Continuous Time Markov Chains

Birth-Death processes

Birth-death processes

- Birth-death processes are CTMC where the transitions are of the form

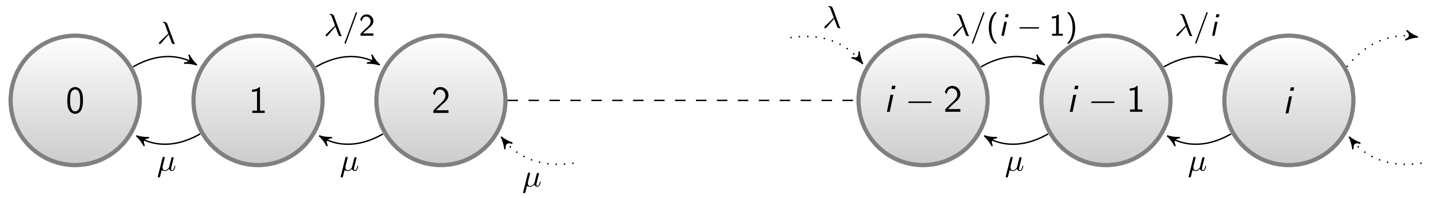
$$i \longrightarrow i+1 \text{ or } i \longrightarrow i-1$$

- Notations :
  - From i to $i+1$: rate λ_i
  - From i to $i-1$: rate μ_i

- Example : Number of customers in M/M/1/K
  - Arrival rate : $\lambda_i = \lambda$ if $i \leq K-1$
  - Departure rate: $\mu_i = \mu$ if $i \geq 1$

It is very often a delicate task to solve balance equations. There is one particular class of continuous-time Markov chains for which it is easily feasible. This class is known under the name of birth and death processes.
A birth and death process is a continuous-time Markov chain where in state i, the only possible transitions are to state $i-1$ and $i+1$. Let's introduce a few notations : λ_i is the transition rate from i to $i+1$ and μ_i is the transition rate from i to $i-1$.
For instance, let's consider the M/M/1/K queue, so? the M/M/1 queue but with a finite buffer of size $K-1$. The arrival rate is λ if i is less than $K-1$ since in state K, the system is full and we cannot accommodate any more customers. The departure rate stays the same equal to μ.

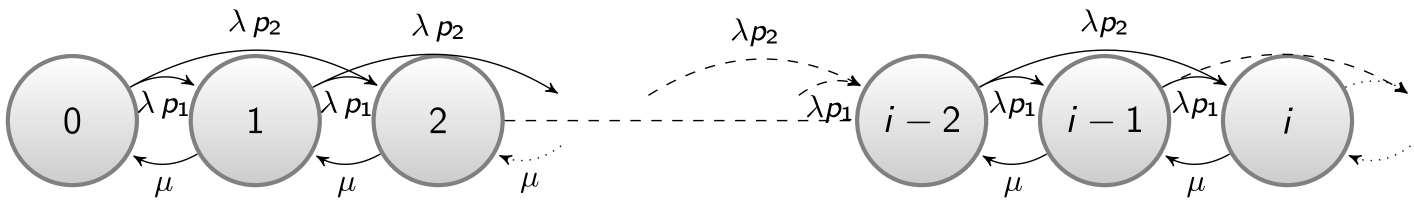
M/M/1 queue with balking



- M/M/1 queue
- If there are i customers, the arriving customer refuses to enter with probability $1 - 1/(i+1)$
- $\lambda_i = \lambda/(i+1)$
- $\mu_i = \mu$

Another interesting example is that of a queue with balking. For call centers, it is a key problem to assess the impact of the impatience of customers. One of the simplest models we can envision to represent this type of impatience is known as an M/M/1 queue with balking. If there are i clients in the system upon the arrival of a new customer, this customer refuses to enter the system with probability $1 - 1/(i+1)$. The probability of balking is there to model the fact that not all people react the same way to long queues.
This means that the arrival rate is no longer λ but only $\lambda/(i+1)$. The departure rate stays the same equal to μ.

