Review from last time:

- Continuous-time discrete state space Markov chains.
- The generator matrix Q, consisting of rates.
- Exponentially distributed holding times.
- Connection with $\tilde{P}$, the embedded discrete-time chain.
- The matrix transition function $P(t)$. $P'(t) = QP(t) = P(t)Q$.
- The exponential matrix e^A for a square matrix A, and its computation.
- $P(t) = e^{tQ}$.

MVE550 2022 Lecture 11
Dobrow Sections 7.4 - 7.7
Absorbing states. Time Reversibility. Queueing theory. Poisson subordination.

Petter Mostad

Chalmers University

December 6, 2022

Limiting and stationary distributions

- A probability vector v represents a *limiting distribution* if, for all states i and j,
$$\lim_{t \to \infty} P_{ij}(t) = v_j.$$
- A probability vector v represents a *stationary distribution*, if, for all $t \geq 0$,
$$v = vP(t)$$
- Note: This happens if and only if $0 = vQ$.
- A limiting distribution is a stationary distribution but a stationary distribution is not necessarily a limiting distribution.
- A continuous-time Markov chain is *irreducible* if for all i and j there exists a $t > 0$ such that $P_{ij}(t) > 0$.
- However, periodic continuous-time Markov chains do not exist: If $P_{ij}(t) > 0$ for some $t > 0$ then $P_{ij}(t) > 0$ for all $t > 0$.

The fundamental limit theorem (for cont. time M. chains)

- An absorbing communication class is one where there is zero probability (i.e., zero rate) of leaving it to other communication classes.
- For a finite-state continuous-time Markov chain (with finite holding time parameters) there are two possibilities:
  - The process is irreducible, and $P_{ij}(t) > 0$ for all $t > 0$ and all i, j.
  - The process contains one or more absorbing communication classes.
- **Fundamental Limit Theorem**: Let $\{X_t\}_{t \geq 0}$ be a finite, irreducible, continuous-time Markov chain with transition function $P(t)$. Then there exists a unique positive stationary distribution vector v which is also the limiting distribution.
- The limiting distribution of such a chain can be found as the unique v satisfying $vQ = 0$.

Absorbing states

- Assume $\{X_t\}_{t\geq 0}$ is a continuous-time Markov chain with k states. Assume the last state is absorbing and the rest are not. (They are then transient).
- We have that $q_k = 0$ and the entire last row must consist of zeros. We get
$$Q = \begin{bmatrix} V & * \\ \mathbf{0} & 0 \end{bmatrix}.$$
- Let F be the $(k-1)\times(k-1)$ matrix so that F_{ij} is the expected time spent in state j when the chain starts in i. We can shown that $F = -V^{-1}$ (see next page).
- Note that, if the chain starts in state i, the expected time until absorbtion is the sum of the i'th row of F. Thus the expected times until absorbtion are given by the matrix product $F1$ of F with a column of 1's.

Outline of proof (different from Dobrow's)

- Generally, define D as the matrix with $(1/q_1, \ldots, 1/q_k)$ along its diagonal, with all other entries zero. If there are no absorbing states

$$\tilde{P} = DQ + I$$

- Write A_- for a square matrix without its last row and column.
- If the last state is absorbing, so that $q_k = 0$, we get

$$\tilde{P}_- = D_- Q_- + I$$

- Let F' be the matrix where F'_{ij} is the expected *number of stays* in state j before absorbtion when starting in state i. As the lengths of stays and changes in states are independent, we get $F = F' D_-$.
- From the theory of Chapter 3, we have that $F' = (I - \tilde{P}_-)^{-1}$.
- We get

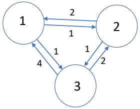
$$F = F' D_- = (I - \tilde{P}_-)^{-1} D_- = (-D_- Q_-)^{-1} D_- = (-Q_-)^{-1}.$$

Stationary distribution of the embedded chain

- The embedded chain of a continuous-time Markov chain: The discrete-time Markov chain where holding times are ignored.
- Stationary distributions for the embedded chain and for the continuous-time chain are generally not the same!
- However, there is a simple relationship: A probability vector π is a stationary distribution for a continuous-time Markov chain if and only if ψ is a stationary distribution for the embedded chain, where $\psi_j = C\pi_j q_j$ for a constant C making the entries of ψ add to 1.

