Markov Chains: An Introduction/Review

David Sirl

dsirl@maths.uq.edu.au

http://www.maths.uq.edu.au/~dsirl/

AUSTRALIAN RESEARCH COUNCIL

Centre of Excellence for Mathematics and Statistics of Complex Systems
Andrei A. Markov (1856 – 1922)
Random Processes

A random process is a collection of random variables indexed by some set $I$, taking values in some set $S$.

- $I$ is the index set, usually time, e.g. $\mathbb{Z}^+$, $\mathbb{R}$, $\mathbb{R}^+$.
- $S$ is the state space, e.g. $\mathbb{Z}^+$, $\mathbb{R}^n$, $\{1, 2, \ldots, n\}$, $\{a, b, c\}$.

We classify random processes according to both the index set (discrete or continuous) and the state space (finite, countable or uncountable/continuous).
A random process is called a *Markov Process* if, conditional on the current state of the process, its future is independent of its past.

More formally, $X(t)$ is Markovian if it has the following property:

$$
P(X(t_n) = j_n \mid X(t_{n-1}) = j_{n-1}, \ldots, X(t_1) = j_1) = P(X(t_n) = j_n \mid X(t_{n-1}) = j_{n-1})$$

for all finite sequences of times $t_1 < \ldots < t_n \in I$ and of states $j_1, \ldots, j_n \in S$. 

A Markov chain \((X(t))\) is said to be *time-homogeneous* if

\[
P(X(s + t) = j \mid X(s) = i)
\]

is independent of \(s\). When this holds, putting \(s = 0\) gives

\[
P(X(s + t) = j \mid X(s) = i) = P(X(t) = j \mid X(0) = i).
\]
Time Homogeneity

A Markov chain \((X(t))\) is said to be *time-homogeneous* if

\[ P(X(s + t) = j \mid X(s) = i) \]

is independent of \(s\). When this holds, putting \(s = 0\) gives

\[ P(X(s + t) = j \mid X(s) = i) = P(X(t) = j \mid X(0) = i). \]

Probabilities depend on elapsed time, not absolute time.
Discrete-time Markov chains

At time epochs \( n = 1, 2, 3, \ldots \) the process changes from one state \( i \) to another state \( j \) with probability \( p_{ij} \).
At time epochs $n = 1, 2, 3, \ldots$ the process changes from one state $i$ to another state $j$ with probability $p_{ij}$.

We write the one-step transition matrix $P = (p_{ij}, \ i, j \in S)$. 

Discrete-time Markov chains

- At time epochs $n = 1, 2, 3, \ldots$ the process changes from one state $i$ to another state $j$ with probability $p_{ij}$.

- We write the one-step transition matrix $P = (p_{ij}, i, j \in S)$.

- Example: a frog hopping on 3 rocks. Put $S = \{1, 2, 3\}$.

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{8} & \frac{1}{2} & \frac{1}{4} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}$$
Example: A frog hopping on 3 rocks. Put \( S = \{1, 2, 3\} \).

\[
P = \begin{pmatrix}
0 & \frac{1}{2} & \frac{1}{2} \\
\frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\
\frac{3}{2} & \frac{1}{3} & 0
\end{pmatrix}
\]

We can gain some insight by drawing a picture:
DTMCs: \( n \)-step probabilities

We have \( P \), which tells us what happens over one time step; let's work out what happens over two time steps:

\[
p^{(2)}_{ij} = \mathbb{P}(X_2 = j \mid X_0 = i) \\
= \sum_{k \in S} \mathbb{P}(X_1 = k \mid X_0 = i) \mathbb{P}(X_2 = j \mid X_1 = k, X_0 = i) \\
= \sum_{k \in S} p_{ik} p_{kj}.
\]
DTMCs: \( n \)-step probabilities

- We have \( P \), which tells us what happens over one time step; lets work out what happens over two time steps:

\[
p_{ij}^{(2)} = \mathbb{P}(X_2 = j \mid X_0 = i)
= \sum_{k \in S} \mathbb{P}(X_1 = k \mid X_0 = i) \mathbb{P}(X_2 = j \mid X_1 = k, X_0 = i)
= \sum_{k \in S} p_{ik} p_{kj}.
\]

- So \( P^{(2)} = PP = P^2 \).
DTMCs: \( n \)-step probabilities

- We have \( P \), which tells us what happens over one time step; lets work out what happens over two time steps:

\[
p_{ij}^{(2)} = P(X_2 = j \mid X_0 = i) \\
= \sum_{k \in S} P(X_1 = k \mid X_0 = i) P(X_2 = j \mid X_1 = k, X_0 = i) \\
= \sum_{k \in S} p_{ik} p_{kj}.
\]

- So \( P^{(2)} = PP = P^2 \).

