Asymptotic frequency and stationary distributions

Theorem (Asymptotic frequency)

Suppose a DTMC is irreducible and recurrent (i.e. all states are recurrent). Then, a.s.

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{1}{\mathbb{E}_y[T_y]}. = \pi(y) \quad \text{IF } \pi \text{ EXISTS}$$

where $1/\infty = 0$.

Ergodic theorem (asymptotic reward)

Theorem

Suppose a DTMC is irreducible and a stationary distribution π *exists. Assume that for some function* f, $\sum_x |f(x)|\pi(x) < \infty$. *Then, a.s.*

$$\frac{1}{n}\sum_{m=1}^{n} f(X_m) \to \sum_x f(x)\pi(x). = E_\pi[g]$$

I.e. time averages (left) equal space averages in stationary regime (right).

The theorem is as a **generalization of the law of large numbers** for discrete random variables, when the random variables we are averaging are not necessarily independent.

Limit distributions and periodicity

Theorem

Suppose a DTMC is irreducible, aperiodic (i.e. all states have period 1), and there is a stationary distribution π. Then, for all $x, y \in S$,

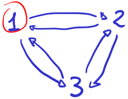
$$\lim_{n\to\infty} P(X_n = y | X_0 = x) = \lim_{n\to\infty} p^{(n)}(x,y) = \pi(y).$$

We will prove $\lim_{n\to\infty} |p^{(n)}(x,y) - \pi(y)| = 0$.

Example

A taxicab driver moves between locations 1, 2, and 3 according to the following treansition probability matrix:

$\pi \cdot P = \pi$



$$P = \begin{pmatrix} 0 & 0.3 & 0.7 \\ 0.2 & 0 & 0.8 \\ 0.7 & 0.3 & 0 \end{pmatrix}$$

$\pi \approx (0.34, 0.23, 0.43)$

IRRIDUCIBILE + FINITO $\Rightarrow$ π ESISTE UNICA

What is the limiting fraction of times the taxi is in location 1?

$$\lim_{n \to +\infty} \frac{N_n(1)}{n} = \pi(1) = 0.34$$

What is the limiting probability the taxi will be in location 1?

$$\lim_{n \to +\infty} P(X_n = 1 \mid X_0 = 2) = \pi(1) = 0.34$$

APERIODICA

$Per(1) = MCD\{2, 3, 4, \dots\} = 1$

Assume that the taxi gets paid \$5, \$10, and \$7 in locations 1, 2, and 3, respectively. What is the average earning per ride?

LONG-TERM

$f(1) = 5 \quad f(2) = 10$

$f(3) = 7$

$$\lim_{n \to +\infty} \frac{\sum_{i=1}^{n} f(X_i)}{n} = f(1)\pi(1) + f(2)\pi(2) + f(3)\pi(3) =$$

Couplings

Definition

Given two probabilities μ and ν on the same probability space $(\Omega, \mathcal{A})$, a coupling of μ and ν is a random variable (X, Y) on $\Omega \times \Omega$ with $X \sim \mu$ and $Y \sim \nu$.

Theorem

Let (X, Y) be a coupling of μ and ν, and let $y \in \Omega$. Then,

$$|\mu(y) - \nu(y)| \leq P(X \neq Y)$$

Proof.

$$\mu(y) - \nu(y) = P(X = y) - P(Y = y) \leq P(X = y, Y \neq y) \leq P(X \neq Y).$$

Similarly, $\nu(y) - \mu(y) \leq P(X \neq Y)$. So, $|\mu(y) - \nu(y)| \leq P(X \neq Y)$. $\square$

Extra

We have (Propositions 4.2 and 4.7 in Levin, Peres, Wilmer's "Markov Chains and Mixing Times") that given two probabilities μ and ν on a discrete set Ω, their **total variation distance** is

$$\begin{aligned}\|\mu - \nu\|_{TV} &= \frac{1}{2}\sum_{y\in\Omega}|\mu(y) - \nu(y)| \\ &= \sup_{A\subseteq\Omega}|\mu(A) - \nu(A)| \\ &= \min\{P(X \neq Y) : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu\}.\end{aligned}$$

Coupling of DTMCs

Consider an irreducible, aperiodic DTMC on S with transition matrix P and stationary distribution π. Let $(X_n)_{n=0}^{\infty}$ and $(Y_n)_{n=0}^{\infty}$ be two *independent* DTMCs with state space S and transition matrix P such that $X_0 = x$ and $Y_0 \sim \pi$.

