Continuous time Markov chains

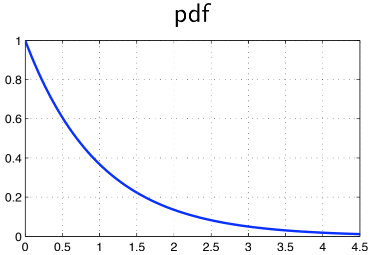
Alejandro Ribeiro
Dept. of Electrical and Systems Engineering
University of Pennsylvania
aribeiro@seas.upenn.edu
http://www.seas.upenn.edu/users/~aribeiro/

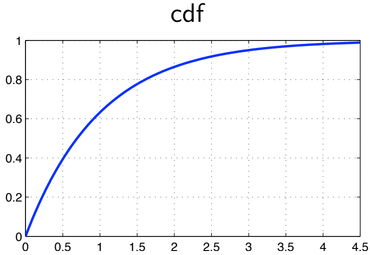
October 16, 2017

- Exponential RVs are used to model times at which events occur
- Or in general time elapsed between occurrence of random events
- RV $T \sim \exp(\lambda)$ is exponential with parameter λ if its pdf is

$$f_T(t) = \lambda e^{-\lambda t}, \qquad \text{for all } t \geq 0$$

- The cdf, integral of the pdf, is $(t \geq 0)$ $\Rightarrow F_T(t) = 1 - e^{-\lambda t}$
- Complementary (c)cdf is $\Rightarrow P(T \geq t) = 1 - F_T(t) = e^{-\lambda t}$





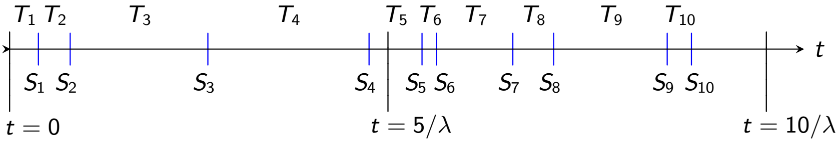
Expected value

- The expected value of time $T \sim \exp(\lambda)$ is

$$\mathbb{E}\left[T\right] = \int_0^\infty t\lambda e^{-\lambda t} = -te^{-\lambda t}\Big|_0^\infty + \int_0^\infty e^{-\lambda t} = 0 + \frac{1}{\lambda}$$

- Integrated by parts with $u = t$, $dv = \lambda e^{-\lambda t}$
- Mean time is inverse of parameter λ

  $\Rightarrow$ λ is rate/frequency of events happening at intervals T
- Average of λt events in time t
- Bigger lambda, smaller expected times, larger frequency of events



Second moment and variance

- For second moment also integrate by parts ($u = t^2$, $dv = \lambda e^{-\lambda t}$)

$$\mathbb{E}\left[T^2\right] = \int_0^\infty t^2 \lambda e^{-\lambda t} = -t^2 e^{-\lambda t}\Big|_0^\infty + \int_0^\infty 2te^{-\lambda t}$$

- Fisrt term is 0, second is, except for a constant, as computed before

$$\mathbb{E}\left[T^2\right] = \frac{2}{\lambda}\int_0^\infty t\lambda e^{-\lambda t} = \frac{2}{\lambda^2}$$

- The variance is computed from the mean and variance

$$\text{var}\left[T\right] = \mathbb{E}\left[T^2\right] - \mathbb{E}\left[T\right]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

- Parameter λ controls mean and variance of exponential RV

Memoryless random times

- Consider random time T. We say time T is memoryless if

$$\mathrm{P}\left[T > s + t \,\middle|\, T > t\right] = \mathrm{P}\left[T > s\right]$$

- Probability of waiting s extra units of time given that we waited t seconds is just the probability of waiting s seconds
  - $\Rightarrow$ System does not remember it has already waited t units
  - $\Rightarrow$ Same probability irrespectively of time already elapsed

Ex: Chemical reaction $A + B \rightarrow AB$ occurs when molecules A and B "collide". A, B move around randomly. Time T until reaction

Ex: Group of molecules of type A and type B. Reaction occurs when any type A encounters any type B. Time T until next reaction

- Write memoryless property in terms of joint pdf

$$P\left[T > s + t \,\middle|\, T > t\right] = \frac{P\left[T > s + t, T > t\right]}{P\left[T > t\right]} = P\left[T > s\right]$$

- Notice that having $T > s + t$ and $T > t$ is equivalent to $T > s + t$
- Replace $P\left[T > s + t, T > t\right] = P\left[T > s + t\right]$ and reorder terms

$$P\left[T > s + t\right] = P\left[T > t\right]P\left[T > s\right]$$

- If T is exponentially distributed ccdf is $P\left[T > t\right] = e^{-\lambda t}$. Then

$$P\left[T > s + t\right] = e^{-\lambda(s+t)} = e^{-\lambda t}e^{-\lambda s} = P\left[T > t\right]P\left[T > s\right]$$

- If random time T is exponential $\Rightarrow$ T is memoryless

Continuous memoryless RVs are exponential

- Consider a function $g(t)$ with the property $g(t+s) = g(t)g(s)$
- Functional form of $g(t)$? Take logarithms

$$\log g(t+s) = \log g(t) + \log g(s)$$

- Can only be true for all t and s if $\log g(t) = ct$ for some constant c
- Which in turn, can only true if $g(t) = e^{ct}$ for some constant c

- Compare observation with statement of memoryless property

$$\mathrm{P}[T > s+t] = \mathrm{P}[T > t]\,\mathrm{P}[T > s]$$

- It must be $\mathrm{P}[T > t] = e^{ct}$ for some constant c
- If T is continuous this can only be true for exponential T
- If T discrete it is geometric $\mathrm{P}[T > t] = (1-p)^t$ with $(1-p) = e^c$

- If continuous random time T is memoryless $\Rightarrow$ T is exponential

Combining results

Theorem

A continuous random variable T is memoryless if and only if it is exponentially distributed. That is

$$P\left[T > s+t \,\middle|\, T > t\right] = P\left[T > s\right]$$

if and only if $f_T(t) = \lambda e^{-\lambda t}$ for some $\lambda > 0$

- Exponential RVs are memoryless. Do not remember elapsed time
- They are the only type of continuous memoryless RV
- Discrete RV T is memoryless if and only of it is geometric
- Geometrics are discrete approximations of exponentials
- Exponentials are continuous limits of geometrics
- Exponential = time until success $\Leftrightarrow$ Geometric = nr. trials until success

Time to first event

- Independent exponential RVs T_1, T_2 with parameters λ_1, λ_2
- Probability distribution of first event, i.e., $T := \min(T_1, T_2)$?
- For having $T > t$ we need both $T_1 > t$ and $T_2 > t$
- Using independence of X and Y we can write

$$\mathrm{P}\left[T > t\right] = \mathrm{P}\left[T_1 > t\right]\mathrm{P}\left[T_2 > t\right] = \big(1 - F_{T_1}(t)\big)\big(1 - F_{T_2}(t)\big)$$

- Substituting expressions of exponential cdfs

$$\mathrm{P}\left[T > t\right] = e^{-\lambda_1 t}e^{-\lambda_2 t} = e^{-(\lambda_1+\lambda_2)t}$$

- T is exponentially distributed with parameter $\lambda_1 + \lambda_2$

- Minimum of exponential variables is exponential
- Given two events happening at random exponential times (T_1, T_2)

  $\Rightarrow$ Time to any of them happening is also exponential

First event to happen

- Prob. $P[T_1 < T_2]$ of $T_1 \sim \exp(\lambda_1)$ happening before $T_2 \sim \exp(\lambda_2)$
- Condition on $T_2 = t$ and integrate over the pdf of T_2

$$P[T_1 < T_2] = \int_0^\infty P\left[T_1 < t \,\middle|\, T_2 = t\right] f_{T_2}(t)\,dt = \int_0^\infty F_{T_1}(t) f_{T_2}(t)\,dt$$