- Similarly, \( P^{(3)} = P^2P = P^3 \) and \( P^{(n)} = P^n \).
We may wish to start the chain according to some initial distribution $\pi^{(0)}$. 
We may wish to start the chain according to some initial distribution $\pi^{(0)}$.

We can then calculate the state probabilities $\pi^{(n)} = (\pi_j^{(n)}, \ j \in S)$ of being in state $j$ at time $n$ as follows:

$$\pi_j^{(n)} = \sum_{k \in S} \mathbb{P}(X_0 = k) \mathbb{P}(X_n = j \mid X_0 = k) = \sum_{k \in S} \pi_j^{(0)} p_{ij}^{(n)}.$$
DTMC: Arbitrary initial distributions

- We may wish to start the chain according to some initial distribution $\pi^{(0)}$.

- We can then calculate the state probabilities $\pi^{(n)} = (\pi_j^{(n)}, j \in S)$ of being in state $j$ at time $n$ as follows:

\[
\pi_j^{(n)} = \sum_{k \in S} \mathbb{P}(X_0 = k) \mathbb{P}(X_n = j \mid X_0 = k) = \sum_{k \in S} \pi_j^{(0)} p_{ij}^{(n)}.
\]

- Or, in matrix notation, $\pi^{(n)} = \pi^{(0)} P^n$; similarly we can show that $\pi^{(n+1)} = \pi^{(n)} P$. 
We say that a state $i$ leads to $j$ (written $i \rightarrow j$) if it is possible to get from $i$ to $j$ in some finite number of jumps: $p_{ij}^{(n)} > 0$ for some $n \geq 0$. 
Class structure

- We say that a state $i$ leads to $j$ (written $i \rightarrow j$) if it is possible to get from $i$ to $j$ in some finite number of jumps: $p_{ij}^{(n)} > 0$ for some $n \geq 0$.

- We say that $i$ communicates with $j$ (written $i \leftrightarrow j$) if $i \rightarrow j$ and $j \rightarrow i$. 
Class structure

- We say that a state $i$ leads to $j$ (written $i \rightarrow j$) if it is possible to get from $i$ to $j$ in some finite number of jumps: $p_{ij}^{(n)} > 0$ for some $n \geq 0$.

- We say that $i$ communicates with $j$ (written $i \leftrightarrow j$) if $i \rightarrow j$ and $j \rightarrow i$.

- The relation $\leftrightarrow$ partitions the state space into *communicating classes*. 
Class structure

- We say that a state $i$ leads to $j$ (written $i \rightarrow j$) if it is possible to get from $i$ to $j$ in some finite number of jumps: $p_{ij}^{(n)} > 0$ for some $n \geq 0$.

- We say that $i$ communicates with $j$ (written $i \leftrightarrow j$) if $i \rightarrow j$ and $j \rightarrow i$.

- The relation $\leftrightarrow$ partitions the state space into communicating classes.

- We call the state space *irreducible* if it consists of a single communicating class.
Class structure

- We say that a state \( i \) leads to \( j \) (written \( i \rightarrow j \)) if it is possible to get from \( i \) to \( j \) in some finite number of jumps: \( p_{ij}^{(n)} > 0 \) for some \( n \geq 0 \).

- We say that \( i \) communicates with \( j \) (written \( i \leftrightarrow j \)) if \( i \rightarrow j \) and \( j \rightarrow i \).

- The relation \( \leftrightarrow \) partitions the state space into communicating classes.

- We call the state space irreducible if it consists of a single communicating class.

- These properties are easy to determine from a transition probability graph.
Classification of states

We call a state $i$ recurrent or transient according as $\mathbb{P}(X_n = i \text{ for infinitely many } n)$ is equal to one or zero.
Classification of states

- We call a state \( i \) recurrent or transient according as \( \mathbb{P}(X_n = i \text{ for infinitely many } n) \) is equal to one or zero.
- A recurrent state is a state to which the process always returns.
- A transient state is a state which the process eventually leaves for ever.
Classification of states

- We call a state \( i \) **recurrent** or **transient** according as \( \Pr(X_n = i \text{ for infinitely many } n) \) is equal to one or zero.
- A recurrent state is a state to which the process always returns.
- A transient state is a state which the process eventually leaves for ever.
- Recurrence and transience are class properties; i.e. if two states are in the same communicating class then they are recurrent/transient together.
- We therefore speak of recurrent or transient classes.
Classification of states

- We call a state \( i \) recurrent or transient according as \( P(X_n = i \text{ for infinitely many } n) \) is equal to one or zero.
- A recurrent state is a state to which the process always returns.
- A transient state is a state which the process eventually leaves for ever.
- Recurrence and transience are class properties; i.e. if two states are in the same communicating class then they are recurrent/transient together.
- We therefore speak of recurrent or transient classes.
- We also assume throughout that no states are periodic.
DTMCs: Two examples

