Time continuous, countable state Markov chain

- Def: Let $\mathcal{S} = \{s_1, s_2, \ldots\}$ be a set of states, and $\{X(t): t \geq 0\}$ be random variables such that $\forall t \geq 0 \ \Pr[X(t) \in \mathcal{S}] = 1$. Suppose also that, for any n and times $t_1 < t_2 < \cdots < t_n$ and states $x_1, x_2, \ldots, x_n$,

$$\Pr\left[\underbrace{X(t_{n+1}) = x_{n+1}}_{future} \middle| \underbrace{X(t_1) = x_1, \ldots,}_{past} \underbrace{X(x_n) = x_n}_{present}\right] = \Pr\left[X(t_{n+1}) = x_{n+1} \middle| X(t_n) = x_n\right]$$

when $\Pr[X(t_1) = x_1, \ldots, X(t_n) = x_n] \neq 0$.

Then, $\{X(t)\}$ is a **Markov chain** (in continuous time).

($\{X(t)\}$ has **Markov property**)

- "Given the present, the past and the future become conditionally independent."
- $P_{X(t_{n+1}), X(t_1), \ldots, X(t_{n-1}) | X(t_n)}(i_{n+1}, i_1, \ldots, i_{n-1} | i_n)$

  $= P_{X(t_{n+1}) | X(t_n)}(i_{n+1} | i_n) \cdot P_{X(t_1), \ldots, X(t_{n-1}) | X(t_n)}(i_1, \ldots, i_{n-1} | i_n)$.

  Hence, to define a continuous-time Markov chain, need $\forall s$ and $\forall t \geq 0 \ \Pr[X(s+t) = j | X(s) = i]$.

- In many application, $\mathcal{S}$ will be the non-negative integers or some subset of them; we assume this holds, unless otherwise specified.
- By convention, $X(t)$ is a right-continuous function.
- Notation:
  - **Transition Probabilities**: $p_{i,j}(s,t) = \Pr[X(t) = j | X(s) = i]$
  - **Time-dependent state probabilities**: $p_i(t) = \Pr[X(t) = i]$
  - $$P(s,t) = [p_{i,j}(s,t)] = \begin{pmatrix} p_{0,0}(s,t) & p_{0,1}(s,t) & p_{0,2}(s,t) & \cdots \\ p_{1,0}(s,t) & p_{1,1}(s,t) & p_{1,2}(s,t) & \cdots \\ p_{2,0}(s,t) & p_{2,1}(s,t) & p_{2,3}(s,t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}_{\mathcal{S} \times \mathcal{S}}$$
  - $\underline{p}(t) = (p_0(t), p_1(t), p_2(t), \ldots)^T$, a row vector.
- $\underline{p}(t) = \underline{p}(s) P(s,t)$
- For $\vec{1} = (1,1,1,\ldots)^T$, $\underline{p}(t)\vec{1} = 1$, and $P(s,t)\vec{1} = \vec{1}$

  Proof.

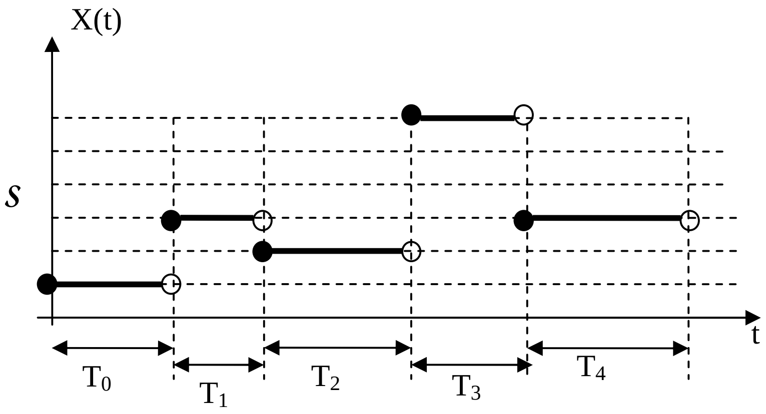
$$\begin{pmatrix} p_{0,0}(s,t) & p_{0,1}(s,t) & p_{0,2}(s,t) & \cdots \\ p_{1,0}(s,t) & p_{1,1}(s,t) & p_{1,2}(s,t) & \cdots \\ p_{2,0}(s,t) & p_{2,1}(s,t) & p_{2,3}(s,t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sum_{j=0}^{\infty} p_{0,j}(s,t) \\ \sum_{j=0}^{\infty} p_{1,j}(s,t) \\ \sum_{j=0}^{\infty} p_{2,j}(s,t) \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{pmatrix}$$

Every row sums up to 1 because If in s at time *s*, then have to be somewhere at time *t*

- Chapman-Kolmogorov equations: $P(s,t) = P(s,u)P(u,t)$ if $s < u < t$.
- **Holding time** / **state occupancy time**:

  $T_0 = \min\{t > 0 : X(t) \neq X(0)\}$

  $T_k = \min\{t > 0 : X(T_0 + \ldots + T_{k-1} + t) \neq X(T_0 + \ldots + T_{k-1})\}$ for $k \geq 1$



- Property of the state occupancy time: $X(t)$ remains at a given value (state) for an exponentially distributed random time.

