Markov Chains IV

Károly Simon

This course is based on the book:
Essentials of Stochastic processes
by R. Durrett

Department of Stochastics
Budapest University of Technology and Economics
www.math.bme.hu/~simonk

Autumn 2025, BME

Barbershop example

The following example is from [1, Section 4.3]

Example 1.1

In a barbershop, a single barber cuts hair. There is also a waiting room with two chairs for two people (not counting the one whose hair is being cut). We know the following:

Barbershop example (cont.)

a. **Customers arrive at times of a rate 2 Poisson process**, where the units are people per hour, but will leave if both chairs in the waiting room are occupied.
b. **The barber can cut hair at rate 3, i.e. each haircut requires an exponentially distributed amount of time with mean 20 minutes**, independently of previous haircuts, and also of the arrivals.

Barbershop example (cont.)

Questions:

a. Find the equilibrium distribution.
b. What fraction of potential customers enter service?
c. What is the average amount of time in the system for a customer who enters service?
d. Which fraction of the time there are no customers in the barbershop?

Some words about the barbershop Example

All of the times are measured in hours. The time of the hair cut is $\mathrm{Exp}(3)$. Let $\Delta t > 0$ be very small.

In a time interval of length Δt:

- with probability $3 \cdot \Delta t + \mathfrak{o}(\Delta t)$ exactly one hair cut will be finished (if there are any costumers in the barbershop),
- with probability $2 \cdot \Delta t + \mathfrak{o}(\Delta t)$ a new costumer arrives.

Some words about the barbershop Example (cont.)

In conclusion:

- At time $t + \Delta t$ there will be one costumer less than at time t with probability $3 \cdot \Delta t + \mathfrak{o}(\Delta t)$, if at time t there were any costumers in the barbershop.
- At time $t + \Delta t$ there will be one costumer more than at time t with probability $2 \cdot \Delta t + \mathfrak{o}(\Delta t)$.

Some words about the barbershop Example (cont.)

- Let $S := \{0, 1, 2, 3\}$ be the state space (the possible number of costumers in the barbershop).
- Let X_t be the number of costumers at time t where $t \in \mathbb{R}^+ := \{t : t \geq 0\}$ non-negative real number it indicates the time measured in hours.

Then for all $0 \leq s_0 < s_1 < \cdots < s_n < s$ and for all $i_0, \dots, i_n, j \in S$ we have

$$(1) \quad \mathbb{P}\left(X_{t+s} = j | X_s = i, X_{s_n} = i_n, \dots, X_{s_0} = i_0\right) = \mathbb{P}\left(X_t = j | X_0 = i\right).$$

Countinuous-time MC, introduction

Definition 1.2

In general, if X_t, $t \geq 0$ takes values from a countable state space S and for all $0 \leq s_0 < s_1 < \cdots < s_n < s$ and for all $i_0, \dots, i_n, j \in S$, (1) holds that is

(2) $\mathbb{P}\left(X_{t+s} = j | X_s = i, X_{s_n} = i_n, \dots, X_{s_0} = i_0\right)$

$$= \mathbb{P}\left(X_t = j | X_0 = i\right) =: p_t(i,j).$$

then we say that X_t is a time homogeneous continuous-time Markov chain (MC).

Countinuous-time MC, introduction (cont.)

Since all of the Markov chains consider in this course are time homogeneous, we simply call them continuous-time Markov chains.

Continuity condition : This is very important!!!

Continuity condition: We always assume that the transition matrix $P_t = (p_t(i,j))_{ij \in S}$, $t > 0$ is continuous at zero. That is:

(3) $$\lim_{t \to 0} p_t(i,j) = \delta_{i,j} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

In this way

(4) $$P_0 = \mathrm{Diag}(1, 1, \dots, 1).$$

Continuity condition (cont.)

Observe that (3) holds for example in the barbershop example:
Namely, for a small $h > 0$,

$$p_h(i, i+1) = 2 \cdot h + \mathfrak{o}(h), \quad p_h(i, i-1) = 3 \cdot h + \mathfrak{o}(h)$$

and $p_h(i,j) = \mathfrak{o}(h)$ if $|i - j| > 1$.

Chapman-Kolmogorov

Lemma 1.3 (Chapman-Kolmogorov equality:)

(5) $$\sum_k p_s(i,k) p_t(k,j) = p_{s+t}(i,j).$$

In other words

(6) $$P_{t+s} = P_t \cdot P_s.$$

Proof.
To get the chain from i to j in time $s + t$, it needs to be somewhere after time s.

Infinitesimal generator

Proposition 1.4

For a general, continuous-time MC with countable state space, the following limits exists:

(7) $$\lim_{h \to 0+} \frac{p_h(i,j)}{h} =: q(i,j), \quad i \neq j \text{ and}$$

(8) $$\lim_{h \to 0+} \frac{1 - p_h(i,i)}{h} =: \lambda(i).$$

Moreover,
$0 \leq q(i,j) < \infty, i \neq j$ but $0 \leq \lambda(i) \leq \infty$.

So $q(i,j)$ is finite, but $\lambda(i)$ can be infinite. If $\#S < \infty$ then of course $\lambda(i)$ is also finite.

In summary

It follows from (7) and (8) that for every $i \in S$

(9) $$\lambda(i) = \sum_{\substack{i \neq j \\ j \in S}} q(i,j).$$

For an $i \neq j$, $i, j \in S$ we have

(10) $$\mathbb{P}\left(X_{t+\Delta t} = j | X_t = i\right) = q(i,j) \cdot \Delta t + \mathfrak{o}(\Delta t).$$

For all $i \in S$

(11) $$\mathbb{P}\left(X_{t+\Delta t} = i | X_t = i\right) = 1 - \lambda(i) \cdot \Delta t + \mathfrak{o}(\Delta t).$$

Infinitesimal generator (cont.)

The proof of the previous Proposition is available in [4, Theorems 1.1 and 1.2]. We define

$$q(i,i) := -\lambda(i).$$

Then we form the matrix called Infinitesimal generator:

$$Q = (q(i,j))_{i,j \in S}.$$

That is

Infinitesimal generator (cont.)

$$Q = \begin{bmatrix} -\lambda_1 & q(1,2) & q(1,3) & \cdots \\ q(2,1) & -\lambda_2 & q(2,3) & \cdots \\ q(3,2) & q(3,2) & -\lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Clearly, $p_h(i,i) - 1 + \sum_{i \neq j} p_h(i,j) = 0$ for all $h > 0$, so

(12) $$\sum_{j \in S} q(i,j) = 0 \qquad \forall i \in S.$$

Infinitesimal generator for the barbershop example

In the barber shop example:by formula (1) on slide 6 and formula (6) on slide 15 of File MC III:

$$q(i,i-1) = 3 \text{ if } i = 1,2,3$$
$$q(i,i+1) = 2 \text{ if } i = 0,1,2.$$

That is:

$$Q = \begin{array}{|c|c|c|c|c|} \hline & \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \hline \mathbf{0} & -2 & 2 & 0 & 0 \\ \hline \mathbf{1} & 3 & -5 & 2 & 0 \\ \hline \mathbf{2} & 0 & 3 & -5 & 2 \\ \hline \mathbf{3} & 0 & 0 & 3 & -3 \\ \hline \end{array}.$$

Infinitesimal generator, a comment

We get from Chapman-Kolmogorov equality, that if we know the transition matrix for small t, then we know it for all t, because $P_{nh} = (P_h)^n$. This gives the idea, that if we know the transition matrices' derivative at 0 then we know the transition matrix P_t for every t.

