Continuous Time Markov Chains

Consider a population of single celled organisms in a stable environment.

Fix short time interval, length h.

Each cell has some prob of dividing to produce 2, some other prob of dying.

We might suppose:

- Different organisms behave independently.
- Probability of division (for specified organism) is λh plus $o(h)$.
- Probability of death is μh plus $o(h)$.
- Prob an organism divides twice (or divides once and dies) in interval of length h is $o(h)$.

Notice tacit assumption: constants of proportionality do not depend on time (that is our interpretation of "stable environment").

Notice too that we have taken the constants not to depend on which organism we are talking about. We are really assuming that the organisms are all similar and live in similar environments.

$Y(t)$: total population at time t.

$\mathcal{H}_t$: history of the process up to time t.

We generally take

$$\mathcal{H}_t = \sigma\{Y(s); 0 \le s \le t\}$$

General definition of a **history** (alternative jargon **filtration**): any family of σ-fields indexed by t satisfying:

- $s < t$ implies $\mathcal{H}_s \subset \mathcal{H}_t$.

- $Y(t)$ is a $\mathcal{H}_t$ measurable random variable.

- $\mathcal{H}_t = \bigcap_{s>t} \mathcal{H}_s$.

The last assumption is a technical detail we will ignore from now on.

Condition on event $Y(t) = n$.

Then the probability of two or more divisions (either more than one division by a single organism or two or more organisms dividing) is $o(h)$ by our assumptions.

Similarly the probability of both a division and a death or of two or more deaths is $o(h)$.

So probability of exactly 1 division by any one of the n organisms is $n\lambda h + o(h)$.

Similarly probability of 1 death is $n\mu h + o(h)$.

We deduce:

$$\begin{aligned}
P(Y(t+h) &= n+1|Y(t)=n,\mathcal{H}_t) \\
&= n\lambda h + o(h) \\
P(Y(t+h) &= n-1|Y(t)=n,\mathcal{H}_t) \\
&= n\mu h + o(h) \\
P(Y(t+h) &= n|Y(t)=n,\mathcal{H}_t) \\
&= 1 - n(\lambda+\mu)h + o(h) \\
P(Y(t+h) &\notin \{n-1,n,n+1\}|Y(t)=n,\mathcal{H}_t) \\
&= o(h)
\end{aligned}$$

These equations lead to:

$$\begin{aligned}
P(Y(t+s) = j|Y(s) = i,\mathcal{H}_s)& \\
&= P(Y(t+s)=j|Y(s)=i) \\
&= P(Y(t)=j|Y(0)=i)
\end{aligned}$$

This is the **Markov Property**.

Def'n: A process $\{Y(t); t \ge 0\}$ taking values in S, a finite or countable state space is a Markov Chain if

$$
\begin{aligned}
P(Y(t+s) = j | Y(s) = i, \mathcal{H}_s) \\
&= P(Y(t+s) = j | Y(s) = i) \\
&\equiv \mathbf{P}_{ij}(s, s+t)
\end{aligned}
$$

Def'n: A Markov chain Y has **stationary transitions** if

$$\mathbf{P}_{ij}(s, s+t) = \mathbf{P}_{ij}(0, t) \equiv \mathbf{P}_{ij}(t)$$

From now on: our chains have stationary transitions.

Summary of Markov Process Results

- Chapman-Kolmogorov equations:

$$\mathbf{P}_{ik}(t+s) = \sum_j \mathbf{P}_{ij}(t)\mathbf{P}_{jk}(s)$$

- Exponential holding times: starting from state i time, T_i, until process leaves i has exponential distribution, rate denoted v_i.

- Sequence of states visited, $Y_0, Y_1, Y_2, \ldots$ is Markov chain – transition matrix has $\mathbf{P}_{ii} = 0$. Y sometimes called **skeleton**.

- **Communicating classes** defined for skeleton chain. Usually assume chain has 1 communicating class.

- Periodicity irrelevant because of continuity of exponential distribution.