- Substitute expressions for exponential pdf and cdf

$$P[T_1 < T_2] = \int_0^\infty (1 - e^{-\lambda_1 t})\lambda_2 e^{-\lambda_2 t}\,dt = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

- Either X comes before Y or vice versa then

$$P[T_2 < T_1] = 1 - P[T_1 < T_2] = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

- Probabilities are relative values of parameters
- Larger parameter $\Rightarrow$ smaller average $\Rightarrow$ larger prob. happening first

Probability of event in infinitesimal time

- Probability of an event happening in infinitesimal time h?
- Want $\text{P}[T < h]$ for small h,

$$\text{P}[T < h] = \int_0^h \lambda e^{-\lambda t}\, dt \approx \lambda h$$

- Equivalent to $\left.\frac{\partial \text{P}[T < t]}{\partial t}\right|_{t=0} = \lambda$
- Sometimes also write $\text{P}[T < h] = \lambda h + o(h)$
- $o(h)$ implies $\lim_{h \to 0} \frac{o(h)}{h} = 0$. Read as "negligible with respect to h"
- Two independent events in infinitesimal time h?

$$\text{P}[T_1 \leq h, T_2 \leq h] \approx (\lambda_1 h)(\lambda_2 h) = \lambda_1 \lambda_2 h^2 = o(h)$$

Counting and Poisson processes

Exponential random variables

Counting processes and definition of Poisson processes

Properties of Poisson processes

Continuous time Markov chains

Transition probability function

Determination of transition probability function

Limit probabilities

Counting processes

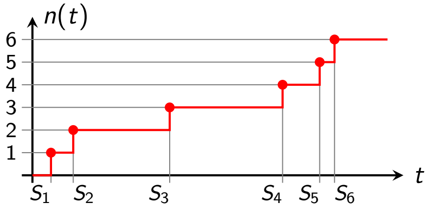
- Stochastic process $N(t)$ taking integer values $0, 1, \ldots, i, \ldots$
- Counting process $N(t)$ counts number of events occurred by time t

- Nonnegative integer valued: $N(0) = 0$, $N(t) = 0, 1, 2, \ldots,$
- Nondecreasing: for $s < t$ it holds $N(s) \leq N(t)$
- Event counter: $N(t) - N(s) =$ number of events in interval $(s, t]$
  - Continuous on the right

Ex.1: # text messages (SMS) typed since beginning of class

Ex.2: # economic crisis since 1900

Ex.3: # customers at Wawa since 8am this morning



Independent increments

- Number of events in disjoint time intervals are independent
- Consider times $s_1 < t_1 < s_2 < t_2$ and intervals $(s_1, t_1]$ and $(s_2, t_2]$
- $N(t_1) - N(s_1)$ events occur in $(s_1, t_1]$. $N(t_2) - N(s_2)$ in $(s_2, t_2]$
- Independent increments implies latter two are independent

$$\mathrm{P}\left[N(t_1) - N(s_1) = k, N(t_2) - N(s_2) = l\right] = \mathrm{P}\left[N(t_1) - N(s_1) = k\right] \mathrm{P}\left[N(t_2) - N(s_2) = l\right]$$

Ex.1: Likely true for SMS, except for "have to send" messages

Ex.2: Most likely not true for economic crisis (business cycle)

Ex.3: Likely true for Wawa, except for unforeseen events (storms)

- Does not mean $N(t)$ independent of $N(s)$
- These events are clearly not independent, since $N(t)$ is at least $N(s)$

Stationary increments

- Prob. dist. of number of events depends on length of interval only
- Consider time intervals $(0, t]$ and $(s, s + t]$
- $N(t)$ events occur in $(0, t]$. $N(s + t) - N(s)$ events in $(s, s + t]$
- Stationary increments implies latter two have same prob. dist.

$$\mathrm{P}\left[N(s+t) - N(s) = k\right] = \mathrm{P}\left[N(t) = k\right]$$

Ex.1: Likely true if lecture is good and you keep interest in the class

Ex.2: Maybe true if you do not believe we are becoming better at preventing economic crisis

Ex.3: Most likely not true because of, e.g., rush hours and slow days

- A counting process is Poisson if it has the following properties
  (a) The process has stationary and independent increments
  (b) The number of events in $(0, t]$ has Poisson distribution with mean λt

$$\mathrm{P}\left[N(t)=n\right]=e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

- An equivalent definition is the following
  (i) The process has stationary and independent increments
  (ii) Prob. of event in infinitesimal time $\Rightarrow \mathrm{P}\left[N(h)=1\right]=\lambda h+o(h)$
  (iii) At most one event in infinitesimal time $\Rightarrow \mathrm{P}\left[N(h)>1\right]=o(h)$
- This is a more intuitive definition (even though difficult to believe now)
- Conditions (i) and (a) are the same
- That (b) implies (ii) and (iii) is obvious.
  - Just substitute small h in Poisson pmf's expression for $\mathrm{P}\left[N(t)=n\right]$
- To see that (ii) and (iii) imply (b) requires some work

Explanation of model (i)-(iii)

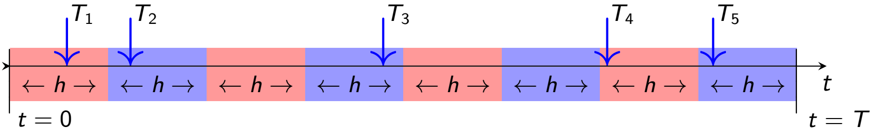
- Consider time T and divide interval $(0, T]$ in n subintervals
- Subintervals are of duration $h = T/n$, h vanishes as n increases
- The $m-th$ subinterval spans $\big((m-1)h, mh\big]$
- Define A_m as the number of events that occur in $m-th$ subinterval

$$A_m = N\big(mh\big) - N\big((m-1)h\big)$$

- The total number of events in (0,T] is the sum of A_ms

$$N(T) = \sum_{m=1}^{n} A_m = \sum_{m=1}^{n} N\big(mh\big) - N\big((m-1)h\big)$$

- In figure, $N(T) = 5$, A_1, A_2, A_4, A_7, A_8 are 1 and A_3, A_5, A_6 are 0



Probability distribution of A_m (intuitive arg.)

- Note first that since increments are stationary as per (i), it holds

$$\text{P}\left[A_m = k\right] = \text{P}\left[N\big(mh\big) - N\big((m-1)h\big) = k\right] = \text{P}\left[N(h) = k\right]$$

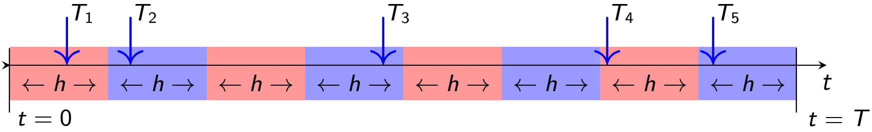
- In particular, using (ii) and (iii)

$$\text{P}\left[A_m = 1\right] = \text{P}\left[N(h) = 1\right] = \lambda h + o(h)$$
$$\text{P}\left[A_m > 1\right] = \text{P}\left[N(h) > 1\right] = o(h)$$

- Set aside $o(h)$ probabilities – They're negligible with respect to λh

$$\text{P}\left[A_m = 1\right] = \lambda h \qquad \text{P}\left[A_m = 0\right] = 1 - \lambda h$$

- A_m is Bernoulli with parameter λh

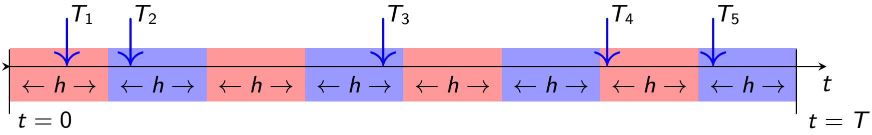


Probability distribution of $N(T)$ (intuitive arg.)