- $S$ irreducible:

\[
P = \begin{pmatrix}
0 & \frac{1}{2} & \frac{1}{2} \\
\frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\
\frac{2}{3} & \frac{1}{3} & 0
\end{pmatrix}
\]

- $S = \{0\} \cup C$, where $C$ is a transient class:

\[
P = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\
0 & \frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\
0 & \frac{2}{3} & \frac{1}{3} & 0
\end{pmatrix}
\]
DTMCs: Quantities of interest

Quantities of interest include:
- Hitting probabilities.
- Expected hitting times.
- Limiting (stationary) distributions.
- Limiting conditional (quasistationary) distributions.
Let $\alpha_i$ be the probability of hitting state 1 starting in state $i$.

Clearly $\alpha_1 = 1$; and for $i \neq 1$,

$$
\alpha_i = \mathbb{P}(\text{hit } 1 \mid \text{start in } i) \\
= \sum_{k \in S} \mathbb{P}(X_1 = k \mid X_0 = i) \cdot \mathbb{P}(\text{hit } 1 \mid \text{start in } k) \\
= \sum_{k \in S} p_{ik} \alpha_k
$$
Let $\alpha_i$ be the probability of hitting state 1 starting in state $i$.

- Clearly $\alpha_1 = 1$; and for $i \neq 1$,

$$
\alpha_i = \mathbb{P}(\text{hit } 1 \mid \text{start in } i) \\
= \sum_{k \in S} \mathbb{P}(X_1 = k \mid X_0 = i) \mathbb{P}(\text{hit } 1 \mid \text{start in } k) \\
= \sum_{k \in S} p_{ik} \alpha_k
$$

- Sometimes there may be more than one solution $\alpha = (\alpha_i, \, i \in S')$ to this system of equations.

If this is the case, then the hitting probabilities are given by the minimal such solution.
Example: Hitting Probabilities

Let $\alpha_i$ be the probability of hitting state 3 starting in state $i$. So $\alpha_3 = 1$ and $\alpha_i = \sum_k p_{ik} \alpha_k$:

\[
\begin{align*}
\alpha_0 &= \alpha_0 \\
\alpha_1 &= \frac{1}{2} \alpha_0 + \frac{1}{4} \alpha_2 + \frac{1}{4} \alpha_3 \\
\alpha_2 &= \frac{5}{8} \alpha_1 + \frac{1}{8} \alpha_2 + \frac{1}{4} \alpha_3
\end{align*}
\]
Example: Hitting Probabilities

Let $\alpha_i$ be the probability of hitting state 3 starting in state $i$.

\[
\alpha = \begin{pmatrix}
0 \\
\frac{9}{23} \\
\frac{13}{23} \\
1
\end{pmatrix} \approx \begin{pmatrix}
0 \\
0.39 \\
0.57 \\
1
\end{pmatrix}.
\]
Let $\beta_i$ be the probability of hitting state 0 before state $N$, starting in state $i$.

- Clearly $\beta_0 = 1$ and $\beta_N = 0$.
- For $0 < i < N$,

$$\beta_i = \mathbb{P}(\text{hit 1 before } n \mid \text{start in } i)$$

$$= \sum_{k \in S} \mathbb{P}(X_1 = k \mid X_0 = i) \mathbb{P}(\text{hit 1 before } n \mid \text{start in } k)$$

$$= \sum_{k \in S} p_{ik} \beta_k$$

- Again, we take the minimal solution of this system of equations.
Example: Hitting Probabilities II

Let $\beta_i$ be the probability of hitting 0 before 3 starting in $i$.

So $\beta_0 = 1$, $\beta_3 = 0$ and $\beta_i = \sum_k p_{ik} \beta_k$:

$$
\beta_1 = \frac{1}{2} \beta_0 + \frac{1}{4} \beta_2 + \frac{1}{4} \beta_3 \\
\beta_2 = \frac{5}{8} \beta_1 + \frac{1}{8} \beta_2 + \frac{1}{4} \beta_3
$$
Example: Hitting Probabilities II

Let $\beta_i$ be the probability of hitting 0 before 3 starting in $i$.

$$\beta = \begin{pmatrix} 1 \\ \frac{14}{23} \\ \frac{10}{23} \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0.61 \\ 0.43 \\ 1 \end{pmatrix}. $$
DTMCs: Expected hitting times

Let $\tau_i$ be the expected time to hit state $1$ starting in state $i$.

- Clearly $\tau_1 = 0$; and for $i \neq 0$,

$$
\tau_i = \mathbb{E}(\text{time to hit } 1 \mid \text{start in } i) = 1 + \sum_{k \in S} \mathbb{P}(X_1 = k \mid X_0 = i) \mathbb{E}(\text{time to hit } 1 \mid \text{start in } k) = 1 + \sum_{k \in S} p_{ik} \tau_k
$$

- If there are multiple solutions, take the minimal one.
Example: Expected Hitting Times

Let $\tau_i$ be the expected time to hit 2 starting in $i$.

