STAT 380
Continuous Time Markov Chains

Richard Lockhart

Simon Fraser University

Spring 2016

Continuous Time Markov Chains

- Population of single celled organisms in a stable environment.
- Fix short time interval, length h.
- Each organism has some probability of dividing to produce two organisms and some other probability of dying.
- Possible assumptions:
  - Different organisms behave independently.
  - Probability of division is λh plus $o(h)$.
  - Probability of death is μh plus $o(h)$.
  - Probability that an organism divides twice (or divides once and dies) in the interval of length h is $o(h)$.

Tacit Assumptions

- Constants of proportionality do not depend on time: "stable environment".
- Constants do not depend on organism: organisms are all similar and live in similar environments.
- $Y(t)$: total population at time t.
- $\mathcal{H}_t$: history of process up to time t.
- Condition on event $Y(t) = n$.
- Probability of two or more divisions (more than one division by a single organism or two or more organisms dividing) is $o(h)$.
- Probability of both a division and a death or of two or more deaths is $o(h)$.

Tacit Assumptions II

- So prob of exactly 1 division by any one of 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**.

Tacit Assumptions III

- **Definition**: A process $\{Y(t); t \geq 0\}$ taking values in S, a finite or countable state space is a Markov Chain if

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

- **Definition**: 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.

Summary of Markov Process Results II

- 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:

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

- For strongly recurrent chains with a single communicating class:

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

- Stationary initial probabilities π_i satisfy:

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

Summary of Markov Process Results III

- 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:

$$P(X(t+h) = i+1 | X_t = i) = \lambda_i h + o(h)$$

and μ_i for the instantaneous "death" rate:

$$P(X(t+h) = i-1 | X_t = i) = \mu_i h + o(h)$$

Summary of Markov Process Results IV

- 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

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

Summary of Markov Process Results V

- Necessary condition for existence of π is

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

- **Linear** birth and death processes have

$$\lambda_n = (n+1)\lambda \qquad \mu_n = n\mu$$

- In this case (note *immigration*)

$$\pi_n = (\lambda/\mu)^n(1 - (\lambda/\mu))$$

provided $\lambda < \mu$.

Details

- Suppose X a Markov Chain with stationary transitions.
- Then

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

- This shows

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

which is the Chapman-Kolmogorov equation.

Holding Times

- 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.
- T_i called a *holding* time.
- Technically, for each t,

$$\{T_i \leq t\} \in \mathcal{H}_t.$$

- Then, by the Markov property,

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

- Given $X(0) = i$, T_i has memoryless property so T_i has exponential distribution.
- Let v_i be the rate parameter.

Embedded Chain: Skeleton

- Let $S_1 < S_2 < \cdots$ be the stopping times at which transitions occur.
- Then $X_n = X(S_n)$.
- The sequence X_n is a Markov chain by the Markov property.
- That $\mathbf{P}_{ii} = 0$ reflects the fact that $P(X(S_{n+1}) = X(S_n)) = 0$ by design.
- As before we say $i \leadsto j$ if $\mathbf{P}_{ij}(t) > 0$ for some t.
- It is fairly clear that $i \leadsto j$ for the $X(t)$ if and only if $i \leadsto j$ for the embedded chain X_n.
- We say $i \leftrightarrow j$ (i and j communicate) if $i \leadsto j$ and $j \leadsto i$.

Towards coupled differential equations

- Now consider
$$P(X(t+h)=j|X(t)=i,\mathcal{H}_t)$$
- Suppose the chain has made n transitions so far so that $S_n < t < S_{n+1}$.
- Then the event $X(t+h)=j$ is, except for possibilities of probability $o(h)$ the event that
$$t < S_{n+1} \le t+h \text{ and } X_{n+1}=j$$
- 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

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

Forward and Backward Equations

- 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

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

Forward and Backward Equations II

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

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

- That is

$$\mathbf{R} = \mathbf{P}'(0).$$

  $\mathbf{R}$ 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**.

Forwards 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;
- Interchange always justified for backward equations but not for forward.