Theorem 1.5

Let X_t be a continuous-time MC with finite state space S. As always, we assume that (3) holds. Then

(a) *the transition matrix $P_t = (p_t(i,j)_{i,j \in S})$ satisfies the so-called Kolmogorov's-forward differential equation:*

(13) $$\frac{d}{dt} P_t = P_t \cdot Q, \qquad t \geq 0.$$

(b) *The solution of (13) is $P_t = \boldsymbol{\alpha} \cdot e^{tQ}$, where $\boldsymbol{\alpha}$ is the initial distribution of the MC at time $t = 0$.*

Proof

We have already used the following notation many times: $\mathbb{P}_x(X_t = y) := \mathbb{P}(X_t = y | X_0 = x)$. Let us fix a small $t > 0$ and $x, y \in S$. Using the Law of Total Probability:

$$\begin{aligned} &\mathbb{P}_x(X_{t+\Delta t} = y) - \mathbb{P}_x(X_t = y) \\ &= \mathbb{P}_x(X_{t+\Delta t} = y | X_t = y) \cdot \mathbb{P}_x(X_t = y) \\ &+ \sum_{u \neq y} \mathbb{P}_x(X_{t+\Delta t} = y | X_t = u) \cdot \mathbb{P}_x(X_t = u) - \mathbb{P}_x(X_t = y) \\ &= \left[1 - \lambda(y)\Delta t + \mathfrak{o}(\Delta t) - 1\right] \cdot \mathbb{P}_x(X_t = y) \\ &+ \sum_{u \neq y} ([q(u,y)\Delta t + \mathfrak{o}(\Delta t)]) \cdot \mathbb{P}_x(X_t = u). \end{aligned}$$

Proof (cont.)

If we divide both sides by Δt, and $\Delta t \to 0$, then

(14) $$\frac{d}{dt}\mathbb{P}_x(X_t = y) = \mathbb{P}_x(X_t = y)(-\lambda(y)) + \sum_{u \neq y} \mathbb{P}_x(X_t = u) \cdot q(u,y).$$

In the equation above, the left-hand side is the (x,y)-element of matrix $\frac{d}{dt}P_t$, and the right-hand side is the (x,y)-element of the matrix $P_t \cdot Q$. Using that $x, y \in S$ and $t > 0$ were arbitrary, we get that $\frac{d}{dt}P_t = P_t \cdot Q$.

Kolmogorov's forward and backward differential equations

Kolmogorov's forward differential equation:

(15) $$\frac{d}{dt}P_t = P_t \cdot Q$$

Kolmogorov backward differential equation:

(16) $$\frac{d}{dt}P_t = Q \cdot P_t.$$

Kolmogorov's forward and backward differential equations (cont.)

These equations have a very important role, but studying them would exceed the limits of this course. Suggested reading: Péter Major's lecture on Continuous-time Markov Chains (A folytonos idejű Markov láncokról): click here We make some comments without proofs.

Kolmogorov's forward and backward differential equations: Conditions

Conditions

(F1) $\lambda(i) < \infty$, $\forall i$ (defined in formula (7)).

(F2) For every fixed j the convergence in formula (7) is uniform in i.

Interestingly, Kolmogorov's backward differential equation can have solutions which are not solutions of Kolmogorov's forward differential equation and which are relevant from probability theory point of view (Satisfy Chapman- Kolmogorov equation).

Kolmogorov's forward and backward differential equations: Conditions

Proposition 1.6

(a) If both of the conditions F1 and F2 hold then P_t satisfies Kolmogorov's forward differential equation.

(b) If we only know that condition F1 holds then P_t satisfies Kolmogorov's backward differential equation.

Recall again that we always assume that (3) holds (we only consider chains with continuous transition matrix in 0).

Exponential waiting times

For all $x \in S$ let T_x be the time that the chain spends at state $x \in S$ after it has arrived at x.

Lemma 1.7

Let us assume that $\lambda_x < \infty$ holds for all $x \in S$. Then

(a) $T_x = \mathrm{Exp}(\lambda_x)$ holds for all $x \in S$ and

(b) $\{T_x\}_{x\in S}$ are independent.

Exponential waiting times (cont.)

Proof of part (a)

Let

$$G_x(t) := \mathbb{P}(T_x \geq t).$$

By the Markov property:

$$\begin{aligned} G_x(t+\Delta t) &= G_x(t)G_x(\Delta t) = \\ &= G_x(t)\left[1-\lambda(x)\Delta t + \mathfrak{o}(\Delta t)\right] \end{aligned}$$

Hence,

$$G_x'(t) = -\lambda(x)G_x(t).$$

Exponential waiting times (cont.)

Proof of part (a) (cont.)

Clearly,

$$1-\mathbb{P}(T_x < t) = G_x(t) = \mathrm{e}^{-t\lambda(x)}.$$

So $T_x = \mathrm{Exp}(\lambda_x)$. □

Proof of part (b) It is obvious from the Markov property, that $\{T_x\}_{x\in S}$ are independent. □

Routing matrix

Definition 1.8

Assume that $\lambda_x < \infty$ holds for all $x \in S$. Now we define the so-called routing matrix: $R = (r(x,y))_{x,y\in S}$ as follows: the diagonal elements are all zeros: $r(x,x) := 0$ for all $x \in S$. Let $x, y \in S$ be arbitrary distinct. Imagine that the chain is in state x and it stays there for a while then it jumps. Let $U(x,y)$ be the event that the chain jumps from x to y when it leaves x and we write $r(x,y)$ for the probability of the event $U(x,y)$. The discrete time MC corresponding to the transition matrix R is called embedded chain.

Lemma 1.9

Assume that $\lambda_x < \infty$ holds for all $x \in S$. Let $R = (r(x,y))_{x,y\in S}$ be the routing matrix. Then

$$r(x,y) = \frac{q(x,y)}{\lambda_x}, \qquad \forall x \neq y. \tag{17}$$

Proof

Let $U(x,y)$ be the event that when the chain jumps from x to y. Let f be the density function of T_x. Then

$$\mathbb{P}(U(x,y)) = \int_{t=0}^{\infty} \mathbb{P}(U(x,y)|T_x = t)\cdot f(t)dt. \tag{18}$$

Proof (cont.)

By definition

$$\begin{aligned} \mathbb{P}(U(x,y)|T_x = t) &= \lim_{\Delta t\to 0} \frac{\mathbb{P}(X_{t+\Delta t} = y|X_t = x)}{\sum\limits_{z\in S\setminus\{x\}} \mathbb{P}(X_{t+\Delta t} = z|X_t = x)} \\ &= \lim_{\Delta t\to 0} \frac{q(x,y)\Delta t + \mathfrak{o}(\Delta t)}{\lambda(x)\Delta t + \mathfrak{o}(\Delta t)} \\ &= \frac{q(x,y)}{\lambda(x)} \quad \forall t\text{-re} \end{aligned}$$

We substitute this back to formula (18) and we obtain the assertion of the Lemma.

Stationary distribution, irreducibility

Like on the previous slides, here we do NOT assume that $\#S < \infty$.