  Proof. Assume, $X(t_0) = i$. The Markov property implies that the past $(\text{time } t < t_0)$ is irrelevant and we can view the system as being restarted in state *i* at time $t = t_0$. Only the exponential random variable satisfies this memoryless property.

- Another way of looking at continuous-time Markov chains:

  Each time a state, say *i*, is entered, an exponentially distributed state occupancy time T_i is selected. When the time is up, the next state *j* is selected according to a discrete-time Markov chain, with transition probabilities $\tilde{p}_{ij}$. Then, the new state occupancy time is selected according to T_j, and so on.

  We call $\tilde{p}_{i,j}$ an embedded Markov chain.

- **Def**: **Transition rate matrix** : $Q(s) = (q_{i,j}(s))$

$a)\ \sum_j q_{i,j}(s) = 0$ for all i and s $b)$ As $h \to 0,\ P_{i,j}(s, s+h) \sim \delta_{i,j} + q_{i,j}(s)h + o(h)$ • $q_{i,j}(s) = \lim_{h\to 0^+} \frac{P_{ij}(s,s+h)}{h} \geq 0$ for $i \neq j$. • $q_{i,i}(s) = -\sum_{\substack{j\\ j\neq i}} q_{i,j}(s) \leq 0$

Assume that there exist a matrix $Q(s) = (q_{i,j}(s))$ of continuous functions of time indexed by pairs (i,j) of states such that

$a)$ Sum in each row = 0: $\sum_j q_{i,j}(s) = 0$ for all i and s

$b)$ As $h \to 0,\ P_{i,j}(s, s+h) \sim \delta_{i,j} + q_{i,j}(s)h + o(h)$

, where function of h o(h) satisfies $\lim_{h\to 0} \frac{o(h)}{h} = 0$ (approach zero faster than h e.g. h^2)

- For $i \neq j,\ \delta_{i,j} = 0$

$$q_{i,j}(s) = \lim_{h\to 0^+}\left(\frac{P_{ij}(s,s+h)}{h} - \frac{o(h)}{h}\right) = \lim_{h\to 0^+}\frac{P_{ij}(s,s+h)}{h} - \lim_{h\to 0^+}\frac{o(h)}{h}$$

(the last term is crossed out)

$$= \lim_{h\to 0^+}\frac{P_{ij}(s,s+h)}{h} \geq 0$$

Therefore, $q_{i,j}(s) \geq 0$ for $i \neq j$. $\Rightarrow$ all non-diagonal element has to be non-negative.

- Since $\sum_j q_{i,j}(s) = 0,\ q_{i,i}(s) = -\sum_{j\neq i} q_{i,j}(s) \leq 0$

- With this assumption, get the **Kolmogorov backward equation**:

$$\frac{\partial}{\partial s}P(s,t) = Q(s)P(s,t)$$

- Under slightly more restrictive technical conditions, designed to prevent infinitely many state transitions from occurring in finite time,

get the **Kolmogorov-Feller forward equation**:

$$\frac{\partial}{\partial t}P(s,t) = P(s,t)Q(t)$$

- $\boxed{\underline{p}'(t) = \underline{p}(t)Q(t)}$

Proof. From $\underline{p}(t) = \underline{p}(s)P(s,t)$, we have $\frac{\partial}{\partial t}\underline{p}(t) = \frac{\partial}{\partial t}\underline{p}(s)P(s,t)$ = $\underline{p}(s)\frac{\partial}{\partial t}P(s,t)$. Applying the K-F forward eqn., $\frac{\partial}{\partial t}P(s,t) = P(s,t)Q(t)$, we then have $\frac{\partial}{\partial t}\underline{p}(t) = \underline{p}(s)P(s,t)Q(t) = \underline{p}(t)Q(t)$.

- $$\boxed{\frac{d}{dt}p_j(t) = \sum_i p_i(t)q_{i,j}(t) = \sum_{\substack{i \\ i\neq j}} p_i(t)q_{i,j}(t) - \sum_{\substack{k \\ k\neq j}} p_j(t)q_{j,k}(s)}$$

Proof.

$$\underline{p}(t)Q(t) = \left(p_0(t), p_1(t), p_2(t), \ldots\right)\begin{pmatrix} q_{0,0}(t) & q_{0,1}(t) & q_{0,2}(t) & \cdots \\ q_{1,0}(t) & q_{1,1}(t) & q_{1,2}(t) & \cdots \\ q_{2,0}(t) & q_{2,1}(t) & q_{2,2}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$= \left(\sum_i p_i(t)q_{i,0}(t), \sum_i p_i(t)q_{i,1}(t), \sum_i p_i(t)q_{i,2}(t) \ldots\right) = \left(\frac{d}{dt}p_j(t)\right)$$

, where the last equality comes from $\underline{p}'(t) = \underline{p}(t)Q(t)$.

