = \frac{q_k}{q}$.

- Set $K = k$ if $T_k < T_j$, for all $j \neq k$, otherwise let K be undefined. Then

$$\begin{aligned} \mathbb{P}(K = k \text{ and } T \geq t) &= \mathbb{P}(T_k \geq t \text{ and } T_j > T_k \text{ for all } j \neq k) \\ &= \int_t^\infty q_k e^{-q_k s} \mathbb{P}(T_j > s \text{ for all } j \neq k) ds \\ &= \int_t^\infty q_k e^{-q_k s} \prod_{j \neq k} e^{-q_j s} ds \\ &= \int_t^\infty q_k e^{-qs} ds = \frac{q_k}{q} e^{-qt}. \end{aligned}$$

Hence, $\mathbb{P}(K = k \text{ for some } k) = 1$.

Moreover, T and K have the claimed joint distribution.

Two Independent Exponential Random Variables

Theorem

For independent random variables $S \sim E(\lambda)$ and $R \sim E(\mu)$ and for $t \geq 0$, we have

$$\mu\mathbb{P}(S \leq t < S + R) = \lambda\mathbb{P}(R \leq t < R + S).$$

- We have

$$\begin{aligned} \mu\mathbb{P}(S \leq t < S + R) &= \mu \int_0^t \int_{t-s}^{\infty} \lambda\mu e^{-\lambda s} e^{-\mu r} dr ds \\ &= \lambda\mu \int_0^t e^{-\lambda s} e^{-\mu(t-s)} ds. \end{aligned}$$

  Symmetrically,

$$\lambda\mathbb{P}(R \leq t < R + S) = \mu\lambda \int_0^t e^{-\mu r} e^{-\lambda(t-r)} dr.$$

  A change of variables shows that the integrals are equal.

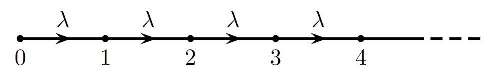
  This establishes the identity.

Subsection 4

Poisson Processes

Poisson Processes

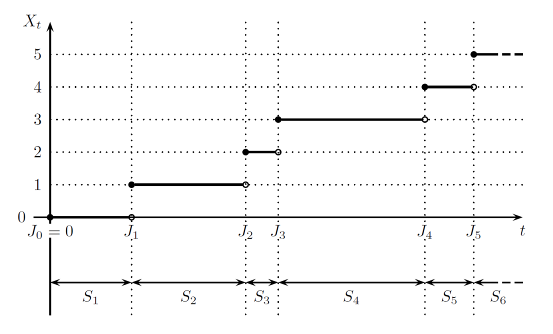
- A right-continuous process $(X_t)_{t\geq 0}$ with values in $\{0,1,2,\ldots\}$ is a **Poisson process of rate** λ, $0<\lambda<\infty$, if its holding times $S_1, S_2, \ldots$ are independent exponential random variables of parameter λ and its jump chain is given by $Y_n = n$.



- The associated Q-matrix is given by $Q = \begin{pmatrix} -\lambda & \lambda & & \\ & -\lambda & \lambda & \\ & & \ddots & \ddots \\ & & & \end{pmatrix}$.
- By a previous theorem (or the Strong Law of Large Numbers) we have $\mathbb{P}(J_n \to \infty) = 1$.
- So there is no explosion and the law of $(X_t)_{t\geq 0}$ is uniquely determined.

Construction

- A simple way to construct a Poisson process of rate λ is to:
  - Take a sequence $S_1, S_2, \ldots$ of independent exponential random variables of parameter λ;
  - Set $J_0 = 0$, $J_n = S_1 + \ldots + S_n$;
  - Set $X_t = n$ if $J_n \leq t < J_{n+1}$.
- The diagram illustrates a typical path.



Markov Property of Poisson Processes

Theorem (Markov Property)

Let $(X_t)_{t\geq 0}$ be a Poisson process of rate λ. Then, for any $s\geq 0$, $(X_{s+t}-X_s)_{t\geq 0}$ is also a Poisson process of rate λ, independent of $(X_r : r\leq s)$.

- It suffices to prove the claim conditional on $X_s=i$, for each $i\geq 0$. Set $\widetilde{X}_t=X_{s+t}-X_s$. We have

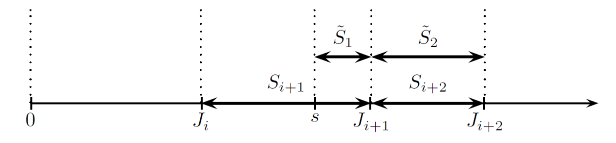
$$\{X_s=i\}=\{J_i\leq s<J_{i+1}\}=\{J_i\leq s\}\cap\{S_{i+1}>s-J_i\}.$$

On this event $X_r=\sum_{j=1}^i 1_{\{S_j\leq r\}}$, for $r\leq s$.

Moreover, the holding times $\widetilde{S}_1,\widetilde{S}_2,\ldots$ of $(\widetilde{X}_t)_{t\geq 0}$ are given by

$$\begin{array}{rcl}\widetilde{S}_1 & = & S_{i+1}-(s-J_i),\\ \widetilde{S}_n & = & S_{i+n},\quad n\geq 2.\end{array}$$

Markov Property of Poisson Processes (Cont'd)



- Recall that the holding times $S_1, S_2, \ldots$ are independent $E(\lambda)$.

  Condition on $S_1, \ldots, S_i$ and $\{X_s = i\}$.

  Take into account:
  - The memoryless property of S_{i+1};
  - Independence.

  Then $\widetilde{S}_1, \widetilde{S}_2, \ldots$ are themselves independent $E(\lambda)$.

  Hence, conditional on $\{X_s = i\}$, $\widetilde{S}_1, \widetilde{S}_2, \ldots$ are independent $E(\lambda)$, and independent of $S_1, \ldots, S_i$.

  So, conditional on $\{X_s = i\}$, $(\widetilde{X}_t)_{t \geq 0}$ is a Poisson process of rate λ and independent of $(X_r : r \leq s)$.

Strong Markov Property

- We shall see later, by an argument in essentially the same spirit, that the result also holds with s replaced by any stopping time T of $(X_t)_{t\geq 0}$.

Theorem (Strong Markov Property)

Let $(X_t)_{t\geq 0}$ be a Poisson process of rate λ and let T be a stopping time of $(X_t)_{t\geq 0}$. Then, conditional on $T<\infty$, $(X_{T+t}-X_T)_{t\geq 0}$ is also a Poisson process of rate λ, independent of $(X_s : s\leq T)$.

Stationary and Independent Increments

- Let $(X_t)_{t\geq 0}$ be a real-valued process.
- Consider its increment $X_t - X_s$ over any interval $(s, t]$.
- We say that $(X_t)_{t\geq 0}$ has **stationary increments** if the distribution of $X_{s+t} - X_s$ depends only on $t \geq 0$.
- We say that $(X_t)_{t\geq 0}$ has **independent increments** if its increments over any finite collection of disjoint intervals are independent.

Fundamental Theorem for Poisson Processes

Theorem

Let $(X_t)_{t \geq 0}$ be an increasing, right-continuous integer-valued process starting from 0. Let $0 < \lambda < \infty$. Then the following three conditions are equivalent:

(a) (**Jump Chain/Holding Time Definition**) The holding times $S_1, S_2, \ldots$ of $(X_t)_{t \geq 0}$ are independent exponential random variables of parameter λ and the jump chain is given by $Y_n = n$ for all n;

(b) (**Infinitesimal Definition**) $(X_t)_{t \geq 0}$ has independent increments and, as $h \searrow 0$, uniformly in t,

$$\begin{array}{rcl} \mathbb{P}(X_{t+h} - X_t = 0) & = & 1 - \lambda h + o(h), \\ \mathbb{P}(X_{t+h} - X_t = 1) & = & \lambda h + o(h); \end{array}$$

(c) (**Transition Probability Definition**) $(X_t)_{t \geq 0}$ has stationary independent increments and, for each t, X_t has Poisson distribution of parameter λ.

Proof of the Theorem ((a)⇒(b))

- Suppose Condition (a) holds.

  By the Markov property, for any $t, h \geq 0$, $X_{t+h} - X_t$ has the same distribution as X_h and is independent of $(X_s : s \leq t)$.

  So $(X_t)_{t \geq 0}$ has independent increments.

  As $h \searrow 0$,

$$\begin{aligned}
\mathbb{P}(X_{t+h} - X_t \geq 1) &= \mathbb{P}(X_h \geq 1) \\
&= \mathbb{P}(J_1 \leq h) \\
&= 1 - e^{-\lambda h} = \lambda h + o(h); \\
\mathbb{P}(X_{t+h} - X_t \geq 2) &= \mathbb{P}(X_h \geq 2) \\
&= \mathbb{P}(J_2 \leq h) \\
&\leq \mathbb{P}(S_1 \leq h \text{ and } S_2 \leq h) \\
&= (1 - e^{-\lambda h})^2 = o(h).
\end{aligned}$$

  This implies Condition (b).