Definition 1.10

X_t is irreducible, if from any state i, any state j can be reached in finitely many steps. In other words, if $\exists k_0 = i, k_1, \dots, k_{n-1}, k_n = j$, that

$$q(k_{\ell-1}, k_\ell) > 0, \quad \forall \ell = 1, \dots, n \tag{19}$$

Stationary distribution, irreducibility (cont.)

Lemma 1.11

If X_t is irreducible, then $\forall t > 0$ and $\forall i, j$, $p_t(i,j) > 0$. (No problem with the period.)

Stationary distribution, irreducibility (cont.)

Proof

Fix an $i, j \in S$ and choose $k_1, k_1, \dots, k_n$ as in Definition 1.10. We obtain from formulas (7) and (19) that $\exists h_0 > 0$, such that for every $0 < h < h_0$, $p_h(k_{\ell-1}, k_\ell) > 0$. From here

$$p_{h'}(i,j) > 0, \quad \forall h' < nh_0 \tag{20}$$

Stationary distribution, irreducibility (cont.)

Proof (cont.)

On the other hand, we know that the waiting time at j has exponential distribution. Then for every $s > 0$:

$$p_s(j,j) \geq \mathbb{P}(T_j > s) = \exp(-s\lambda_j) > 0. \tag{21}$$

Let $0 < h < h_0$ and $s > 0$ s.t. $t = s + nh$. Then from formulas (20) and (21):

$$p_t(i,j) \geq p_{nh}(i,j) \cdot p_s(j,j) > 0. \quad \square$$

Definition 1.12

Probability vector $\boldsymbol{\pi}$ is called stationary distribution, if

$$\forall t > 0 : \quad \boldsymbol{\pi}^T \cdot P_t = \boldsymbol{\pi}^T, \quad \forall t > 0. \tag{22}$$

Because it is hard to check such a condition simultaneously for every t, the following Lemma will be useful:

Lemma 1.13

The probability vector $\boldsymbol{\pi}$ is the stationary distribution iff

$$\boldsymbol{\pi}^T \cdot Q = \mathbf{0}. \tag{23}$$

Stationary distribution

Proof

Assume, that $\boldsymbol{\pi}^T \cdot P_t = \boldsymbol{\pi}^T$ holds for all $t > 0$. By Kolmogorov's forward differential equation:

$$\begin{aligned} 0 &= \frac{d}{dt}\left(\boldsymbol{\pi}^T \cdot P_t\right)(j) \\ &= \sum_i \pi(i) \sum_k p_t(i,k) \cdot q(k,j) \\ &= \sum_k \underbrace{\sum_i \pi(i) p_t(i,k)}_{\pi(k)} q(k,j). \end{aligned}$$

Proof (cont.)

So, the j^{th} component of the vector $\boldsymbol{\pi}^T \cdot Q$ is 0 for every j. This means that $\boldsymbol{\pi}^T \cdot Q = \mathbf{0}$.

The other direction: Assume, that $\boldsymbol{\pi}^T \cdot Q = \mathbf{0}$. Using Kolmogorov backward differential equation in the second step and the fact that $P_0 = \mathrm{Diag}(1, \dots, 1)$ we get

Proof (cont.)

$$\begin{aligned} \frac{d}{dt}\left(\sum_i \pi(i) p_t(i,j)\right) &= \sum_i \pi(i) p_t'(i,j) \\ &= \sum_i \pi(i) \sum_k q(i,k) p_t(k,j) \\ &= \sum_k \underbrace{\sum_i \pi(i) q(i,k)}_{0} p_t(k,j) = 0. \end{aligned}$$

Hence, $\boldsymbol{\pi}^T P_t$ is constant. So, it is equal to $\boldsymbol{\pi}^T P_0 = \boldsymbol{\pi}^T \cdot \mathrm{Diag}(1, \dots, 1) = \boldsymbol{\pi}^T$. $\square$

Limiting behavior

Theorem 1.14

Consider a continuous-time and irreducible MC for which there exists a stationary distribution $\boldsymbol{\pi}$. Then

$$\lim_{t\to\infty} p_t(i,j) = \pi(j), \quad \forall i \in S. \tag{24}$$

Proof.
Because of Lemma 1.11 for every $h > 0$ matrix P_h is irreducible and aperiodic. Thus using Theorem 6.2 from file MC I: we get $\lim_{n\to\infty} p_{nh}(i,j) = \pi(j)$. □

Detailed balance condition

Extending the notion for discrete-time MC, we say that detailed balance condition holds if:

Definition 1.15

$$\pi(k)q(k,j) = \pi(j)q(j,k), \quad \forall j \neq k. \tag{25}$$

Detailed balance condition (cont.)

Theorem 1.16

Let π be a probability vector ($\sum_{i\in S} \pi_i = 1$ and $\pi_i \geq 0$). If π satisfies (25) then $\boldsymbol{\pi}$ is stationary distribution.

Detailed balance condition (cont.)

Proof.
Fix an arbitrary $j \in S$

$$\sum_{k\neq j, k\in S} \pi(k)q(k,j) = \pi(j) \sum_{k\neq j, k\in S} q(k,j) = \pi(j)\lambda_j,$$

in other words, $\forall j$:

$$\sum_{k\neq j, k\in S} \pi(k)q(k,j) - \pi(j)\lambda_j = 0.$$

Observe that the left-hand side is the j^{th} component of vector $\boldsymbol{\pi}^T \cdot Q$. □

The chain from given rates if $\#S < \infty$

Informal construction of the chain:

Let us assume, that the chain is at state i at a given time $t \geq 0$. If $\lambda_i = 0$, then it remains in i forever, if $\lambda_i > 0$, then the chain remains in i for time $\mathrm{Exp}(\lambda_i)$ and then it jumps to j with probability $r(i,j)$, where $r(i,j)$ was defined on slide 30.

Now we give another description of the continuous time finite sate MC. To understand it recall part (e) on slide 7 from File MC III.

The chain from given rates if $\#S < \infty$ (cont.)

The same in other words:

Assume that the chain now is at state i. Imagine that at every state $j \neq i$ there is a clock with parameter $\mathrm{Exp}(q(i,j))$. The chain jumps:

- when the first clock rings,
- to the state where the first clock rings.

The equivalence of this characterization follows from Lemmas 1.7 and 1.9 (see slides 27 and 31).

Lemma 2.1

Let X_t be an irreducible, continuous MC with finite state space. We denote the infinitesimal generator by Q, as usual. Then

(a) *There exists a unique probability vector $\boldsymbol{\pi}$ which is the left eigenvector of Q with eigenvalue 0.*

(b) *The real part of any non-zero eigenvalues of Q is negative.*

Proof of part (a):

Let $a > |\max_{i,j} q(i,j)|$. Then

$$P := (1/a)Q + I$$

is an irreducible stochastic matrix. Let $\boldsymbol{\pi}$ be the left eigenvector of P for eigenvalue 1. Obviously, $\boldsymbol{\pi}^T \cdot Q = \mathbf{0}$ if and only if $\boldsymbol{\pi}^T \cdot P = \boldsymbol{\pi}^T$. This yields existence and uniqueness of $\boldsymbol{\pi}$.

For the proof of Part (b) see [5, Exercise 3.4].