$Y_1 \sim \pi$

$Y_2 \sim \pi \quad \dots \quad Y_n \sim \pi$

Let $\tau = \min\{n \geq 0 : X_n = Y_n\}$ and define

$$\hat{X}_n = \begin{cases} X_n & \text{if } n < \tau \\ Y_n & \text{if } n \geq \tau \end{cases}.$$

Then, $(\hat{X}_n)_{n=0}^{\infty}$ is still a DTMCs with state space S and transition matrix P, and $\hat{X}_0 = X_0 = x$.

Hence, for any n the random variable $(\hat{X}_n, Y_n)$ is a coupling of $p^{(n)}(x, \cdot)$ and π, so for all $y \in S$

$$|p^{(n)}(x,y) - \pi(y)| \leq P(\hat{X}_n \neq Y_n) = P(\tau > n).$$

Coupling of DTMCs

- We can prove the following lemma (the proof is in the course material):

Lemma

If the original DTMC is irreducible and aperiodic, then $((X_n, Y_n))_{n=0}^{\infty}$ is irreducible.

- $\tilde{\pi}(x, x') = \pi(x)\pi(x')$ is a stationary distribution for $((X_n, Y_n))_{n=0}^{\infty}$ (check it!). Hence, all states of $S \times S$ are positive recurrent.

Then,

$$|p^{(n)}(x, y) - \pi(y)| \le P(\tau > n) \le P(T_{(x,x)} > n) \xrightarrow[n\to\infty]{} 0$$

because, by continuity of the probability measure,

$$\lim_{n\to\infty} P(T_{(x,x)} \le n) = P(T_{(x,x)} < \infty) = \sum_{x'\in S} \underbrace{P_{(x,x')}(T_{(x,x)} < \infty)}_{=1}\pi(x') = \sum_{x'\in S} \pi(x') = 1.$$

(x, x) POS. REC.
IRR.

Time Reversibility and Discrete Time Birth and Death Chains

Detailed balanced distributions

$$\begin{tikzcd} x \arrow[r, bend left, "\pi(x)p(x,y)"] & y \arrow[l, bend left, "\pi(y)p(y,x)"] \end{tikzcd}$$

Definition

A distribution π is said to be detailed balanced if for any $x, y \in S$

$$\pi(x)p(x,y) = \pi(y)p(y,x)$$

Summing over y,

$$(\pi \cdot P)(x) = \sum_{y \in S} \pi(y)p(y,x) \overset{\text{d.b.}}{=} \sum_{y \in S} \boxed{\pi(x)}\, p(x,y) = \underline{\pi(x)} \overbrace{\sum_{y \in S} p(x,y)}^{1} = \pi(x),$$

$$\pi = \pi \cdot P$$

hence a detailed balanced distribution is a **stationary distribution**. The converse does not hold in general.