Hence, $\frac{d}{dt}p_j(t) = \sum_i p_i(t)q_{i,j}(t) \Rightarrow$ weighted sum in Q's column

$$\frac{d}{dt}p_j(t) = \sum_{\substack{i \\ i\neq j}} p_i(t)q_{i,j}(t) + p_j(t)q_{j,j}(t) = \sum_{\substack{i \\ i\neq j}} p_i(t)q_{i,j}(t) + p_j(t)\left(-\sum_{\substack{k \\ k\neq j}} q_{j,k}(s)\right)$$
$$= \sum_{\substack{i \\ i\neq j}} p_i(t)q_{i,j}(t) - \sum_{\substack{k \\ k\neq j}} p_j(t)q_{j,k}(s)$$

So we have

- $$\boxed{\begin{pmatrix}\text{Instantaneous rate of} \\ \text{change of probability} \\ \text{of state } j\end{pmatrix} = \begin{pmatrix}\text{Instantaneous flow of} \\ \text{probability into state } j\end{pmatrix} - \begin{pmatrix}\text{Instantaneous flow of} \\ \text{probability out of state } j\end{pmatrix}}$$

This is why we call $Q(t)$ the transition rate matrix.

So, each column of Q represents the flow-in weight. The j^{th} position of the i^{th} column represents the flow-in weight from state j to state i. Note that the i^{th} position of the i^{th} column is negative. Hence, it represents the flow-out from state i to other states. This makes sense because the sum in each row = 0. The rest of row i shows how much flow-out weight from state i to each individual states.

Homogeneous, time continuous, countable state Markov chain

- **Def**: A **continuous-time Markov chains** $\{X(t), t \geq 0\}$ where all $X(t)$'s take values in a countable state space $\mathcal{S}$ (e.g., $\mathcal{S}$ = {0, 1, 2, …}) has **homogeneous transition probabilities** iff

  ≡ $\forall s \ \ \forall t \geq 0$ $\boxed{\Pr\left[X(s+t)=j \middle| X(s)=i\right] = \Pr\left[X(t)=j \middle| X(0)=i\right] := p_{ij}(t)}$.

  ≡ The holding times $\{T_{i,1}, T_{i,2}, \ldots\}$ of visits to a particular state *i* are i.i.d. exponential random variable with a parameter α_i that depends on the state.

- **Def**: $\boxed{P(t) = \left[p_{ij}(t)\right]}$ = the matrix of transition probabilities in an interval of length *t*.

  Note that $\boxed{P(0) = I}$ (initial condition) because $p_{ij}(t) = \delta(i, j)$.

- $P(s,t) = P(t-s)$

- We consider on "standard" Markov chains, i.e. those in which every p_{ij} is a **continuous function** of *t*, with $\lim_{t \to 0^+} P(t) = I$.

  It turns out that the functions p_{ij} are

  - uniformly continuous

    Proof. First we will show that $\forall h \ \left|p_{ij}(t+h) - p_{ij}(t)\right| \leq 1 - p_{ii}(h)$.

    Note that $p_{ij}(t+h) = \sum_k p_{ik}(h) p_{kj}(t)$. Hence,

$$
\begin{aligned}
p_{ij}(t+h) - p_{ij}(t) &= \left(\sum_k p_{ik}(h) p_{kj}(t)\right) - p_{ij}(t) \\
&= \left(\sum_{k \neq i} p_{ik}(h) p_{kj}(t)\right) + p_{ii}(h) p_{ij}(t) - p_{ij}(t) \\
&= \left(\sum_{k \neq i} p_{ik}(h) p_{kj}(t)\right) + \left(p_{ii}(h) - 1\right) p_{ij}(t) \\
&\leq \left(\sum_{k \neq i} p_{ik}(h)\right) + \left(p_{ii}(h) - 1\right) p_{ij}(t) \\
&= \left(1 - p_{ii}(h)\right) + \left(p_{ii}(h) - 1\right) p_{ij}(t) \\
&= \left(1 - p_{ii}(h)\right)\left(1 - p_{ij}(t)\right) \\
&\leq 1 - p_{ii}(h)
\end{aligned}
$$

Also, $p_{ij}(t+h) - p_{ij}(t) = \underbrace{\left(\sum_{k \neq i} p_{ik}(h)\, p_{kj}(t)\right)}_{>0} - \left|p_{ii}(h) - 1\right| p_{ij}(t)$

$$\geq -\left|p_{ii}(h) - 1\right| p_{ij}(t) \geq -\left|p_{ii}(h) - 1\right|$$

Therefore, $\left|p_{ij}(t+h) - p_{ij}(t)\right| \leq 1 - p_{ii}(h)$. So, $\forall x\ \forall y$ $\left|p_{ij}(y) - p_{ij}(x)\right| \leq 1 - p_{ii}(|y-x|)$. By the continuity of $p_{ii}(t)$ and $p_{ii}(0) = 1$, given any $\frac{1}{m}$, we can find $\frac{1}{n}$ such that $|y-x| \leq \frac{1}{n}$ implies $1 - p_{ii}(|y-x|) = \left|1 - p_{ii}(|y-x|)\right| \leq \frac{1}{m}$.

- continuously differentiable (C^1) at all points $t > 0$.
  At $t = 0$, where their left-hand derivatives are not defined, their right-hand derivatives, denoted by $q_{ij} = p'_{ij}(0)$, also exist.

- $\forall t \geq 0\ \ \forall j\ ,\ p_{jj}(t) > 0$.