- Instantaneous transition rates from i to j:

$$q_{ij} = v_i \mathbf{P}_{ij}$$

- Kolmogorov backward equations:

$$\mathbf{P}'_{ij}(t) = \sum_{k \neq i} q_{ik} \mathbf{P}_{kj}(t) - v_i \mathbf{P}_{ij}(t)$$

- Kolmogorov forward equations:

$$\mathbf{P}'_{ij}(t) = \sum_{k \neq j} q_{kj} \mathbf{P}_{ik}(t) - v_i \mathbf{P}_{ij}(t)$$

- For strongly recurrent chains with a single communicating class:

$$\mathbf{P}_{ij}(t) \to \pi_j$$

- Stationary initial probabilities π_i satisfy:

$$v_j \pi_j = \sum_{k \neq j} q_{kj} \pi_k$$

- Transition probabilities given by

$$\mathbf{P}(t) = e^{\mathbf{R}t}$$

where $\mathbf{R}$ has entries

$$\mathbf{R}_{ij} = \begin{cases} q_{ij} & i \neq j \\ -v_i & i = j \end{cases}$$

- Process is a **Birth and Death** process if

$$\mathbf{P}_{ij} = 0 \text{ if } |i - j| > 1$$

In this case we write λ_i for the instantaneous "birth" rate:

$$P(Y(t+h) = i+1 | Y_t = i) = \lambda_i h + o(h)$$

and μ_i for the instantaneous "death" rate:

$$P(Y(t+h) = i-1 | Y_t = i) = \mu_i h + o(h)$$

We have

$$q_{ij} = \begin{cases} 0 & |i - j| > 1 \\ \lambda_i & j = i+1 \\ \mu_i & j = i-1 \end{cases}$$

- If all $\mu_i = 0$ then process is a **pure birth** process. If all $\lambda_i = 0$ a **pure death** process.

- Birth and Death process have stationary distribution

$$\pi_n = \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}\right)}$$

  Necessary condition for existence of π is

$$\sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} < \infty$$

Detailed development

Suppose X a Markov Chain with stationary transitions. Then

$$
\begin{aligned}
P(X(t+s) = k|X(0) = i)& \\
&= \sum_j P(X(t+s) = k, X(t) = j|X(0) = i) \\
&= \sum_j P(X(t+s) = k|X(t) = j, X(0) = i) \\
&\qquad \times P(X(t) = j|X(0) = i) \\
&= \sum_j P(X(t+s) = k|X(t) = j) \\
&\qquad \times P(X(t) = j|X(0) = i) \\
&= \sum_j P(X(s) = k|X(0) = j) \\
&\qquad \times P(X(t) = j|X(0) = i)
\end{aligned}
$$

This shows

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s) = \mathbf{P}(s)\mathbf{P}(t)$$

which is the Chapman-Kolmogorov equation.

Now consider the chain starting from i and let T_i be the first t for which $X(t) \neq i$. Then T_i is a stopping time.

[Technically:

$$\{T_i \le t\} \in \mathcal{H}_t$$

for each t.] Then

$$\begin{aligned} P(T_i > t + s | T_i > s, X(0) = i) \\ &= P(T_i > t + s | X(u) = i; 0 \le u \le s) \\ &= P(T_i > t | X(0) = i) \end{aligned}$$

by the Markov property. Note: we actually are asserting a generalization of the Markov property: If f is some function on the set of possible paths of X then

$$\begin{aligned} \mathsf{E}(f(X(s+\cdot)) | X(u) = x(u), 0 \le u \le s) \\ &= \mathsf{E}\left[f(X(\cdot)) | X(0) = x(s)\right] \\ &= \mathsf{E}^{x(s)}\left[f(X(\cdot))\right] \end{aligned}$$

The formula requires some sophistication to appreciate. In it, f is a function which associates a sample path of X with a real number. For instance,

$$f(x(\cdot)) = \sup\{t : x(u) = i, 0 \le u \le t\}$$

is such a functional. Jargon: **functional** is a function whose argument is itself a function and whose value is a scalar.

FACT: Strong Markov Property – for a stopping time T

$$\mathsf{E}\left[f\left\{X(T+\cdot)\right\}|\mathcal{F}_T\right] = \mathsf{E}^{X(T)}\left[f\left\{X(\cdot)\right\}\right]$$

with suitable fix on event $T = \infty$.

Conclusion: given $X(0) = i$, T_i has memoryless property so T_i has an exponential distribution. Let v_i be the rate parameter.

Embedded Chain: Skeleton

Let $T_1 < T_2 < \cdots$ be the stopping times at which transitions occur.

Then $X_n = X(T_n)$.