So $\tau_2 = 0$ and $\tau_i = 1 + \sum_k p_{ik} \tau_k$:

$\tau_1 = 1 + \frac{1}{2} \tau_2 + \frac{1}{2} \tau_3$

$\tau_3 = 1 + \frac{2}{3} \tau_1 + \frac{1}{3} \tau_2$
Example: Expected Hitting Times

Let $\tau_i$ be the expected time to hit 2 starting in $i$.

$$
\tau = \begin{pmatrix}
\frac{9}{4} \\
0 \\
\frac{5}{2}
\end{pmatrix}
= \begin{pmatrix}
2.25 \\
0 \\
2.5
\end{pmatrix}.
$$

$$
P = \begin{pmatrix}
0 & \frac{1}{2} & \frac{1}{2} \\
\frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\
\frac{2}{3} & \frac{1}{3} & 0
\end{pmatrix}
$$
DTMCs: Hitting Probabilities and Times

- Just systems of linear equations to be solved.
- In principle can be solved analytically when $S$ is finite.
- When $S$ is an infinite set, if $P$ has some regular structure ($p_{ij}$ same/similar for each $i$) the resulting systems of difference equations can sometimes be solved analytically.
- Otherwise we need numerical methods.
DTMCs: The Limiting Distribution

Assume that the state space is irreducible, aperiodic and recurrent.

What happens to the state probabilities \( \pi_j^{(n)} \) as \( n \to \infty \)?
DTMCs: The Limiting Distribution

Assume that the state space is irreducible, aperiodic and recurrent.

- What happens to the state probabilities $\pi_j^{(n)}$ as $n \to \infty$?
- We know that $\pi^{(n+1)} = \pi^{(n)} P$. 
DTMCs: The Limiting Distribution

Assume that the state space is irreducible, aperiodic and recurrent.

- What happens to the state probabilities \( \pi_j^{(n)} \) as \( n \to \infty \)?
- We know that \( \pi^{(n+1)} = \pi^{(n)} P \).
- So if there is a limiting distribution \( \pi \), it must satisfy

\[
\pi = \pi P \quad \text{(and } \sum_i \pi_i = 1). \]

(Such a distribution is called *stationary*.)
DTMCs: The Limiting Distribution

Assume that the state space is irreducible, aperiodic and recurrent.

- What happens to the state probabilities $\pi_j^{(n)}$ as $n \to \infty$?
- We know that $\pi^{(n+1)} = \pi^{(n)} P$.
- So if there is a limiting distribution $\pi$, it must satisfy

$$\pi = \pi P \quad \text{(and } \sum_i \pi_i = 1).$$

(Such a distribution is called stationary.)

- This limiting distribution does not depend on the initial distribution.
DTMCs: The Limiting Distribution

Assume that the state space is irreducible, aperiodic and recurrent.

- What happens to the state probabilities \( \pi_j^{(n)} \) as \( n \to \infty \)?

- We know that \( \pi^{(n+1)} = \pi^{(n)} P \).

- So if there is a limiting distribution \( \pi \), it must satisfy

  \[
  \pi = \pi P \quad \text{(and } \sum_i \pi_i = 1)\]

  (Such a distribution is called \textit{stationary}.)

- This limiting distribution does not depend on the initial distribution.

- When the state space is infinite, it may happen that \( \pi_j^{(n)} \to 0 \) for all \( j \).
Example: The Limiting Distribution

\[ P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix} \]

Substituting \( P \) into \( \pi = \pi P \) gives

\[
\begin{align*}
\pi_1 &= \frac{5}{8} \pi_2 + \frac{2}{3} \pi_3, \\
\pi_2 &= \frac{1}{2} \pi_1 + \frac{1}{8} \pi_2 + \frac{1}{3} \pi_3, \\
\pi_3 &= \frac{1}{2} \pi_1 + \frac{1}{4} \pi_2,
\end{align*}
\]

which together with \( \sum_i \pi_i = 1 \) yields

\[ \pi = \left( \frac{38}{97} \frac{32}{97} \frac{27}{97} \right) \approx \left( 0.39 \ 0.33 \ 0.28 \right). \]
DTMCs: The Limiting Conditional Dist’n

Assume that the state space is consists of an absorbing state and a transient class \((S = \{0\} \cup C)\).

The limiting distribution is \((1, 0, 0, \ldots)\).
DTMCs: The Limiting Conditional Dist’n

Assume that the state space is consists of an absorbing state and a transient class \( S = \{0\} \cup C \).