Time reversibility

Assume a DTMC has a detailed balanced distribution π, and $X_0 \sim \pi$. Then,

$$\begin{aligned}
P(X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x_n)& \\
&= \boxed{\pi(x_0)p(x_0, x_1)}p(x_1, x_2)\cdots p(x_{n-1}, x_n) \\
&= \boxed{p(x_1, x_0)}\boxed{\pi(x_1)p(x_1, x_2)}\cdots p(x_{n-1}, x_n) \\
&= p(x_1, x_0)\boxed{p(x_2, x_1)\pi(x_2)}\cdots p(x_{n-1}, x_n) \\
&= \cdots = p(x_1, x_0)p(x_2, x_1)\cdots p(x_n, x_{n-1})\pi(x_n) \\
&= \underline{P(X_0 = x_n, X_1 = x_{n-1}, \dots, X_{n-1} = x_1, X_n = x_0)}
\end{aligned}$$

Time reversibility

Assume a DTMC has a detailed balanced distribution π, and $X_0 \sim \pi$. Then,

$$P(X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x_n) \\ = P(X_0 = x_n, X_1 = x_{n-1}, \dots, X_{n-1} = x_1, X_n = x_0)$$

At stationary regime, a trajectory or its reverse have the same probability. By observing a stationary chain we cannot guess whether the time is moving forward or backwards!

Definition

An irreducible DTMC with a detailed balanced distribution is called time reversible.

Important concept in Physics and will also be discussed in the other portion of the course.

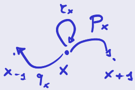
Discrete time birth and death chains

Definition

A DTMC is called a (discrete time) birth and death chain if the following conditions hold.

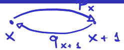
- The state space is a sequence of consecutive integers, bounded from below. That is either $S = \{x \in \mathbb{Z} : a \leq x \leq b\}$ for some integers $a < b$, or $S = \{x \in \mathbb{Z} : a \leq x\}$ for some integer a.
- The size of the allowed jumps is at most one, that is, for every $x, y \in S$

$$
\begin{aligned}
p(x,y) &= 0 \qquad \text{if } |x-y| > 1. \\
p(x,x+1) &= p_x \\
p(x,x-1) &= q_x \\
p(x,x) &= r_x \qquad (\text{therefore } p_x + q_x + r_x = 1).
\end{aligned}
$$



E.g. the gambler model is a birth and death chain.

Discrete time birth and death chains

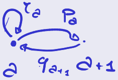


Theorem

If a discrete time birth and death chain has a stationary distribution π, then π is detailed balanced. Namely $\pi(x)p_x = \pi(x+1)q_{x+1}$

Proof.

We prove it by induction. If π is a stationary distribution, then $\boxed{\pi(a) = \pi(a)r_a + \pi(a+1)q_{a+1}}$ and



$(\pi \cdot P)(a) = \pi(a) r_a + \pi(a+1) \cdot q_{a+1}$

$$\pi(a)\underbrace{(1 - r_a)}_{p_a} = \pi(a+1)q_{a+1}.$$

Since $q_a = 0$ then $p_a = 1 - r_a$ and

$$\pi(a)p_a = \pi(a+1)q_{a+1}.$$

□

Discrete time birth and death chains

Theorem

If a discrete time birth and death chain has a stationary distribution π*, then* π *is detailed balanced. Namely* $\pi(x)p_x = \pi(x+1)q_{x+1}$

Proof.

Assume that $\pi(x)p_x = \pi(x+1)q_{x+1}$. Since π is a stationary distribution,

$$\pi(x+1) = \underbrace{\pi(x)p_x}_{\pi(x+1)q_{x+1}} + \pi(x+1)r_{x+1} + \pi(x+2)q_{x+2}. = (\pi \cdot P)(x+1)$$

Hence,

$$\pi(x+1)\underbrace{p_{x+1}}_{1-r_{x+1}-q_{x+1}} = \pi(x+1)(1-r_{x+1}-q_{x+1}) = \pi(x+2)q_{x+2}.$$

$\square$

Discrete time birth and death chains

This gives us an easier way to calculate the stationary distribution of a B&D chain, especially useful when we have infinite states: if a stationary distribution π exists, then

$\pi(x-1)p_{x-1} = \pi(x)q_x$

$$\pi(x) = \frac{p_{x-1}}{q_x}\pi(x-1) = \frac{p_{x-1}}{q_x}\frac{p_{x-2}}{q_{x-1}}\pi(x-2) = \cdots = \left(\prod_{i=a}^{x-1}\frac{p_i}{q_{i+1}}\right)\underline{\pi(a)}.$$