Proof of the Theorem ((b)⇒(c))

- Suppose Condition (b) holds.
  For $i = 2, 3, \ldots$, $\mathbb{P}(X_{t+h} - X_t = i) = o(h)$ as $h \searrow 0$, uniformly in t.
  Set $p_j(t) = \mathbb{P}(X_t = j)$.
  Then, for $j = 1, 2, \ldots$,

$$\begin{array}{rcl} p_j(t+h) & = & \mathbb{P}(X_{t+h} = j) \\ & = & \sum_{i=0}^{j} \mathbb{P}(X_{t+h} - X_t = i)\mathbb{P}(X_t = j - i) \\ & = & (1 - \lambda h + o(h))p_j(t) + (\lambda h + o(h))p_{j-1}(t) + o(h). \end{array}$$

  So

$$\frac{p_j(t+h) - p_j(t)}{h} = -\lambda p_j(t) + \lambda p_{j-1}(t) + O(h).$$

  This estimate is uniform in t.
  So we can put $t = s - h$ to obtain, for all $s \geq h$,

$$\frac{p_j(s) - p_j(s-h)}{h} = -\lambda p_j(s-h) + \lambda p_{j-1}(s-h) + O(h).$$

Proof of the Theorem ((b)⇒(c) Cont'd)

- We found:
  - $\frac{p_j(t+h)-p_j(t)}{h} = -\lambda p_j(t) + \lambda p_{j-1}(t) + O(h)$;
  - $\frac{p_j(s)-p_j(s-h)}{h} = -\lambda p_j(s-h) + \lambda p_{j-1}(s-h) + O(h)$, for $s \geq h$.

  Now let $h \searrow 0$ to see that:
  - $p_j(t)$ is continuous;
  - $p_j(t)$ is differentiable and satisfies

$$p_j'(t) = -\lambda p_j(t) + \lambda p_{j-1}(t).$$

  By a simpler argument we also find

$$p_0'(t) = -\lambda p_0(t).$$

  Since $X_0 = 0$, we have initial conditions

$$p_0(0) = 1, \quad p_j(0) = 0, \text{ for } j = 1, 2, \ldots.$$

Proof of the Theorem ((b)⇒(c) Cont'd)

- As we saw in a previous example, the preceding system of equations has a unique solution given by

$$p_j(t) = e^{-\lambda t}\frac{(\lambda t)^j}{j!}, \quad j = 0, 1, 2, \ldots.$$

Hence $X_t \sim P(\lambda t)$.

If $(X_t)_{t \geq 0}$ satisfies Condition (b), then certainly $(X_t)_{t \geq 0}$ has independent increments.

Also $(X_{s+t} - X_s)_{t \geq 0}$ satisfies Condition (b).

So the above argument shows $X_{s+t} - X_s \sim P(\lambda t)$, for any s.

This implies Condition (c).

Proof of the Theorem ((c)⇒(a))

- There is a process satisfying Condition (a).

  Moreover, we have shown that it must then satisfy Condition (c).

  But Condition (c) determines the finite dimensional distributions of $(X_t)_{t\geq 0}$.

  Hence it determines the distribution of jump chain and holding times.

  So, if one process satisfying Condition (c) also satisfies Condition (a), so must every process satisfying Condition (c).

The Forward Equations for the Poisson Process

- Consider the possibility of starting the process from i at time 0.
- We write $\mathbb{P}_i$ as a reminder.
- Set $p_{ij}(t) = \mathbb{P}_i(X_t = j)$.
- By spatial homogeneity, $p_{ij}(t) = p_{j-i}(t)$.
- So we could rewrite the differential equations as

$$\begin{array}{ll} p'_{i0}(t) = -\lambda p_{i0}(t), & p_{i0}(0) = \delta_{i0}, \\ p'_{ij}(t) = \lambda p_{i,j-1}(t) - \lambda p_{ij}(t), & p_{ij}(0) = \delta_{ij}. \end{array}$$

- In matrix form, for Q as above,

$$P'(t) = P(t)Q, \quad P(0) = I.$$

Sum of Independent Poisson Processes

Theorem

If $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are independent Poisson processes of rates λ and μ, respectively, then $(X_t + Y_t)_{t \geq 0}$ is a Poisson process of rate $\lambda + \mu$.

- We shall use the infinitesimal definition, according to which:
  - $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ have independent increments;
  - As $h \searrow 0$, uniformly in t,

$$\begin{array}{rcl} \mathbb{P}(X_{t+h} - X_t = 0) & = & 1 - \lambda h + o(h), \\ \mathbb{P}(X_{t+h} - X_t = 1) & = & \lambda h + o(h), \\ \mathbb{P}(Y_{t+h} - Y_t = 0) & = & 1 - \mu h + o(h), \\ \mathbb{P}(Y_{t+h} - Y_t = 1) & = & \mu h + o(h). \end{array}$$

  Set $Z_t = X_t + Y_t$.

  By hypothesis, $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are independent.

  So $(Z_t)_{t \geq 0}$ has independent increments.

Sum of Independent Poisson Processes (Cont'd)

- As $h \searrow 0$, uniformly in t,

$$\begin{aligned}
\mathbb{P}(Z_{t+h} - Z_t = 0) &= \mathbb{P}(X_{t+h} - X_t = 0)\mathbb{P}(Y_{t+h} - Y_t = 0) \\
&= (1 - \lambda h + o(h))(1 - \mu h + o(h)) \\
&= 1 - (\lambda + \mu)h + o(h); \\
\mathbb{P}(Z_{t+h} - Z_t = 1) &= \mathbb{P}(X_{t+h} - X_t = 1)\mathbb{P}(Y_{t+h} - Y_t = 0) \\
&\quad + \mathbb{P}(X_{t+h} - X_t = 0)\mathbb{P}(Y_{t+h} - Y_t = 1) \\
&= (\lambda h + o(h))(1 - \mu h + o(h)) \\
&\qquad + (1 - \lambda h + o(h))(\mu h + o(h)) \\
&= (\lambda + \mu)h + o(h).
\end{aligned}$$

Hence $(Z_t)_{t \geq 0}$ is a Poisson process of rate $\lambda + \mu$.

Jumps of Poisson Process and Uniform Distribution

Theorem

Let $(X_t)_{t\geq 0}$ be a Poisson process. Then, conditional on $(X_t)_{t\geq 0}$ having exactly one jump in the interval $[s, s+t]$, the time at which that jump occurs is uniformly distributed on $[s, s+t]$.

- We shall use the finite-dimensional distribution definition.

  By stationarity of increments, it suffices to consider the case $s = 0$.

  Then, for $0 \leq u \leq t$,

$$\begin{aligned}\mathbb{P}(J_1 \leq u | X_t = 1) &= \frac{\mathbb{P}(J_1 \leq u \text{ and } X_t = 1)}{\mathbb{P}(X_t = 1)} \\ &= \frac{\mathbb{P}(X_u = 1 \text{ and } X_t - X_u = 0)}{\mathbb{P}(X_t = 1)} \\ &= \frac{\lambda u e^{-\lambda u} e^{-\lambda(t-u)}}{\lambda t e^{-\lambda t}} \\ &= \frac{u}{t}.\end{aligned}$$

Joint Density Function of Jump Times

Theorem

Let $(X_t)_{t \geq 0}$ be a Poisson process. Then, conditional on the event $\{X_t = n\}$, the jump times $J_1, \ldots, J_n$ have joint density function

$$f(t_1, \ldots, t_n) = n! t^{-n} 1_{\{0 \leq t_1 \leq \cdots \leq t_n \leq t\}}.$$

Thus, conditional on $\{X_t = n\}$, the jump times $J_1, \ldots, J_n$ have the same distribution as an ordered sample of size n from the uniform distribution on $[0, t]$.