- The limiting distribution is \((1, 0, 0, \ldots)\).
- Instead of looking at the limiting behaviour of

\[
P(X_n = j \mid X_0 = i) = p_{ij}^{(n)},
\]

we need to look at

\[
P(X_n = j \mid X_n \neq 0, X_0 = i) = \frac{p_{ij}^{(n)}}{1 - p_{i0}^{(n)}}
\]

for \( i, j \in C \).
DTMCs: The Limiting Conditional Dist’n

It turns out we need a solution $m = (m_i, i \in C)$ of

$$mP_C = rm,$$

for some $r \in (0, 1)$. 
It turns out we need a solution $m = (m_i, \ i \in C)$ of

$$mP_C = rm,$$

for some $r \in (0, 1)$.

If $C$ is a finite set, there is a unique such $r$. 
DTMCs: The Limiting Conditional Dist’n

- It turns out we need a solution \( m = (m_i, \ i \in C) \) of
  \[
  mP_C = rm,
  \]
  for some \( r \in (0, 1) \).

- If \( C \) is a finite set, there is a unique such \( r \).

- If \( C \) is infinite, there is \( r^* \in (0, 1) \) such that all \( r \) in the interval \([r^*, 1)\) are admissible; and the solution corresponding to \( r = r^* \) is the LCD.
Example: Limiting Conditional Dist’n

\[ P = \begin{pmatrix}
1 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\
0 & \frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\
0 & \frac{2}{3} & \frac{1}{3} & 0
\end{pmatrix} \]
Example: Limiting Conditional Dist’n

\[ P_C = \begin{pmatrix}
0 & \frac{1}{4} & \frac{1}{4} \\
\frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\
\frac{2}{3} & \frac{1}{3} & 0
\end{pmatrix} \]
Example: Limiting Conditional Dist’n

\[
P_C = \begin{pmatrix}
0 & \frac{1}{4} & \frac{1}{4} \\
\frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\
\frac{2}{3} & \frac{1}{3} & 0
\end{pmatrix}
\]

Solving \( mP_C = rm \), we get

\[ r_1 \approx 0.773 \quad \text{and} \quad m \approx (0.45, 0.30, 0.24) \]
DTMCs: Summary

From the one-step transition probabilities we can calculate:

- $n$-step transition probabilities,
- hitting probabilities,
- expected hitting times,
- limiting distributions, and
- limiting conditional distributions.
Continuous Time

In the real world, time is continuous — things do not happen only at prescribed, equally spaced time points.
Continuous Time

- In the real world, time is continuous — things do not happen only at prescribed, equally spaced time points.
- Continuous time is slightly more difficult to deal with as there is no real equivalent to the one-step transition matrix from which one can calculate all quantities of interest.
Continuous Time

- In the real world, time is continuous — things do not happen only at prescribed, equally spaced time points.

- Continuous time is slightly more difficult to deal with as there is no real equivalent to the one-step transition matrix from which one can calculate all quantities of interest.

- The study of continuous-time Markov chains is based on the transition function.
CTMCs: Transition Functions

If we denote by $p_{ij}(t)$ the probability of a process starting in state $i$ being in state $j$ after elapsed time $t$, then we call $P(t) = (p_{ij}(t), i, j \in S, t > 0)$ the transition function of that process.
CTMCs: Transition Functions

If we denote by $p_{ij}(t)$ the probability of a process starting in state $i$ being in state $j$ after elapsed time $t$, then we call $P(t) = (p_{ij}(t), i, j \in S, t > 0)$ the transition function of that process.

$P(t)$ is difficult/impossible to write down in all but the simplest of situations.
If we denote by \( p_{ij}(t) \) the probability of a process starting in state \( i \) being in state \( j \) after elapsed time \( t \), then we call \( P(t) = (p_{ij}(t), \, i, j \in S, \, t > 0) \) the transition function of that process.

\( P(t) \) is difficult/impossible to write down in all but the simplest of situations.

However all is not lost: we can show that there exist quantities \( q_{ij}, \, i, j \in S \) satisfying

\[
q_{ij} = p'_{ij}(0^+) = \begin{cases} 
\lim_{t \downarrow 0} \frac{p_{ij}(t)}{t}, & i \neq j, \\
\lim_{t \downarrow 0} \frac{1 - p_{ii}(t)}{t}, & i = j.
\end{cases}
\]
CTMCs: The q-matrix

- We call the matrix \( Q = (q_{ij}, i, j \in S) \) the *q-matrix* of the process and can interpret it as follows:
  - For \( i \neq j \), \( q_{ij} \in [0, \infty) \) is the instantaneous rate the process moves from state \( i \) to state \( j \), and
  - \( q_i = -q_{ii} \in [0, \infty] \) is the rate at which the process leaves state \( i \).
  - We also have \( \sum_{j \neq i} q_{ij} \leq q_i \).
We call the matrix \( Q = (q_{ij}, i, j \in S) \) the \textit{q-matrix} of the process and can interpret it as follows:

- For \( i \neq j \), \( q_{ij} \in [0, \infty) \) is the instantaneous rate the process moves from state \( i \) to state \( j \), and
- \( q_i = -q_{ii} \in [0, \infty] \) is the rate at which the process leaves state \( i \).
- We also have \( \sum_{j \neq i} q_{ij} \leq q_i \).