It follows that

$$1 = \sum_{x=a}^{b}\pi(x) = \pi(a)\sum_{x=a}^{b}\prod_{i=a}^{x-1}\frac{p_i}{q_{i+1}}$$

$$\boxed{\prod_{i=a}^{a-1}\frac{p_i}{q_{i+1}} = 1}$$

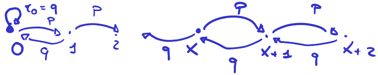
Hence we must have

$$\pi(a) = \frac{1}{M} \qquad \text{where } M = \sum_{x=a}^{b}\prod_{i=a}^{x-1}\frac{p_i}{q_{i+1}}$$

$\pi(x) = \frac{1}{M}\prod_{i=a}^{x-1}\frac{p_i}{q_{i+1}}$

A stationary distribution for the B&D chain exists and is unique if and only if M is finite and greater than 0. Note that checking that $M > 0$ is never an issue, but in many cases M could be infinite.

1d partially reflected Random Walk



$p + q = 1$

Consider a birth and death chain on the natural numbers $\mathbb{N}$ including 0, with transition probabilities

$$
\begin{aligned}
p(x, x+1) &= p_x = p, \text{ for any } \underline{x \geq 0}, \\
p(x, x-1) &= q_x = 1 - p = q, \text{ for any } x \geq 1, \quad (\text{hence } r_x = 0) \\
p(0, 0) &= r_0 = 1 - p = q, \quad (\text{hence } q_0 = 0).
\end{aligned}
$$

Is there a stationary distribution? Since this is a B&D Markov chain, we know that it has a unique stationary distribution if and only if

$$M = \sum_{x=0}^{\infty} \prod_{i=0}^{x-1} \frac{p_i}{q_{i+1}} < \infty.$$

In our case,

$$M = \sum_{x=0}^{\infty} \prod_{i=0}^{x-1} \frac{p_i}{q_{i+1}} = \sum_{x=0}^{\infty} \prod_{i=0}^{x-1} \frac{p}{q} = \sum_{x=0}^{\infty} \left(\frac{p}{q}\right)^x = \frac{1}{1-\frac{p}{q}} \text{ se } p<q$$

and the latter is finite if and only if $p < q$, that is if and only if $p < 1/2$. In this case, a unique stationary distribution exists and it is given by

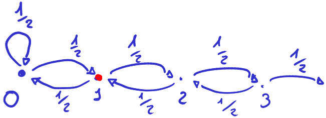
$$\frac{1}{M} = \pi(0) = \frac{1}{\sum_{x=0}^{\infty} \left(\frac{p}{q}\right)^x} = 1 - \frac{p}{q}$$

and

$$\pi(x) = \pi(0) \prod_{i=0}^{x-1} \frac{p_i}{q_{i+1}} = \left(1 - \frac{p}{q}\right)\left(\frac{p}{q}\right)^x.$$

Hence, if $p < 1/2$ then all the states are positive recurrent and we know what the stationary distribution is. Specifically, π is a *shifted geometric distribution* with success parameter p/q.

What happens if $p = \frac{1}{2}$? Is the process transient or null recurrent?



$$P_0(T_0 < \infty) = p(0,1) \cdot P(T_0 < \infty \mid X_0 = 0, X_1 = 1) + p(0,0) \cdot P(T_0 < \infty \mid X_0 = 0, X_1 = 0)$$

// MARKOV

$$= \frac{1}{2} \cdot \underbrace{P_1(T_0 < \infty)}_{\rho_{10} = \rho_{10}^{RW} = 1} + \frac{1}{2} \cdot 1 = \frac{1}{2} + \frac{1}{2} = 1$$

+ STAZ. NON ESISTE

$\Downarrow$

$\boxed{\text{NULL REC.}}$

1d partially reflected Random Walk

What happens if $p = \frac{1}{2}$? Is the process transient or null recurrent?

$$P_0^{\overbrace{RRW}^{\text{REFLECTED RANDOM WALK}}}(T_0 < \infty) = q + pP_1^{RRW}(T_0 < \infty) = q + pP_1^{\overbrace{RW}^{\text{RANDOM WALK}}}(T_0 < \infty)$$

1d partially reflected Random Walk

What happens if $p > \frac{1}{2}$? Is the process transient or null recurrent?