- The holding times $S_1, \ldots, S_{n+1}$ have joint density function
$$\lambda^{n+1} e^{-\lambda(s_1 + \cdots + s_{n+1})} 1_{\{s_1, \ldots, s_{n+1} \geq 0\}}.$$
So the jump times $J_1, \ldots, J_{n+1}$ have joint density function
$$\lambda^{n+1} e^{-\lambda t_{n+1}} 1_{\{0 \leq t_1 \leq \cdots \leq t_{n+1}\}}.$$

Joint Density Function of Jump Times (Cont'd)

- The jump times $J_1, \ldots, J_{n+1}$ have joint density function

$$\lambda^{n+1} e^{-\lambda t_{n+1}} 1_{\{0 \leq t_1 \leq \cdots \leq t_{n+1}\}}.$$

So, for $A \subseteq \mathbb{R}^n$, we have

$$\begin{aligned} & \mathbb{P}((J_1, \ldots, J_n) \in A \text{ and } X_t = n) \\ & = \mathbb{P}((J_1, \ldots, J_n) \in A \text{ and } J_n \leq t < J_{n+1}) \\ & = e^{-\lambda t} \lambda^n \int_{(t_1, \ldots, t_n) \in A} 1_{\{0 \leq t_1 \leq \cdots \leq t_n \leq t\}} dt_1 \cdots dt_n. \end{aligned}$$

Now $\mathbb{P}(X_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$.
So we obtain

$$\mathbb{P}((J_1, \ldots, J_n) \in A | X_t = n) = \int_A n! t^{-n} 1_{\{0 \leq t_1 \leq \cdots \leq t_n \leq t\}} dt_1 \cdots dt_n.$$

So $f(t_1, \ldots, t_n)$ is as claimed.

Example: Robins and Blackbirds

- Robins and blackbirds make brief visits to my birdtable.

  The probability that in any small interval of duration h a robin will arrive is found to be $\rho h + o(h)$.

  The corresponding probability for blackbirds is $\beta h + o(h)$.

  What is the probability that the first two birds I see are both robins?

  What is the distribution of the total number of birds seen in time t?

  Given that this number is n, what is the distribution of the number of blackbirds seen in time t?

Example: Robins and Blackbirds (Solution)

- By the infinitesimal characterization:
  - The number of robins seen by time t is a Poisson process $(R_t)_{t \geq 0}$ of rate ρ;
  - The number of blackbirds is a Poisson process $(B_t)_{t \geq 0}$ of rate β.

  The times spent waiting for the first robin or blackbird are independent exponential random variables:
  - S_1 of parameter ρ;
  - T_1 of parameter β.

  So a robin arrives first with probability $\frac{\rho}{\rho+\beta}$.

  By the memoryless property of T_1, the probability that the first two birds are robins is $\frac{\rho^2}{(\rho+\beta)^2}$.

  By a previous theorem,the total number of birds seen in an interval of duration t has Poisson distribution of parameter $(\rho + \beta)t$.

Example: Robins and Blackbirds (Solution Cont'd)

- Finally

$$\begin{aligned}\mathbb{P}(B_t = k | R_t + B_t = n) &= \frac{\mathbb{P}(B_t{=}k \text{ and } R_t{=}n{-}k)}{\mathbb{P}(R_t{+}B_t{=}n)} \\ &= \frac{\frac{e^{-\beta}\beta^k}{k!}\frac{e^{-\rho}\rho^{n-k}}{(n-k)!}}{\frac{e^{-(\rho+\beta)}(\rho+\beta)^n}{n!}} \\ &= \binom{n}{k}\left(\frac{\beta}{\rho+\beta}\right)^k\left(\frac{\rho}{\rho+\beta}\right)^{n-k}.\end{aligned}$$

So if n birds are seen in time t, then the distribution of the number of blackbirds is binomial of parameters n and $\frac{\beta}{\rho+\beta}$.

Subsection 5

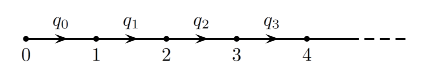
Birth Processes

Birth Processes

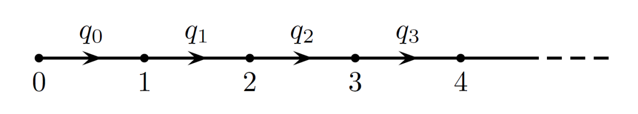
- A birth process is a generalization of a Poisson process in which the parameter λ is allowed to depend on the current state of the process.
- The data for a birth process consist of birth rates

$$0 \leq q_j < \infty, \quad j = 0, 1, 2, \ldots.$$

- We begin with a definition in terms of jump chain and holding times.
- A minimal right-continuous process $(X_t)_{t \geq 0}$ with values in $\{0, 1, 2, \ldots\} \cup \{\infty\}$ is a **birth process of rates** $(q_j : j \geq 0)$ if, conditional on $X_0 = i$:
  - Its holding times $S_1, S_2, \ldots$ are independent exponential random variables of parameters $q_i, q_{i+1}, \ldots$, respectively;
  - Its jump chain is given by $Y_n = i + n$.



Birth Processes (Q-Matrix)



- The Q-matrix is

$$Q = \begin{pmatrix} -q_0 & q_0 & & & \\ & -q_1 & q_1 & & \\ & & -q_2 & q_2 & \\ & & & \ddots & \ddots \\ & & & & \end{pmatrix}.$$

Example (Simple Birth Process)

- Consider a population in which each individual gives birth after an exponential time of parameter λ, all independently.

  If i individuals are present then the first birth will occur after an exponential time of parameter $i\lambda$.

  Then we have $i+1$ individuals and, by the memoryless property, the process begins afresh.

  Thus the size of the population performs a birth process with rates

$$q_i = i\lambda.$$

  Let X_t denote the number of individuals at time t.

  Suppose $X_0 = 1$.

  Write T for the time of the first birth.

Example (Simple Birth Process Cont'd)

- Now we have

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}(X_t 1_{T \le t}) + \mathbb{E}(X_t 1_{T > t}) \\ &= \int_0^t \lambda e^{-\lambda s} \mathbb{E}(X_t | T = s) ds + \int_t^\infty \lambda e^{-\lambda s} \mathbb{E}(X_t | T = s) ds \\ &= \int_0^t \lambda e^{-\lambda s} \mathbb{E}(X_t | T = s) ds + \int_t^\infty \lambda e^{-\lambda s} ds \\ &= \int_0^t \lambda e^{-\lambda s} \mathbb{E}(X_t | T = s) ds + e^{-\lambda t}. \end{aligned}$$

Put $\mu(t) = \mathbb{E}(X_t)$.

Then

$$\mathbb{E}(X_t | T = s) = 2\mu(t - s).$$

So

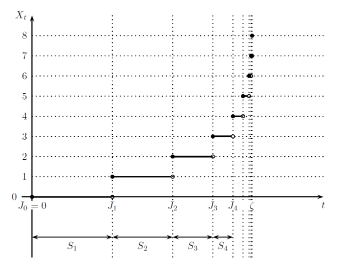
$$\mu(t) = \int_0^t 2\lambda e^{-\lambda s} \mu(t - s) ds + e^{-\lambda t}.$$

Example (Simple Birth Process Cont'd)

- We found $\mu(t) = \int_0^t 2\lambda e^{-\lambda s}\mu(t-s)ds + e^{-\lambda t}$.
  Setting $r = t - s$,
  $$\begin{aligned} \mu(t) &= \int_0^t 2\lambda e^{\lambda(r-t)}\mu(r)dr + e^{-\lambda t} \\ &= 2\lambda e^{-\lambda t}\int_0^t e^{\lambda r}\mu(r)dr + e^{-\lambda t}. \end{aligned}$$
  So
  $$e^{\lambda t}\mu(t) = 2\lambda \int_0^t e^{\lambda r}\mu(r)dr + 1.$$
  By differentiating we obtain
  $$\lambda e^{\lambda t}\mu(t) + e^{\lambda t}\mu'(t) = 2\lambda e^{\lambda t}\mu(t)$$
  $$\mu'(t) = \lambda\mu(t).$$
  So the mean population size grows exponentially,
  $$\mathbb{E}(X_t) = e^{\lambda t}.$$

Explosion in Birth Processes

- Much of the theory associated with the Poisson process goes through for birth processes with little change.
- But some calculations can no longer be made so explicitly.
- The most interesting new phenomenon present in birth processes is the possibility of explosion.
- For certain choices of birth rates, a typical path will make infinitely many jumps in a finite time.



- The convention of setting the process to equal ∞ after explosion is particularly appropriate for birth processes!

Explosion Time of Birth Processes and Markov Property

Theorem

Let $(X_t)_{t\geq 0}$ be a birth process of rates $(q_j : j \geq 0)$, starting from 0.

(i) If $\sum_{j=0}^{\infty} \frac{1}{q_j} < \infty$, then $\mathbb{P}(\zeta < \infty) = 1$.

(ii) If $\sum_{j=0}^{\infty} \frac{1}{q_j} = \infty$, then $\mathbb{P}(\zeta = \infty) = 1$.

- We apply a previous theorem to the sequence of holding times $S_1, S_2, \ldots$.