Global Balance

- Let $v = (v_1, v_2, v_3)$ be the stationary distribution. At v the flow into a state must be equal to the flow out of that state.



- Get: $2v_1 = 2v_2 + 4v_3$, $3v_2 = 1v_1 + 2v_3$, and $6v_3 = 1v_1 + 1v_2$.
- Note that these are exactly the equations we get from $vQ = 0$:

$$(v_1, v_2, v_3) \begin{bmatrix} -2 & 1 & 1 \\ 2 & -3 & 1 \\ 4 & 2 & -6 \end{bmatrix} = 0$$

- This happens because $vQ = 0$ gives for each state j

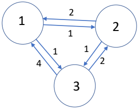
$$\sum_{i \neq j} v_i q_{ij} = v_j q_j$$

- These are called the global balance equations.
- Generalization: If A is a set of states, then the long term rates of movement *into* and *out of* A are the same:

$$\sum_{i \notin A} \sum_{j \in A} v_i q_{ij} = \sum_{i \notin A} \sum_{j \in A} v_j q_{ji}$$

Local balance and time reversibility

- A stronger condition: The flow between *every pair* of states is balanced. This is *not* true for all models!



- From model at right we now get the equations $1v_1 = 2v_2$, $1v_2 = 2v_3$, and $4v_3 = 1v_1$.
- An irreducible continuous-time Markov chain with stationary distribution v is said to be *time reversible* if for all i, j,

$$v_i q_{ij} = v_j q_{ji}$$

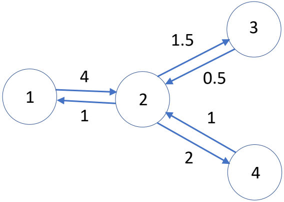
- This is called the *local balance* condition.
- Note: The rate of observed changes from i to j is the same as the rate of observed changes from j to i. Thus this is also called *time reversibility*.
- Note that (similar to discrete chains): If a probability vector v satisfies local balance condition, then v is a stationary distribution. (Easy to show).

Markov processes with transition graphs that are trees

- A *tree* is a graph that does not contain cycles.
- Assume the transition graph of an irreducible continuous-time Markov chain is a tree.
- In a tree, any edge between two states divides all states into two groups (each on each side of the edge). Thus, the flow must be balanced across each edge.
- It follows that the Markov chain must satisfy the local balance condition, i.e., be time reversible, i.e., $v_i q_{ij} = v_j q_{ji}$ for all i and j.
- More formally, this can be proved using the generalized global balance property.
- Note that the process can be time reversible even if the transition graph is not a tree.

Example

- Consider the continuous-time Markov chain with transition graph



- As the transition graph is a tree, the chain is necessarily time reversible. We can find the stationary distribution by considering the local balance equations:

$$4v_1 = 1v_2, \qquad 1.5v_2 = 0.5v_3, \qquad 2v_2 = 1v_4$$

- Together with the equation $v_1 + v_2 + v_3 + v_4 = 1$ we easily get the limiting distribution

$$v = \left(\frac{1}{25}, \frac{4}{25}, \frac{12}{25}, \frac{8}{25} \right)$$

Birth-and-death processes

- A birth-and-death process is a continuous-time Markov chain where the state space is the set of nonnegative integers and transitions only occur to neighbouring integers.
- The process is necessarily time-reversible, as the transition graph is a tree (in fact, a line).
- We denote the rate of *births* from i to $i+1$ with λ_i, and the rate of *deaths* from i to $i-1$ with μ_i.
- The generator matrix is

$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ \mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & -(\mu_2+\lambda_2) & \lambda_2 & \dots \\ 0 & 0 & \mu_3 & -(\mu_3+\lambda_3) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