Counter-example: M/M/1 with batch arrivals



- Arrivals: Poisson process of intensity λ
- Batch arrivals : one arrival = 1 or 2 customers with probability p_1, p_2

$$
\begin{array}{c c}
 & \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & \dots \end{array} \\
\begin{array}{c} 0 \\ 1 \\ 2 \\ \\ \vdots \end{array} &
\begin{pmatrix}
-\lambda & \lambda\, p_1 & \lambda\, p_2 & & & \\
\mu & -(\lambda+\mu) & \lambda\, p_1 & \lambda\, p_2 & & \\
 & \mu & -(\lambda+\mu) & \lambda\, p_1 & \lambda\, p_2 & \\
 & & & & & \\
 & \ddots & \ddots & \ddots & \ddots &
\end{pmatrix}
\end{array}
$$

However, not all continuous-time Markov chains are birth-death processes, even if they are representing queuing systems. For instance, we can have what are called batch arrivals. Imagine a desk at which a group of customers may arrive together, they create the same number of tasks as there are members of the group. Consider that the arrivals occur according to a Poisson process and that for the model to be simple, we may have 1 customer with probability p_1 and two customers with probability p_2. So, there may be three types of transitions: i to $i-1$, i to $i+1$ or i to $i+2$ This can be represented either by the transition rate matrix or by the transition diagram. The transition rate from i to $i+1$ is λp_1, and from i to $i+2$ is λp_2.

Local balance

- Flow from i to j = Flow from j to i

$$\pi_i \, q_{i,j} = \pi_j \, q_{j,i}$$

- For birth-death processes

$$\text{local balance} \iff \text{global balance equations}$$

The global balance equations are equivalent to the equality of the global input flow into state i and its global output flow. For local balance equations, we only consider local equilibrium between two states: the local balance equations are satisfied if for any pair of states i and j the flow from i to j is equal to the flow from j to i.
There is no reason why the global balance equations should imply the local balance equations. But for birth death processes, they are equivalent. This means that if we can solve the local balance equations then we also find a solution to the global balance equations. The local balance equations are far easier to handle since they only relate to pairs of states.

Triangular system

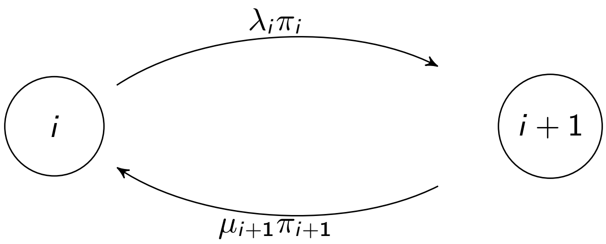
$$\begin{pmatrix}\pi_0 & \pi_1 & \dots \end{pmatrix}\begin{pmatrix} -\lambda_0 & \lambda_0 & & & \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & & \\ & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \\ & & \ddots & \ddots & \ddots \end{pmatrix} = 0$$

$$\iff \begin{cases} -\lambda_0\,\pi_0 + \mu_1\,\pi_1 & = 0 \\ \lambda_0\pi_0 - \mu_1\,\pi_1 - \lambda_1\,\pi_1 + \mu_2\pi_2 & = 0 \\ \lambda_1\,\pi_1 - \mu_2\,\pi_2 - \lambda_2\,\pi_2 + \mu_3\,\pi_3 & = 0 \\ \vdots & \vdots \end{cases}$$

The main thing to notice is that the transition rate matrix of a birth-death process is always tri-diagonal. We have non-zero coefficients on the diagonal (red coefficients), the subdiagonal (in blue) and the superdiagonal (in orange), and that's all. This is of course due to the fact that the only possible transitions are from i to $i+1$ or $i-1$. When we write the system $\pi Q = 0$ we see that at each line, only π_i, π_{i-1} and π_{i+1} are involved.

Local balance equations

$$\begin{cases} \boxed{-\lambda_0\,\pi_0 + \mu_1\,\pi_1} - \lambda_0\,\pi_0 + \mu_1\,\pi_1 & = 0 \\ \lambda_0\,\pi_0 - \mu_1\,\pi_1 \boxed{\lambda_0\pi_0 - \mu_1\,\pi_1} - \lambda_1\,\pi_1 + \mu_2\pi_2 \boxed{-\lambda_1\,\pi_1 + \mu_2\pi_2} & = 0 \\ \lambda_1\,\pi_1 - \mu_2\,\pi_2 \boxed{\lambda_1\,\pi_1 - \mu_2\,\pi_2} - \lambda_2\,\pi_2 + \mu_3\,\pi_3 & = 0 \\ \vdots & \vdots \end{cases}$$



We can go even further and greatly simplify this system. We see in line 1 that $-\lambda\pi_0 + \mu_1\pi_1$ is zero but we notice that in line 2, we retrieve this very same quantity up to a change of sign, so we can replace this block in line 2 by 0. Line 2 now becomes $-\lambda_1\pi_1 + \mu_2\pi_2 = 0$, but once again this block appears almost as it is in line 3, so we can again remove that block from line 3. And we may proceed likewise forever and we have a series of equations involving only π_i and π_{i+1}. A small picture shows that these equations mean exactly that the local balance equations are satisfied: the flow between i and $i+1$ is equal to the flow between $i+1$ and i.