When we formulate a model, it is \( Q \) that we can write down; so the question arises, can we recover \( P(\cdot) \) from \( Q = P'(0) \)?
If we are given a conservative q-matrix $Q$, then a $Q$-function $P(t)$ must satisfy the backward equations

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

and may or may not satisfy the forward (or master) equations

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

with the initial condition $P(0) = I$. 
CTMCs: The Kolmogorov DEs

If we are given a conservative q-matrix $Q$, then a $Q$-function $P(t)$ must satisfy the backward equations

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

and may or may not satisfy the forward (or master) equations

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

with the initial condition $P(0) = I$.

There is always one such $Q$-function, but there may also be infinitely many such functions — so $Q$ does not necessarily describe the whole process.
Suppose $X(0) = i$:

- The holding time $H_i$ in state $i$ is exponentially distributed with parameter $q_i$, i.e.

$$\mathbb{P}(H_i \leq t) = 1 - e^{-q_i t}, \quad t \geq 0.$$
CTMCs: Interpreting the q-matrix

Suppose $X(0) = i$:

- The holding time $H_i$ in state $i$ is exponentially distributed with parameter $q_i$, i.e.

  $$\mathbb{P}(H_i \leq t) = 1 - e^{-q_i t}, \quad t \geq 0.$$ 

- After this time has elapsed, the process jumps to state $j$ with probability $q_{ij}/q_i$. 
Suppose $X(0) = i$:

- The holding time $H_i$ in state $i$ is exponentially distributed with parameter $q_i$, i.e.
  \[ \mathbb{P}(H_i \leq t) = 1 - e^{-q_i t}, \quad t \geq 0. \]

- After this time has elapsed, the process jumps to state $j$ with probability $q_{ij}/q_i$.

- Repeat...
CTMCs: Interpreting the q-matrix

Suppose \( X(0) = i \):

- The holding time \( H_i \) in state \( i \) is exponentially distributed with parameter \( q_i \), i.e.
  \[
  \mathbb{P}(H_i \leq t) = 1 - e^{-q_it}, \quad t \geq 0.
  \]

- After this time has elapsed, the process jumps to state \( j \) with probability \( q_{ij}/q_i \).

- Repeat...

- Somewhat surprisingly, this recipe does not always describe the whole process.
Consider a process described by the q-matrix

\[ q_{ij} = \begin{cases} 
\lambda_i & \text{if } j = i + 1, \\
-\lambda_i & \text{if } j = i, \\
0 & \text{otherwise.}
\end{cases} \]

Assume \( \lambda_i > 0, \quad \forall i \in S. \)
CTMCs: An Explosive Process

Consider a process described by the q-matrix

\[
q_{ij} = \begin{cases} 
\lambda_i & \text{if } j = i + 1, \\
-\lambda_i & \text{if } j = i, \\
0 & \text{otherwise.}
\end{cases}
\]

- Assume \( \lambda_i > 0, \quad \forall i \in S. \)
- Suppose we start in state \( i_0. \)
CTMCs: An Explosive Process

Consider a process described by the q-matrix

\[ q_{ij} = \begin{cases} 
\lambda_i & \text{if } j = i + 1, \\
-\lambda_i & \text{if } j = i, \\
0 & \text{otherwise.}
\end{cases} \]

- Assume \( \lambda_i > 0, \quad \forall i \in S. \)
- Suppose we start in state \( i_0. \)
- Stay for time \( H_{i_0} \sim \exp(\lambda_{i_0}) \) then move to state \( i_0 + 1, \)
Consider a process described by the q-matrix

\[ q_{ij} = \begin{cases} 
\lambda_i & \text{if } j = i + 1, \\
-\lambda_i & \text{if } j = i, \\
0 & \text{otherwise}. 
\end{cases} \]

Assume \( \lambda_i > 0, \quad \forall i \in S. \)

Suppose we start in state \( i_0. \)

Stay for time \( H_{i_0} \sim \exp(\lambda_{i_0}) \) then move to state \( i_0 + 1, \)

Stay for time \( H_{i_0+1} \sim \exp(\lambda_{i_0+1}) \) then move to \( i_0 + 2, \ldots. \)
CTMCs: An Explosive Process

Consider a process described by the q-matrix

\[
q_{ij} = \begin{cases} 
\lambda_i & \text{if } j = i + 1, \\
-\lambda_i & \text{if } j = i, \\
0 & \text{otherwise.}
\end{cases}
\]