  Proof. The statement is trivial for $t = 0$ because $p_{jj}(0) = 1 > 0$. Now, let $t > 0$ be given. Because $p_{jj}(t)$ is continuous everywhere and $p_{jj}(0) = 1$, $\exists n \in \mathbb{N}$ such that $p_{jj}\left(\frac{t}{n}\right) \geq \frac{1}{2}$. The Chapman-Kolmogorov equation then tells us that $p_{jj}(t) \geq \left(p_{jj}\left(\frac{t}{n}\right)\right)^n \geq \frac{1}{2}^n > 0$.

  - We do not have to consider the possibility of periodicity because the function $p_{jj}(t)$ is always strictly positive.

- If $i \neq j$ and $p_{ij}(t_0) > 0$, then $\forall t \geq t_0\ \ p_{ij}(t) > 0$.

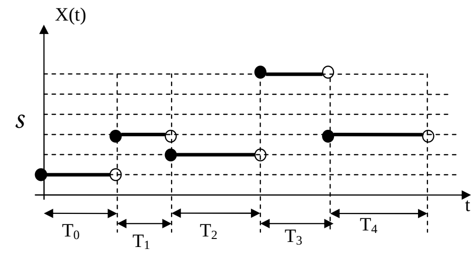
  Proof. $p_{ij}(t) \geq \underbrace{p_{ij}(t_0)}_{>0} \underbrace{p_{jj}(t - t_0)}_{>0} > 0$.

  - Hence, if $i \neq j$ and p_{ij} is not identically zero, then $\exists t_0 > 0$ such that, starting from state i, it is possible to be in state j at all times subsequent to t_0.
  - Actually, a sharper result holds:
    If p_{ij} is not identically zero, it is strictly positive $\forall t > 0$.

    Even if we cannot move directly from i to j, if we can get there somehow, we can get there in an arbitrary short time.

- Ex. Poisson process.

- Ex. Here, T_1 and T_4 are i.i.d. exponential random variable.



- Chapman-Kolmogorov equation:

$$p_{ij}(s+t)=\sum_k p_{ik}(s)\,p_{kj}(t),$$

$$P(t+s)=P(s)P(t)=P(t)P(s).$$

Proof. From Chapman-Kolmogorov equations for general time continuous Markov chain, $P(t_1,t_2)=P(t_1,u)P(u,t_2)$. For homogeneous case, by letting $t_1=0, u=s, t_2=t+s$, we have

$$P(t+s)\overset{\text{hom.}}{=}P(0,t+s)\overset{\text{C-K}}{=}P(0,s)P(s,t+s)\overset{\text{hom.}}{=}P(s)P(t)$$

- If we observe the chain at times 0, 1, 2, … only, then $P:=P(1)$ can be thought of as the one-step transition matrix of a discrete time chain;

Proof. Use induction and $P(n+1)=P(1)P(n)$.

- For a chain currently in state i, let T denote the time it stays in that state before moving to a different one. Then, $\Pr[T\ge t]=e^{-\alpha_i}t$ for some α_i.

Proof. Given $X(t_0)=i$. Then, for $s\ge 0$ and $t\ge 0$, we have

$$\begin{aligned}
&\Pr\left[T\ge s+t \middle| T\ge s\right]\\
&=\Pr\left[X(\tau)=i \text{ for } t_0\le\tau\le t_0+s+t \middle| X(\tau)=i \text{ for } t_0\le\tau\le t_0+s\right]\\
&=\Pr\left[X(\tau)=i \text{ for } \underbrace{t_0+s\le\tau\le t_0+s+t}_{t\text{ more}} \middle| X(\tau)=i \text{ for } t_0\le\tau\le t_0+s\right]\\
&=\Pr\left[X(\tau)=i \text{ for } t_0+s\le\tau\le t_0+s+t \middle| X(t_0+s)=i\right] \text{ (Markov)}\\
&=\Pr\left[X(\tau)=i \text{ for } t_0+s\le\tau\le t_0+s+t \middle| X(t_0+s)=i\right] \text{ (Homogeneity)}\\
&=\Pr[T\ge t].
\end{aligned}$$

So we have $\Pr[T\ge s+t] = \Pr\left[T\ge s+t \middle| T\ge s\right]\Pr[T\ge s] =$ $\Pr[T\ge t]\Pr[T\ge s]$, the defining property of the exponential distribution. Thus, $\Pr[T\ge t]=e^{-\alpha_i}t$ for some α_i.

- Consider an interval Δ. Pr[at least one jump occurs before Δ] = Pr[1st jump occurs before Δ] = $1-e^{-\alpha_i\Delta} \le \max_i\left(1-e^{-\alpha_i\Delta}\right)$.

$\Pr\left[2^{\text{nd}} \text{ jump before } \Delta\right]$

$= P\left[1^{\text{st}} \text{ jump before } \Delta\right] P\left[2^{\text{nd}} \text{ jump before } \Delta \middle| 1^{\text{st}} \text{ jump before } \Delta\right]$.

$\Pr\left[2^{\text{nd}} \text{ jump before } \Delta \middle| 1^{\text{st}} \text{ jump before } \Delta\right] = \max_j\left(1-e^{-\alpha_j(\Delta-\tau)}\right) \le \max_j\left(1-e^{-\alpha_j\Delta}\right)$, where τ is the time till the first jump.

Hence, $\Pr\left[2^{\text{nd}} \text{ jump before } \Delta\right] \le \left(\max_i\left(1-e^{-\alpha_i\Delta}\right)\right)^2 = \left(1-e^{-\alpha\Delta}\right)^2$ where $\alpha = \max_i \alpha_i$.