- Since increments are also independent as per (i), A_m are independent
- $N(T)$ is sum of n independent Bernoulli RVs with parameter λh
- $N(T)$ is binomial with parameters $(n, \lambda h) = (n, \lambda T/n)$

- As interval length $h \to 0$, number of intervals $n \to \infty$

  $\Rightarrow$ The product $n(\lambda h) = \lambda T$ stays constant
- $N(T)$ is Poisson with parameter λT
- Then (ii)-(iii) imply (b) and definitions are equivalent
- Not a proof because we neglected $o(h)$ terms. But explains what a Poisson process is



What is a Poisson process?

- Events happen in small interval h with probability λh proportional to h
- Whether event happens in an interval has no effect on other intervals
- Modeling questions
  - $\Rightarrow$ Expect probability of event proportional to length of interval?
  - $\Rightarrow$ Expect subsequent intervals to behave independently?
- Then a Poisson process model is appropriate
- Typically arise in a large population of agents acting independently
  - $\Rightarrow$ Larger interval, larger chance an agent takes an action
  - $\Rightarrow$ Action of one agent has no effect on action of other agents
  - $\Rightarrow$ Has therefore negligible effect on action of group

Examples of Poisson processes

Ex.1: Number of people arriving at subway station. Number of cars arriving at a highway entrance. Number of bids in an auction. Number of customers entering a store ... Large number of agents (people, drivers, bidders, customers) acting independently.

Ex.2: SMS generated by all students in the class. Once you send a SMS you are likely to stay silent for a while. But in a large population this has a minimal effect in the probability of someone generating a SMS

Ex.3: Count of molecule reactions. Molecules are "removed" from pool of reactants once they react. But effect is negligible in large population. Eventually reactants are depleted, but in small time scale process is approximately Poisson

Formal argument

- Define $A_{\max} = \max_{m=1,\ldots,n} (A_m)$, maximum nr. of events in one interval

- If $A_{\max} \leq 1$ all intervals have 0 or 1 events $\Rightarrow$ Easy (binomial)
- Consider conditional probability

$$\text{P}\left[N(T) = k \,\middle|\, A_{\max} \leq 1\right]$$

- For given h, $N(T)$ conditioned on $A_{\max} \leq 1$ is binomial
- Parameters are $n = T/h$ and $p = \lambda h + o(h)$
- Interval length $h \to 0$ $\Rightarrow$ parameter $p \to 0$, nr. of intervals $n \to \infty$

  $\Rightarrow$ Product np $\Rightarrow \lim_{h\to 0} np = \lim_{h \to 0} (T/h)(\lambda h + o(h)) = \lambda T$
- $N(T)$ conditioned on $A_{\max} \leq 1$ is Poisson with parameter λT

$$\text{P}\left[N(T) = k \,\middle|\, A_{\max} \leq 1\right] = e^{-\lambda T} \frac{\lambda t^k}{k!}$$

Formal argument (continued)

- Separate study in $A_{\max} \leq 1$ and $A_{\max} > 1$. That is, condition

$$\mathrm{P}\left[N(T)=k\right] = \mathrm{P}\left[N(T)=k \,\middle|\, A_{\max} \leq 1\right] \mathrm{P}\left[A_{\max} \leq 1\right] + \mathrm{P}\left[N(T)=k \,\middle|\, A_{\max} > 1\right] \mathrm{P}\left[A_{\max} > 1\right]$$

- Property (iii) implies that $\mathrm{P}\left[A_{\max} > 1\right]$ vanishes as $h \to 0$

$$\mathrm{P}\left[A_{\max} > 1\right] \leq \sum_{m=1}^{n} \mathrm{P}\left[A_m > 1\right] = no(h) = T\frac{o(h)}{h} \to 0$$

- Thus, as $h \to 0$, $\mathrm{P}\left[A_{\max} > 1\right] \to 0$ and $\mathrm{P}\left[A_{\max} \leq 1\right] \to 1$. Then

$$\lim_{h\to 0} \mathrm{P}\left[N(T)=k\right] = \lim_{h\to 0} \mathrm{P}\left[N(T)=k \,\middle|\, A_{\max} \leq 1\right]$$

- Latter is, as already seen, Poisson
  $\Rightarrow$ Then $N(T)$ Poisson with parameter λT

Properties of Poisson processes

Interarrival times

- Let $T_1, T_2, \ldots$ be sequence of times between events
- T_1 is time until first event (arrival). T_2 is time between second and first event. T_i is time between $i-1$-st and i-th event
- Ccdf of T_1 $\Rightarrow \text{P}\left[T_1 > t\right] = \text{P}\left[N(t) = 0\right] = e^{-\lambda t}$
- T_1 has exponential distribution with parameter λ
- Since we have independent increments this is likely true for all T_i

Theorem

Interarrival times T_i of a Poisson process are independent identically distributed exponential random variables with parameter λ, i.e.,

$$\text{P}\left[T_i > t\right] = e^{-\lambda t}$$

- Have already proved for T_1. Let us see the rest.

Proof.

- Let S_i be absolute time of i-th event. Condition on S_i

$$\mathrm{P}\left[T_{i+1} > t\right] = \int \mathrm{P}\left[T_{i+1} > t \,\middle|\, S_i = s\right] \mathrm{P}\left[S_i = s\right] ds$$

- To have $T_{i+1} > t$ given that $S_i = s$ it must be $N(s+t) = N(s)$

$$\mathrm{P}\left[T_{i+1} > t \,\middle|\, S_i = s\right] = \mathrm{P}\left[N(t+s) - N(s) = 0 \,\middle|\, N(s)\right]$$

- Since increments are independent conditioning on $N(s)$ is moot

$$\mathrm{P}\left[T_{i+1} > t \,\middle|\, S_i = s\right] = \mathrm{P}\left[N(t+s) - N(s) = 0\right]$$

- Since increments are also stationary the latter is

$$\mathrm{P}\left[T_{i+1} > t \,\middle|\, S_i = s\right] = \mathrm{P}\left[N(t) = 0\right] = e^{-\lambda t}$$

- Substituting into integral yields $\Rightarrow \mathrm{P}\left[T_{i+1} > t\right] = e^{-\lambda t}$ $\square$

Alternative definition of Poisson process

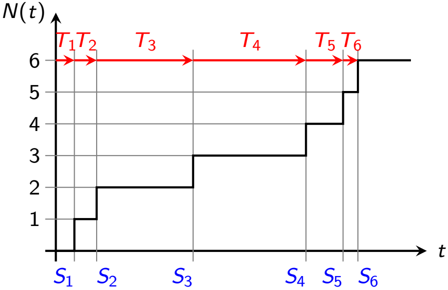
- Start with sequence of independent random times $T_1, T_2, \ldots$
- Times $T_i \sim \exp(\lambda)$ have exponential distribution with parameter λ
- Define time of i-th event S_i

$$S_i = T_1 + T_2 + \ldots + T_i$$

- Define counting process of events happening at S_i

$$N(t) = \max_i (S_i \leq t)$$

- $N(t)$ is a Poisson process



- If $N(t)$ is a Poisson process interarrival times T_i are exponential
- To show that definition is equivalent have to show the converse
- I.e., if interarrival times are exponential, process is Poisson

Alternative definition of Poisson process (cont.)