- Assume \( \lambda_i > 0, \quad \forall i \in S \).
- Suppose we start in state \( i_0 \).
- Stay for time \( H_{i_0} \sim \exp(\lambda_{i_0}) \) then move to state \( i_0 + 1 \),
- Stay for time \( H_{i_0+1} \sim \exp(\lambda_{i_0+1}) \) then move to \( i_0 + 2, \ldots \)
- Define \( T_n = \sum_{i=i_0}^{i_0+n-1} H_i \) to be the time of the \( n \)th jump. We would expect \( T := \lim_{n \to \infty} T_n = \infty \).
**Lemma:** Suppose $\{S_n, \; n \geq 1\}$ is a sequence of independent exponential rv’s with respective rates $a_i$, and put $S = \sum_{n=1}^{\infty} S_n$.

Then either $S = \infty$ a.s. or $S < \infty$ a.s., according as $\sum_{i=1}^{\infty} \frac{1}{a_i}$ diverges or converges.

- We identify $S_n$ with the holding times $H_{i_0+n}$ and $S$ with $T$. 

Lemma: Suppose \( \{S_n, \ n \geq 1 \} \) is a sequence of independent exponential rv’s with respective rates \( a_i \), and put \( S = \sum_{n=1}^{\infty} S_n \).

Then either \( S = \infty \) a.s. or \( S < \infty \) a.s., according as \( \sum_{i=1}^{\infty} \frac{1}{a_i} \) diverges or converges.

- We identify \( S_n \) with the holding times \( H_{i_0+n} \) and \( S \) with \( T \).
- If, for example, \( \lambda_i = i^2 \), we have

\[
\sum_{i=i_0}^{\infty} \frac{1}{\lambda_i} = \sum_{i=i_0}^{\infty} \frac{1}{i^2} < \infty,
\]

so \( \mathbb{P}(T < \infty) = 1 \).
CTMCs: Reuter’s Uniqueness Condition

For there to be no explosion possible, we need the backward equations to have a unique solution.
CTMCs: Reuter’s Uniqueness Condition

For there to be no explosion possible, we need the backward equations to have a unique solution.

When $Q$ is conservative, this is equivalent to

$$\sum_{j \in S} q_{ij} x_j = \nu x_i \quad i \in S$$

having no bounded non-negative solution $(x_i, i \in S)$ except the trivial solution $x_i \equiv 0$ for some (and then all) $\nu > 0$. 
CTMCs: Ruling Out Explosion

- Analysis of a continuous-time Markov process is greatly simplified if it is *regular*, that is non-explosive.
CTMCs: Ruling Out Explosion

- Analysis of a continuous-time Markov process is greatly simplified if it is *regular*, that is non-explosive.

- A process is regular if
CTMCs: Ruling Out Explosion

Analysis of a continuous-time Markov process is greatly simplified if it is *regular*, that is non-explosive.

A process is regular if
- The state space is finite.
CTMCs: Ruling Out Explosion

- Analysis of a continuous-time Markov process is greatly simplified if it is regular, that is non-explosive.
- A process is regular if
  - The state space is finite.
  - The q-matrix is bounded, that is \( \sup_i q_i < \infty \).
CTMCs: Ruling Out Explosion

Analysis of a continuous-time Markov process is greatly simplified if it is regular, that is non-explosive.

A process is regular if

- The state space is finite.
- The $q$-matrix is bounded, that is $\sup_i q_i < \infty$.
- $X_0 = i$ and $i$ is recurrent.
CTMCs: Ruling Out Explosion

- Analysis of a continuous-time Markov process is greatly simplified if it is regular, that is non-explosive.

- A process is regular if
  - The state space is finite.
  - The q-matrix is bounded, that is $\sup_i q_i < \infty$.
  - $X_0 = i$ and $i$ is recurrent.

- Reuter’s condition simplifies considerably for a birth-death process, a process where from state $i$, the only possible transitions are to $i - 1$ or $i + 1$.

We now assume that the process we are dealing with is non-explosive, so $Q$ is enough to completely specify the process.
A Birth-Death Process on \( \{0, 1, 2, \ldots\} \) is a CTMC with q-matrix of the form

\[
q_{ij} = \begin{cases} 
\lambda_i & \text{if } j = i + 1 \\
\mu_i & \text{if } j = i - 1, \ i \geq 1 \\
-(\lambda_i + \mu_i) & \text{if } j = i \geq 1 \\
-\lambda_0 & \text{if } j = i = 0 \\
0 & \text{otherwise}
\end{cases}
\]

where \( \lambda_i, \mu_i > 0, \ \forall i \in S \).

