Markov processes (Continuous time Markov chains)

Consider (stationary) Markov processes with a continuous parameter space (the parameter usually being time). Transitions from one state to another can occur at any instant of time.

- Due to the Markov property, the time the system spends in any given state is memoryless: the distribution of the remaining time depends solely on the state but not on the time already spent in the state $\Rightarrow$ the time is exponentially distributed.

A Markov process X_t is completely determined by the so called generator matrix or transition rate matrix

$$q_{i,j} = \lim_{\Delta t \to 0} \frac{\mathrm{P}\{X_{t+\Delta t} = j \,|\, X_t = i\}}{\Delta t} \qquad i \neq j$$

- probability per time unit that the system makes a transition from state i to state j

- transition rate or transition intensity

The total transition rate out of state i is

$$q_i = \sum_{j \neq i} q_{i,j} \qquad | \text{ lifetime of the state} \sim \mathrm{Exp}(q_i)$$

This is the rate at which the probability of state i decreases. Define

$$q_{i,i} = -q_i$$

Transition rate matrix and time dependent state probability vector

The transition rate matrix in full is

$$\mathbf{Q} = \begin{pmatrix} q_{0,0} & q_{0,1} & \cdots \\ q_{1,0} & q_{1,1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} -q_0 & q_{0,1} & \cdots \\ q_{1,0} & -q_1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

row sums equal zero: the probability mass flowing out of state i will go to some other states (is conserved)

State probability vector $\boldsymbol{\pi}(t)$ is now a function of time evolving as follows

$$\boxed{\frac{d}{dt}\boldsymbol{\pi}(t) = \boldsymbol{\pi}(t) \cdot \mathbf{Q}} \quad \Rightarrow \quad \boldsymbol{\pi}(t+\Delta t) = \boldsymbol{\pi}(t) + \boldsymbol{\pi}(t) \cdot \mathbf{Q}\,\Delta t + o(\Delta t) = \boldsymbol{\pi}(t)(\mathbf{I} + \mathbf{Q}\,\Delta t) + o(\Delta t)$$

Transition probability matrix over time interval Δt is $\mathbf{P}(\Delta t) = \mathbf{I} + \mathbf{Q}\,\Delta t$

- tends to the identity matrix $\mathbf{I}$ as $\Delta t \to 0$
- $\mathbf{Q} = \mathbf{P}'(0)$ is the time derivative of the transition prob. matrix (transition rate matrix)

A formal solution to the time dependent state probability vector is

$$\boxed{\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) \cdot e^{\mathbf{Q}\,t}}$$

The matrix exponent function $e^{\mathbf{A}}$ can be defined by means of a power series: $e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \cdots$

Global balance conditions

The stationary solution $\boldsymbol{\pi} = \lim_{t\to\infty} \boldsymbol{\pi}(t)$ is independent of time and thus satisfies

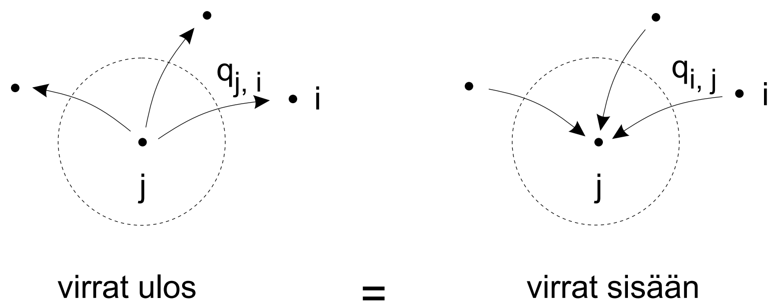
$$\boxed{\boldsymbol{\pi} \cdot \mathbf{Q} = \mathbf{0}}$$

Global balance condition which expresses the balance of probability flows.