TRANSIENTE!

1d partially reflected Random Walk

What happens if $p > \frac{1}{2}$? Is the process transient or null recurrent?

For the (not reflected!) random walk we have

$$\begin{aligned}
1 > P_0^{RW}(T_0 < \infty) &= qP_{-1}^{RW}(T_0 < \infty) + pP_1^{RW}(T_0 < \infty) \\
&= q\underbrace{P_{-1}^{RW}(T_0 < \infty)}_{1} + pP_1^{RRW}(T_0 < \infty)
\end{aligned}$$

and $P_{-1}^{RW}(T_0 < \infty) = P_1^{RRW}(T_0 < \infty)$ with p and q interchanged.
Hence $\boxed{P_{-1}^{RW}(T_0 < \infty) = 1}$. In conclusion,

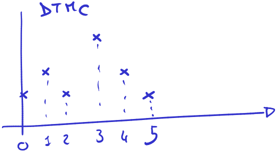
$$1 > q + pP_1^{RRW}(T_0 < \infty) \Rightarrow pP_1^{RRW}(T_0 < \infty) < 1$$

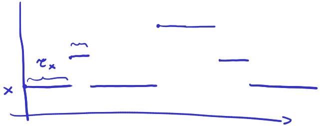
and

$$P_0^{RRW}(T_0 < \infty) = q + pP_1^{RRW}(T_0 < \infty) < 1.$$

Continuous time Markov chains

In short





Continuous time Markov chains (CTMCs) incorporate the structure of a discrete time Markov chain in a continuous time Markov process. Basically, when thinking about a Continuous time Markov chain you should think of a process that "jumps" between different states (exactly as in a discrete time Markov chain), and, before jumping, remains in a given state x for a time τ_x, which is a continuous random variable.

Construction 1

1. We let $(Y_n)_{n\geq 0}$ be a discrete time Markov chain on a state space S. We denote its transition probabilities by

$$r(i,j) = P(Y_{n+1} = j \mid Y_n = i),$$

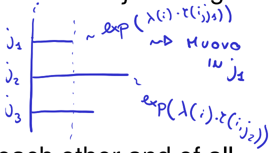
with the condition that $r(i,i) = 0$ for all $i \in S$.

2. After arriving (or starting the process at time zero) at state $i \in S$, we let the amount of time we spend in state i, the *holding time*, to be an exponential random variable with rate $\lambda(i)$. We denote this random variable by τ_i.
INDEPENDENT OF THE REST
3. After the holding time τ_i we transition away from state i according to the probabilities associated with the discrete time Markov chain $(Y_n)_{n\geq 0}$.
4. Finally, we assume that all exponential random variables utilized are independent of each other and of the discrete time Markov chain $(Y_n)_{n\geq 0}$.

Construction 2

1. We suppose that the process has just arrived (or is starting) in state $i \in S$. For each $j \in S$ with $r(i,j) > 0$ we place an alarm clock on state j set to go off after an amount of time $\boxed{\tau_{ij} \sim \text{Exp}(q(i,j))}$ where

$$q(i,j) = \lambda(i) \cdot r(i,j)$$



All exponential random variables are independent of each other and of all previous random variables.

2. When the first alarm goes off, we move to the state associated with that alarm. Formally, we do the following:
   - We let $\tau_i = \min_j \{\tau_{ij}\}$, and let $y \in S$ be the index of the minimum.
   - We then move to state y after a holding time equal to τ_i.