Note that $\left(1-e^{-\alpha\Delta}\right)^2 = \left(\alpha\Delta + o(\Delta)\right)^2 = (\alpha\Delta)^2 + 2\alpha\Delta o(\Delta) + o\left(\Delta^2\right) = o\left(\Delta^2\right)$. Hence, $\Pr\left[2^{\text{nd}} \text{ jump before } \Delta\right] = o\left(\Delta^2\right)$. Thus, for small Δ, can disregard the second jump.

- **Def**: $q_{ij} = p'_{ij}(0) = \begin{cases} \lim_{t\to 0^+} \dfrac{p_{ij}(0+t) - \overset{0}{\cancel{p_{ij}(0)}}}{t} & i \ne j \\ \lim_{t\to 0^+} \dfrac{p_{ii}(0+t) - \overset{1}{\cancel{p_{ii}(0)}}}{t} & i = j \end{cases} = \begin{cases} \lim_{t\to 0^+} \dfrac{p_{ij}(t)}{t} & i \ne j \\ \lim_{t\to 0^+} \dfrac{p_{ii}(t)-1}{t} & i = j \end{cases}$.

- It can be shown that every q_{ij} with $i \ne j$ is automatically finite,

  but if the state space S does not have finitely many states, it may happen that $q_{ii} = -\infty$. (We shall not consider such models.)

- Properties:

• $0 \le p_{ij}(t) \le 1$, $\sum_j p_{ij}(t) = 1$, $p_{jj}(0) = 1$, $P(s,t) = P(t-s)$. • Chapman-Kolmogorov equation: $p_{ij}(s+t) = \sum_k p_{ik}(s)\, p_{kj}(t)$, $P(t+s) = P(s)P(t) = P(t)P(s)$. • $P(t_1+t_2+\cdots+t_n) = P(t_1)P(t_2)\cdots P(t_n)$, $p_{ii}(t_1+t_2+\cdots+t_n) \ge p_{ii}(t_1)\, p_{ii}(t_2)\cdots p_{ii}(t_n)$
• $Q(t) = Q = \left[q_{i,j}\right] \Rightarrow$ a constant matrix. $q_{ij} = p'_{ij}(0) = \begin{cases} \lim_{t\to 0^+} \dfrac{p_{ij}(t)}{t} & i \ne j \\ \lim_{t\to 0^+} \dfrac{p_{ii}(t)-1}{t} & i = j \end{cases}$. $Q = \lim_{t\to 0^+} \dfrac{P(t)-I}{t}$.

- $\sum_j q_{ij} = 0$, i.e., $q_{ii} = -\sum_{j \neq i} q_{ij}$.
- $p_{ij}(t) = \begin{cases} q_{ij}t + o(t) & i \neq j \\ 1 + q_{ii}t + o(t) & i = j \end{cases}$
- $P'(t) = QP(t) = P(t)Q$, $\frac{d}{dt} p_j(t) = \sum_i q_{ij} p_i(t)$
- $q_{ii} = -\alpha_i$.
- $\tilde{p}_{ij} = \frac{q_{ij}}{\alpha_i} = \frac{q_{ij}}{\sum_{k \neq i} q_{ik}}$
- $P(t) = \exp(Qt) = I + Qt + \frac{Q^2t^2}{2!} + \frac{Q^3t^3}{3!} + \cdots$

- $\sum_j q_{ij} = 0$

  Proof. Note that $\sum_j p_{ij}(\delta) = 1$. Hence, $\lim_{\delta \to 0^+} \left(\sum_{j \neq i} \frac{p_{ij}(\delta)}{\delta} + \frac{p_{ii}(\delta) - 1}{\delta} \right) = 0$. This implies $\sum_{j \neq i} q_{ij} + q_{ii} = \sum_j q_{ij} = 0$. (When the sum has finitely many terms, we can always interchange the order of limit and sum. Even when S is infinite, we consider only chains in which this "conservative" condition holds.

- **Derivation 1**

  By definition of derivative,

$$p_{ij}(t) = p_{ij}(0) + p'_{ij}(0)(t-0) + o(|t-0|) = \begin{cases} q_{ij}t + o(t) & i \neq j \\ 1 + q_{ii}t + o(t) & i = j \end{cases}.$$

The quantity then can be interpreted as $-q_{ii}$ the rate at which a process now in state *i* departs from that sate. Similarly, for $i \neq j$, q_{ij} is the rate at which we jump to *j*, when we are now in *i*.

We have shown that for a chain currently in state *i*, the time it stays in that state before moving to a different one is $\mathcal{E}(\alpha_i)$. We will show later that $p'_{\bar{i}\bar{i}}(0) = p'_{ii}(0)$; hence, $p'_{ii}(0) = \frac{d}{dt} e^{-t\alpha_i} \Big|_{t=0} = -\alpha_i$. Therefore, we must have $-\alpha_i = q_{ii}$.

The chain develops by remaining in its current state *i* for a random time , then jumping to $j \neq i$ with probability $\frac{q_{ij}}{-q_{ii}}$. If we ignore how long is spent in each state, and just look at the sequence of movements from state to state, the process we observe is called the **jump chain**.