Example: a special chain with two states

Let $S = \{1,2\}$ and we know, that $q(1,2) = 1$ and $q(2,1) = 2$. Then $\lambda(1) = 1$ and $\lambda(2) = 2$. In other words,

$$Q = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$$

We know, that

$$P_t = e^{tQ}. \tag{26}$$

Example: a special chain with two states (cont.)

To compute this, we must diagonalize Q:

$$D = \begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix}, R = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}, R^{-1} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}$$

So

$$Q = R \cdot D \cdot R^{-1}$$

From here

$$e^{tQ} = R \cdot \begin{bmatrix} 1 & 0 \\ 0 & e^{-3t} \end{bmatrix} \cdot R^{-1}$$

Example: a special chain with two states (cont.)

In other words:

$$e^{tQ} = \begin{bmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{bmatrix} + e^{-3t} \begin{bmatrix} 1/3 & -1/3 \\ -2/3 & 2/3 \end{bmatrix}$$

Obviously for $\boldsymbol{\pi}^T = (2/3, 1/3)$,

$$\lim_{t\to\infty} P_t = \begin{bmatrix} \boldsymbol{\pi}^T \\ \boldsymbol{\pi}^T \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{bmatrix}.$$

Chains with two states in general

In general: let us assume that for some $\lambda, \mu > 0$

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

The one can prove, like above, that

$$P_t = \begin{bmatrix} \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \\ \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \end{bmatrix} + e^{-t(\mu+\lambda)} \begin{bmatrix} \frac{\lambda}{\lambda+\mu} & -\frac{\lambda}{\lambda+\mu} \\ -\frac{\mu}{\lambda+\mu} & \frac{\mu}{\lambda+\mu} \end{bmatrix} \tag{27}$$

In other words, for $\boldsymbol{\pi}^T := (\frac{\mu}{\lambda+\mu}, \frac{\lambda}{\lambda+\mu})$

$$\lim_{t\to 0} P_t = \begin{bmatrix} \boldsymbol{\pi}^T \\ \boldsymbol{\pi}^T \end{bmatrix}.$$

A chain with four states

Example 2.2

Let us consider the continuous MC, whose infinitesimal generator is

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

Compute the stationary distribution for this chain.

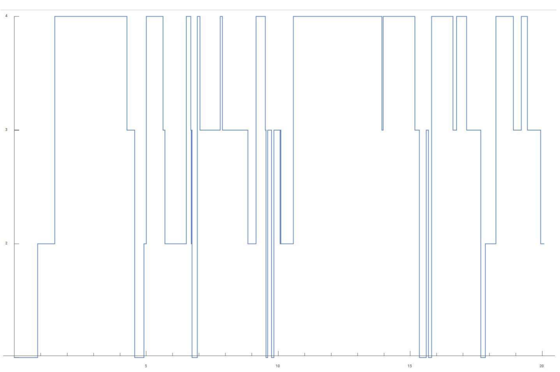

Figure: Simulation for Example 2.2

Birth and death Chains

The state space S may be finite or countably infinite: $S=\{0,1,2,\dots N\}$, where $N \leq \infty$ and we are allowed to make only one step ahead (birth) with rate λ_n or one step back one step (death) with rate μ_n. That is

$$q(n,n+1)=\lambda_n \text{ for } n<N \tag{28}$$
$$q(n,n-1)=\mu_n \text{ for } n>0. \tag{29}$$

This means that

$$\begin{aligned}\mathbb{P}(X_{t+\Delta t}=n|X_t=n) &= 1-(\mu_n+\lambda_n)\Delta t+\mathfrak{o}(\Delta t)\\ \mathbb{P}(X_{t+\Delta t}=n+1|X_t=n) &= \lambda_n\Delta t+\mathfrak{o}(\Delta t)\\ \mathbb{P}(X_{t+\Delta t}=n-1|X_t=n) &= \mu_n\Delta t+\mathfrak{o}(\Delta t).\end{aligned}$$

Barbershop again

Recall from slide 18 that in the barbershop example $S=\{0,1,2,3\}$ and the infinitesimal generator:

$$Q=\begin{array}{c|cccc} & 0 & 1 & 2 & 3\\ \hline 0 & -2 & 2 & 0 & 0\\ 1 & 3 & -5 & 2 & 0\\ 2 & 0 & 3 & -5 & 2\\ 3 & 0 & 0 & 3 & -3\end{array}. \tag{30}$$

This is a birth and death chain with

$$\lambda_0=\lambda_1=\lambda_2=2 \text{ and } \mu_1=\mu_2=\mu_3=3. \tag{31}$$

Stationary distribution

Theorem 3.1

Let X_n be a birth and death chain with: $S=\{0,1,\dots,N\}$, *where* $N\leq\infty$. $q(n,n+1)=\lambda_n$ *if* $n<N$ *and* $q(n,n-1)=\mu_n$ *if* $n>0$. $\mu_0=0$ *and* $\lambda_N=0$, *if* $N<\infty$. *Then*

$$\pi(n)=\frac{\lambda_{n-1}\lambda_{n-2}\cdots\lambda_0}{\mu_n\mu_{n-1}\cdots\mu_1}\pi(0) \tag{32}$$

satisfies detaliled balance condition, so it gives stationary distribution, if $\sum_{n=1}^{N}\frac{\lambda_{n-1}\lambda_{n-2}\cdots\lambda_0}{\mu_n\mu_{n-1}\cdots\mu_1}<\infty$ *(which is always satisfied, if* $N<\infty$*).*

Stationary distribution for the barbershop

$S:=\{0,1,2,3\}$ using (31):

$$\mu_i=3,\ i=1,2,3 \text{ and } \lambda_i=2,\ i=0,1,2.$$

If $\pi(0)=c$, then repeated applications of (32) gives:

$$\pi(1)=\frac{2c}{3},\ \pi(2)=\frac{2^2}{3^2}c,\ \pi(3)=\frac{2^3}{3^3}c. \tag{33}$$

$\sum_{i=0}^{3}\pi(i)=1$ yields $c\left(1+\frac{2}{3}+\left(\frac{2}{3}\right)^2+\left(\frac{2}{3}\right)^3\right)=1$. From this we get c and substitute it back to (33). We get

$$\pi(0)=\tfrac{27}{65},\ \pi(1)=\tfrac{18}{65},\ \pi(2)=\tfrac{12}{65},\ \pi(3)=\tfrac{8}{65}. \tag{34}$$

Conclusion

This gives answer to the question (a) asked on slide 3 The answer of question (b) (from the same place) is as follows: there are three customers at $\pi(3)=\frac{8}{65}$ part of the time.This means that $57/65=87.7\%$ of potential customers who enter the barbershop have eventually get their haircut. We will answer question (c) later.

M/M/s queuing

Example 3.2 (M/M/s queuing)

Let us imagine a bank, where customers are being served by $s\leq\infty$ servers, and they are waiting in one queue if there are more customers than servers. It is reasonable to assume, that customers arrive by a $\mathrm{Poisson}(\lambda)$ process and the serving times are independent $\mathrm{Exp}(\mu)$.

Jump rates:

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

Stationary distribution for $M/M/\infty$ queuing

Example 3.3 ($M/M/\infty$ queuing)

$$q(n,n+1)=\lambda \text{ and } q(n,n-1)=n\mu.$$

Then $\pi(n)=\frac{(\lambda/\mu)^n}{n!}\pi(0)$. So, we choose $\pi(0)=e^{-\lambda/\mu}$ and then we see that the stationary distribution is $\mathrm{Poi}(\lambda/\mu)$.