Example

- 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

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

Example: Kolmogorov equations

- 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

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

Solving Kolmogorov's equations

- Add $v_1 \times$first plus $v_0 \times$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

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

Solving Kolmogorov's equations

- 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

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

Matrix exponentials

- Alternative calculation:

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

can be written as

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

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

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

Matrix exponentials

- Then

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

- Now

$$
\sum_0^\infty \boldsymbol{\Lambda}^n \frac{t^n}{n!} = \left[\begin{array}{cc} 1 & 0 \\ 0 & e^{-(v_0+v_1)t} \end{array}\right]
$$

Matrix exponentials

- So

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

  where

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

Matrix exponentials

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

$$\mathbf{P}^\infty = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \pi \equiv \mathbf{1}\pi$$

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.

Birth and Death Processes

- Consider a population of $X(t)$ individuals.
- Suppose in next time interval $(t, t+h)$ probability of population increase of 1 (called a birth) is $\lambda_i h + o(h)$
- Suppose 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.
  - $\lambda_n \equiv \lambda$, $\mu_n \equiv 0$: Poisson Process.
  - $\lambda_n = n\lambda + \theta$ and $\mu_n = n\mu$ is a **linear** birth and death process with immigration.

Example: Queuing Theory

Ingredients of Queuing Problem:

- Queue input process.
- Number of servers
- Queue discipline: first come first serve? last in first out? pre-emptive priorities?
- Service time distribution.

Example

- Customers arrive at a facility at times of a Poisson Process
- Call this *input* process N, with rate λ.
- Denote Poisson inputs with M (for Markov) in queueing literature.
- Single server case:
- Service distribution: exponential service times, rate μ.
- Queue discipline: first come first serve.

Analysis via Continuous Time Markov Chain

- Let $X(t)$ = number of customers in line at time t.
- Then X is a Markov process called $M/M/1$ queue.
- Departure rate from a state:

$$v_i = \lambda + \mu$$

- Skeleton chain transition probabilities:

$$\mathbf{P}_{ij} = \begin{cases} \frac{\mu}{\mu+\lambda} & j = i - 1 \geq 0 \\ \frac{\lambda}{\mu+\lambda} & j = i + 1 \\ 0 & \text{otherwise} \end{cases}$$

Example: $M/M/\infty$ queue

- Customers arrive according to PP rate λ.
- Each customer begins service immediately.
- $X(t)$ is number being served at time t.
- X is a birth and death process with

$$v_n = \lambda + n\mu$$

and

$$\mathbf{P}_{ij} = \begin{cases} \frac{i\mu}{i\mu+\lambda} & j = i-1 \geq 0 \\ \frac{\lambda}{i\mu+\lambda} & j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

Stationary Initial Distributions

- Recall: stationary initial distribution π is a probability vector solving

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

- Rewrite this as

$$v_j \pi_j = \sum_{i \neq j} q_{ij} \pi_i$$

- Interpretation: LHS is rate at which process leaves state j.
- Process is in state j a fraction π_j of time, then makes transition at rate v_j.
- RHS is total rate of arrival in state j.
- For each state $i \neq j$ π_i is fraction of time spent in state i and then q_{ij} the instantaneous rate of transition from i to j.

Interpretation of equation

- So equation says:
- Rate of departure from j balances rate of arrival to j.
- This is called *balance*.
- Application to birth and death processes:
- Equation is

$$(\lambda_j + \mu_j)\pi_j = \lambda_{j-1}\pi_{j-1} + \mu_{j+1}\pi_{j+1}$$

for $j \geq 1$ and

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

Stationary Distribution for Birth and Death

- Equations permit recursion:

$$\pi_1 = \frac{\lambda_0}{\mu_1}\pi_0$$

$$\pi_2 = \frac{\lambda_1 + \mu_1}{\mu_2}\pi_1 - \frac{\lambda_0}{\mu_2}\pi_0 = \frac{\lambda_0\lambda_1}{\mu_1\mu_2}\pi_0$$

- Extend by induction to

$$\pi_k = \frac{\lambda_0 \cdots \lambda_{k-1}}{\mu_1 \cdots \mu_k}\pi_0$$

- Apply $\sum \pi_k = 1$ to get

$$\pi_0 \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}\right) = 1$$

- This gives the formula announced:

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

Interpretation – Richard should add warnings

- If
$$\sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} < \infty$$
then we have defined a probability vector which solves
$$\pi \mathbf{R} = \mathbf{0}$$
- Since
$$\mathbf{P}' = \mathbf{R}\mathbf{P}$$
we see that
$$\{\pi \mathbf{P}(t)\}' = 0$$
so that $\pi \mathbf{P}(t)$ is constant.
- Put $t = 0$ to discover that the constant is π.