Theorem (Markov Property)

Let $(X_t)_{t\geq 0}$ be a birth process of rates $(q_j : j \geq 0)$. Then, conditional on $X_s = i$, $(X_{s+t})_{t\geq 0}$ is a birth process of rates $(q_j : j \geq 0)$ starting from i and independent of $(X_r : r \leq s)$.

Setting for the Fundamental Theorem of Birth Processes

- We shall shortly prove a theorem on birth processes which generalizes the key theorem on Poisson processes.
- The Poisson probabilities arose as the unique solution of a system of differential equations, essentially the forward equations.
- Now we can still write down the forward equation

$$P'(t) = P(t)Q, \quad P(0) = I.$$

In components

$$p'_{i0}(t) = -p_{i0}(t)q_0, \quad p_{i0}(0) = \delta_{i0};$$

For $j = 1, 2, \ldots,$

$$p'_{ij}(t) = p_{i,j-1}(t)q_{j-1} - p_{ij}(t)q_j, \quad p_{ij}(0) = \delta_{ij}.$$

- These equations still have a unique solution.
- But it is not as explicit as before.

Setting for the Fundamental Theorem (Cont'd)

- We must have
$$p_{i0}(t) = \delta_{i0} e^{-q_0 t}.$$
- This can be substituted in the equation
$$p'_{i1}(t) = p_{i0}(t) q_0 - p_{i1}(t) q_1, \quad p_{i1}(0) = \delta_{i1}.$$
- This equation can be solved to give
$$p_{i1}(t) = \delta_{i1} e^{-q_1 t} + \delta_{i0} \int_0^t q_0 e^{-q_0 s} e^{-q_1 (t-s)} ds.$$
- Now we can substitute for $p_{i1}(t)$ in the next equation up the hierarchy and find an explicit expression for $p_{i2}(t)$, and so on.

Key Theorem of Birth Processes

Theorem

Let $(X_t)_{t\geq 0}$ be an increasing, right-continuous process with values in $\{0,1,2,\ldots\}\cup\{\infty\}$. Let $0\leq q_j<\infty$, for all $j\geq 0$. Then the following three conditions are equivalent:

(a) (**Jump Chain/Holding Time Definition**) Conditional on $X_0=i$, the holding times $S_1,S_2,\ldots$ are independent exponential random variables of parameters $q_i,q_{i+1},\ldots$, respectively, and the jump chain is given by $Y_n=i+n$ for all n;

(b) (**Infinitesimal Definition**) For all $t,h\geq 0$, conditional on $X_t=i$, X_{t+h} is independent of $(X_s:s\leq t)$ and, as $h\searrow 0$, uniformly in t,

$$\begin{aligned}P(X_{t+h}=i|X_t=i) &= 1-q_ih+o(h),\\ P(X_{t+h}=i+1|X_t=i) &= q_ih+o(h);\end{aligned}$$

Key Theorem of Birth Processes (Cont'd)

Theorem

(c) (**Transition Probability Definition**) For all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq \cdots \leq t_{n+1}$ and all states $i_0, \ldots, i_{n+1}$,

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n),$$

where $(p_{ij}(t) : i, j = 0, 1, 2, \ldots)$ is the unique solution of the forward equations.

If $(X_t)_{t \geq 0}$ satisfies any of these conditions then it is called a **birth process of rates** $(q_j : j \geq 0)$.

- Suppose Condition (a) holds.

  By the Markov Property, for any $t, h \geq 0$, conditional on $X_t = i$, X_{t+h} is independent of $(X_s : s \leq t)$.

Proof ((a)⇒(b))

As $h \searrow 0$, uniformly in t,

$$\begin{aligned}\mathbb{P}(X_{t+h} \geq i+1 | X_t = i) &= \mathbb{P}(X_h \geq i+1 | X_0 = i) \\ &= \mathbb{P}(J_1 \leq h | X_0 = i) \\ &= 1 - e^{-q_i h} \\ &= q_i h + o(h).\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{P}(X_{t+h} \geq i+2 | X_t = i) &= \mathbb{P}(X_h \geq i+2 | X_0 = i) \\ &= \mathbb{P}(J_2 \leq h | X_0 = i) \\ &\leq \mathbb{P}(S_1 \leq h \text{ and } S_2 \leq h | X_0 = i) \\ &= (1 - e^{-q_i h})(1 - e^{-q_{i+1} h}) \\ &= o(h).\end{aligned}$$

This implies Condition (b).

Proof ((b)⇒(c))

- If (b) holds, then, for $k = i+2, i+3, \ldots$, as $h \searrow 0$, uniformly in t

$$\mathbb{P}(X_{t+h} = k | X_t = i) = o(h).$$

Set

$$p_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i).$$

Then, for $j = 1, 2, \ldots$,

$$\begin{aligned} p_{ij}(t+h) &= \mathbb{P}(X_{t+h} = j | X_0 = i) \\ &= \textstyle\sum_{k=i}^{j} \mathbb{P}(X_t = k | X_0 = i) \mathbb{P}(X_{t+h} = j | X_t = k) \\ &= p_{ij}(t)(1 - q_j h + o(h)) + p_{i,j-1}(t)(q_{j-1} h + o(h)) + o(h). \end{aligned}$$

So

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = p_{i,j-1}(t) q_{j-1} - p_{ij}(t) q_j + O(h).$$

Proof ((b)⇒(c) Cont'd)

- As in the proof of a previous theorem, we can deduce that:
  - $p_{ij}(t)$ is differentiable;
  - Satisfies the differential equation

$$p'_{ij}(t) = p_{i,j-1}(t)q_{j-1} - p_{ij}(t)q_j.$$

By a simpler argument we also find

$$p'_{i0}(t) = -p_{i0}(t)q_0.$$

Thus

$$(p_{ij}(t) : i, j = 0, 1, 2, \ldots)$$

must be the unique solution to the forward equations.

Proof ((b)⇒(c) Cont'd)

- If $(X_t)_{t\geq 0}$ satisfies Condition (b), then certainly

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_0 = i_0, \ldots, X_{t_n} = i_n) = \mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n).$$

But also $(X_{t_n+t})_{t\geq 0}$ satisfies Condition (b).

So, by uniqueness for the forward equations, we have

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n).$$

Hence $(X_t)_{t\geq 0}$ satisfies Condition (c).

(c)⇒(a) This mimics the proof of the implication (c)⇒(a) of the theorem for Markov processes.

Subsection 6

Jump Chain and Holding Times

Q-Matrices Revisited

- Let I be a countable set.
- The basic data for a continuous-time Markov chain on I are given in the form of a Q-matrix.
- Recall that a Q-matrix on I is any matrix $Q = (q_{ij} : i, j \in I)$ which satisfies the following conditions:
  (i) $0 \leq -q_{ii} < \infty$, for all i;
  (ii) $q_{ij} \geq 0$, for all $i \neq j$;
  (iii) $\sum_{j \in I} q_{ij} = 0$, for all i.
- We will sometimes find it convenient to write q_i or $q(i)$ as an alternative notation for $-q_{ii}$.

From a Q-Matrix to a Stochastic Matrix

- We are going to describe a simple procedure for obtaining from a Q-matrix Q a stochastic matrix Π.
- The **jump matrix** $\Pi = (\pi_{ij} : i, j \in I)$ of Q is defined by

$$\pi_{ij} = \begin{cases} \frac{q_{ij}}{q_i}, & \text{if } j \neq i \text{ and } q_i \neq 0, \\ 0, & \text{if } j \neq i \text{ and } q_i = 0, \end{cases} \qquad \pi_{ii} = \begin{cases} 0, & \text{if } q_i \neq 0, \\ 1, & \text{if } q_i = 0. \end{cases}$$

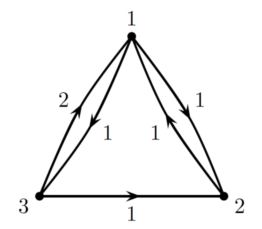
- This procedure is best thought of row by row:
  - For each $i \in I$, we take, where possible, the off-diagonal entries in the i-th row of Q and scale them so they add up to 1, putting a 0 on the diagonal.
  - This is only impossible when the off-diagonal entries are all 0. Then we leave them alone and put a 1 on the diagonal.

Example

- The Q-matrix

$$Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$$

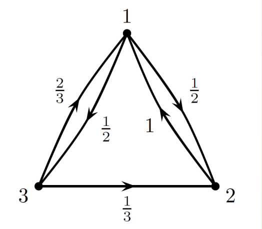
has the diagram on the right.



The jump matrix Π of Q is given by

$$\Pi = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

It has the diagram on the right.