The j^{th} row is

$$\underbrace{q_j}_{\sum_{i\neq j} q_{j,i}} \pi_j = \sum_{i\neq j} \pi_i q_{i,j}$$

$$\boxed{\sum_{i\neq j} \pi_j q_{j,i} = \sum_{i\neq j} \pi_i q_{i,j}} \qquad \begin{aligned} \pi_i q_{i,j} \;=\;& \text{probability flow from state } i \text{ to state } j \\ & \text{(transition frequency from state } i \text{ to state } j) \end{aligned}$$



Global balance conditions (continued)

- The equations are linearly dependent: any given equation is automatically satisfied if the other ones are satisfied ("conservation of probability").
- The solution is unique up to a constant factor.
- The solution is uniquely determined by the normalization condition.

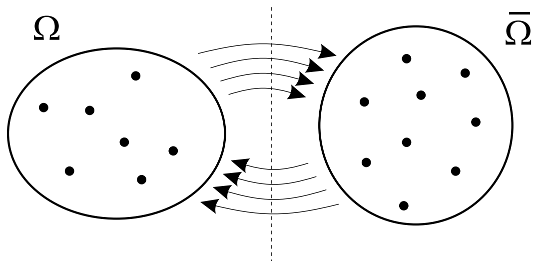
$$\boldsymbol{\pi} \cdot \mathbf{e}^{\mathrm{T}} = 1 \quad \text{or} \quad \sum_j \pi_j = 1$$

- $\boldsymbol{\pi}$ is the (left) eigenvector belonging to the eigenvalue 0.

Global balance condition applies also to any set of states.

In stationarity, the probability flows between two sets constituting a partition of the state space are in balance: Let Ω and $\bar{\Omega}$ be the complementary sets of the partition. Then

$$\sum_{i\in\Omega, j\in\bar{\Omega}} \pi_j q_{j,i} = \sum_{i\in\Omega, j\in\bar{\Omega}} \pi_i q_{i,j}$$



Solving the balance equations

In the same way as in the case of a Markov chain the solution to the (homogeneous) balance equation

$$\boldsymbol{\pi} \cdot \mathbf{Q} = \mathbf{0}$$

satisfying the normalization condition $\boldsymbol{\pi} \cdot \mathbf{e}^{\mathrm{T}} = 1$, is expediently obtained by writing n copies of the normalization condition

$$\boldsymbol{\pi} \cdot \mathbf{E} = \mathbf{e}$$

where $\mathbf{E}$ is an $n \times n$ matrix with all elements equal to one, $\mathbf{E} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$,

by summing the equations, $\boldsymbol{\pi} \cdot (\mathbf{Q} + \mathbf{E}) = \mathbf{e}$, and by solving the inhomogeneous equation thus obtained

$$\boxed{\boldsymbol{\pi} = \mathbf{e} \cdot (\mathbf{Q} + \mathbf{E})^{-1}}$$

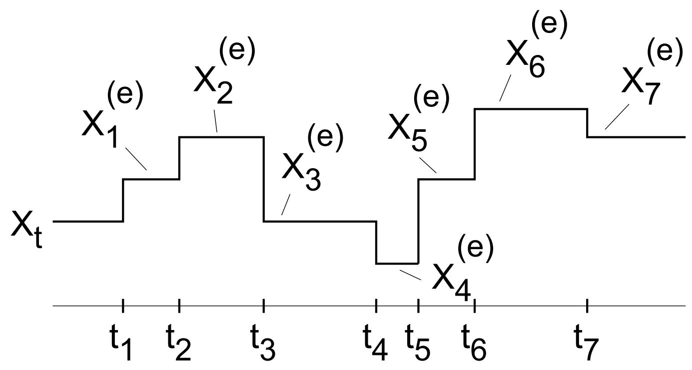
Embedded Markov chain

With every continuous time Markov process X_t we can associate a discrete time Markov chain, so called embedded Markov chain or jump chain $X_n^{(e)}$.