- **Note**: $\boxed{p_{\bar{i}\bar{i}}(t) - p_{ii}(t) = o(t^2)}$

  Given $X(t_0) = i$, let $p_{\bar{i}\bar{i}}(t)$ = $\Pr\left[X(t) \text{ is till in state } i \text{ for at least } t \text{ more time unit}\right]$ = $\Pr\left[X(t_0+\tau) = i \text{ for } 0 \le \tau \le t \middle| X(t_0) = i\right]$ = $e^{-\alpha_i t}$ = $1 - \alpha_i t + o(t)$. However, $p_{ii}(t)$ = $\Pr\left[X(t) \text{ is in state } i \text{ after } t \text{ time unit}\right]$ = $1 + q_{ii}t + o(t)$. If there is any jump at all in the process of going from $i \to i$ in the interval *t*, then, there has to be at least two jumps (*i* to a state, and then back.) The probability of at least two jumps in the interval *t* is $o(t^2)$. The probability that those jumps come back to *i* is even less than this. Hence, the probability of going from $i \to i$ with any jump is also $o(t^2)$. The probability of going from $i \to i$ without any jump is then $p_{\bar{i}\bar{i}}(t) = p_{ii}(t) - o(t^2)$. So, $p_{\bar{i}\bar{i}}(t) - p_{ii}(t) = o(t^2)$. This implies $p'_{\bar{i}\bar{i}}(0) = p'_{ii}(0)$.

- **Derivation 2**

  Consider the transition probabilities in a very short time interval of duration δ. Assume that $X(t_0) = i$. We know that the state occupancy time for all continuous-time Markov chains are exponentially distributed; hence the probability that the process remains in state *i* during the interval is $\Pr\left[X(t) = i, t_0 \le t < t_0 + \delta \middle| X(t_0) = i\right]$ = $p_{ii}(\delta)$ = $e^{-\alpha_i \delta}$ = $1 - \alpha_i \delta + o(\delta)$ for some α_i. $\left(\lim_{\delta \to 0} \frac{o(\delta)}{\delta} = 0\right)$. Therefore, $1 - p_{ii}(\delta) = \alpha_i \delta + o(\delta)$. We then regard α_i as the rate at which the process $X(t)$ leaves state *i*.

  Once the process leaves state *i*, it will enter state *j* with probability $\tilde{p}_{ij}$. Thus,

  $p_{ij}(\delta) = (1 - p_{ii}(\delta))\tilde{p}_{ij} = (\alpha_i \delta + o(\delta))\tilde{p}_{ij} = \alpha_i \delta \tilde{p}_{ij} + o(\delta)$.

  Define, for $i \ne j$, $q_{ij} = \alpha_i \tilde{p}_{ij}$ as the rate at which the process $X(t)$ enters state *j* from state *i*. For completeness, define $q_{ii} = -\alpha_i$. Then,

  $$p_{ii}(\delta) - 1 = q_{ii}\delta + o(\delta), \text{ and } p_{ij}(\delta) = q_{ij}\delta + o(\delta).$$

  And

  $$\lim_{\delta \to 0} \frac{p_{ii}(\delta) - 1}{\delta} = q_{ii}, \text{ and } \lim_{\delta \to 0} \frac{p_{ij}(\delta)}{\delta} = q_{ij}.$$

- Kolmogorov backward and forward equations: $P'(t) = QP(t) = P(t)Q$

  Proof. By Chapman-Kolmogorov equation

  $$P(t+h) - P(t) = P(h)P(t) - P(t) = (P(h) - I)P(t).$$

  Hence, $\lim_{h \to 0^+} \frac{P(t+h) - P(t)}{h} = \left(\lim_{h \to 0^+} \frac{P(h) - I}{h}\right)P(t) = QP(t)$. (backward)

("Backward" because we look back in time to the rates of transition at time zero.)

Similarly,

$$P(t+h)-P(t)=P(t)P(h)-P(t)=P(t)\big(P(h)-I\big).$$

Hence, $\lim_{h\to 0^+}\dfrac{P(t+h)-P(t)}{h}=P(t)\left(\lim_{h\to 0^+}\dfrac{P(h)-I}{h}\right)=P(t)Q$. (forward)

Note that we assume the interchange of limit and sum is justifiable. (Always in the case of finite number of states.)

Note that the formal justification needs careful analysis.

- The formal solution is $P(t)=\exp(Qt)=I+Qt+\dfrac{Q^2t^2}{2!}+\dfrac{Q^3t^3}{3!}+\cdots$.

- $\underline{p}'(t)=\underline{p}(t)Q$. Equivalently, $\dfrac{d}{dt}p_j(t)=\sum_i q_{ij}p_i(t)$.