Recursion

$$\begin{cases} \lambda_0 \, \pi_0 & = \mu_1 \, \pi_1 \\ \lambda_1 \, \pi_1 & = \mu_2 \, \pi_2 \\ \lambda_2 \, \pi_2 & = \mu_3 \, \pi_3 \\ \vdots & = \vdots \end{cases}$$

Hence

$$\begin{aligned} \pi_1 &= \frac{\lambda_0}{\mu_1} \, \pi_0 \\ \pi_2 &= \frac{\lambda_1}{\mu_2} \, \pi_1 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} \, \pi_0 \\ \ldots \end{aligned}$$

Now we can solve the balance equations recursively. Rewrite the system by putting π_i and π_{i+1} on the two sides of the equal sign. From line 1, we get π_1 as a function of π_0. From line 2, we get π_2 as a function of π_1 and then combining with the previous identity, we obtain π_2 as a function of π_0. And we can proceed likewise forever.

Stationary distribution

For birth-death processes,

$$\boxed{\pi_i = \pi_0 \, \frac{\lambda_0 \ldots \lambda_{i-1}}{\mu_1 \ldots \mu_i} \text{ for } i \geq 1}$$

with

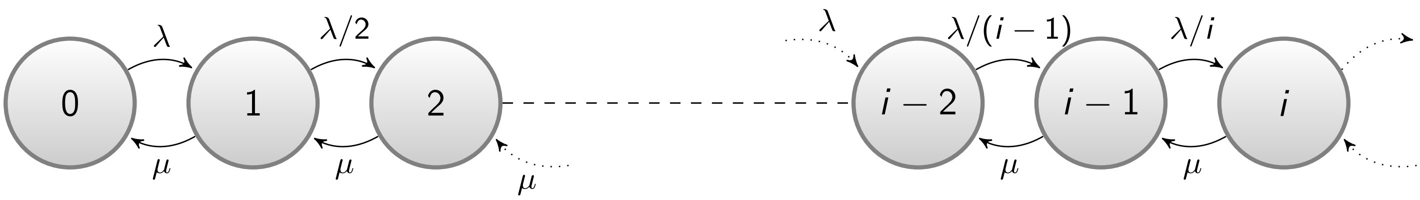
$$\begin{array}{ll} \pi_0 = \frac{1}{G} & \text{with} \quad G = 1 + \sum_{i=1}^{\infty} \frac{\lambda_0 \ldots \lambda_{i-1}}{\mu_1 \ldots \mu_i} \\ \text{since} & \sum_{i=0}^{\infty} \pi_i = 1 \end{array}$$

Stability condition:

$$G = 1 + \sum_{i=1}^{\infty} \frac{\lambda_0 \ldots \lambda_{i-1}}{\mu_1 \ldots \mu_i} < \infty$$

The result is that we have a rather simple expression of π_i which is given by $\lambda_0 \times \lambda_1 ... \times \lambda_i$ divided by the analog products of the μ_i. Since we have an infinite state space, we have a stability condition which only means that we can find π_0 such that the sum of all π_i is equal to 1.

M/M/1 with balking (continued)



- $\lambda_i = \lambda/(i+1)$, $\mu_i = \mu$, hence

$$\pi_i = \pi_0 \, \frac{\lambda}{\mu} \, \frac{\lambda}{2\mu} \, \ldots \, \frac{\lambda}{i\mu} = \pi_0 \, \frac{\rho^i}{i!}$$

  with $\rho = \lambda/\mu$ and $i! = i(i-1)\ldots 2$
- Using the normalizing condition

$$\pi_i = e^{-\rho} \, \frac{\rho^i}{i!}$$

  since $e^\rho = 1 + \rho + \rho^2/2 + \rho^3/3! + \ldots$

We may apply these computations to our model of M/M/1 with balking.
Recall that $\lambda_i = \lambda/(i+1)$ and $\mu_i = \mu$. Using the previous result, we have that π_i is proportional to $\frac{\lambda}{\mu} \times \frac{?}{2?}$.... This is equal to $\rho^i/i!$ with, as usual, $\rho = \lambda/\mu$. It is known that $1 + \rho + \rho^2/2 + \rho^3/3! + ...$ is equal to e^ρ, so the normalizing constant should be $e^{-\rho}$. This means that the stationary distribution of the M/M/1 queue with balking is a Poisson distribution of parameter ρ, whatever the value of ρ.
There is no condition on ρ to ensure stability, as opposed to the standard M/M/1 queue. Actually, when the number of customers is high, the rejection rate is close to 1 and thus very few customers are likely to enter the system. There is a sort of feedback which prevents the system from exploding.