- Exponential i.i.d interarrival times $\Rightarrow$ Poisson process?
- Show that implies definition (i)-(iii)

- Stationary true because all T_i have same distribution
- Independent increments true because interarrival times are independent and exponential RVs are memoryless
- Can have more than two events in $(0, h]$ only if $T_1 < h$ and $T_2 < h$

$$\mathrm{P}\left[N(h) > 1\right] \leq \mathrm{P}\left[T_1 \leq h\right] \mathrm{P}\left[T_2 \leq h\right]$$
$$= (1 - e^{-\lambda h})^2 = (\lambda h)^2 + o(h^2) = o(h)$$

- We have no event in $(0, h]$ if $T_1 > h$

$$\mathrm{P}\left[N(h) = 0\right] = \mathrm{P}\left[T_1 \geq h\right] = e^{-\lambda h} = 1 - \lambda h + o(h)$$

- The remaining case is $N(h) = 1$ whose probability is

$$\mathrm{P}\left[N(h) = 1\right] = 1 - \mathrm{P}\left[N(h) = 0\right] - \mathrm{P}\left[N(h) > 1\right] = \lambda h + o(h)$$

Three definitions of Poisson processes

Def. 1: Prob. of event proportional to interval width. Intervals independent

- Physical model definition.
- Can a phenomenon be reasonably modeled as a Poisson process?
- The other two definitions are used for analysis and/or simulation

Def. 2: Prob. distribution of events in $(0, t]$ is Poisson

- Event centric definition. Nr. of events in given time intervals
- Allows analysis and simulation
- Used when information about nr. of events in given time is desired

Three definitions of Poisson processes (continued)

Def. 3: Prob. distribution of interarrival times is exponential

- Time centric definition. Times at which events happen
- Allows analysis and simulation.
- Used when information about event times is of interest

Obs: Restrictions in Def. 1 are mild, yet they impose a lot of structure as implied by Defs. 2 & 3

Example: Number of visitors to a web page

- Nr. of unique visitors to a webpage between 6:00pm to 6:10pm

Def 1: Poisson process? Probability proportional to time interval and independent intervals seem reasonable assumptions

- Model as Poisson process with rate λ visits/second (v/s)

Def 2: Arrivals in interval of duration t are Poisson with parameter λt

- Expected nr. of visits in 10 minutes? $\Rightarrow \mathbb{E}\left[N(600)\right] = 600\lambda$
- Prob. of exactly 10 visits in 1 sec? $\Rightarrow \text{P}\left[N(1) = 10\right] = e^{-\lambda}\lambda/10!$
- Data shows N average visits in 10 minutes (600s). Approximate λ

  $\Rightarrow$ Since $\mathbb{E}\left[N(600)\right] = 600\lambda$ can make $\hat{\lambda} = N/600$

Def 3: Interarrival times T_i are exponential with parameter λ

- Expected time between visitors? $\Rightarrow \mathbb{E}[T_i] = 1/\lambda$
- Expected arrival time S_n of n-th visitor?
  - $\Rightarrow$ Can write time S_n as sum of T_i, i.e., $S_n = \sum_{i=1}^{n} T_i$.
  - $\Rightarrow$ Taking expected value $\mathbb{E}[S_n] = \sum_{i=1}^{n} \mathbb{E}[T_i] = n/\lambda$

Continuous time Markov chains

Definition

- Continuous time positive variable $t \in \mathbb{R}^+$
- States $X(t)$ taking values in countable set, e.g., $0, 1, \ldots, i, \ldots$
- Stochastic process $X(t)$ is a continuous time Markov chain (CTMC) if

$$\begin{aligned} & \mathrm{P}\left[X(t+s)=j \,\middle|\, X(s)=i, X(u)=x(u), u<s\right] \\ = & \mathrm{P}\left[X(t+s)=j \,\middle|\, X(s)=i\right] \end{aligned}$$

- Memoryless property $\Rightarrow$ The future $X(t+s)$ given the present $X(s)$ is independent of the past $X(u)=x(u), u<s$

- In principle need to specify functions $\mathrm{P}\left[X(t+s)=j \,\middle|\, X(s)=i\right]$
  $\Rightarrow$ For all times t, for all times s and for all pairs of states (i, j)

Notation, homogeneity

- Notation: $X[s:t]$ state values for all times $s \leq u \leq t$, includes borders
- $X(s:t)$ values for all times $s < u < t$, borders excluded
- $X(s:t]$ values for all times $s < u \leq t$, exclude left, include right
- $X[s:t)$ values for all times $s \leq u < t$, include left, exclude right

- Homogeneous CTMC if $\text{P}\left[X(t+s)=j \,\middle|\, X(s)=i\right]$ constant for all s
- Still need $\text{P}\left[X(t+s)=j \,\middle|\, X(s)=i\right]$ for all t and pairs (i,j)
- We restrict consideration to homogeneous CTMCs

- Memoryless property makes it somewhat simpler

Transition times

- T_i = time until transition out of state i into any other state j
- T_i is a RV called transition time with cdf

$$\mathrm{P}\left[T_i > t\right] = \mathrm{P}\left[X(0:t] = i \,\middle|\, X(0) = i\right]$$

- Probability of $T_i > t + s$ given that $T_i > s$? Use cdf expression

$$\begin{aligned}
\mathrm{P}\left[T_i > t+s \,\middle|\, T_i > s\right] &= \mathrm{P}\left[X(0:t+s] = i \,\middle|\, X[0:s] = i\right] \\
&= \mathrm{P}\left[X(s:t+s] = i \,\middle|\, X[0:s] = i\right] \\
&= \mathrm{P}\left[X(s:t+s] = i \,\middle|\, X(s) = i\right] \\
&= \mathrm{P}\left[X(0:t] = i \,\middle|\, X(0) = i\right]
\end{aligned}$$

- Equalities true because: Already observed that $X[0:s] = i$. Memoryless property. Homogeneity
- From cdf expression $\Rightarrow \mathrm{P}\left[T_i > t+s \,\middle|\, T_i > s\right] = \mathrm{P}\left[T_i > t\right]$
- Transition times are exponential RVs

Alternative definition

- Exponential transition times is a fundamental property of CTMCs
- Can be used as "algorithmic" definition of CTMCs

- Continuous times stochastic process $X(t)$ is a CTMC if
- Transition times T_i are exponential RVs with mean $1/\nu_i$
- When they occur, transitions out of i are into j with probability P_{ij}

$$\sum_{j=1}^{\infty} P_{ij} = 1, \qquad P_{ii} = 0$$

- Transition times T_i and transitioned state j are independent
- Define matrix $\mathbf{P}$ grouping transition probabilities P_{ij}

- CTMC states evolve as a discrete time MC
- State transitions occur at random exponential intervals $T_i \sim \exp(\nu_i)$
- As opposed to occurring at fixed intervals

Definition

Consider a CTMC with transition probs. **P** and rates ν_i. The (discrete time) MC with transition probs. **P** is the CTMC's embedded MC

- Transition probabilities **P** describe a discrete time MC
- States do not transition into themselves ($P_{ii} = 0$, **P**'s diagonal null)
- Can use underlying discrete time MCs to understand CTMCs

- Accesibility: State j is accessible from state i if it is accessible in the discrete time MC with transition probability matrix **P**
- Communication: States i and j communicate if they communicate in the discrete time MC with transition probability matrix **P**
- Communication is a class property. Proof: It is for discrete time MC

- Recurrent/Transient. Recurrence is class property. Etc. More later

Transition rates

- Expected value of transition time T_i is $\mathbb{E}[T_i] = 1/\nu_i$
- Can interpret ν_i as the rate of transition out of state i
- Of these transitions, P_{ij} of them are into state j
- Can interpret $q_{ij} := \nu_i P_{ij}$ as transition rate from i to j. Define so

- Transition rates are yet another specification of CTMC
- If q_{ij} are given can recover ν_i as

$$\nu_i = \nu_i \sum_{j=1, j\neq i}^{\infty} P_{ij} = \sum_{j=1, j\neq i}^{\infty} \nu_i P_{ij} = \sum_{j=1, j\neq i}^{\infty} q_{ij}$$

- And can recover P_{ij} as $\Rightarrow P_{ij} = q_{ij}/\nu_i = q_{ij}\left(\sum_{j=1, j\neq i}^{\infty} q_{ij}\right)^{-1}$

Example: Birth and death processes

- State $X(t) = 0, 1, \ldots$ Interpret as number of individuals

- Birth and deaths occur at state-dependent rates. When $X(t) = i$
- Births $\Rightarrow$ individuals added at exponential times with mean $1/\lambda_i$
  $\Rightarrow$ birth or arrival rate $= \lambda_i$ births per unit of time
- Deaths $\Rightarrow$ individuals removed at exponential times with rate $1/\mu_i$
  $\Rightarrow$ death or departure rate $= \mu_i$ births per unit of time