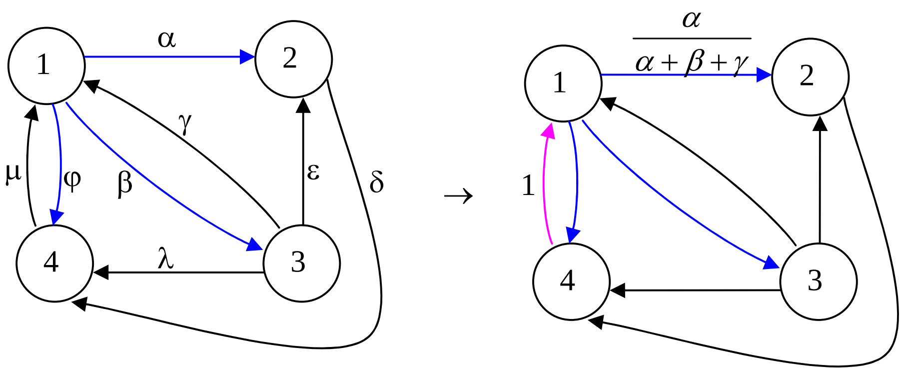
  Proof. $\underline{p}'(t)=\underline{p}(0)P'(t)\overset{\text{forward}}{=}\underline{p}(0)P(t)Q=\underline{p}(t)Q$

  Alternative proof.

  $$\underline{p}(t+h)-\underline{p}(t)=\underline{p}(t)P(h)-\underline{p}(t)=\underline{p}(t)\big(P(h)-I\big)$$

  $$\underline{p}'(t)=\lim_{h\to 0^+}\frac{\underline{p}(t+h)-\underline{p}(t)}{h}=\underline{p}(t)\left(\lim_{h\to 0^+}\frac{P(h)-I}{h}\right)=\underline{p}(t)Q$$

- The Q is analogous to the one-step transition matrix P of homogeneous time-discrete Markov chains.
- **Jump Chains**
  - Time-discrete Markov chain embedded in a time-continuous Markov chain
  - The jump chain has same state space ***S*** as corresponding continuous chain does, and its state diagram has the same arrows, but not the same arrow labels.
  - Jump chain's discrete time advances by 1 every time the corresponding continuous-time changes state.
  - $\tilde{p}_{ij}=\begin{cases}\dfrac{q_{ij}}{-q_{ii}} & \text{if } i\neq j\\ 0 & \text{otherwise}\end{cases}$
  - Example



$$P\left(X(k)=2 \middle| X(k-1)=1\right)=P\left(T_{12}<T_{13},T_{12}<T_{14}\right)$$
$$=\int_0^\infty \alpha e^{-\alpha t} P\left(T_{1\to 3}>t,T_{1\to 4}>t\right)dt$$
$$=\int_0^\infty \alpha e^{-\alpha t} P\left(T_{1\to 3}>t\right)P\left(T_{1\to 4}>t\right)dt$$
$$=\int_0^\infty \alpha e^{-\alpha t} e^{-\gamma t} e^{-\varphi t}dt=\frac{\alpha}{\alpha+\beta+\gamma}$$

$P\left(X(k)=1 \middle| X(k-1)=4\right)=1$ since only one arrow out of state 4

- **Def**: $\{p_i\}$ = the stationary state pmf of the Markov chain. $\underline{p}=(p_1,p_2,\ldots)^T$.
- **Global balance equations**: At "equilibrium" (or "steady state")

  $\forall j$ $\boxed{\sum_{i\neq j} q_{ij}p_i=-q_{jj}p_j}$, or equivalently, $\sum_{i\neq j} q_{ij}p_i=\sum_{k\neq j} q_{jk}p_j$.

  At "equilibrium" (or "steady state"), $p_j(t)\to p_j$ and $p_j'(t)\to 0$; hence, from $\frac{d}{dt}p_j(t)=\sum_i q_{ij}p_i(t)$, we have $\sum_i q_{ij}p_i=0$. From $q_{jj}=-\sum_{k\neq j}q_{jk}$, we have $\sum_i q_{ij}p_i=\sum_{i\neq j}q_{ij}p_i+q_{jj}p_j=\sum_{i\neq j}q_{ij}p_i-\sum_{k\neq j}q_{jk}p_j=0$.

  "the rate of probability flow out of state *j*, namely "$-q_{jj}p_j$", is equal to the rate of flow into state *j*, "$\sum_{i\neq j}q_{ij}p_i$".

  By solving this set of linear equations $\left(\text{with condition } \sum_j p_j=1\right)$ we can obtain the stationary state pmf of the system (when it exists.)

- If we start the Markov chain with initial state pmf given by $\underline{p}$, then the state probabilities will be $\forall i\ \forall t\ \ p_i(t) = p_i$. The resulting process is a stationary random process.

  Let $\vec{t} = (t_1, t_2, \ldots, t_n)$ where $t_i < t_{i+1}$. Then,

  $$\Pr\left[\vec{X}(\vec{t}+\tau) = \vec{x}\right] \stackrel{\text{Markov}}{=} \Pr\left[X(t_0+\tau) = x_0\right] \prod_{i=1}^{n} \Pr\left[X(t_i+\tau) = x_i \middle| X(t_{i-1}+\tau) = x_{i-1}\right].$$

  By homogeneity, $\Pr\left[X(t_i+\tau) = x_i \middle| X(t_{i-1}+\tau) = x_{i-1}\right] = p_{x_{i-1},x_i}(t_i - t_{i-1})$.

  Since $\forall i\ \forall t\ \ p_i(t) = p_i$, we have $\Pr\left[X(t_0+\tau) = x_0\right] = p_{x_0}$.

  Hence, $\Pr\left[\vec{X}(\vec{t}+\tau) = \vec{x}\right] = p_{x_0} \prod_{i=1}^{n} p_{x_{i-1},x_i}(t_i - t_{i-1}) = \Pr\left[\vec{X}(\vec{t}) = \vec{x}\right]$.