$M/M/s$ queuing with balking I

Recall example 3.2 about $M/M/s$ queuing (on slide 62):

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

We modify it slightly: Customers arrive at times of a Poisson process with rate λ but only join the queue with probability a_n if there are n customers in line. and with probability $1-a_n$ the customers leave. So it is a birth and death process with the following rates:

$$\lambda_n=\lambda a_n \text{ and } \mu_n=\begin{cases} n\mu, & \text{if } 0\leq n\leq s;\\ s\mu, & \text{if } n\geq s.\end{cases}$$

$M/M/s$ queuing with balking II

Theorem 3.4

If $a_n \to 0$, then there exists stationary distribution.

Proof.

By (32), $\pi(n+1) = \frac{a_n\lambda}{s\mu}\pi(n)$ holds for $n \geq s$. There exists an N, s.t. if $n > N$, then $\frac{a_n\lambda}{s\mu} < \frac{1}{2}$. Thus for all $n > \max\{N, s\}$ we have $\pi(n+1) < \left(\frac{1}{2}\right)^{n-N}\pi(N)$. Thus $\sum_{n\geq 1}\pi(n) < \infty$. By Theorem 3.1 there exists stationary distribution. □

If $s = 1$ and $a_n = 1/(n+1)$, then $\boldsymbol{\pi} = \mathrm{Poi}(\lambda/\mu)$.

Branching processes

Example 3.5 (Branching processes)

In this example each individual dies with rate μ and gives birth to a new individual with rate λ and we start with one individual. So, the state space is $S = \{0, 1, 2, 3, \dots\}$ that is, the set of the non-negative integers and the rates are

$$q(n, n+1) = \lambda n \text{ and } q(n, n-1) = \mu n \quad \text{if } n \geq 1.$$

We start with one individual.

Branching process with imigration

Example 3.6 (Branching process with immigration)

Let us assume, that every individual dies with rate μ, and new children are born with rate λ as above. Furthermore, there are incoming members with rate ν. Then

$$q(n, n+1) = n\lambda + \nu \text{ and } q(n, n-1) = n\mu.$$

Example: fast growing population model

Example 3.7

Let

$$\mu_n \equiv 0 \text{ and } \lambda_n = \lambda \cdot n^2, \ \lambda > 0$$

In this case the population growths very fast and it becomes infinite in finite time. We study this phenomenon in the next few slides:

Pure birth processes

Definition 3.8

Pure birth processes are such birth and death processes, that $\forall n : \mu_n = 0$.

Theorem 3.9

(a) If $\sum_{n=0}^{\infty}\frac{1}{\lambda_n} = \infty$, then $\sum_{j=i}^{\infty} p_t(i,j) = 1$, $\forall t \geq 0$.

(b) If $\sum_{n=0}^{\infty}\frac{1}{\lambda_n} < \infty$, then $\sum_{j=i}^{\infty} p_t(i,j) < 1$, $\forall t > 0$.

Pure birth processes (cont.)

Explanation: Let X_n be the waiting time for jump from n to $n+1$. We have learned that $X_n \sim \mathrm{Exp}(\lambda_n)$. The r.v. $\{X_n\}_{n=1}^{\infty}$ are independent and $\mathbb{E}[X_n] = 1/\lambda_n$. The time of the n-th jump is $T_n := \sum_{i=1}^{n} X_n$. Then $\mathbb{E}[T_n] = \sum_{n=1}^{N} 1/\lambda_n$. When $\sum_{n=1}^{\infty} 1/\lambda_n = \infty$, then from Kolmogorov's Three-Series Theorem (next slide) $T_n \to \infty$ almost surely, but if $\sum_{n=1}^{\infty} 1/\lambda_n < \infty$, then $\{T_n\}_{n=1}^{\infty}$ is bounded, so $\exists T < \infty$, that the population

Pure birth processes (cont.)

growths to infinity before time T. Now we explain this with details:

Pure birth processes (cont.)

Theorem 3.10 (Kolmogorov's Three-Series Theorem)

Let $X_1, X_2, \dots$ be independent r.v.. The Random series $\sum_{i=1}^{\infty} X_i$ converges a.s. iff all of the following three series are convergent. If at least one of these series is not convergent, then $\sum_{i=1}^{\infty} X_i$ is divergent a.s..

1. $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > 1) < \infty$.
2. $\sum_{n=1}^{\infty} \mathbb{E}\left[X_n \cdot \mathbb{1}_{\{|X_n| \leq 1\}}\right]$ *is convergent.*
3. $\sum_{n=1}^{\infty} \mathrm{Var}\left(X_n \cdot \mathbb{1}_{\{|X_n| \leq 1\}}\right) < \infty$.

Explosion in the pure birth process

Proof of part (a) of Theorem 3.9

Let us assume, that $\sum_{n=1}^{\infty} 1/\lambda_n = \infty$. Let

$$X_n \sim \text{Exp}(\lambda_n),\ Y_n = X_n \cdot \mathbb{1}_{X_n \le 1},\ Z_n = X_n \cdot \mathbb{1}_{X_n > 1}.$$

Using that $\mathbb{E}[X_n] = 1/\lambda_n$

(35) $$\mathbb{E}[Y_n] = 1/\lambda_n - \mathbb{E}[Z_n].$$

Now we compute $\mathbb{E}[Z_n]$:

Explosion in the pure birth process (cont.)

(36) $$\begin{aligned}\mathbb{E}[Z_n] &= \int_0^\infty \mathbb{P}(Z_n \ge t)\,dt \\ &= \int_0^1 \mathbb{P}(Z_n \ge t)\,dt + \int_1^\infty \mathbb{P}(Z_n \ge t)\,dt \\ &= e^{-\lambda_n} + \frac{e^{-\lambda_n}}{\lambda_n}.\end{aligned}$$

From here and formula (35):

(37) $$\mathbb{E}[Y_n] = \frac{1 - e^{-\lambda_n}}{\lambda_n} - e^{-\lambda_n}.$$

Explosion in the pure birth process (cont.)

First observe that the first sum in Kolmogorov's Three-Series Theorem is

(38) $$\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > 1) = \sum_{n=1}^{\infty} e^{-\lambda_n}.$$

Assume that

(39) $$\sum_{n=1}^{\infty} e^{-\lambda_n} = \infty$$

Then $\sum_{n=1}^{\infty} X_n$ is divergent almost surely by Kolmogorov's Three-Series Theorem. Observe that (39) can happen only if $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$.

Explosion in the pure birth process (cont.)

Now assume that

(40) $$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty \text{ but } \sum_{n=1}^{\infty} e^{-\lambda_n} < \infty.$$

Then it follows from (37) that the second series in Kolmogorov's Three-Series Theorem is divergent so in this case also $\sum_{n=1}^{\infty} X_n$ is divergent almost surely. This and the argument on the previous slide together implies that part (a) of Theorem 3.9 holds. Now to prove part (b), we assume that

(41) $$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty.$$

Explosion in the pure birth process (cont.)