- Birth and death times are independent
- As are subsequent birth and death processes

- Birth and death (BD) processes are then CTMC

Transition times and probabilities

- Transition times. Leave state $i \neq 0$ when birth or death occur
- If T_B and T_D are times to next birth and death, $T_i = \min(T_B, T_D)$
- Since T_B and T_D are exponential, so is T_i with rate

$$\nu_i = \lambda_i + \mu_i$$

- When leaving state i can go to $i+1$ (birth first) or $i-1$ (death first)
- Birth occurs before death with probability $\Rightarrow \dfrac{\lambda_i}{\lambda_i + \mu_i} = P_{i,i+1}$
- Death occurs before birth with probability $\Rightarrow \dfrac{\mu_i}{\lambda_i + \mu_i} = P_{i,i-1}$
- Leave state 0 only if a birth occurs (it at all) then

$$\nu_0 = \lambda_0 \qquad P_{01} = 1$$

- If leaves 0, goes to 1 with probability 1. Might not leave 0 if $\lambda_0 = \infty$

Transition rates

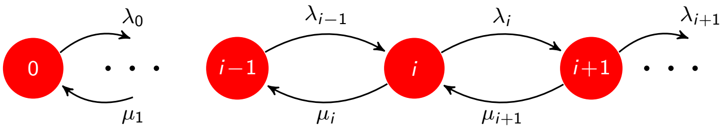
- Rate of transition from i to $i+1$ is (recall definition, $q_{ij} = \nu_i P_{ij}$)

$$q_{i,i+1} = \nu_i P_{i,i+1} = (\lambda_i + \mu_i) \frac{\lambda_i}{\lambda_i + \mu_i} = \lambda_i$$

- Likewise, rate of transition from i to $i-1$ is

$$q_{i,i-1} = \nu_i P_{i,i-1} = (\lambda_i + \mu_i) \frac{\mu_i}{\lambda_i + \mu_i} = \mu_i$$

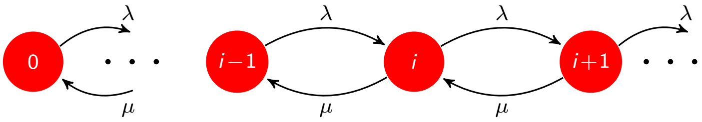
- For $i = 0$, $\Rightarrow q_{01} = \nu_1 P_{01} = \lambda_0$



- Somewhat more natural representation. More similar to discrete MC

M/M/1 queue

- A M/M/1 queue is a BD process with $\lambda_i = \lambda$ and $\mu_i = \mu$ constant
- Customers arrive for service at a rate of λ per unit of time
- They are serviced at a rate of μ customers per time unit



- The M/M is for Markov arrivals / Markov departures
- The 1 is because there is only one server

Transition probability function

Exponential random variables

Counting processes and definition of Poisson processes

Properties of Poisson processes

Continuous time Markov chains

Transition probability function

Determination of transition probability function

Limit probabilities

Definition

- Two equivalent ways of specifying a CTMC

- Transition time averages $1/\nu_i$ + Transition probabilities P_{ij}
- Easier description. Typical starting point for CTMC modeling

- Transition prob. function $\Rightarrow P_{ij}(t) := \mathrm{P}\left[X(t+s) = j \,\middle|\, X(s) = i\right]$
- More complete description

- Compute transition prob. function from transition times and probs.
- Notice two obvious properties $P_{ij}(0) = 1$, $P_{ii}(0) = 1$

Roadmap to determine $P_{ij}(t)$

- Goal is to obtain a differential equation whose solution is $P_{ij}(t)$
- For that, need to study change in $P_{ij}(t)$ when time changes slightly
- Separate in two subproblems
  - $\Rightarrow$ Transition probabilities for small time h, $P_{ij}(h)$
  - $\Rightarrow$ Transition probabilities in $t+h$ as function of those in t and h
- We can combine both results in two different ways
- Jump from 0 to t then to $t+h$ $\Rightarrow$ Process runs a little longer
- Changes where the process is going to $\Rightarrow$ Forward equations
- Jump from 0 to h then to $t+h$ $\Rightarrow$ Process starts a little later
- Changes where the process comes from $\Rightarrow$ Backward equations

Transition probability in infinitesimal time

Theorem

The transition probability functions $P_{ii}(t)$ and $P_{ij}(t)$ satisfy the following limits for time t going to 0

$$\lim_{t\to 0} \frac{P_{ij}(t)}{t} = q_{ij}, \qquad \lim_{t\to 0} \frac{1 - P_{ii}(t)}{t} = \nu_i$$

- Since $P_{ij}(0) = 0$, $P_{ii}(0) = 1$ above limits are derivatives at $t = 0$

$$\left.\frac{\partial P_{ij}(t)}{\partial t}\right|_{t=0} = q_{ij}, \qquad \left.\frac{\partial P_{ii}(t)}{\partial t}\right|_{t=0} = -\nu_i$$

- Limits also imply that for small h

$$P_{ij}(h) = q_{ij} h + o(h), \qquad P_{ii}(h) = 1 - \nu_i h + o(h)$$

- Transition rates q_{ij} are "instantaneous transition probabilities"

  $\Rightarrow$ Transition probability coefficient for small time h

Proof.

- Consider a small time h
- Since $1 - P_{ii}$ is the probability of transitioning out of state i

$$1 - P_{ii}(h) = \mathrm{P}\left[T_i < h\right] = \nu_i h + o(h)$$

- Divide by h and take limit
- For $P_{ij}(t)$ notice that since two or more transitions have $o(h)$ prob

$$P_{ij}(h) = \mathrm{P}\left[X(h) = j \,\middle|\, X(0) = i\right] = P_{ij}\mathrm{P}\left[T_i \leq h\right] + o(h)$$

- Since T_i is exponential $\mathrm{P}\left[T_i \leq h\right] = \nu_i h + o(h)$. Then

$$P_{ij}(h) = \nu_i P_{ij} + o(h) = q_{ij} h + o(h)$$

- Divide by h and take limit □

Chapter: Chapman-Kolmogorov equations

Theorem

For all times s and t the transition probability functions $P_{ij}(t+s)$ *are obtained from* $P_{ik}(t)$ *and* $P_{kj}(s)$ *as*

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s)$$

- As for discrete time MC, to go from i to j in time $t+s$
  - $\Rightarrow$ Go from i to a certain k at time t $\Rightarrow P_{ik}(t)$
  - $\Rightarrow$ In the remaining time s go from k to j $\Rightarrow P_{kj}(s)$
  - $\Rightarrow$ Sum over all possible intermediate steps k

Proof.

$$
\begin{aligned}
& P_{ij}(t+s) \\
&= \mathrm{P}\left[X(t+s)=j \,\middle|\, X(0)=i\right] && \text{Definition of } P_{ij}(t+s) \\
&= \sum_{k=0}^{\infty} \mathrm{P}\left[X(t+s)=j \,\middle|\, X(t)=k, X(0)=i\right] \mathrm{P}\left[X(t)=k \,\middle|\, X(0)=i\right] \\
& && \text{Sum of conditional probabilities} \\
&= \sum_{k=0}^{\infty} \mathrm{P}\left[X(t+s)=j \,\middle|\, X(t)=k\right] P_{ik}(t) && \text{Memoryless property of CTMC} \\
& && \text{and definition of } P_{ik}(t) \\
&= \sum_{k=0}^{\infty} \mathrm{P}\left[X(s)=j \,\middle|\, X(0)=k\right] P_{ik}(t) && \text{Time invariance / homogeneity} \\
&= \sum_{k=0}^{\infty} P_{kj}(s) P_{ik}(t) && \text{Definition of } P_{kj}(s)
\end{aligned}
$$