Sequence X_n is a Markov chain by the strong Markov property.

That $\mathbf{P}_{ii} = 0$ reflects fact that $P(X(T_{n+1}) = X(T_n)) = 0$ by design.

As before we say $i \rightsquigarrow j$ if $\mathbf{P}_{ij}(t) > 0$ for some t. It is fairly clear that $i \rightsquigarrow j$ for the $X(t)$ if and only if $i \rightsquigarrow j$ for the embedded chain X_n.

We say $i \leftrightarrow j$ if $i \rightsquigarrow j$ and $j \rightsquigarrow i$.

Now consider

$$P(X(t+h) = j | X(t) = i, \mathcal{H}_t)$$

Suppose the chain has made n transitions so far so that $T_n < t < T_{n+1}$. Then the event $X(t+h) = j$ is, except for possibilities of probability $o(h)$ the event that

$$t < T_{n+1} \le t + h \text{ and } X_{n+1} = j$$

The probability of this is

$$(v_i h + o(h))\mathbf{P}_{ij} = v_i \mathbf{P}_{ij} h + o(h)$$

Kolmogorov's Equations

The Chapman-Kolmogorov equations are

$$\mathbf{P}(t+h) = \mathbf{P}(t)\mathbf{P}(h)$$

Subtract $\mathbf{P}(t)$ from both sides, divide by h and let $h \to 0$. Remember that $\mathbf{P}(0)$ is the identity. We find

$$\frac{\mathbf{P}(t+h) - \mathbf{P}(t)}{h} = \frac{\mathbf{P}(t)(\mathbf{P}(h) - \mathbf{P}(0))}{h}$$

which gives

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{P}'(0)$$

The Chapman-Kolmogorov equations can also be written

$$\mathbf{P}(t+h) = \mathbf{P}(h)\mathbf{P}(t)$$

Now subtracting $\mathbf{P}(t)$ from both sides, dividing by h and letting $h \to 0$ gives

$$\mathbf{P}'(t) = \mathbf{P}'(0)\mathbf{P}(t)$$

Look at these equations in component form:

$$\mathbf{P}'(t) = \mathbf{P}'(0)\mathbf{P}(t)$$

becomes

$$\mathbf{P}'_{ij}(t) = \sum_k \mathbf{P}'_{ik}(0)\mathbf{P}_{kj}(t)$$

For $i \neq k$ our calculations of instantaneous transition rates gives

$$\mathbf{P}'_{ik}(0) = v_i \mathbf{P}_{ik}$$

For $i = k$ we have

$$P(X(h) = i | X(0) = i) = e^{-v_i h} + o(h)$$

($X(h) = i$ either means $T_i > h$ which has probability $e^{-v_i h}$ or there have been two or more transitions in $[0, h]$, a possibility of probability $o(h)$.) Thus

$$\mathbf{P}'_{ii}(0) = -v_i$$

Let $\mathbf{R}$ be the matrix with entries

$$\mathbf{R}_{ij} = \begin{cases} q_{ij} \equiv v_i \mathbf{P}_{ij} & i \neq j \\ -v_i & i = j \end{cases}$$

$\mathbf{R}$ is the **infinitesimal generator** of the chain.

Thus

$$\mathbf{P}'(t) = \mathbf{P}'(0)\mathbf{P}(t)$$

becomes

$$\begin{aligned} \mathbf{P}'_{ij}(t) &= \sum_k \mathbf{R}_{ik} \mathbf{P}_{kj}(t) \\ &= \sum_{k \neq i} q_{ik} \mathbf{P}_{kj}(t) - v_i \mathbf{P}_{ij}(t) \end{aligned}$$

Called **Kolmogorov's backward equations**.

On the other hand

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{P}'(0)$$

becomes

$$\begin{aligned}\mathbf{P}'_{ij}(t) &= \sum_k \mathbf{P}_{ik}(t)\mathbf{R}_{kj} \\ &= \sum_{k \neq j} q_{kj}\mathbf{P}_{ik}(t) - v_j \mathbf{P}_{ij}(t)\end{aligned}$$

These are **Kolmogorov's forward equations**.

Remark: When the state space is infinite the forward equations may not be justified. In deriving them we interchanged a limit with an infinite sum; the interchange is always justified for the backward equations but not for forward.

