1) Birth-death processes where $A$ is tridiagonal

\[ A_{i,i+1} = \lambda_i = \text{aggregate birth rate} \]
\[ A_{i,i-1} = \mu_i = \text{aggregate death rate} \]
\[ A_{i,i} = -\mu_i - \lambda_i \quad \text{(row sums = 0)} \]
\[ A_{i,j} = 0 \quad \text{for } |i-j| \geq 2 \]

d) Poisson process (counts/arrivals)

\[ \lambda_i = \lambda, \quad \mu_i = 0 \]

b) Queueing theory with $k$ servers

\[ \lambda_i = \lambda \quad \text{(arrival rate)} \]
\[ M_i = \frac{i}{\mu} \]
\[ M_i = k\mu \quad \text{for } i \geq k \]

where $\mu = \frac{1}{T_{\text{serve}}} \rightarrow \text{average time for service}$

c) Population models

\[ \lambda_i = i\lambda \quad \text{or} \quad \lambda_i = i^2\lambda \quad \text{or} \quad \lambda_i = \lambda_i (k-i) \]
\[ M_i = i\mu \quad \text{or} \quad M_i = i^2\lambda \quad \text{or} \quad \lambda_i = i\lambda + \mu \]
Annoying technicality: continuous-time MC is not necessarily well-defined by the matrix $A$ if state space is infinite.

- no problem if $A$ is bounded
- have to possibly worry if $A$ has unbounded entries (KTT p. 135)
  - may be problems if $\mu_i \propto i^2$ and $\lambda_i \propto i^2$
  
  \[
  \frac{\lambda_i}{\mu_i} \rightarrow 0 \text{ as } i \rightarrow \infty
  \]

2) Finite-state processes with linear ordering
- transitions between energy levels in QM
- machine switching between modes of operation
- biomolecular switch between conformational states
How do we compute statistical properties of the continuous-time MC from its infinitesimal generator $A$?

First we'll look at $P_{ij}(t) = \text{Prob}(X(t+s) = j | X(t) = i)$,

Recall $A = \frac{dP}{dt}\bigg|_{t=0^+}$

If $A$ exists, then

$P_{ij}(\Delta t) = \delta_{ij} + \Delta t A_{ij} + o_{ij}(\Delta t)$

$\uparrow$

higher order terms

$\frac{o_{ij}(\Delta t)}{\Delta t} \to 0$ as $\Delta t \to 0$

What about $P_{ij}(t)$ when $t$ is not small?

Chapman-Kolmogorov equation in continuous-time:

$P_{ij}(t+s) = \sum_{k \in S} P_{ik}(s) P_{kj}(t)$

\[ i \quad j \quad k \]

\[ t \quad s \quad t+s \quad t \quad t+s+t \]
We'll use Chapman-Kolmogorov to recursively extend the time interval over which we know $P_{ij}(t)$.

\[ P_{ij}(t + \delta t) = \sum_{k \in S} P_{ik}(\delta t) P_{kj}(t) \]

\[ = \sum_{k \in S} \left( \delta_{ik} + \delta t A_{ik} + o_{ik}(\delta t) \right) P_{kj}(t) \]

\[ P_{ij}(t + \delta t) = P_{ij}(t) + \delta t \sum_{k \in S} A_{ik} P_{kj}(t) \]

\[ + \sum_{k \in S} o_{ik}(\delta t) P_{kj}(t) \]

\[ \frac{P_{ij}(t + \delta t) - P_{ij}(t)}{\delta t} = \sum_{k \in S} A_{ik} P_{kj}(t) + \sum_{k \in S} o_{ik}(\delta t) \frac{P_{kj}(t)}{\delta t} \]
Formal \( \Delta t \to 0 \) limit:

\[
\frac{d P_{ij}(t)}{d t} = \sum_{k \in S} A_{ik} P_{kj}(t)
\]

\[
\frac{d P}{d t} = AP
\]

\[
P(t=0) = I
\]

Backward Kolmogorov equation

Rigorous handling of the \( \Delta t \to 0 \) limit:
- Use dominated convergence theorem with \( P_{kj}(t) \leq 1 \) and \( \sum_{k \in S} A_{ik}(\Delta t) = 0 \), \( A_{ik}(\Delta t) < \infty \)
- Karlin and Taylor Vol. I + II

One could also repeat the argument starting from

\[
P_{ij}(t) = \sum_{k \in S} P_{ik}(t) P_{kj}(\Delta t)
\]

One gets from the formal calculation

\[
\frac{d P_{ij}(t)}{d t} = \sum_{k \in S} P_{ik}(t) A_{kj}
\]

Forward Kolmogorov Equation

The forward equation does not always make rigorous sense!
The solution to these equations (when they make sense) can be formally expressed as

\[ P(t) = e^{At} \]

Makes perfect sense for finite state space.