Equivalence

Theorem

The two constructions are equivalent, in the sense that the distribution of the process $(X_t)_{t\geq 0}$ associated with any of the two constructions is the same.

Proof.

This follows from the properties of independent exponential random variables. E.g. the minimum time $\tau_i = \min_j\{\tau_{ij}\}$ is exponential with parameter

$$\sum_{j\in S-\{i\}} \underbrace{q(i,j)}_{\lambda(i)\cdot r(i,j)} = \lambda(i) \overbrace{\sum_{j\in S\setminus\{i\}} r(i,j)}^{1} = \lambda(i).$$

Moreover, the probability that the minimum is achieved at state y is precisely

$$\frac{q(i,y)}{\sum_j q(i,j)} = \frac{\cancel{\lambda(i)}\cdot r(i,y)}{\cancel{\lambda(i)}} = r(i,y),$$

The time spent in one state and which state is next are independent r. v. □

Definitions

The quantities introduced introduce so far are very important and they have a name: $q(i,j)$ are called the transition rates of the process.
It was helpful to put the transition probabilities of a discrete time Markov chain into matrix form. We do the same here. We define the matrix Q to have entries

$$Q(i,j) = \begin{cases} q(i,j) & i \neq j \\ -\sum_{\ell \neq i} q(i,\ell) & i = j \end{cases}$$

for all i,j in S. Note that the diagonal is the *negative* of the sum of the other terms in its row (this choice will turn out to be clever later on). This matrix is also called *transition rate matrix* or *generator*.

Remarks

1. The values $q(i,j)$ are *not* probabilities. Specifically, they can take values larger than 1. Still they are non-negative
2. The row sums of Q are zero.
3. Such matrices (row sums equal to zero, non-negative off-diagonal entries) are sometimes termed *generator matrices* in Linear Algebra.
4. Why did we choose to put the q's in the matrix and not the values $r(i,j)$? The reason is that from the q's we can recover the values $r(i,j)$ and the values $\lambda(i)$. Specifically,

$$\lambda(i) = \sum_{j \neq i} q(i,j) \qquad r(i,j) = \frac{q(i,j)}{\lambda(i)} = \frac{q(i,j)}{\sum_{\ell \neq i} q(i,\ell)}.$$

From the matrix Q we derive all the quantities of interest of our process.

The embedded DTMC

Let $(X_t)_{t\geq 0}$ be a process following one of the above constructions and let Y_n be the discrete time Markov chain giving the (ordered) sequence of states visited by $(X_t)_{t\geq 0}$ (I.e., the DTMC given in Construction 1.) Y_n is called *the embedded discrete time Markov chain* associated with X_t. Moreover, the transition probabilities of Y_n are given by the values $r(i,j)$. That is

$$\begin{aligned} P(Y_{n+1}=j \mid Y_n=i) &= P(\text{next state visited by } X \text{ is } j \mid X \text{ is currently in state } i) \\ &= r(i,j). \end{aligned}$$

Note that the embedded discrete time Markov chain keeps track of the "changes of state" of the continuous time Markov chain. Therefore it never stays put in the same state: that is why transition probabilities $r(i,i)$ are set to 0 for every state i - unless the state is absorbing (we'll get back to it in a few slides).

Example 1: Poisson process

Let $N(t)$ be a rate λ Poisson process. Note that this process follows the above construction with

$$r(i, i+1) = 1, \text{ and } \tau_i \sim \text{Exp}(\lambda).$$

The Q matrix is below as well as the transition matrix P for the embedded DTMC

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & & \ddots & \ddots \end{pmatrix} \qquad P = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}.$$

Note that we could have $\lambda > 1$.