Markov Chains

- Recall that a **minimal process** is one which is set equal to ∞ after any explosion.
- A minimal right-continuous process $(X_t)_{t\geq 0}$ on I is a **Markov chain with initial distribution** λ **and generator matrix** Q if:
  - Its jump chain $(Y_n)_{n\geq 0}$ is discrete-time Markov(λ, Π);
  - For each $n \geq 1$, conditional on $Y_0, \ldots, Y_{n-1}$, its holding times $S_1, \ldots, S_n$ are independent exponential random variables of parameters $q(Y_0), \ldots, q(Y_{n-1})$, respectively.
- We say $(X_t)_{t\geq 0}$ is Markov(λ, Q) for short.

Construction of a Markov Chain

- We can construct such a process as follows:
  - Let $(Y_n)_{n\geq 0}$ be discrete-time Markov(λ, Π);
  - Let $T_1, T_2, \ldots$ be independent exponential random variables of parameter 1, independent of $(Y_n)_{n\geq 0}$.
- Set:
  - $S_n = \frac{T_n}{q(Y_{n-1})}$;
  - $J_n = S_1 + \cdots + S_n$;
  - $X_t = \begin{cases} Y_n, & \text{if } J_n \leq t < J_{n+1} \text{ for some } n, \\ \infty, & \text{otherwise.} \end{cases}$.
- Then $(X_t)_{t\geq 0}$ has the required properties.

A Second Construction

- Begin with:
  - An initial state $X_0 = Y_0$ with distribution λ;
  - An array $(T_n^j : n \geq 1, j \in I)$ of independent exponential random variables of parameter 1.
- Then, inductively for $n = 0, 1, 2, \ldots$, if $Y_n = i$, set:
  - $S_{n+1}^j = \frac{T_{n+1}^j}{q_{ij}}$, for $j \neq i$;
  - $S_{n+1} = \inf_{j \neq i} S_{n+1}^j$;
  - $Y_{n+1} = \begin{cases} j, & \text{if } S_{n+1}^j = S_{n+1} < \infty, \\ i, & \text{if } S_{n+1} = \infty. \end{cases}$

A Second Construction (Cont'd)

- Conditional on $Y_n = i$, the random variables S^j_{n+1} are independent exponentials of parameter q_{ij} for all $j \neq i$.
- So, by a previous theorem, conditional on $Y_n = i$:
  - S_{n+1} is exponential of parameter $q_i = \sum_{j \neq i} q_{ij}$;
  - Y_{n+1} has distribution $(\pi_{ij} : j \in I)$;
  - S_{n+1} and Y_{n+1} are independent, and independent of $Y_0, \ldots, Y_n$ and $S_1, \ldots, S_n$.
- This construction presents a justification for calling:
  - q_i the **rate of leaving** i;
  - q_{ij} the **rate of going from** i **to** j.

Introducing a Third Construction

- Our third construction of a Markov chain with generator matrix Q and initial distribution λ is based on the Poisson process.
- Imagine the state-space I as a labyrinth of chambers and passages.
- Each passage is shut off by a single door which opens briefly from time to time to allow us through in one direction only.
- Suppose the door giving access to chamber j from chamber i opens at the jump times of a Poisson process of rate q_{ij}.
- We take every chance we can to move.
- Then we will perform a Markov chain with Q-matrix Q.

A Third Construction

- We begin with:
  - An initial state $X_0 = Y_0$ with distribution λ;
  - A family of independent Poisson processes $\{(N_t^{ij})_{t\geq 0} : i, j \in I, i \neq j\}$, $(N_t^{ij})_{t\geq 0}$ having rate q_{ij}.
- We set $J_0 = 0$.
- We define inductively for $n = 0, 1, 2, \ldots$,

$$\begin{aligned} J_{n+1} &= \inf\left\{t > J_n : N_t^{Y_n j} \neq N_{J_n}^{Y_n j} \text{ for some } j \neq Y_n\right\}; \\ Y_{n+1} &= \begin{cases} j, & \text{if } J_{n+1} < \infty \text{ and } N_{J_{n+1}}^{Y_n j} \neq N_{J_n}^{Y_n j}, \\ i, & \text{if } J_{n+1} = \infty. \end{cases} \end{aligned}$$

- The first jump time of $(N_t^{ij})_{t\geq 0}$ is exponential of parameter q_{ij}.
- So, by a previous theorem, conditional on $Y_0 = i$:
  - J_1 is exponential of parameter $q_i = \sum_{j\neq i} q_{ij}$;
  - Y_1 has distribution $(\pi_{ij} : j \in I)$;
  - J_1 and Y_1 are independent.

A Third Construction (Cont'd)

- Now suppose T is a stopping time of $(X_t)_{t\geq 0}$.
- Suppose we condition on X_0 and on the processes $(N_t^{k\ell})_{t\geq 0}$ for $(k,\ell)\neq(i,j)$, which are independent of N_t^{ij}.
- Then $\{T\leq t\}$ depends only on $(N_s^{ij}:s\leq t)$.
- By the Strong Markov Property of the Poisson process,

$$\widetilde{N}_t^{ij}:=N_{T+t}^{ij}-N_T^{ij}$$

  is a Poisson process of rate q_{ij} independent of $(N_s^{ij}:s\leq T)$, and independent of X_0 and $(N_t^{k\ell})_{t\geq 0}$ for $(k,\ell)\neq(i,j)$.
- Hence, conditional on $T<\infty$ and $X_T=i$, $(X_{T+t})_{t\geq 0}$ has the same distribution as $(X_t)_{t\geq 0}$ and is independent of $(X_s:s\leq T)$.

A Third Construction (Cont'd)

- In particular, we can take $T = J_n$.
- We see that, conditional on $J_n < \infty$ and $Y_n = i$:
  - S_{n+1} is exponential of parameter q_i;
  - Y_{n+1} has distribution $(\pi_{ij} : j \in I)$;
  - S_{n+1} and Y_{n+1} are independent, and independent of $Y_0, \ldots, Y_n$ and $S_1, \ldots, S_n$.
- Hence, $(X_t)_{t \geq 0}$ is Markov(λ, Q).
- Moreover, $(X_t)_{t \geq 0}$ has the Strong Markov Property.
- The conditioning on which this argument relies requires some further justification, especially when the state-space is infinite.
- So we avoid relying on this third construction in the development of the theory.

Subsection 7

Explosion

Explosion Time

- Consider a process with:
  - Jump times $J_0, J_1, J_2, \ldots$;
  - Holding times $S_1, S_2, \ldots$.
- The **explosion time** ζ is given by

$$\zeta = \sup_n J_n = \sum_{n=1}^{\infty} S_n.$$

Sufficient Conditions for Non-Explosion

Theorem

Let $(X_t)_{t\geq 0}$ be Markov(λ, Q). Then $(X_t)_{t\geq 0}$ does not explode if any one of the following conditions holds:

(i) I is finite;

(ii) $\sup_{i\in I} q_i < \infty$;

(iii) $X_0 = i$, and i is recurrent for the jump chain.

- Set $T_n = q(Y_{n-1})S_n$.

  Then $T_1, T_2, \ldots$ are independent $E(1)$ and independent of $(Y_n)_{n\geq 0}$.

  In Cases (i) and (ii), we have:

  - $q = \sup_i q_i < \infty$;
  - $q\zeta \geq \sum_{n=1}^{\infty} T_n = \infty$ with probability 1.

  In Case (iii), we know that $(Y_n)_{n\geq 0}$ visits i infinitely often, at times $N_1, N_2, \ldots$, say. Then $q_i\zeta \geq \sum_{m=1}^{\infty} T_{N_m+1} = \infty$ with probability 1.

Explosive Q-Matrices

- We denote by $\mathbb{P}_i$ the conditional probability

$$\mathbb{P}_i(A) = \mathbb{P}(A | X_0 = i).$$

- It is a simple consequence of the Markov property for $(Y_n)_{n \geq 0}$ that, under $\mathbb{P}_i$, the process $(X_t)_{t \geq 0}$ is Markov(δ_i, Q).
- We say that a Q-matrix Q is **explosive** if, for the associated Markov chain

$$\mathbb{P}_i(\zeta < \infty) > 0, \text{for some } i \in I.$$

- Otherwise Q is **non-explosive**.
- The result just proved gives simple conditions for non-explosion and covers many cases of interest.

Expectation of the Exponential of Explosion Time

Theorem

Let $(X_t)_{t\geq 0}$ be a continuous-time Markov chain with generator matrix Q and write ζ for the explosion time of $(X_t)_{t\geq 0}$. Fix $\theta > 0$ and set $z_i = \mathbb{E}_i(e^{-\theta\zeta})$. Then $z = (z_i : i \in I)$ satisfies:

(i) $|z_i| \leq 1$, for all i;

(ii) $Qz = \theta z$.

Moreover, if $\widetilde{z}$ also satisfies (i) and (ii), then $\widetilde{z}_i \leq z_i$, for all i.