- Focus is on the transitions of X_t (when they occur), i.e. on the sequence of (different) states visited by X_t.
- Let the state transitions of X_t occur at instants $t_0, t_1, \ldots$
- Define $X_n^{(e)}$ to be the value of X_t immediately after the transition at time t_n (at the instant t_n^+) or the value of X_t in (t_n, t_{n+1}).

$$\boxed{X_n^{(e)} = X_{t_n^+}}$$

Since X_t is a Markov process, the embedded chain $X_n^{(e)}$ constitutes a Markov chain.

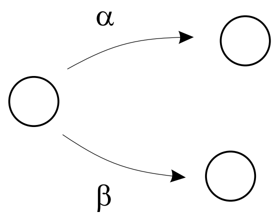


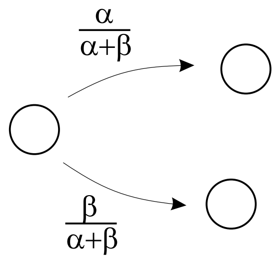
Embedded Markov chain (continued)

The states of a Markov process can be classified by the classification provided by the embedded Markov chain (transient, absorbing, recurrent,...).

The transition probabilities of the embedded chain

$$\begin{aligned} p_{i,j} &= \lim_{\Delta t \to 0} \mathrm{P}\{X_{t+\Delta t} = j \,|\, X_{t+\Delta t} \neq i, X_t = i\} \\ &= \lim_{\Delta t \to 0} \frac{\mathrm{P}\{X_{t+\Delta t} = j, X_{t+\Delta t} \neq i \,|\, X_t = i\}}{\mathrm{P}\{X_{t+\Delta t} \neq i \,|\, X_t = i\}} \\ &= \begin{cases} \dfrac{q_{i,j}}{\Sigma_j\, q_{i,j}} & i \neq j \quad \text{cf. } \mathrm{P}\{\min(X_1, \ldots, X_n) = X_i\} = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n}, \text{ when } X_i \sim \mathrm{Exp}(\lambda_i) \\ 0 & i = j \end{cases} \end{aligned}$$





Markov process, transition rates $q_{i,j}$
equilibrium probabilities π_i

Embedded Markov chain, transition probabilities $p_{i,j}$
equilibrium probabilities $\pi_i^{(e)}$

Equilibrium probabilities of the embedded Markov chain

$$\boxed{\pi_i = \frac{\pi_i^{(e)}\mathrm{E}[T_i]}{\sum_j \pi_j^{(e)}\mathrm{E}[T_j]}} \quad \Leftrightarrow \quad \boxed{\pi_i^{(e)} = \frac{\pi_i q_i}{\sum_j \pi_j q_j}} \qquad \mathrm{E}[T_i] = 1/q_i, \qquad q_i = \sum_{j \neq i} q_{i,j}$$

π_i = proportion of time that the X_t spends in state i (weight $\mathrm{E}[T_i]$)

$\pi_i^{(e)}$ = relative frequency with which state i occurs in the jump chain $X_n^{(e)}$ (weight 1)

Note $\pi_i q_i$ is the frequency with which the Markov chain X_t makes transitions out of state i. In equilibrium, this equals the frequency with which the system jumps into state i.

- Now we have considered the sequence $X_n^{(e)}$ of all different states visited by X_t
- Sometimes it is possible to pick a subsequence of this chain which again is an embedded Markov chain.
  - later we will base the analysis of so called $M/G/1$ queue on the consideration of an appropriately chosen embedded Markov chain (a subsequence of the full jump chain)

Semi-Markov processes

Conversely, with every Markov chain $Z_n,\ n = 1, 2, \ldots$ we can associate a continuous time stochastic process X_t by drawing the time T_i spent by X_t in state i from some distribution

- every time the value is drawn independently
- different states can have different lifetime distributions

and then drawing the new state Z_n according to the state transition probabilities.

The process X_t thus obtained is called a semi-Markov process

- generally is not a Markov process
- is a Markov process if and only if $T_i \sim \mathrm{Exp}(\lambda_i)$
- it has the same stationary distribution as the corresponding Markov process