- Provided $\sum_{k=1}^{\infty}\prod_{i=1}^{k}\frac{\lambda_{i-1}}{\mu_i} < \infty$, the unique stationary distribution is given by

$$v_k = v_0 \prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i} \text{for } k = 1, 2, \dots, \qquad v_0 = \left(1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\frac{\lambda_{i-1}}{\mu_i}\right)^{-1}$$

Example

- The simplest example of a birth-and-death process is one where all birth rates λ_i and all death rates μ_i are the same values λ and μ, respectively.
- We get that

$$
\begin{aligned}
v_k &= v_0 \prod_{i=1}^{k} \frac{\lambda}{\mu} = v_0 \left(\frac{\lambda}{\mu}\right)^k \\
v_0 &= \left(\sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^k\right)^{-1} = \frac{1}{1 + \frac{\lambda}{\mu} + (\frac{\lambda}{\mu})^2 + \dots} = \frac{1}{1/(1 - \frac{\lambda}{\mu})} = 1 - \frac{\lambda}{\mu}
\end{aligned}
$$

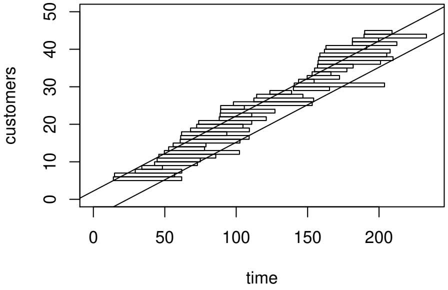
- We see that the limiting distribution is Geometric $\left(1 - \frac{\lambda}{\mu}\right)$.
- For example, the long-term average value of X_t will be

$$
\frac{\lambda/\mu}{1 - \lambda/\mu} = \frac{\lambda}{\mu - \lambda}
$$

Queueing theory

- Birth-and-death processes are special cases of queues.
- In the more general theory of queues:
  - The arrival process ("births") need not be a Poisson process, with exponentially distributed inter-arrival times.
  - The service times in the system need not be exponentially distributed.
  - There can be many other generalizations, such as how many servers there are, how the servers work, how the line works, etc.
- One can use notation $A/B/n$ where A denotes arrival process, B denotes service process, and n the number of servers.
- With this notation, our birth-and-death model above with constant birth and death rates is denoted $M/M/1$. (M means Markov).
- Our formulas above also give the limiting distribution for an $M/M/c$ queue, where there are c different servers.

Little's formula



The boxes represent customers arriving at a rate λ and staying for an average time W. The left line represents the average arrival times of customers: It has slope λ. The right line represents the average departure time of customers. The horizontal distance between the lines is W. The vertical distance between the lines will be L, the average number of customers in the system. Thus

$$\lambda = \frac{L}{W}$$

Poisson subordination

- We may simulate from a continuous time finite state Markov chain by drawing the holding times from distributions Exponential(q_i), where q_i depends on the state i.
- INSTEAD, simulate a holding time from Exponential(λ) where λ is large, and allow movement back to the same state.
- Matematical formulation: Given generator matrix Q. If $\lambda \geq \max(q_, \dots, q_k)$ then
  - $R = \frac{1}{\lambda} Q + I$ is a stochastic matrix.
  - We can write

$$P(t) = e^{tQ} = e^{-t\lambda I} e^{t\lambda R} = e^{-t\lambda} \sum_{k=0}^{\infty} \frac{(t\lambda R)^k}{k!} = \sum_{k=0}^{\infty} R^k \frac{e^{-\lambda t}(\lambda t)^k}{k!}.$$

- Thus: To find the probability of going from i to j during time t:
  1. Simulate the number of changes occurring $k \sim \text{Poisson}(\lambda t)$.
  2. Move the discrete chain with transition matrix R k steps.
- This provides a good way to compute e^{tQ}: Throw away terms where k is over some limit. Better accuracy than using definition of exponential matrix!
- Ths discrete chain has the same stationary distribution as the continuous chain.