- Condition on $X_0 = i$.

  The time and place of the first jump are independent.

  Also, J_1 is $E(q_i)$ and $\mathbb{P}_i(X_{J_1} = k) = \pi_{ik}$.

  By the Markov Property of the jump chain at time $n = 1$, conditional on $X_{J_1} = k$, $(X_{J_1+t})_{t\geq 0}$ is Markov(δ_k, Q) and independent of J_1.

Exponential of Explosion Time (Cont'd)

- So we have

$$\begin{aligned}\mathbb{E}_i(e^{-\theta\zeta}|X_{J_1}=k) &= \mathbb{E}_i(e^{-\theta J_1}e^{-\theta\sum_{n=2}^{\infty}S_n}|X_{J_1}=k) \\ &= \int_0^{\infty}e^{-\theta t}q_ie^{-q_it}dt\mathbb{E}_k(e^{-\theta\zeta}) \\ &= \frac{q_iz_k}{q_i+\theta}.\end{aligned}$$

So

$$z_i=\sum_{k\neq i}\mathbb{P}_i(X_{J_1}=k)\mathbb{E}_i(e^{-\theta\zeta}|X_{J_1}=k)=\sum_{k\neq i}\frac{q_i\pi_{ik}z_k}{q_i+\theta}.$$

Recall that $q_i=-q_{ii}$ and $q_i\pi_{ik}=q_{ik}$.

Then

$$(\theta-q_{ii})z_i=(\theta+q_i)\sum_{k\neq i}\frac{q_i\pi_{ik}z_k}{q_i+\theta}=\sum_{k\neq i}q_{ik}z_k.$$

So $\theta z_i=q_{ii}z_i+\sum_{k\neq i}q_{ik}z_k=\sum_{k\in I}q_{ik}z_k$.

So z satisfies (i) and (ii).

Exponential of Explosion Time (Cont'd)

- Note that the same argument also shows that

$$\mathbb{E}_i(e^{-\theta J_{n+1}}) = \sum_{k \neq i} \frac{q_i \pi_{ik}}{q_i + \theta} \mathbb{E}_k(e^{-\theta J_n}).$$

Suppose that $\widetilde{z}$ also satisfies Conditions (i) and (ii).

Then, in particular, $\widetilde{z}_i \leq 1 = \mathbb{E}_i(e^{-\theta J_0})$, for all i.

Suppose inductively that $\widetilde{z}_i \leq \mathbb{E}_i(e^{-\theta J_n})$.

Then, since $\widetilde{z}$ satisfies Condition (ii),

$$\widetilde{z}_i = \sum_{k \neq i} \frac{q_i \pi_{ik}}{q_i + \theta} \widetilde{z}_k \leq \sum_{k \neq i} \frac{q_i \pi_{ik}}{q_i + \theta} \mathbb{E}_i(e^{-\theta J_n}) = \mathbb{E}_i(e^{-\theta J_{n+1}}).$$

Hence, $\widetilde{z}_i \leq \mathbb{E}_i(e^{-\theta J_n})$, for all n.

By Monotone Convergence, $\mathbb{E}_i(e^{-\theta J_n}) \to \mathbb{E}_i(e^{-\theta \zeta})$ as $n \to \infty$.

So $\widetilde{z}_i \leq z_i$, for all i.

Characterization of Non-Explosiveness

Corollary

For each $\theta > 0$, the following are equivalent:

(a) Q is non-explosive;

(b) $Qz = \theta z$ and $|z_i| \leq 1$, for all i, imply $z = 0$.

- Suppose Condition (a) holds.
  Then $\mathbb{P}_i(\zeta = \infty) = 1$.
  So $\mathbb{E}_i(e^{-\theta\zeta}) = 0$.
  By the theorem, $Qz = \theta z$ and $|z| \leq 1$ imply $z_i \leq \mathbb{E}_i(e^{-\theta\zeta})$.
  Hence $z \leq 0$. By symmetry $z \geq 0$. Hence (b) holds.
  Conversely, suppose Condition (b) holds.
  Then, by the theorem, $\mathbb{E}_i(e^{-\theta\zeta}) = 0$, for all i.
  So $\mathbb{P}_i(\zeta = \infty) = 1$.
  This proves (a).

Subsection 8

Forward and Backward Equation

Strong Markov Property for Birth Processes

- Recall that a random variable T with values in $[0, \infty]$ is a **stopping time of** $(X_t)_{t \geq 0}$ if, for each $t \in [0, \infty)$, the event $\{T \leq t\}$ depends only on $(X_s : s \leq t)$.

Theorem (Strong Markov Property)

Let $(X_t)_{t \geq 0}$ be Markov(λ, Q) and let T be a stopping time of $(X_t)_{t \geq 0}$. Then, conditional on $T < \infty$ and $X_T = i$, $(X_{T+t})_{t \geq 0}$ is Markov(δ_i, Q) and independent of $(X_s : s \leq T)$.

Key Theorem of Birth Processes

Theorem

Let $(X_t)_{t\geq 0}$ be a right-continuous process with values in a finite set I. Let Q be a Q-matrix on I with jump matrix Π. Then the following are equivalent:

(a) (**Jump Chain/Holding Time Definition**) Conditional on $X_0 = i$:
- The jump chain $(Y_n)_{n\geq 0}$ of $(X_t)_{t\geq 0}$ is discrete-time Markov(δ_i, Π);
- For each $n \geq 1$, conditional on $Y_0, \ldots, Y_{n-1}$, the holding times $S_1, \ldots, S_n$ are independent exponential random variables of parameters $q(Y_0), \ldots, q(Y_{n-1})$, respectively;

(b) (**Infinitesimal Definition**) For all $t, h \geq 0$, conditional on $X_t = i$, X_{t+h} is independent of $(X_s : s \leq t)$ and, as $h \searrow 0$, uniformly in t, for all j,

$$\mathbb{P}(X_{t+h} = j | X_t = i) = \delta_{ij} + q_{ij} h + o(h);$$

Key Theorem of Birth Processes (Cont'd)

Theorem (Cont'd)

(c) (**Transition Probability Definition**) For all $n = 0, 1, 2, \ldots$, all times $0 \le t_0 \le t_1 \le \cdots \le t_{n+1}$ and all states $i_0, \ldots, i_{n+1}$,

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n),$$

where $(p_{ij}(t) : i, j \in I, t \ge 0)$ is the solution of the forward equation

$$P'(t) = P(t)Q, \quad P(0) = I.$$

If $(X_t)_{t \ge 0}$ satisfies any of these conditions, then it is called a **Markov chain with generator matrix** Q. We say that $(X_t)_{t \ge 0}$ is Markov(λ, Q) for short, where λ is the distribution of X_0.

Proof ((a)⇒(b))

- Suppose (a) holds. Then, as $h \searrow 0$,

  $$\mathbb{P}_i(X_h = i) \geq \mathbb{P}_i(J_1 > h) = e^{-q_i h} = 1 + q_{ii} h + o(h).$$

  For $j \neq i$, we have

  $$\begin{aligned} \mathbb{P}_i(X_h = j) &\geq \mathbb{P}_i(J_1 \leq h, Y_1 = j, S_2 > h) \\ &= (1 - e^{-q_i h})\pi_{ij} e^{-q_j h} \\ &= q_{ij} h + o(h). \end{aligned}$$

  Thus, for every state j, $\mathbb{P}_i(X_h = j) \geq \delta_{ij} + q_{ij} h + o(h)$.
  By taking the finite sum over j, we see that these must be equalities.
  By the Markov Property, for any $t, h \geq 0$, conditional on $X_t = i$, X_{t+h} is independent of $(X_s : s \leq t)$.
  As $h \searrow 0$, uniformly in t,

  $$\mathbb{P}(X_{t+h} = j | X_t = i) = \mathbb{P}_i(X_h = j) = \delta_{ij} + q_{ij} h + o(h).$$

Proof ((b)⇒(c))

- Set $p_{ij}(t) = \mathbb{P}_i(X_t = j) = \mathbb{P}(X_t = j | X_0 = i)$.
  If (b) holds, then for all $t, h \geq 0$, as $h \searrow 0$, uniformly in t,
$$\begin{aligned} p_{ij}(t+h) &= \sum_{k \in I} \mathbb{P}_i(X_t = k)\mathbb{P}(X_{t+h} = j | X_t = k) \\ &= \sum_{k \in I} p_{ik}(t)(\delta_{kj} + q_{kj}h + o(h)). \end{aligned}$$
  Since I is finite, we have
$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_{k \in I} p_{ik}(t)q_{kj} + O(h).$$
  So, letting $h \searrow 0$, we see that $p_{ij}(t)$ is differentiable on the right.
  By uniformity, we can replace t by $t - h$ and let $h \searrow 0$ to get:
  - $p_{ij}(t)$ is continuous on the left;
  - $p_{ij}(t)$ is differentiable on the left, hence differentiable;
  - $p_{ij}(t)$ satisfies the forward equations
$$p'_{ij}(t) = \sum_{k \in I} p_{ik}(t)q_{kj}, \quad p_{ij}(0) = \delta_{ij}.$$

Proof ((b)⇒(c) Cont'd)

- By a previous theorem, since I is finite, $p_{ij}(t)$ is the unique solution of

$$p'_{ij}(t) = \sum_{k \in I} p_{ik}(t) q_{kj}, \quad p_{ij}(0) = \delta_{ij}.$$

Also, if (b) holds, then

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = \mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n).$$

Moreover, (b) holds for $(X_{t_n+t})_{t \geq 0}$.