□

Combining both results

- Let us combine the last two results to express $P_{ij}(t+h)$
- Use Chapman-Kolmogorov's equations for $0 \to t \to h$

$$P_{ij}(t+h) = \sum_{k=0}^{\infty} P_{ik}(t)P_{kj}(h) = P_{ij}(t)P_{jj}(h) + \sum_{k=0, k\neq i}^{\infty} P_{ik}(t)P_{kj}(h)$$

- Use infinitesimal time expression

$$P_{ij}(t+h) = P_{ij}(t)(1-\nu_j h) + \sum_{k=0, k\neq i}^{\infty} P_{ik}(t)q_{kj}h + o(h)$$

- Subtract $P_{ij}(t)$ from both sides and divide by h

$$\frac{P_{ij}(t+h) - P_{ij}(t)}{h} = -\nu_j P_{ij}(t) + \sum_{k=0, k\neq i}^{\infty} P_{ik}(t)q_{kj} + \frac{o(h)}{h}$$

- The right hand side is a "derivative" ratio. Let $h \to 0$

Kolmogorov's forward equations

Theorem

The transition probability functions $P_{ij}(t)$ of a CTMC satisfy the system of differential equations (for all pairs i, j)

$$\frac{\partial P_{ij}(t)}{\partial t} = \sum_{k=0,k\neq j}^{\infty} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t)$$

- $\partial P_{ij}(t)/\partial t$ = rate of change of $P_{ij}(t)$
- $q_{kj} P_{ik}(t)$ = (transition into k in $0 \to t$) $\times$ (rate of moving into j in next instant)
- $\nu_j P_{ij}(t)$ = (transition into j in $0 \to t$) $\times$ (rate of leaving j in next instant)
- Change in $P_{ij}(t) = \sum_k$ (moving into j from k) $-$ (leaving j)
- Kolmogorov's forward equations valid in most cases, but not always

Kolmogorov's backward equations

- For forward equations used Chapman-Kolmogorov's for $0 \to t \to h$
- For backward equations we use $0 \to h \to t$ to express $P_{ij}(t+h)$ as

$$P_{ij}(t+h) = \sum_{k=0}^{\infty} P_{ik}(h)P_{kj}(t) = P_{ii}(h)P_{ij}(t) + \sum_{k=0,k\neq i}^{\infty} P_{ik}(h)P_{kj}(t)$$

- Use infinitesimal time expression

$$P_{ij}(t+h) = (1-\nu_i h)P_{ij}(t) + \sum_{k=0,k\neq i}^{\infty} q_{ik} h P_{kj}(t) + o(h)$$

- Subtract $P_{ij}(t)$ from both sides and divide by h

$$\frac{P_{ij}(t+h) - P_{ij}(t)}{h} = -\nu_i P_{ij}(t) + \sum_{k=0,k\neq i}^{\infty} q_{ik} P_{kj}(t) + \frac{o(h)}{h}$$

Kolmogorov's backward equations

Theorem

The transition probability functions $P_{ij}(t)$ of a CTMC satisfy the system of differential equations (for all pairs i, j)

$$\frac{\partial P_{ij}(t)}{\partial t} = \sum_{k=0, k\neq i}^{\infty} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t)$$

- $\nu_i P_{ij}(t) =$ (transition into j in $h \to t$) $\times$
  (do not do that if leave i in initial instant)
- $q_{ik} P_{kj}(t) =$ (rate of transition into k in $0 \to h$) $\times$
  (transition from k into j in $h \to t$)

- Forward equations $\Rightarrow$ change in $P_{ij}(t)$ if finish h later
- Backward equations $\Rightarrow$ change in $P_{ij}(t)$ if start h later
- Where process goes (forward) $\Leftrightarrow$ where process comes from (backward)

Determination of transition probability function

Exponential random variables

Counting processes and definition of Poisson processes

Properties of Poisson processes

Continuous time Markov chains

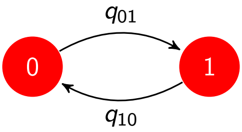
Transition probability function

Determination of transition probability function

Limit probabilities

A CTMC with two states

- Simplest possible CTMC has only two states. Say 0 and 1
- Transition rates are q_{01} and q_{10}
- Given transition rates can find mean transition times as



$$\nu_0 = \sum_{j \neq 0} q_{0i} = q_{01} \qquad \nu_1 = \sum_{j \neq 1} q_{1i} = q_{10}$$

- Use Kolmogorov's equations to find transition probability functions

$$P_{00}(t), \quad P_{01}(t), \quad P_{10}(t), \quad P_{11}(t)$$

- Note that transition probs. out of each state sum up to one

$$P_{00}(t) + P_{01}(t) = 1 \qquad P_{10}(t) + P_{11}(t) = 1$$

Kolmogorov's forward equations

- Kolmogorov's forward equations (process runs a little longer)

$$P'_{ij}(t) = \sum_{k=0,k\neq j}^{\infty} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t)$$

- For the two node CTMC

$$P'_{00}(t) = q_{10} P_{01}(t) - \nu_0 P_{00}(t) \qquad P'_{01}(t) = q_{01} P_{00}(t) - \nu_1 P_{01}(t)$$
$$P'_{10}(t) = q_{10} P_{11}(t) - \nu_0 P_{10}(t) \qquad P'_{11}(t) = q_{01} P_{10}(t) - \nu_1 P_{11}(t)$$

- Probabilities out of 0 sum up to 1 $\Rightarrow$ eqs. in first row are equivalent
- Probabilities out of 1 sum up to 1 $\Rightarrow$ eqs. in second row are equivalent
- Pick the equations for $P'_{00}(t)$ and P'_{11}(t)

Solution of backward equations

- Use $\Rightarrow$ Relation between transition rates: $\nu_0 = q_{01}$ and $\nu_1 = q_{10}$
  $\Rightarrow$ Probs. sum 1: $P_{01}(t) = 1 - P_{00}(t)$ and $P_{01}(t) = 1 - P_{00}(t)$

$$P_{00}^{'}(t) = q_{10}\big[1 - P_{00}(t)\big] - q_{01}P_{00}(t) = q_{10} - (q_{10} + q_{01})P_{00}(t)$$
$$P_{11}^{'}(t) = q_{01}\big[1 - P_{11}(t)\big] - q_{10}P_{11}(t) = q_{01} - (q_{10} + q_{01})P_{11}(t)$$

- Can obtain exact same pair of equations from backward equations

- First order linear differential equations. Solutions are exponential
- For $P_{00}(t)$ propose candidate solution (just take derivate to check)

$$P_{00}(t) = \frac{q_{10}}{q_{10} + q_{01}} + ce^{-(q_{10}+q_{01})t}$$

- To determine c use initial condition $P_{00}(t) = 1$

Solution of backward equations (continued)

- Evaluation of candidate solution at initial condition $P_{00}(t) = 1$ yields

$$1 = \frac{q_{10}}{q_{10} + q_{01}} + c \Rightarrow c = \frac{q_{01}}{q_{10} + q_{01}}$$

- Finally transition probability function $P_{00}(t)$

$$P_{00}(t) = \frac{q_{10}}{q_{10} + q_{01}} + \frac{q_{01}}{q_{10} + q_{01}} e^{-(q_{10}+q_{01})t}$$

- Repeat for $P_{11}(t)$. Same exponent, different constants

$$P_{11}(t) = \frac{q_{01}}{q_{10} + q_{01}} + \frac{q_{10}}{q_{10} + q_{01}} e^{-(q_{10}+q_{01})t}$$

- As time goes to infinity, $P_{00}(t)$ and $P_{11}(t)$ converge exponentially
- Convergence rate depends on magnitude of $q_{10} + q_{01}$

Convergence of transition probabilities

- Recall $P_{01} = 1 - P_{00}$ and $P_{10} = 1 - P_{11}$. Steady state distributions

$$\lim_{t\to\infty} P_{00}(t) = \frac{q_{10}}{q_{10}+q_{01}} \qquad \lim_{t\to\infty} P_{01}(t) = \frac{q_{01}}{q_{10}+q_{01}}$$
$$\lim_{t\to\infty} P_{11}(t) = \frac{q_{01}}{q_{10}+q_{01}} \qquad \lim_{t\to\infty} P_{10}(t) = \frac{q_{10}}{q_{10}+q_{01}}$$