Then clearly $\sum_{n=1}^{\infty} e^{-\lambda_n} < \infty$, so the first and the second series are convergent in the Kolmogorov's Three-Series Theorem. Now we prove that the third series is also convergent. For this, we observe that

(42)
$$Var(Y_n) \le Var(X_n) + \mathbb{E}[Y_n]\,\mathbb{E}[Z_n] = \frac{1}{\lambda_n^2} + \mathbb{E}[Y_n]\,\mathbb{E}[Z_n].$$

The fact that the right hand side is summable follows from (41), (37) and (36). □

Embedded MC

Recall that on slide (30) we introduced the routing matrix $r(i,j) := q(i,j)/\lambda_i$, if $i \ne j$ and $r(i,i) = 0$, where $\lambda_i = \sum_{j \ne i} q(i,j)$. This is a stochastic matrix which determines a discrete-time MC, called embedded MC. Let

$$V_k := \min\{t \ge 0 : X_t = k\}$$

and

$$T_k := \min\{t > 0 : X_t = k \text{ and } \exists s < t,\ X_s \ne k\}.$$

Embedded MC (cont.)

Example 3.11 ($M/M/1$ queuing)

$q(i, i+1) = \lambda$, if $i \ge 0$ and $q(i, i-1) = \mu$ if $i \ge 1$. The embedded MC: $r(0,1) = 1$ and

$$r(i, i+1) = \frac{\lambda}{\lambda + \mu},\ i \ge 1,\ r(i, i-1) = \frac{\mu}{\lambda + \mu},\ i \ge 1.$$

It is a random walk with partly reflective bounds. So, as seen

- is positive recurrent, if $\lambda < \mu$.
- is null recurrent, if $\lambda = \mu$.
- is transient, if $\lambda > \mu$.

Example 3.12 (Branching processes)

$q(i, i+1) = \lambda i$ and $q(i, i-1) = \mu i$. State zero is an absorbing one, but for $i \ge 1$:

$$r(i, i+1) = \frac{\lambda}{\lambda + \mu} \text{ and } r(i, i-1) = \frac{\mu}{\lambda + \mu}.$$

If $\lambda < \mu$, then absorbing happens at zero almost surely, but

(43) $$\text{if } \lambda > \mu \text{ then } \rho := \mathbb{P}_1(T_0 < \infty) = \tfrac{\mu}{\lambda} < 1.$$

So for $x \ge 1$: $\mathbb{P}_x(T_0 < \infty) = \left(\frac{\mu}{\lambda}\right)^x$.

Proving this:

$$\rho = \frac{\mu}{\lambda+\mu}\cdot 1 + \frac{\lambda}{\lambda+\mu}\cdot \rho^2.$$

So when the chain leaves state 1, then either it goes to 0 and then dies out with probability 1 or goes to 2 and then branches of both children should die out, which has probability ρ^2. From here $\rho = \frac{\mu}{\lambda}$. The last statement comes from that if we want to go from x to 0, then first we must reach $x-1$, $x-2$.

Exit distributions with embedded MC

Question: if there are some absorbing states (we denote it by A), then what is the probability that the chain gets to $a \in A$?
Let $A \subset S$ and $a \in A$.

$$V_A := \min\{t \geq 0 : X_t \in A\}, h(i) := \mathbb{P}_i(X_{V_A} = a).$$

Then if $b \in A \setminus \{a\}$:

$$h(a) = 1, \;\; h(b) = 0.$$

Exit distributions with embedded MC (cont.)

So we only need to specify $h(i)$ for $\forall i \notin A$. To do this, we must see, that: $\forall i \notin A$:

$$h(i) = \sum_{j\neq i} \frac{q(i,j)}{\lambda_i} \cdot h(j) \text{ where } \lambda_i = \sum_{j \neq i} q(i,j). \tag{44}$$

Hence $\forall i \notin A$:

$$\boxed{\sum_j q(i,j)h(j) = 0}, \text{ where } q(i,i) = -\lambda_i. \tag{45}$$

Exit distributions with embedded MC (cont.)

So for all $i \notin A$ we have an equation, from what we can determine $h(i)$, $i \notin A$.

Expected time of exit: theory

We write the analogue of (45) for the expected exit time.

$$V_A := \min\{t \geq 0 : X_t \in A\}, g(i) := \mathbb{E}_i[V_A].$$

So $g(i) = 0$, if $i \in A$. As usual

$$\lambda_i = \sum_{j \neq i} q(i,j) \text{ and } r(i,j) := \frac{q(i,j)}{\lambda_i}.$$

We know, that the chain in the i^{th} state remains for time $\mathrm{Exp}(\lambda_i)$ and then jumps into state $j \neq i$ with probability

Expected time of exit: theory (cont.)

$r(i,j)$. Using the fact that $\mathbb{E}[\mathrm{Exp}(\lambda_i)] = 1/\lambda_i$ we get, that:

$$i \notin A: \quad g(i) = \frac{1}{\lambda_i} + \sum_{j \neq i} \frac{q(i,j)}{\lambda_i} g(j).$$

By rearranging it and using that $q(i,i) = -\lambda_i$:

$$\boxed{i \notin A: \quad \sum_j q(i,j)g(j) = -1.} \tag{46}$$

If S is finite, these are $\#S - \#A$ equations for $\#S - \#A$ unknowns $g(i)$, $i \notin A$.

Expected time of exit: At the barber's

Recall Barbershop Example: Customers are served by rate 3 and they arrive by rate 2, but they leave, if both chairs are occupied on: In other words

$$q(i, i-1) = 3 \text{ if } i = 1,2,3$$
$$q(i, i+1) = 2 \text{ if } i = 0,1,2.$$

Transition matrix for the embedded MC:

	0	**1**	**2**	**3**
0	0	1	0	0
1	3/5	0	2/5	0
2	0	3/5	0	2/5
3	0	0	1	0

Expected time of exit: At the barber's (cont.)

So now $A = \{0\}$, $g(0) = 0$, $g(i) = \mathbb{E}_i[V_0]$. Let

$$\mathbf{g} := \begin{bmatrix} g(1) \\ g(2) \\ g(3) \end{bmatrix} \text{ and } \mathbb{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Then equation system (46) is equivalent with:

$$\boxed{\widetilde{Q} \cdot \mathbf{g} = -\mathbb{1}}, \tag{47}$$

Expected time of exit: At the barber's (cont.)

where $\widetilde{Q}$ is the restriction of matrix Q for columns belonging to $S \setminus A$ (now those who are not 0). This equivalence comes from that know that$g(i) = 0$, if $i \in A$. So columns $i \in A$ add zero to all equations.

$$\widetilde{Q} = \begin{bmatrix} -5 & 2 & 0 \\ 3 & -5 & 2 \\ 0 & 3 & -3 \end{bmatrix}$$

Expected time of exit: At the barber's (cont.)

and

$$-(\widetilde{Q})^{-1} = \begin{bmatrix} 1/3 & 2/9 & 4/27 \\ 1/3 & 5/9 & 10/27 \\ 1/3 & 5/9 & 19/27 \end{bmatrix}$$

From formula (45):

$$\mathbf{g} = -(\widetilde{Q})^{-1} \cdot \mathbb{1} = \begin{bmatrix} 19/27 \\ 34/27 \\ 43/27 \end{bmatrix},$$

Expected time of exit: At the barber's (cont.)

so i^{th} element of $\mathbf{g}$ is given by i^{th} row sum of matrix $-(\widetilde{Q})^{-1}$.

Expected time of exit: When can the kindergarten teacher go home?