So, by the above argument,

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n).$$

This proves (c).

Proof ((c)⇒(a))

- (c)⇒(a) again mimics the one for Poisson processes.

  There is a process satisfying Part (a).

  We have shown that it must then satisfy Part (c).

  But Condition (c) determines the finite-dimensional distributions of

$$(X_t)_{t\geq 0}.$$

  Hence it determines the distribution of jump chain and holding times.

  So if a process satisfying Part (c) also satisfies Part (a), so must every process satisfying Part (c).

Infinite State Spaces

- For infinite state space, the backward equation may still be written in the form
$$P'(t) = QP(t), \quad P(0) = I.$$
- We have an infinite system of differential equations
$$p'_{ij}(t) = \sum_{k \in I} q_{ik} p_{kj}(t), \quad p_{ij}(0) = \delta_{ij}.$$
- So the results on matrix exponentials no longer apply.
- A solution to the backward equation is any matrix
$$(p_{ij}(t) : i, j \in I)$$
of differentiable functions satisfying this system of differential equations.

Infinite State Spaces

Theorem

Let Q be a Q-matrix. Then the backward equation

$$P'(t) = QP(t), \quad P(0) = I,$$

has a minimal non-negative solution $(P(t) : t \geq 0)$. This solution forms a matrix semigroup

$$P(s)P(t) = P(s+t), \quad \text{for all } s, t \geq 0.$$

- We shall prove this result by a probabilistic method in combination with the following result.
- Note that, if I is finite, we must have $P(t) = e^{tQ}$.
- We call $(P(t) : t \geq 0)$ the **minimal non-negative semigroup** associated to Q, or simply the **semigroup** of Q.

Markov Chains with Infinite State Spaces

Theorem

Let $(X_t)_{t\geq 0}$ be a minimal right continuous process with values in I. Let Q be a Q-matrix on I with jump matrix Π and semigroup $(P(t) : t \geq 0)$. Then the following conditions are equivalent:

(a) (**Jump Chain/Holding Time Definition**) Conditional on $X_0 = i$:
- The jump chain $(Y_n)_{n\geq 0}$ of $(X_t)_{t\geq 0}$ is discrete time Markov(δ_i, Π);
- For each $n \geq 1$, conditional on $Y_0, \ldots, Y_{n-1}$, the holding times $S_1, \ldots, S_n$ are independent exponential random variables of parameters $q(Y_0), \ldots, q(Y_{n-1})$ respectively;

(b) (**Transition Probability Definition**) For all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n+1}$ and all states $i_0, i_1, \ldots, i_{n+1}$,

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n).$$

If $(X_t)_{t\geq 0}$ satisfies any of these conditions, it is called a **Markov chain with generator matrix** Q. We say $(X_t)_{t\geq 0}$ is Markov(λ, Q) for short, where λ is the distribution of X_0.

Combined Proof of the Theorems (Step 1)

- We know that there exists a process $(X_t)_{t\geq 0}$ satisfying (a). Define $P(t)$ by
$$p_{ij}(t) = \mathbb{P}_i(X_t = j).$$
Step 1: $P(t)$ satisfies the backward equation.
Conditional on $X_0 = i$ we have:
  - $J_1 \sim E(q_i)$;
  - $X_{J_1} \sim (\pi_{ik} : k \in I)$.

Conditional on $J_1 = s$ and $X_{J_1} = k$, $(X_{s+t})_{t\geq 0} \sim \text{Markov}(\delta_k, Q)$. So
$$\mathbb{P}_i(X_t = j, t < J_1) = e^{-q_i t}\delta_{ij};$$
$$\mathbb{P}_i(J_1 \leq t, X_{J_1} = k, X_t = j) = \int_0^t q_i e^{-q_i s}\pi_{ik} p_{kj}(t-s)ds.$$
Therefore,
$$\begin{aligned} p_{ij}(t) &= \mathbb{P}_i(X_t = j, t < J_1) + \sum_{k\neq i}\mathbb{P}_i(J_1 \leq t, X_{J_1} = k, X_t = j) \\ &= e^{-q_i t}\delta_{ij} + \sum_{k\neq i}\int_0^t q_i e^{-q_i s}\pi_{ik} p_{kj}(t-s)ds. \end{aligned}$$

Combined Proof of the Theorems (Step 1 Cont'd)

- We derived

$$p_{ij}(t) = e^{-q_i t}\delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} p_{kj}(t-s) ds.$$

Change variable $u = t - s$ in each of the integrals.

Interchange sum and integral by Monotone Convergence.

Multiply by $e^{q_i t}$ to obtain

$$e^{q_i t} p_{ij}(t) = \delta_{ij} + \int_0^t \sum_{k \neq i} q_i e^{q_i u} \pi_{ik} p_{kj}(u) du.$$

This equation shows that:

  - $p_{ij}(t)$ is continuous in t for all i, j;
  - The integrand is a uniformly converging sum of continuous functions. So it is continuous.

Combined Proof of the Theorems (Step 1 Cont'd)

- Hence, $p_{ij}(t)$ is differentiable in t and satisfies

$$e^{q_i t}(q_i p_{ij}(t) + p'_{ij}(t)) = \sum_{k \neq i} q_i e^{q_i t} \pi_{ik} p_{kj}(t).$$

Recall that $q_i = -q_{ii}$ and $q_{ik} = q_i \pi_{ik}$, for $k \neq i$.

Then, on rearranging, we obtain

$$p'_{ij}(t) = \sum_{k \in I} q_{ik} p_{kj}(t).$$

So $P(t)$ satisfies the backward equation.

The integral equation

$$p_{ij}(t) = e^{-q_i t}\delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} p_{kj}(t-s) ds$$

is called the **integral form of the backward equation**.

Combined Proof of the Theorems (Step 2)

Step 2: If $\widetilde{P}(t)$ is another non-negative solution of the backward equation, then $P(t) \leq \widetilde{P}(t)$, hence $P(t)$ is the minimal non-negative solution.

The argument used to prove the integral form also shows that

$$\begin{aligned}&\mathbb{P}_i(X_t = j, t < J_{n+1}) \\ &= e^{-q_i t}\delta_{ij} + \textstyle\sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} \mathbb{P}_k(X_{t-s} = j, t - s < J_n) ds.\end{aligned}$$

If $\widetilde{P}(t)$ satisfies the backward equation, then, by reversing the steps in the last part of Step 1, it also satisfies the integral form:

$$\widetilde{p}_{ij}(t) = e^{-q_i t}\delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} \widetilde{p}_{kj}(t - s) ds.$$

Combined Proof of the Theorems (Step 2 Cont'd)

- If $\widetilde{P}(t) \geq 0$, then $\mathbb{P}_i(X_t = j, t < J_0) = 0 \leq \widetilde{p}_{ij}(t)$, for all i, j and t.
  Suppose inductively that, for all i, j and t,
  $$\mathbb{P}_i(X_t = j, t < J_n) \leq \widetilde{p}_{ij}(t).$$
  Then by comparing the preceding equations, for all i, j and t,
  $$\mathbb{P}_i(X_t = j, t < J_{n+1}) \leq \widetilde{p}_{ij}(t).$$
  So the induction proceeds.
  Hence, for all i, j and t,
  $$p_{ij}(t) = \lim_{n \to \infty} \mathbb{P}_i(X_t = j, t < J_n) \leq \widetilde{p}_{ij}(t).$$

Combined Proof of the Theorems (Step 3)

Step 3: Since $(X_t)_{t\geq 0}$ does not return from ∞, we have,

$$\begin{aligned} p_{ij}(s+t) &= \mathbb{P}_i(X_{s+t}=j) \\ &= \sum_{k\in I}\mathbb{P}_i(X_{s+t}=j|X_s=k)\mathbb{P}_i(X_s=k) \\ &= \sum_{k\in I}\mathbb{P}_i(X_s=k)\mathbb{P}_k(X_t=j) \\ &\quad \text{(Markov Property)} \\ &= \sum_{k\in I}p_{ik}(s)p_{kj}(t). \end{aligned}$$

Hence $(P(t):t\geq 0)$ is a matrix semigroup.