Example 2

Consider the 3 state model with $S = \{A, B, C\}$

$$Q = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} -3 & 1 & 2 \\ 0 & -4 & 4 \\ 2 & 6 & -8 \end{pmatrix} \end{array}.$$

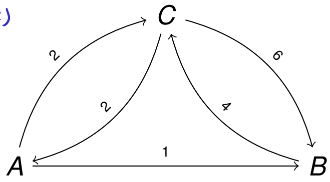
$-3 = -\lambda(A)$

$\lambda(C) = 8$

We can draw the situation by means of a *transition graph*:

$$\tau(A, C) = \frac{q(A, C)}{q(A, B) + q(A, C)} = \frac{2}{1+2} = \frac{2}{3}$$

$$\lambda(A) = q(A, B) + q(A, C) = 1 + 2 = 3$$



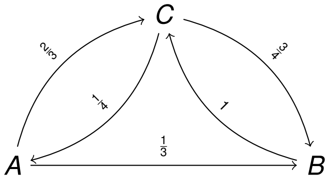
where the numbers on top of the arrows describe the transition rates. The holding times in states A, B, and C are exponentially distributed with parameters $3, 4$ and 8, respectively.

Example 2

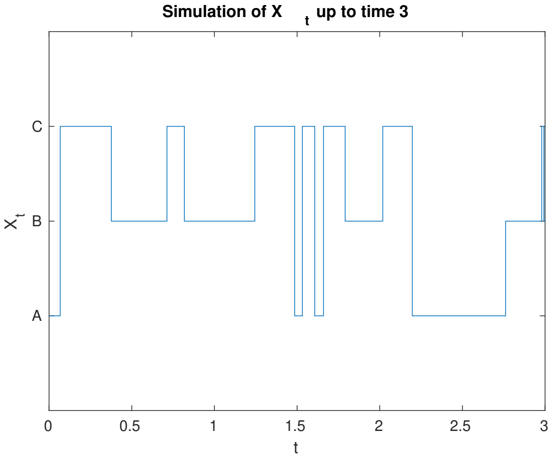
The embedded chain has the same state space and transition matrix

$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 1 \\ \frac{1}{4} & \frac{3}{4} & 0 \end{pmatrix}$$

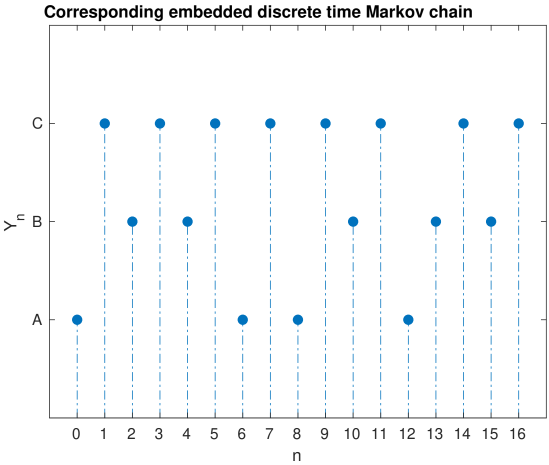
For comparison, the transition graph of the embedded DTMS:



Example 2



Example 2


Corresponding embedded discrete time Markov chain
C
B
A
Y_n
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
n

Absorbing states

A problem with Construction 1 is that in the presence of absorbing states we should redefine the embedded DTMC. Indeed after the system reaches an absorbing state, there is no *next state*. To be consistent we set the following agreement $r(i,i) = 1$

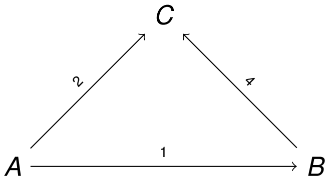
- if i is absorbing, then $\lambda(i) = 0$, meaning that to exit i (towards any other state) we need to wait a degenerate exponential time with an infinite mean.
- an absorbing state i in a DTMC is such that the transition matrix $p(i,j) = \delta_{i,j}$. Analogously for the embedded DTMC, we set $r(i,j) = \delta_{i,j}$ for all absorbing states i. Therefore for absorbing states (and only for them) the transition matrix of the embedded DTMC has non-null diagonal entries equal to 1.