Example: $S = \{0, 1\}$. Then

$$\mathbf{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and the chain is otherwise specified by v_0 and v_1. The matrix $\mathbf{R}$ is

$$\mathbf{R} = \begin{bmatrix} -v_0 & v_0 \\ v_1 & -v_1 \end{bmatrix}$$

The backward equations become

$$\begin{aligned}
\mathbf{P}'_{00}(t) &= v_0 \mathbf{P}_{10}(t) - v_0 \mathbf{P}_{00}(t) \\
\mathbf{P}'_{01}(t) &= v_0 \mathbf{P}_{11}(t) - v_0 \mathbf{P}_{01}(t) \\
\mathbf{P}'_{10}(t) &= v_1 \mathbf{P}_{00}(t) - v_1 \mathbf{P}_{10}(t) \\
\mathbf{P}'_{11}(t) &= v_1 \mathbf{P}_{01}(t) - v_1 \mathbf{P}_{11}(t)
\end{aligned}$$

while the forward equations are

$$\begin{aligned}
\mathbf{P}'_{00}(t) &= v_1 \mathbf{P}_{01}(t) - v_0 \mathbf{P}_{00}(t) \\
\mathbf{P}'_{01}(t) &= v_0 \mathbf{P}_{00}(t) - v_1 \mathbf{P}_{01}(t) \\
\mathbf{P}'_{10}(t) &= v_1 \mathbf{P}_{11}(t) - v_0 \mathbf{P}_{10}(t) \\
\mathbf{P}'_{11}(t) &= v_0 \mathbf{P}_{10}(t) - v_1 \mathbf{P}_{11}(t)
\end{aligned}$$

Add v_1 times first and v_0 times third backward equations to get

$$v_1 \mathbf{P}'_{00}(t) + v_0 \mathbf{P}'_{10}(t) = 0$$

so

$$v_1 \mathbf{P}_{00}(t) + v_0 \mathbf{P}_{10}(t) = c.$$

Put $t = 0$ to get $c = v_1$. This gives

$$\mathbf{P}_{10}(t) = \frac{v_1}{v_0}\{1 - \mathbf{P}_{00}(t)\}$$

Plug this back in to the first equation and get

$$\mathbf{P}'_{00}(t) = v_1 - (v_1 + v_0)\mathbf{P}_{00}(t)$$

Multiply by $e^{(v_1+v_0)t}$ and get

$$\left\{e^{(v_1+v_0)t}\mathbf{P}_{00}(t)\right\}' = v_1 e^{(v_1+v_0)t}$$

which can be integrated to get

$$\mathbf{P}_{00}(t) = \frac{v_1}{v_0 + v_1} + \frac{v_0}{v_0 + v_1} e^{-(v_1+v_0)t}$$

Alternative calculation:

$$\mathbf{R} = \begin{bmatrix} -v_0 & v_0 \\ v_1 & -v_1 \end{bmatrix}$$

can be written as

$$\mathbf{M\Lambda M}^{-1}$$

where

$$\mathbf{M} = \begin{bmatrix} 1 & v_0 \\ & \\ 1 & -v_1 \end{bmatrix}$$

$$\mathbf{M}^{-1} = \begin{bmatrix} \frac{v_1}{v_0+v_1} & \frac{v_0}{v_0+v_1} \\ & \\ \frac{1}{v_0+v_1} & \frac{-1}{v_0+v_1} \end{bmatrix}$$

and

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 0 \\ & \\ 0 & -(v_0 + v_1) \end{bmatrix}$$

Then

$$\begin{aligned} e^{\mathbf{R}t} &= \sum_0^\infty \mathbf{R}^n t^n / n! \\ &= \sum_0^\infty \left(\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1}\right)^n \frac{t^n}{n!} \\ &= \mathbf{M}\left(\sum_0^\infty \mathbf{\Lambda}^n \frac{t^n}{n!}\right)\mathbf{M}^{-1} \end{aligned}$$