- Limit distribution exists and is independent of initial condition

  $\Rightarrow$ Compare across diagonals

Kolmogorov's forward equations in matrix form

- Restrict attention to finite CTMCs with N states
- Define matrix $\mathbf{R} \in \mathbb{R}^{N\times N}$ with elements $r_{ij} = q_{ij},\ r_{ii} = -\nu_i$.
- Rewrite Kolmogorov's forward eqs. as (process runs a little longer)

$$P'_{ij}(t) = \sum_{k=1,k\neq j}^{N} q_{kj}P_{ik}(t) - \nu_j P_{ij}(t) = \sum_{k=1}^{N} r_{kj}P_{ik}(t)$$

- Right hand side defines elements of a matrix product

$$\begin{pmatrix} r_{11} & \cdot & r_{1j} & \cdot & r_{1N} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{k1} & \cdot & r_{kj} & \cdot & r_{kN} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{N1} & \cdot & r_{Nj} & \cdot & r_{NN} \end{pmatrix} = \mathbf{R}$$

$$\mathbf{P}(t) = \begin{pmatrix} P_{11}(t) & \cdot & P_{1k}(t) & \cdot & P_{1N}(t) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ P_{i1}(t) & \cdot & P_{ik}(t) & \cdot & P_{iN}(t) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ P_{N1}(t) & \cdot & P_{Nk}(t) & \cdot & P_{JN}(t) \end{pmatrix} \qquad \begin{pmatrix} s_{11} & \cdot & s_{1j} & \cdot & s_{1N} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{i1} & \cdot & s_{ij} & \cdot & s_{iN} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{N1} & \cdot & s_{Nk} & \cdot & s_{NN} \end{pmatrix} = \mathbf{P}(t)\mathbf{R} = \mathbf{P}'(t)$$

Kolmogorov's backward equations in matrix form

- Similarly, Kolmogorov's backward eqs. (process starts a little later)

$$P'_{ij}(t) = \sum_{k=0,k\neq i}^{\infty} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t) = \sum_{k=1}^{N} r_{ik} P_{kj}(t)$$

- Right hand side also defines a matrix product

$$\begin{pmatrix} P_{11}(t) & \cdot & P_{1j}(t) & \cdot & P_{1N}(t) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ P_{k1}(t) & \cdot & P_{kj}(t) & \cdot & P_{kN}(t) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ P_{N1}(t) & \cdot & P_{Nj}(t) & \cdot & P_{NN}(t) \end{pmatrix} = \mathbf{P}(t)$$

$$\mathbf{R} = \begin{pmatrix} r_{11} & \cdot & r_{1k} & \cdot & r_{1N} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{i1} & \cdot & r_{ik} & \cdot & r_{iN} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{N1} & \cdot & r_{Nk} & \cdot & r_{JN} \end{pmatrix} \quad \begin{pmatrix} s_{11} & \cdot & s_{1j} & \cdot & s_{1N} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{i1} & \cdot & s_{ij} & \cdot & s_{iN} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{N1} & \cdot & s_{Nk} & \cdot & s_{NN} \end{pmatrix} = \mathbf{RP}(t) = \mathbf{P}'(t)$$

Kolmogorov's equations in matrix form

- Matrix form of Kolmogorov's forward equation $\Rightarrow \mathbf{P}'(t) = \mathbf{P}(t)\mathbf{R}$
- Matrix form of Kolmogorov's backward equation $\Rightarrow \mathbf{P}'(t) = \mathbf{R}\mathbf{P}(t)$

  $\Rightarrow$ More similar than apparent

  $\Rightarrow$ But not equivalent because matrix product not commutative

- Notwithstanding both equations have to accept the same solution

Matrix exponential

- Kolmogorov's equations are first order linear differential equations
- They are coupled, though. $P'_{ij}(t)$ depends on $P_{kj}(t)$ for all k
- Still accept exponential solution. Need to define matrix exponential

Definition

The matrix exponential $e^{\mathbf{A}t}$ of matrix $\mathbf{A}t$ is defined as the series

$$e^{\mathbf{A}t} = \sum_{n=0}^{\infty} \frac{(\mathbf{A}t)^n}{n!} = \mathbf{I} + \mathbf{A}t + \frac{(\mathbf{A}t)^2}{2} + \frac{(\mathbf{A}t)^3}{2 \times 3} + \ldots$$

- Derivative of matrix exponential with respect to t

$$\frac{\partial e^{\mathbf{A}t}}{\partial t} = \mathbf{0} + \mathbf{A} + \mathbf{A}^2 t + \frac{\mathbf{A}^3 t^2}{2} + \ldots = \mathbf{A}\left(\mathbf{I} + \mathbf{A}t + \frac{(\mathbf{A}t)^2}{2} + \ldots\right) = \mathbf{A}e^{\mathbf{A}t}$$

- Putting $\mathbf{A}$ on right side of product shows that $\Rightarrow \frac{\partial e^{\mathbf{A}t}}{\partial t} = e^{\mathbf{A}t}\mathbf{A}$

Solution of Kolmogorov's equations

- Propose solution of the form $\mathbf{P}(t) = e^{\mathbf{R}t}$
- $\mathbf{P}(t)$ solves backward equations. Derivative with respect to time is

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

- It also solves forward equations

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

- Also notice that $\mathbf{P}(0) = \mathbf{I}$. As it should ($P_{ii}(0) = 1$, and $P_{ij}(0) = 0$)

Unconditional probabilities

- $\mathbf{P}(t)$ is transition prob. from states at time 0 to states at time t
- Define unconditional probs. at time t, $p_j(t) := \text{P}\left[X(t) = i\right]$
- Group in vector $\mathbf{p}(t) = [p_1(t), p_2(t), \ldots, p_j(t), \ldots]^T$
- Given initial distribution $\mathbf{p}(0)$ find $p_j(t)$ conditioning on initial state

$$p_j(t) = \sum_{i=0}^{\infty} \text{P}\left[X(t) = j \,\middle|\, X(0) = i\right] \text{P}\left[X(0) = i\right] = \sum_{i=0}^{\infty} P_{ij}(t) p_i(0)$$

- Using matrix-vector notation $\Rightarrow \mathbf{p}(t) = \mathbf{P}^T(t)\mathbf{p}(0)$

Recurrent and transient states

- Recall the embedded discrete time MC associated with any CTMC
- Transition probs. of MC form the matrix $\mathbf{P}$ of the CTMC
- States do not transition into themselves ($P_{ii} = 0$, $\mathbf{P}$'s diagonal null)

- States $i \leftrightarrow j$ communicate in the CTMC if $i \leftrightarrow j$ in the MC
- Communication partitions MC in classes $\Rightarrow$ induces CTMC partition
- CTMC is irreducible if embedded MC contains a single class

- State i is recurrent if i is recurrent in the embedded MC
- State i is transient if i is transient in the embedded MC
- State i is positive recurrent if i is so in the embedded MC
- Transience and recurrence shared by elements of a MC class

  $\Rightarrow$ Transience and recurrence are class properties of CTMC

- Periodicity not possible in CTMCs

Limit probabilities

Theorem

Consider irreducible positive recurrent CTMC with transition rates ν_i *and* q_{ij}. *Then,* $\lim_{t\to\infty} P_{ij}(t)$ *exists and is independent of the initial state* i*, i.e.,*

$$P_j = \lim_{t\to\infty} P_{ij}(t) \quad \textit{exists for all } i$$

Furthermore, steady state probabilities $P_j \geq 0$ *are the unique nonnegative solution of the system of linear equations*

$$\nu_j P_j = \sum_{k=0,k\neq j}^{\infty} q_{kj} P_k, \qquad \sum_{j=0}^{\infty} P_j = 1$$