We also set \( \pi_1 = 1 \), and \( \pi_i = \frac{\lambda_1 \lambda_2 \cdots \lambda_{i-1}}{\mu_2 \mu_3 \cdots \mu_i} \).
CTMCs: Quantities of interest

Again we look at

- Hitting probabilities.
- Expected hitting times.
- Limiting (stationary) distributions.
- Limiting conditional (quasistationary) distributions.
CTMCs: Hitting Probabilities

Using the same reasoning as for discrete-time processes, we can show that the hitting probabilities $\alpha_i$ of a state $\kappa$, starting in state $i$, are given by the minimal non-negative solution to the system $\alpha_\kappa = 1$ and, for $i \neq \kappa$,

$$\sum_{j \in S} q_{ij} \alpha_j = 0.$$
CTMCs: Hitting Probabilities

Using the same reasoning as for discrete-time processes, we can show that the hitting probabilities \( \alpha_i \) of a state \( \kappa \), starting in state \( i \), are given by the minimal non-negative solution to the system \( \alpha_\kappa = 1 \) and, for \( i \neq \kappa \),

\[
\sum_{j \in S} q_{ij} \alpha_j = 0.
\]

For a BDP, we can show that the probability of hitting 0 is one if and only if

\[
\mathcal{A} := \sum_{i=1}^{\infty} \frac{1}{\lambda_i \pi_i} = \infty.
\]
Again, we can use an argument similar to that for discrete-time processes to show that the expected hitting times $\tau_i$ of state $\kappa$, starting in $i$, are given by the minimal non-negative solution of the system $\tau_\kappa = 0$ and, for $i \neq \kappa$,

$$\sum_{j \in S} q_{ij} \tau_j = -1.$$
CTMCs: Hitting times

Again, we can use an argument similar to that for discrete-time processes to show that the expected hitting times $\tau_i$ of state $\kappa$, starting in $i$, are given by the minimal non-negative solution of the system $\tau_\kappa = 0$ and, for $i \neq \kappa$,

$$
\sum_{j \in S} q_{ij} \tau_j = -1.
$$

For a BDP, the expected time to hit zero, starting in state $i$ is given by

$$
\tau_i = \sum_{j=1}^{i} \frac{1}{\mu_j \pi_j} \sum_{k=j}^{\infty} \pi_k.
$$
CTMCs: Limiting Behaviour

As with discrete-time chains, the class structure is important in determining what tools are useful for analysing the long term behaviour of the process.
As with discrete-time chains, the class structure is important in determining what tools are useful for analysing the long term behaviour of the process.

If the state space is irreducible and positive recurrent, the limiting distribution is the most useful device.
CTMCs: Limiting Behaviour

As with discrete-time chains, the class structure is important in determining what tools are useful for analysing the long term behaviour of the process.

- If the state space is irreducible and positive recurrent, the limiting distribution is the most useful device.
- If the state space consists of an absorbing state and a transient class, the limiting conditional distribution is of most use.
Assume that the state space $S$ is irreducible and recurrent. Then there is a unique (up to constant multiples) solution

$$\pi = (\pi_i, \ i \in S)$$

such that

$$\pi Q = 0,$$

where $0$ is a vector of zeros. If $\sum_i \pi_i < \infty$, then $\pi$ is can be normalised to give a probability distribution which is the limiting distribution. (If $\pi$ is not summable then there is no proper limiting distribution.)
Assume that the state space \( S \) is irreducible and recurrent. Then there is a unique (up to constant multiples) solution \( \pi = (\pi_i, \ i \in S) \) such that

\[
\pi Q = 0,
\]

where \( 0 \) is a vector of zeros. If \( \sum_i \pi_i < \infty \), then \( \pi \) is can be normalised to give a probability distribution which is the limiting distribution. (If \( \pi \) is not summable then there is no proper limiting distribution.)

For the BDP, the potential coefficients \( \pi_1 = 1, \ \pi_i = \frac{\lambda_1 \lambda_2 \cdots \lambda_{i-1}}{\mu_2 \mu_3 \cdots \mu_i} \) are the essentially unique solution of \( \pi Q = 0 \).
If the $S = \{0\} \cup C$ and the absorbing state zero is reached with probability one, the limiting conditional distribution is given by $m = (m_i, \ i \in C)$ such that

$$mQ_C = -\nu m,$$

for some $\nu > 0$. 
CTMCs: Limiting Conditional Dist’ns

If the \( S = \{0\} \cup C \) and the absorbing state zero is reached with probability one, the limiting conditional distribution is given by \( m = (m_i, \ i \in C) \) such that

\[
m_QC = -\nu m,
\]

for some \( \nu > 0 \).

When \( C \) is a finite set then there is a unique such \( \nu \).
CTMCs: Summary

- Countable state Markov chains are stochastic modelling tools which have been analysed extensively.

- Where closed form expressions are not available there are accurate numerical methods for approximating quantities of interest.

- They have found application in fields as diverse as ecology, physical chemistry and telecommunications systems modelling.