Example 3

Consider the 3 state model with $S = \{A, B, C\}$

$$Q = \begin{pmatrix} -3 & 1 & 2 \\ 0 & -4 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

We can draw the situation by means of a *transition graph*:



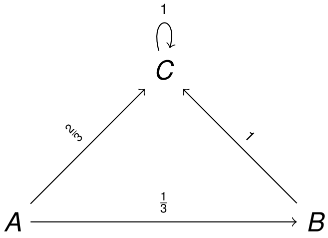
where the numbers on top of the arrows describe the transition rates. The holding times in states A and B are exponentially distributed with parameters 3 and 4, respectively. The state C is absorbing.

Example 2

The embedded chain has the same state space and transition matrix

$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

For comparison, the transition graph of the embedded DTMS:



The Markov property

Definition (Alternative definition of CTMCs)

A continuous-time stochastic process $(X_t)_{t\in[0,\infty)}$ with discrete state space S is a CTMC if for all $t, s \in [0,\infty)$ and all $x \in S$ we have

$$P(X_{t+s} = x \mid \mathcal{F}_t^X) = P(X_{t+s} = x \mid X_t)$$

Moreover, as in the discrete time case, we will only consider time homogeneous cases, that is we always assume that

$$P(X_{t+s} = y \mid X_t = x) = P(X_s = y \mid X_0 = x).$$

Construction 1 is Markovian

x t $\sim \exp(\lambda(x))$

INDEPENDE OF THE PAST.

The idea is the following. Consider Construction 1, then

- If you know that $X_t = x$, the residual holding time at x from time s on will still be $\sim \text{Exp}(\lambda(x))$ by the memoryless property.
- Then by construction the process will jump to a target state that is independent from the holding time and reiterate this behavior, independently on anything that happened before time t.

Therefore, after time t the process will start anew from state x with the same law that it would have had if x had been the initial state, irrespectively of all the history of the process before time t.

Markovianity + Discrete state space $\Longrightarrow$ Construction 1

the proof mainly boils down to prove that

- the holding times of $(X_t)_{t\in[0,\infty)}$ must be memoryless, hence exponentially distributed;
- the jumps directions are independent on the holding times.

so $(X_t)_{t\in[0,\infty)}$ follows Construction 1.
More gory math details are addressed in Norris, *Markov Chains*, CUP 1997

Example: M/M/s queue

In queueing theory, the notation M/M/s means:

- the arrivals are Markovian, meaning that customers arrive according to a Poisson process: the time between two arrivals is $\sim$ Exp(λ)
- the service times are Markovian, meaning that they are independent and exponentially distributed with a common rate μ;
- there are s servers that works in parallel.

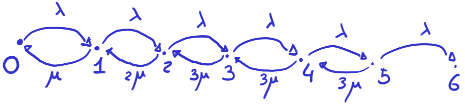
X_t = # PEOPLE IN THE BANK (BOTH SERVED AND NOT)

To make it more concrete you can think at a bank with s teller stations.
Try to guess the generator.

$S = \{0, 1, 2, 3, 4, 5, \dots\} = \mathbb{N}$

$q(i,j) = ?$

$s = 3$



Example: M/M/s queue

In queueing theory, the notation M/M/s means:

- the arrivals are Markovian, meaning that customers arrive according to a Poisson process: the time between two arrivals is $\sim \text{Exp}(\lambda)$
- the service times are Markovian, meaning that they are independent and exponentially distributed with a common rate μ;
- there are s servers that works in parallel.

To make it more concrete you can think at a bank with s teller stations.

Try to guess the generator.

The answer is

$$q(n, n-1) = \begin{cases} n\mu & \text{if } 0 \leq n \leq s \\ s\mu & \text{if } n \geq s \end{cases} \qquad q(n, n+1) = \lambda$$