Several papers in SIAM Review

"Nineteen Ways to Compute Matrix Exponentials"

Clive Moler and...
Forward vs. Backward Equation

A) Consider first \( \phi_j(t) = \text{Prob}(X(t) = j) \)

\[
\phi_j(t) = \text{Prob}(X(t) = j) = \sum_{k \in S} \text{Prob}(X(t) = j \text{ and } X(0) = k) \\
= \sum_{k \in S} \text{Prob}(X(t) = j | X(0) = k) \text{Prob}(X(0) = k) \\
\phi_j(t) = \sum_{k \in S} p_{kj}(t) \phi_k(0)
\]

\( \Phi(t) = \Phi(0) \cdot P(t) \)

\( \frac{d\Phi}{dt} = \Phi \cdot A \)

\( = \Phi(0) \cdot P(t) \cdot A \)

Can avoid solving eqn for large matrix \( P \) by noting that \( \Phi \) itself satisfies

\( \sum \text{Kolmogorov forward eqn} \)

\( \text{but not Kolmogorov backward eqn} \)
What is the Kolmogorov backward equation good for?

Expectations starting from given state.

Let \( u_i(t) = \mathbb{E}[f(X(t)) | X(0) = i] \)

\[ = \sum_{j \in S} f(j) \text{Prob}(X(t) = j | X(0) = i) \]

\[ = \sum_{j \in S} f(j) \pi_{ij}(t) \]

\[ \overrightarrow{u}(t) = \pi(t) \cdot \overrightarrow{f} \]

This satisfies the Kolmogorov backward equation:

\[ \frac{d \overrightarrow{u}}{dt} = \overrightarrow{A} \cdot \overrightarrow{u} \]

\[ \overrightarrow{u}(t=0) = \overrightarrow{f} \]

but not the forward eqn.

Forward equation describes how probabilities of \textit{a future} time evolve.

Backward equation describes how expected "payoffs" evolve given an initial state.
What is the Kolmogorov backward equation good for?

Expectations starting from given state:

Let $u_i(t) = \mathbb{E}\left[f(X(t)) \mid X(0) = i \right]$

$$= \sum_{j \in S} f(j) \text{Prob}(X(t) = j \mid X(0) = i)$$

$$= \sum_{j \in S} f(j) \, p_{ij}(t)$$

$$\vec{u}(t) = \vec{p}(t) \cdot \vec{f}$$

This satisfies Kolmogorov backward equation:

$$\frac{d\vec{u}}{dt} = A \vec{u}$$

$$\vec{u}(t=0) = \vec{f}$$

but not the forward equation.

Forward equation describes how probabilities at a future time evolve.

Backward equation describes how expected "payoffs" evolve given an initial state.
Poisson process

Poisson counting process $X(t)$ is a birth-death process with $X(0) = 0$, $\lambda_i = \lambda$, $\mu_i = 0$.

Infinitesimal generator

$$A = \begin{pmatrix}
-\lambda & \lambda & 0 & 0 \\
0 & -\lambda & \lambda & 0 \\
0 & 0 & -\lambda & \lambda \\
0 & 0 & 0 & -\lambda
\end{pmatrix}$$

What is the probability distribution for the amount of time $X(t)$ spends in a given state?

Consider $\Phi_0(t) = \text{Prob}(X(t) = 0)$

Initialize: $\Phi_0(0) = 1$. 
\[ \frac{d \Phi}{dt} = \vec{\Phi} \cdot \vec{A} \quad \text{Kolmogorov forward eqn} \]

0th component:
\[ \frac{d \Phi_0}{dt} = -\lambda \Phi_0 \]
\[ \Phi_0(0) = 1 \]
\[ \Phi_0(t) = e^{-\lambda t} \]

Let \( T \) be the random time at which \( X(t) \) leaves state 0:
\[ T = \inf \left\{ t \geq 0 \mid X(t) \neq 0 \right\} \]

\[ \text{Prob} \left( T > t \right) = \Phi_0(t) = \text{Prob} \left( X(t) = 0 \right) \]
\[ = e^{-\lambda t} \]

So, if we write \( \text{Prob} \left( T \in B \right) = \int_B p_T(t) \, dt \)
then \[ p_T(t) = \lambda e^{-\lambda t} \]

The transition time in a Poisson process has exponential distribution, same for any state.
There is a deeper reason for why the time to transition, $T$, has exponential distribution.

- Markov property that determines it,

$$\text{Prob}(T > t | T > s) = \text{Prob}(T > t - s)$$

for $t > s$, because given that $T > s$, there is no memory about what happened before time $s$.

$$\text{Prob}(X(t) = 0 | X(s) = 0, X(0) = 0) = \text{Prob}(X(t) = 0 | X(s) = 0)$$

By definition of conditional prob,

$$\frac{\text{Prob}(T > t)}{\text{Prob}(T > s)} = \text{Prob}(T > t - s)$$

Let $G(t) = \text{Prob}(T > t)

\Rightarrow \frac{G(t)}{G(s)} = G(t - s)$

Clearly one solution is $G(t) = e^{-\lambda t}$ for some $\lambda > 0$.

This is the only class of solutions.
If \( g(t) \) is differentiable then

if differentiable \( \forall \) with respect to \( t \),

and write \( u = t - s \):

\[
G(t) = G(t-s) \cdot G(s)
\]

\[
0 = -G'(u) \cdot G(s) + G(u) \cdot G'(s)
\]

\[
\frac{G'(s)}{G(s)} = \frac{G'(u)}{G(u)} \quad \text{for all } s, u \geq 0
\]

\[
\frac{G'(s)}{G(s)} > \text{constant} +
\]

There are also no nonsmooth solutions

Karlin + Taylor Th. 4.2.2.