Example: In a nursury school at closing time parents haven't come for three children Anne (A), Bella (B) and Charlie (C). Kindergarten teacher stays as long as all the children go home. Parents phoned that they would arrive by time $\mathrm{Exp}(1), \mathrm{Exp}(2)$ and $\mathrm{Exp}(3)$ after close time. (So expectedly they will fetch their child 1, $1/2$ and $1/3$ hours after close time, independently of each other.) Question is when can the kindergarten teacher go home?

Expected time of exit: When can the kindergarten teacher go home? (cont.)

Solution: States of MC are the names of remaining children and $\emptyset$ when no child is left:

Q	**ABC**	**AB**	**AC**	**BC**	**A**	**B**	**C**	$\emptyset$
ABC	−6	3	2	1	0	0	0	0
AB	0	−3	0	0	2	1	0	0
AC	0	0	−4	0	3	0	1	0
BC	0	0	0	−5	0	3	2	0
A	0	0	0	0	−1	0	0	1
B	0	0	0	0	0	−2	0	2
C	0	0	0	0	0	0	−3	3
$\emptyset$	0	0	0	0	0	0	0	0

Expected time of exit: When can the kindergarten teacher go home? (cont.)

Let us use the notation and method of the previous example:

Now $A := \emptyset$. So $\widetilde{Q}$ is the above matrix restricted to the first 7 rows and columns. Then the first row vector of matrix $-\left(\widetilde{Q}\right)^{-1}$:

$$(1/6, 1/6, 1/2, 1/30, 7/12, 2/15, 1/20).$$

Expected time of exit: When can the kindergarten teacher go home? (cont.)

Sum of them is: 63/60. So kindergarten teacher can go home 63 minutes after close time.
Note: This can be seen from the fact, that for every number a, b, c:

$$\begin{aligned} \max\{a,b,c\} =& a+b+c-\min\{a,b\}-\min\{a,c\} \\ & -\min\{b,c\}+\min\{a,b,c\}. \end{aligned}$$

Expected time of exit: When can the kindergarten teacher go home? (cont.)

We can use this and part (d2) of MC III slide 7 , if $T_i = \mathrm{Exp}(\lambda_i)$, $i = 1,2,3$ are independent for determining $\max\{T_1, T_2, T_3\}$.

$M/M/1$ queuing again

- $q(n, n+1) = \lambda$, if $n \geq 0$,
- $q(n, n-1) = \mu$ if $n \geq 1$.

We assume, that

(48) $$\lambda < \mu.$$

As we have seen, this is a birth and death process in which
$$\lambda_n = \lambda \text{ and } \mu_n = \mu.$$

$M/M/1$ queuing again (cont.)

Because of condition (48) we can use Theorem 3.1.
From here:

(49) $$\pi(n) = \left(\frac{\lambda}{\mu}\right)^n \cdot \pi(0).$$

For this to give a measure, we need: $\pi(0) := 1 - \lambda/\mu$. So

(50) $$\pi(n) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n, \quad n \geq 0.$$

Let us assume, that the system is in a stationary state.
Then let

$M/M/1$ queuing again (cont.)

- X_s the number of customers at time s in the system.
- Q be the length of the queue,
- T_Q be the time spent in the queue, $W_Q = \mathbb{E}[T_Q]$ and $W = W_Q + \mathbb{E}[\text{serving time}]$
- L the long time average a customer sepends in the system. $L = \lim_{t\to\infty} \frac{1}{t}\int_0^\infty X_s$.

$M/M/1$ queuing again (cont.)

- λ_a the long time average rate at which arriving customers join the system. $\lambda_a = \lim_{t\to\infty} \frac{N_a(t)}{t}$, $N_a(t)$the number of customers who joined the system befor time t.

Obviously

(51) $$\mathbb{P}(T_Q = 0) = \pi(0) = 1 - \frac{\lambda}{\mu}.$$

$M/M/1$ queuing again (cont.)

Let $f(x)$ be the conditional density function of T_Q on $(0, \infty)$ assuming that $T_Q > 0$. Note that because of (51): $\mathbb{P}(T_Q = 0) > 0$.
Assuming, that at the arrival of a customer there are already n customers in the system, (whose probability if given in (50)). Conditioned on this, the conditional density function of T_Q is $Gamma(n, \mu)$. Using this we get:

(52) $$f(x) = \frac{\mu}{\lambda} \cdot \sum_{n=1}^{\infty} \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n e^{-\mu x} \frac{\mu^n x^{n-1}}{(n-1)!}.$$

$M/M/1$ queuing again (cont.)

After trivial rearrangement we get, that

(53) $$f(x) = (\mu - \lambda)e^{-(\mu-\lambda)x}.$$

We have proven by this, that

$M/M/1$ queuing again (cont.)

Lemma 4.1

- *The conditional distribution of T_Q for $T_Q > 0$ is* $\mathrm{Exp}(\mu - \lambda)$.
- $W_Q = \mathbb{E}[T_Q] = \frac{\lambda}{\mu}\frac{1}{\mu-\lambda}$.
- $\mathbb{E}[W] = W_Q + \frac{1}{\mu} = \frac{\lambda}{\mu}\frac{1}{\mu-\lambda} + \frac{1}{\mu} = \frac{1}{\mu-\lambda}$.
- $L = \frac{1}{1-\lambda/\mu} - 1 = \frac{\lambda}{\mu-\lambda}$.

$M/M/1$ queue finite waiting room

- There is one server and serving a customer takes time $\mathrm{Exp}(\mu)$.
- Customers arrive by $\mathrm{Poisson}(\lambda)$.
- In the waiting room during 1 serving there is place for $N-1$ waiting customers. Customers, who arrive when there is no empty seat, leave at once and will never return.

Lemma 4.2

- *Let X_t be a MC, for which there exists stationary distribution π and it satisfies detailed balance condition. Infinitesimal generator of chain X_t is Q.*
- *Let $A \subset S$ and Y_t be the restriction of X_t to A. In other words, Y_t's infinitesimal generator is $\widetilde{Q}$, where for distinct x, y:*

$$\tilde{q}(x,y) = \begin{cases} q(x,y), & \text{if } x, y \in A,\ x \neq y; \\ 0, & \text{otherwise.} \end{cases}$$

- *Let $C := \sum_{x \in A} \pi(x)$.*

Then $\nu := \boldsymbol{\pi}/C$ is the stationary state of Y_t.

$M/M/1$ queue with finite waiting room III

Proof.
Using, that $\boldsymbol{\pi}$ satisfies detailed balance condition, it clearly comes, that ν also satisfies it, so ν is stationary distribution for chain Y_t. □

$M/M/1$ queue with finite waiting room IV

From here and from (50) comes, that for the $M/M/1$ queue with waiting room of space N introduced above, the stationary state:

$$\pi(n) := \frac{1-\lambda/\mu}{1-(\lambda/\mu)^{N+1}} \left(\frac{\lambda}{\mu}\right)^n \text{ if } 0 \leq n \leq N. \tag{54}$$

With finite state space it is also true, if $\lambda > \mu$. It is only false, if $\lambda = \mu$. In this case:

$$\pi(n) = \frac{1}{N+1} \text{ if } 0 \leq n \leq N.$$

At the barber's for the last time

Review: We have introduced barber shopb example on slide 3 and on slide 60 we have computed its stationary distribution:

$$\boldsymbol{\pi}^T = \left(\tfrac{27}{65}, \tfrac{18}{65}, \tfrac{12}{65}, \tfrac{8}{65}\right),$$

which is the same as what comes from formula (54).