Now

$$\sum_0^\infty \mathbf{\Lambda}^n \frac{t^n}{n!} = \begin{bmatrix} 1 & 0 \\ 0 & e^{-(v_0+v_1)t} \end{bmatrix}$$

so we get

$$\begin{aligned} \mathbf{P}(t) = e^{\mathbf{R}t} &= \mathbf{M}\begin{bmatrix} 1 & 0 \\ 0 & e^{-(v_0+v_1)t} \end{bmatrix}\mathbf{M}^{-1} \\ &= \mathbf{P}^\infty - \frac{e^{-(v_0+v_1)t}}{v_0+v_1}\mathbf{R} \end{aligned}$$

where

$$\mathbf{P}^\infty = \begin{bmatrix} \frac{v_1}{v_0+v_1} & \frac{v_0}{v_0+v_1} \\ \frac{v_1}{v_0+v_1} & \frac{v_0}{v_0+v_1} \end{bmatrix}$$

Linear Algebra

Theorem: If $\mathbf{A}$ is a $n \times n$ matrix then there are matrices $\mathbf{D}$ and $\mathbf{N}$ such that

1. $\mathbf{D}$ is diagonalizable:

   $$\mathbf{D} = \mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1}$$

   for some invertible $\mathbf{M}$ and diagonal $\mathbf{\Lambda}$.

2. $\mathbf{N}$ is **nilpotent**: there is $r \le n$ such that $\mathbf{N}^r = \mathbf{0}$.

3. $\mathbf{N}$ and $\mathbf{D}$ commute: $\mathbf{N}\mathbf{D} = \mathbf{D}\mathbf{N}$.

In this case

$$\begin{aligned}\mathbf{A}^k &= (\mathbf{D}+\mathbf{N})^k \\ &= \sum_{j=0}^{k} \binom{k}{j} \mathbf{D}^{k-j}\mathbf{N}^j\end{aligned}$$

Thus

$$
\begin{aligned}
e^{\mathbf{A}} &= \sum_{k=0}^{\infty} \mathbf{A}^k/k! \\
&= \sum_{k=0}^{\infty} (\mathbf{D}+\mathbf{N})^k/k! \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{(k-j)!j!}\mathbf{D}^{k-j}\mathbf{N}^j \\
&= \sum_{j=0}^{r-1} \frac{\mathbf{N}^j}{j!} \sum_{k=j}^{\infty} \frac{\mathbf{D}^{k-j}}{(k-j)!} \\
&= \sum_{j=0}^{r-1} \frac{\mathbf{N}^j}{j!} \sum_{k=0}^{\infty} \frac{\mathbf{D}^{k}}{k!} \\
&= e^{\mathbf{D}} \sum_{j=0}^{r-1} \frac{\mathbf{N}^j}{j!}
\end{aligned}
$$

Notice: rows of $\mathbf{P}^\infty$ are a stationary initial distribution. If rows are π then

$$\mathbf{P}^\infty = \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \pi \equiv \mathbf{1}\pi$$

so

$$\pi\mathbf{P}^\infty = (\pi\mathbf{1})\pi = \pi$$

Moreover

$$\pi\mathbf{R} = \mathbf{0}$$

Fact: $\pi_0 = v_1/(v_0+v_1)$ is long run fraction of time in state 0.

Fact:

$$\frac{1}{T}\int_0^T f(X(t))dt \to \sum_j \pi_j f(j)$$

Ergodic Theorem in continuous time.

Potential Pathologies

Suppose that for each k you have a sequence

$$T_{k,1}, T_{k,2}, \cdots$$

such that all T_{ij} are independent exponential random variables and T_{ij} has rate parameter λ_j. We can use these times to make a Markov chain with state space $S = \{1, 2, \ldots\}$:

Start the chain in state 1. At time $T_{1,1}$ move to 2, $T_{1,2}$ time units later move to 3, etc. Chain progresses through states in order 1,2,.... We have

$$v_i = \lambda_i$$

and

$$\mathbf{P}_{ij} = \begin{cases} 0 & j \neq i+1 \\ 1 & j = i+1 \end{cases}$$

Does this define a process?

Depends on $\sum \lambda_i^{-1}$.

Case 1: if $\sum \lambda_i^{-1} = \infty$ then

$$P(\sum_1^\infty T_{1,j} = \infty) = 1$$

(converse to Borel Cantelli) and our construction defines a process $X(t)$ for all t.

Case 2: if $\sum \lambda_j^{-1} < \infty$ then for each k

$$P(\sum_{j=1}^\infty T_{kj} < \infty) = 1$$

In this case put $T_k = \sum_{j=1}^\infty T_{kj}$. Our definition above defines a process $X(t)$ for $0 \le t < T_1$. We put $X(T_1) = 1$ and then begin the process over with the set of holding times $T_{2,j}$. This defines X for $T_1 \le t < T_1 + T_2$. Again we put $X(T_2) = 1$ and continue the process.