This completes the proof of the first theorem.

Combined Proof of the Theorems (Step 4)

Step 4: Suppose, as we have throughout, that $(X_t)_{t\geq 0}$ satisfies (a). Then, by the Markov Property

$$\begin{array}{rcl}\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ldots, X_{t_n} = i_n) & = & \mathbb{P}_{i_n}(X_{t_{n+1}-t_n} = i_{n+1}) \\ & = & p_{i_n i_{n+1}}(t_{n+1} - t_n).\end{array}$$

So $(X_t)_{t\geq 0}$ satisfies (b).

We complete the proof of the second theorem by the usual argument that (b) must now imply (a) (as done in a previous proof).

Time Reversal Identity

Lemma

We have

$$q_{i_n}\mathbb{P}(J_n \le t < J_{n+1}|Y_0 = i_0, Y_1 = i_1, \ldots, Y_n = i_n)$$
$$= q_{i_0}\mathbb{P}(J_n \le t < J_{n+1}|Y_0 = i_n, \ldots, Y_{n-1} = i_1, Y_n = i_0).$$

- Conditional on $Y_0 = i_0, \ldots, Y_n = i_n$, the holding times $S_1, \ldots, S_{n+1}$ are independent with $S_k \sim E(q_{i_{k-1}})$.
  So the left-hand side is given by
  $$\int_{\Delta(t)} q_{i_n} \exp\{-q_{i_n}(t - s_1 - \cdots - s_n)\} \prod_{k=1}^{n} q_{i_{k-1}} \exp\{-q_{i_{k-1}} s_k\} ds_k,$$
  where
  $$\Delta(t) = \{(s_1, \ldots, s_n) : s_1 + \cdots + s_n \le t \text{ and } s_1, \ldots, s_n \ge 0\}.$$

Time Reversal Identity (Cont'd)

- We have

$$\int_{\Delta(t)} q_{i_n} \exp\{-q_{i_n}(t - s_1 - \cdots - s_n)\} \prod_{k=1}^{n} q_{i_{k-1}} \exp\{-q_{i_{k-1}} s_k\} ds_k.$$

Substitute

$$\begin{aligned} u_1 &= t - s_1 - \cdots - s_n, \\ u_k &= s_{n-k+2}, \quad k = 2, \ldots, n. \end{aligned}$$

We get

$$\begin{aligned} & q_{i_n} \mathbb{P}(J_n \leq t < J_{n+1} | Y_0 = i_0, \ldots, Y_n = i_n) \\ = & \int_{\Delta(t)} q_{i_0} \exp\{-q_{i_0}(t - u_1 - \cdots - u_n)\} \\ & \qquad \prod_{k=1}^{n} q_{i_{n-k+1}} \exp\{-q_{i_{n-k+1}} u_k\} du_k \\ = & \, q_{i_0} \mathbb{P}(J_n \leq t < J_{n+1} | Y_0 = i_n, \ldots, Y_{n-1} = i_1, Y_n = i_0). \end{aligned}$$

The Forward Equation

Theorem

The minimal non-negative solution $(P(t) : t \geq 0)$ of the backward equation is also the minimal non-negative solution of the forward equation

$$P'(t) = P(t)Q, \quad P(0) = I.$$

- Let $(X_t)_{t \geq 0}$ be the minimal Markov chain with generator matrix Q.
  By the previous theorem, we know that

$$\begin{aligned} p_{ij}(t) &= \mathbb{P}_i(X_t = j) \\ &= \sum_{n=0}^{\infty} \sum_{k \neq j} \mathbb{P}_i(J_n \leq t < J_{n+1}, Y_{n-1} = k, Y_n = j). \end{aligned}$$

The Forward Equation (Cont'd)

- By the preceding lemma, for $n \geq 1$, we have

$$\begin{aligned}
&\mathbb{P}_i(J_n \leq t < J_{n+1} | Y_{n-1} = k, Y_n = j) \\
&= \tfrac{q_i}{q_j} \mathbb{P}_j(J_n \leq t < J_{n+1} | Y_1 = k, Y_n = i) \\
&= \tfrac{q_i}{q_j} \int_0^t q_j e^{-q_j s} \mathbb{P}_k(J_{n-1} \leq t - s < J_n | Y_{n-1} = i) ds \\
&\qquad \text{(by the Markov Property of } (Y_n)_{n \geq 0}) \\
&= q_i \int_0^t e^{-q_j s} \tfrac{q_k}{q_i} \mathbb{P}_i(J_{n-1} \leq t - s < J_n | Y_{n-1} = k) ds.
\end{aligned}$$

The Forward Equation (Cont'd)

- Now we get

$$\begin{aligned} p_{ij}(t) &= \delta_{ij}e^{-q_i t} + \textstyle\sum_{n=1}^{\infty}\sum_{k\neq j}\int_0^t \mathbb{P}_i(J_{n-1} \leq t-s < J_n | Y_{n-1}=k) \\ &\qquad \times \mathbb{P}_i(Y_{n-1}=k, Y_n=j) q_k e^{-q_j s} ds \\ &= \delta_{ij}e^{-q_i t} + \\ &\quad \textstyle\sum_{n=1}^{\infty}\sum_{k\neq j}\int_0^t \mathbb{P}_i(J_{n-1} \leq t-s < J_n, Y_{n-1}=k) q_k \pi_{kj} e^{-q_j s} ds \\ &= \delta_{ij}e^{-q_i t} + \textstyle\int_0^t \sum_{k\neq j} p_{ik}(t-s) q_{kj} e^{-q_j s} ds, \end{aligned}$$

  the interchange of sum and integral by Monotone Convergence.
  This is the **integral form of the forward equation**.
  Make a change of variable $u = t - s$ in the integral.
  Then multiply by $e^{q_j t}$ to obtain

$$p_{ij}(t)e^{q_j t} = \delta_{ij} + \int_0^t \sum_{k\neq j} p_{ik}(u) q_{kj} e^{q_j u} du.$$

The Forward Equation (Cont'd)

- We have seen that $e^{q_i t} p_{ik}(t)$ is increasing for all i, k.
  Hence, one of the following occurs:
  - $\sum_{k \neq j} p_{ik}(u) q_{kj}$ converges uniformly, for $u \in [0, t]$;
  - $\sum_{k \neq j} p_{ik}(u) q_{kj} = \infty$, for all $u \geq t$.

  However, the left-hand side in the previous equation is finite for all t.

  So the last option would contradict the preceding equation.

  So it is the former option that holds.

  From the backward equation, $p_{ij}(t)$ is continuous for all i, j.

  By uniform convergence, the integrand is continuous.

  So we may differentiate to obtain

$$p'_{ij}(t) + p_{ij}(t) q_j = \sum_{k \neq j} p_{ik}(t) q_{kj}.$$

  Hence, $P(t)$ solves the forward equation.

The Forward Equation (Minimality)

- To establish minimality let us suppose that $\widetilde{p}_{ij}(t)$ is another solution of the forward equation.

  Then we also have

  $$\widetilde{p}_{ij}(t) = \delta_{ij} e^{-q_i t} + \sum_{k \neq j} \int_0^t \widetilde{p}_{ik}(t-s) q_{kj} e^{-q_j s} ds.$$

  A similar argument leading to the formula for $p_{ij}(t)$ shows that, for $n \geq 0$,

  $$\mathbb{P}_i(X_t = j, t < J_{n+1}) = \delta_{ij} e^{-q_i t} + \sum_{k \neq j} \int_0^t \mathbb{P}_i(X_t = j, t < J_n) q_{kj} e^{-q_j s} ds.$$

The Forward Equation (Minimality Cont'd)

- If $\widetilde{P}(t) \geq 0$, then, for all i, j and t,

$$\mathbb{P}(X_t = j, t < J_0) = 0 \leq \widetilde{p}_{ij}(t).$$

Suppose inductively that, for all i, j and t,

$$\mathbb{P}_i(X_t = j, t < J_n) \leq \widetilde{p}_{ij}(t).$$

Then by comparing the formulas on the preceding slide, we obtain, for all i, j and t,

$$\mathbb{P}_i(X_t = j, t < J_{n+1}) \leq \widetilde{p}_{ij}(t).$$

So the induction proceeds.

Hence, for all i, j and t,

$$p_{ij}(t) = \lim_{n \to \infty} \mathbb{P}_i(X_t = j, t < J_n) \leq \widetilde{p}_{ij}(t).$$