On slide 87 we have computed, that if there are $i = 1, 2, 3$ customers at the barber's, then how much time should we wait till no costumer are in the barber shop.

At the barber's for the last time (cont.)

Clearly,

$$L = 1 \cdot \frac{18}{65} + 2 \cdot \frac{12}{65} + 3 \cdot \frac{8}{65} = \frac{66}{65}. \tag{55}$$

Let λ_a be the long run rate of customers at the barber's who have their haircut (who don't leave) because of the occupied waiting room. That is let $N_a(t)$ be the number of customers who have arrived before time t and did not leave immediately because of the busy waiting room but who stayed at the barber shop and eventually got served by the barber. More precisely: $\lambda_a := \lim_{n\to\infty} \frac{N_a(t)}{t}$.

Finding λ_a: We know, that customers arrive by $\mathrm{Poisson}(2)$ process. This means that during a time interval of length Δt, the probability that a customer enters into the barbershop is $2 \cdot \Delta t$ (plus $o(\Delta t)$ what we will suppress below for the sake of simpler presentation). But if there are already 3 customers, the newly arrived customer leaves. This results, that with probability $2 \cdot \Delta t \cdot \pi(3)$ a potential customers is lost. **We have to subtract this**. So, during a time interval of length Δt there will be a new costumer who enters the service and who remains inside the system with probability $2(1-\pi(3))\Delta t$. Hence

$$\lambda_a = 2(1-\pi(3)) = \frac{114}{65}. \tag{56}$$

Little's Formula

The following formula holds in general for GI/G/1 (general input /general service/ one server) queues.

Theorem 4.3 (Little's Formula)

$$L = W \cdot \lambda_a.$$

The sketch of the proof is available in Durrett's book p. 107.
Using Little's formula, (55) and (56) we get

$$W = \frac{66/65}{114/65} = \frac{33}{57} = 0.579 \text{ hours} = 34.74 \text{ mins}$$

We can also compute this, as when I get inside, there can be $i = 0, 1, 2, 3$ customers inside. In the case of $i = 3$ I go home. In the case of $i = 0, 1, 2$ I spend time $(i+1) \cdot \frac{1}{3}$ inside (because people before me and I also have a haircut in time $\mathrm{Exp}(3)$, which requires $1/3$ hours.) Regarding these, expected value of my time W spent inside:

$$\begin{aligned} W &= \frac{1}{1-\pi(3)} \left[\pi(0) \cdot \frac{1}{3} + \pi(1) \cdot \frac{2}{3} + \pi(2) \cdot 1 \right] \\ &= \frac{33}{57}. \end{aligned}$$

So, the expectation of my waiting time in the queue:

$$W_Q = W - \frac{1}{3} = \frac{14}{57} = 0.2456 \text{ hours} = 14.736 \text{ mins.}$$

$M/M/s$ queue

We have introduced $M/M/s$ queue in slide 62

- In a bank, customers are being served by s servers, and they are waiting in one queue if there are more customers than servers.
- Customers arrive by a $\mathrm{Poisson}(\lambda)$ process.
- Serving times are independent times of $\mathrm{Exp}(\mu)$.

$M/M/s$ queue (cont.)

Now, $S = 0, 1, 2, \dots$ is the number of customers in the bank. As we have seen, this is a birth and death process with the following rates:

$$q(n, n+1) = \lambda, \quad n \geq 0.$$

and

$$q(n, n-1) = \begin{cases} n\mu, & \text{if } 1 \leq n \leq s; \\ s\mu, & \text{if } n \geq s. \end{cases}$$

$M/M/s$ queue (cont.)

Lemma 4.4

If $\lambda < s\mu$, then there exists a π stationary state, which satisfies detailed balance condition.

Proof If we write down detailed balance condition, we get the following conditions:

$$\begin{aligned} \lambda\pi(j-1) &= \mu j \pi(j) && \text{if } j \leq s \\ \lambda\pi(j-1) &= \mu s \pi(j) && \text{if } j \geq s \end{aligned}$$

$M/M/s$ queue (cont.)

From here

$$\pi(k) = \begin{cases} \frac{c}{k!} \left(\frac{\lambda}{\mu}\right)^k, & \text{if } k \leq s; \\ \frac{c}{s! s^{k-s}} \left(\frac{\lambda}{\mu}\right)^k, & \text{if } k \geq s. \end{cases} \tag{57}$$

where we would like to choose c s.t. π be stationary measure. It is possible, if $\lambda < s\mu$. □

$M/M/s$ queue (cont.)

Lemma 4.5

If $\lambda > s\mu$, chain $M/M/s$ is transient., If $\lambda < s\mu$, chain $M/M/s$ is recurrent.

$M/M/s$ queue (cont.)

Proof.

If $\lambda > s\mu$, then the $M/M/1$ queue with serving time $n\mu$ is obviously transient. This is from that for the $M/M/1$ queue there is stationary state π (so it is recurrent) if $\lambda < \mu$. The $M/M/s$ queue with serving time μ is less efficient, so it is also transient. The other direction is from the existence of stationary state. □

$M/M/s$ queue (cont.)

Example 4.6

Compute the stationary measure for the

(a) $M/M/s$ queue, if $\mu = 1$, $\lambda = 2$, $s = 3$,

(b) $M/M/1$ queue, if $\mu = 3$, $\lambda = 2$, $s = 1$.

And compare the chains by this in view of efficiency.

$M/M/s$ queue (cont.)

Solution (a):
$\sum_{k=2}^{\infty} \pi(k) = \frac{c}{2} \cdot 2^2 \cdot \sum_{j=0}^{\infty} (2/3)^j = 6c$, $\pi(0) = c$,
$\pi(1) = \frac{\lambda}{\mu} c = 2c$. In other words $9c = 1$, from which $c = 1/9$. So
(58)

$$\pi(0) = \frac{1}{9}, \quad \pi(1) = \frac{2}{9} \text{ and } \pi(k) = \frac{2}{9}\left(\frac{2}{3}\right)^k \text{ if } k \geq 3.$$

Solution (b): from formula (50):

$$\pi(n) = \frac{1}{3} \cdot \left(\frac{2}{3}\right)^n, \quad n \geq 0,$$

$M/M/s$ queue (cont.)

So $\pi(0) = \frac{1}{3}$ and $\pi(1) = \frac{2}{9}$.

Examples

[1] R. Durrett
Essentials of Stochastic Processes, Second edition
Springer, 2012.
http://www.math.duke.edu/ rtd/EOSP/EOSP2E

[2] R. Durrett
Probability Theory with examples, Second edition
Duxbury Press, 1996

[3] S. Karlin, H.M. Taylor
A first course in stochastic processes
Academic Press,2014

[4] S. Karlin, H.M. Taylor
A second course in stochastic processes
Academic Press, 1981

[5] G. Lawler
Intoduction to Stochastic Processes
Chapmann & Hall 1995.

[6] D.A. Levin, Y. Peres,E.L. Wilmer
Markov chains and mixing times
American Mathematical Society, 2009.

[7] Piet Van Mieghem
The Poisson process
http://www.nas.its.tudelft.nl/people/Piet/CUPbookChapters/PACUP_Poisson.pdf