Result: X is a Markov Chain with specified transition rates.

Problem: what if we put $X(T_1) = 2$ and continued?

What if we used probability vector $\alpha_1, \alpha_2, \ldots$ to pick a value for $X(T_1)$ and continued?

All yield Markov Processes with the same infinitesimal generator $\mathbf{R}$.

Point of all this: gives example of non-unique solution of differential equations!

Birth and Death Processes

Consider a population of $X(t) = i$ individuals. Suppose in next time interval $(t, t+h)$ probability of population increase of 1 (called a birth) is $\lambda_i h + o(h)$ and probability of decrease of 1 (death) is $\mu_i h + o(h)$.

Jargon: X is a birth and death process.

Special cases:

All $\mu_i = 0$; called a **pure birth** process.

All $\lambda_i = 0$ (0 is absorbing): **pure death** process.

$\lambda_n = n\lambda$ and $\mu_n = n\mu$ is a **linear** birth and death process.

$\lambda_n \equiv 1$, $\mu_n \equiv 0$: Poisson Process.

$\lambda_n = n\lambda + \theta$ and $\mu_n = n\mu$ is a **linear** birth and death process with immigration.

Applications:

1) cable strength: Cable consists of n fibres.

$X(t)$ is number which have *not* failed up to time t.

Pure death process: μ_i will be large for small i, small for large i.

2) Chain reactions. $X(t)$ is number of free neutrons in lump of uranium.

Births produced as sum of: spontaneous fission rate (problem — I think each fission produces 2 neutrons) plus rate of collision of neutron with nuclei.

Ignore: neutrons leaving sample and decay of free neutrons.

Get $\lambda_n = n\lambda + \theta$

(At least in early stages where decay has removed a negligible fraction of atoms).

Stationary initial distributions.

As in discrete time an initial distribution is probability vector π with

$$P(X(0) = i) = \pi_i$$

An initial distribution π is **stationary** if

$$\pi = \pi \mathbf{P}(t)$$

or

$$P(X(t) = i) = \pi_i$$

for all $t \geq 0$.

If so take derivative wrt t to get

$$0 = \pi \mathbf{P}'(t)$$

or

$$\pi \mathbf{R} = 0$$

Conversely: if

$$\pi \mathbf{R} = 0$$

then

$$\begin{aligned} \pi \mathbf{P}(t) &= \pi e^{\mathbf{R}t} \\ &= \pi \left(\mathbf{I} + \mathbf{R}t + \mathbf{R}^2 t^2 / 2 + \cdots \right) \\ &= \pi \end{aligned}$$

So a probability vector π such that

$$\pi \mathbf{R} = 0$$

is a stationary initial distribution.

NOTE: π is a left eigenvector of $\mathbf{P}(t)$.

Perron-Frobenius theorem asserts that 1 is the largest (in modulus) eigenvalue of $\mathbf{P}(t)$, that this eigenvalue has multiplicity 1, that the corresponding eigenvector has all positive entries. So: can prove every row of $\mathbf{P}(t)$ converges to π.

Conditions for stationary initial distribution:

1) $v_n = \lambda_n + \mu_n$.

2) $P_{n,n+1} = \lambda_n / v_n = 1 - P_{n,n-1}$.

3) From $\pi \mathbf{R} = 0$:

$$v_n \pi_n = \lambda_{n-1} \pi_{n-1} + \mu_{n+1} \pi_{n+1}$$

4) Start at $n = 0$:

$$\lambda_0 \pi_0 = \mu_1 \pi_1$$

so $\pi_1 = (\lambda_0 / \mu_1) \pi_0$.

5) Now look at $n = 1$.

$$(\lambda_1 + \mu_1) \pi_1 = \lambda_0 \pi_0 + \mu_2 \pi_2$$

Solve for π_2 to get

$$\pi_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} \pi_0$$

And so on. Then use $\sum \pi_n = 1$.

Relation of π to stationary initial distribution of skeleton chain.

Let α be stationary initial dist of skeleton.

Heuristic: fraction of time in state j proportional to

fraction of skeleton visits to state j times average time spent in state j:

$$\pi_j \propto \alpha_j \times (1/v_j)$$