- Limit distribution exists and is independent of initial condition
- Obtained as solution of system of linear equations
- Analogous to MCs. Algebraic equations slightly different

Algebraic relation to determine limit probabilities

- As with MCs difficult part is to prove that $P_j = \lim_{t\to\infty} P_{ij}(t)$ exists
- Algebraic relations obtained from Kolmogorov's forward equation

$$\frac{\partial P_{ij}(t)}{\partial t} = \sum_{k=0,k\neq j}^{\infty} q_{kj} P_{ik}(t) - \nu_j P_{ij}(t)$$

- If limit distribution exists we have, independent of initial state i

$$\lim_{t\to\infty} \frac{\partial P_{ij}(t)}{\partial t} = 0, \qquad \lim_{t\to\infty} P_{ij}(t) = P_j(t)$$

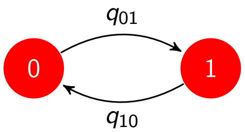
- Considering the limit of Kolomogorov's forward equation then yields

$$0 = \sum_{k=0,k\neq j}^{\infty} q_{kj} P_k(t) - \nu_j P_j(t)$$

- Reordering terms limit distribution equations follow

Example: Two state CTMC

- Simplest CTMC with two states 0 and 1
- Transition rates are q_{01} and q_{10}



- From transition rates find mean transition times $\nu_0 = q_{01}$, $\nu_1 = q_{10}$
- Stationary distribution equations

$$\nu_0 P_0 = q_{10} P_1, \qquad \nu_1 P_1 = q_{01} P_0, \qquad P_0 + P_1 = 1$$
$$q_{01} P_0 = q_{10} P_1, \qquad q_{10} P_1 = q_{01} P_0,$$

- Solution yields $\Rightarrow P_0 = \dfrac{q_{10}}{q_{10} + q_{01}} \qquad P_1 = \dfrac{q_{01}}{q_{10} + q_{01}}$
- Larger prob. q_{10} of entering 0 $\Rightarrow$ larger prob. P_0 of being at 0
- Larger prob. q_{01} of entering 1 $\Rightarrow$ larger prob. P_1 of being at 1

Ergodicity

- Consider the fraction $T_i(t)$ of time spent in state i up to time t

$$T_i(t) := \frac{1}{t} \int_0^t \mathbb{I}\left\{X(t) = i\right\}$$

- $T_i(t)$ = time/ergodic average. Its limit $\lim_{t\to\infty} T_i(t)$ is ergodic limit
- If CTMC is irreducible, positive recurrent ergodicity holds

$$P_i = \lim_{t\to\infty} T_i(t) = \lim_{t\to\infty} \frac{1}{t} \int_0^t \mathbb{I}\left\{X(t) = i\right\} \qquad \text{a.s.}$$

- Ergodic limit coincides with limit probabilities (almost surely)

Ergodicity (continued)

- Consider function $f(i)$ associated with state i. Can write $f\big(X(t)\big)$ as

$$f\big(X(t)\big) = \sum_{i=1}^{\infty} f(i)\mathbb{I}\left\{X(t) = i\right\}$$

- Consider the time average of the function of the state $f\big(X(t)\big)$

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f\big(X(t)\big) = \lim_{t\to\infty} \frac{1}{t}\int_0^t \sum_{i=1}^{\infty} f(i)\mathbb{I}\left\{X(t) = i\right\}$$

- Interchange summation with integral and limit to say

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f\big(X(t)\big) = \sum_{i=1}^{\infty} f(i) \lim_{t\to\infty} \frac{1}{t}\int_0^t \mathbb{I}\left\{X(t) = i\right\} = \sum_{i=1}^{\infty} f(i)P_i$$

- Function's ergodic limit = functions limit average value

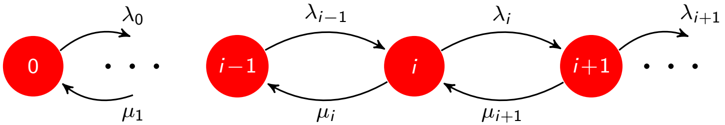
Limit distribution equations as balance equations

- Reconsider limit distribution equations $\Rightarrow \nu_j P_j = \sum_{i=0, k\neq j}^{\infty} q_{kj} P_k$
- P_j = fraction of time spent in state j
- ν_j = rate of transition out of state j given CTMC is in state j
  $\Rightarrow \nu_j P_j$ = rate of transition out of state j (unconditional)
- q_{kj} = rate of transition from k to j given CTMC is in state k
  $\Rightarrow q_{kj} P_k$ = rate of transition from k to j (unconditional)
  $\Rightarrow \sum_{i=0, k\neq j}^{\infty} q_{kj} P_k$ = rate of transition into j, from all states
- Rate of transition out of state j = rate of transition into j
- Balance equations $\Rightarrow$ Balance nr. of transitions in and out of state j

- Birth/deaths occur at state-dependent rates. When $X(t) = i$
- Births $\Rightarrow$ individuals added at exponential times with mean $1/\lambda_i$

  $\Rightarrow$ birth rate = upward transition rate = $q_{i,i+1} = \lambda_i$
- Deaths $\Rightarrow$ individuals removed at exponential times with rate $1/\mu_i$

  $\Rightarrow$ birth rate = downward transition rate = $q_{i,i-1} = \mu_i$
- Mean transition times $\Rightarrow \nu_i = \lambda_i + \mu_i$



- Limit distribution/balance equations: Rate out of j = rate into j

$$(\lambda_i + \mu_i)P_i = \lambda_{i-1}P_{i-1} + \mu_{i+1}P_{i+1}$$
$$\lambda_0 P_0 = \mu_1 P_1$$

Finding solution of balance equations

- Start expressing all probabilities in terms of P_0
- Equation for P_0

$$\lambda_0 P_0 = \mu_1 P_1$$

- Sum eqs. for P_1 and P_0

$$\begin{aligned} \lambda_0 P_0 &= \mu_1 P_1 \\ (\lambda_1 + \mu_1) P_1 &= \lambda_0 P_0 + \mu_2 P_2 \end{aligned}$$

$$\lambda_1 P_1 = \mu_2 P_2$$

- Sum result and eq. for P_2

$$\begin{aligned} \lambda_1 P_1 &= \mu_2 P_2 \\ (\lambda_2 + \mu_2) P_2 &= \lambda_1 P_1 + \mu_3 P_3 \end{aligned}$$

$$\lambda_2 P_2 = \mu_3 P_3$$

$\vdots$

- Sum result and eq. for P_i

$$\begin{aligned} \lambda_{i-1} P_{i-1} &= \mu_i P_i \\ (\lambda_i + \mu_i) P_i &= \lambda_{i-1} P_{i-1} + \mu_{i+1} P_{i+1} \end{aligned}$$

$$\lambda_i P_i = \mu_{i+1} P_{i+1}$$

Finding solution of balance equations (continued)

- Recursive substitutions on equations on the left

$$
\begin{aligned}
P_1 &= \frac{\lambda_0}{\mu_1} P_0 \\
P_2 &= \frac{\lambda_1}{\mu_2} P_1 \quad = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} P_0 \\
P_3 &= \frac{\lambda_2}{\mu_3} P_2 \quad = \frac{\lambda_2 \lambda_1 \lambda_0}{\mu_3 \mu_2 \mu_1} P_0 \\
&\vdots \\
P_{i+1} &= \frac{\lambda_i}{\mu_{i+1}} P_i = \frac{\lambda_i \lambda_{i-1} \ldots \lambda_0}{\mu_{i+1} \mu_i \ldots \mu_0} P_0
\end{aligned}
$$

- To find P_0 use $\sum_{i=0}^{\infty} P_i = 1 \;\Rightarrow 1 = P_0 + \displaystyle\sum_{i=0}^{\infty} \frac{\lambda_i \lambda_{i-1} \ldots \lambda_0}{\mu_{i+1} \mu_i \ldots \mu_0} P_0$