- A subset B of the state space $\mathcal{S}$ is **closed** for $\{X(t)\}$ if

  $\forall s\ \ p_{i,j}(s) = 0$ if $i \in B$ and $j \notin B$

  $\Rightarrow$ can't get out once you're in

- $\{X(t)\}$ is **irreducible / indecomposable** if $\mathcal{S}$ itself is the smallest non-empty set closed for $\{X(t)\}$

  - In that case, the entire state space $\mathcal{S}$ is a communicating class in the sense that $\forall i \in \mathcal{S}\ \forall j \in \mathcal{S}\ \exists s$ such that $p_{i,j}(s) > 0$.

- A probability measure $\underline{p}$ on $\mathcal{S}$ is **invariant** (is an "equilibrium measure") if

  $\underline{p} = \underline{p}P(\tau)$ for all $\tau \Rightarrow \underline{p}(t)$ not depends on (t)

  where $P(\tau) = P(s, s+\tau)$.

- If $\{X(t)\}$ is irreducible,

  there exists at most one invariant measure.

- $\underline{p}$ is invariant in the homogeneous case if and only if $\underline{p}Q = \underline{0}$

  $\equiv \forall j \sum_{i \neq j} p_i q_{i,j} - \sum_{k \neq j} p_j q_{j,k} = 0$.

  Proof. We know that $\frac{d}{dt}\underline{p}(t) = \underline{p}(t)Q(t) \stackrel{\text{homogeneous}}{=} \underline{p}(t)Q$. For $\underline{p}(t)$ to be invariant, must have $\frac{d}{dt}\underline{p}(t) = 0$. So, $\underline{p}(t)Q = \underline{p}Q = 0$

  - This is analogous to $\underline{p} = \underline{p}P$ where P is a 1-step transition matrix of homogeneous discrete-time Markov chain.

Limiting Probabilities for Continuous-Time Homogeneous Markov Chains

- A continuous-time Markov chain $X(t)$ can be viewed as consisting of a sequence of states determined by 1) some discrete-time Markov chain X_n with transition probabilities $\tilde{p}_{ij}$ and 2) a corresponding sequence of exponentially distributed $(\mathcal{E}(\alpha_i))$ state occupancy times.

- If the associated discrete-time chain X_n (governed by $\left[\tilde{p}_{ij}\right]$) is irreducible and positive recurrent with stationary pmf π_j, then the long-term proportion of time spent by $X(t)$ in state *i* is

$$p_i = \frac{\frac{\pi_i}{\alpha_i}}{\sum_j \frac{\pi_i}{\alpha_j}},$$

where $\frac{1}{\alpha_i}$ is the mean occupancy time in state *i*.

Further more, the p_i's are the unique solution to the global balance equations $\forall j$ $\sum_{i \neq j} q_{ij} p_i = -q_{jj} p_j$.

Proof.

Suppose that the embedded Markov chain X_n is irreducible and positive recurrent. Then we can find π_i's, the unique solution of $\forall j\ \pi_j = \sum_i \pi_i \tilde{p}_{ij}$, and $1 = \sum_i \pi_i$. Note that one time step of the embedded Markov chain denote one transition (jump) of the original continuous-time Markov process.

Let $N_i(n)$ denote the number of times state *i* occurs in the first *n* transitions. Then the portion of "jump" to state *i* is $\frac{N_i(n)}{n}$. Because the process is irreducible and recurrent, proportion of "jump" to state $i \to \pi_i$. So we have, with probability one, $\lim_{n\to\infty} \frac{N_i(n)}{n} = \pi_i$. Note also that $\lim_{n\to\infty} N_i(n) = \infty$.

Let $T_i(j)$ denote the occupancy time the j^{th} time state i occurs. $\left(T_i(j)\right)_{j=1}^{\infty}$ are then an iid sequence ~ $\mathcal{E}(\alpha_i)$. Hence, by the strong law of large numbers, with probability one, $\lim_{n\to\infty}\frac{1}{N_i(n)}\sum_{j=1}^{N_i(n)}T_i(j)=E\left[T_i(j)\right]=\frac{1}{\alpha_i}$.

$$p_i=\lim_{n\to\infty}\frac{\text{time spent in state } i}{\text{time spent in all states}}=\lim_{n\to\infty}\frac{\sum_{j=1}^{N_i(n)}T_i(j)}{\sum_i\sum_{j=1}^{N_i(n)}T_i(j)}$$

$$=\lim_{n\to\infty}\frac{\frac{N_i(n)}{n}\frac{1}{N_i(n)}\sum_{j=1}^{N_i(n)}T_i(j)}{\sum_i\frac{N_i(n)}{n}\frac{1}{N_i(n)}\sum_{j=1}^{N_i(n)}T_i(j)}$$

We know that, with probability one, $\lim_{n\to\infty}\frac{N_i(n)}{n}=\pi_i$ and $\lim_{n\to\infty}\frac{1}{N_i(n)}\sum_{j=1}^{N_i(n)}T_i(j)=\frac{1}{\alpha_i}$. Hence, $p_i=\frac{\pi_i\frac{1}{\alpha_i}}{\sum_i\pi_i\frac{1}{\alpha_i}}$.

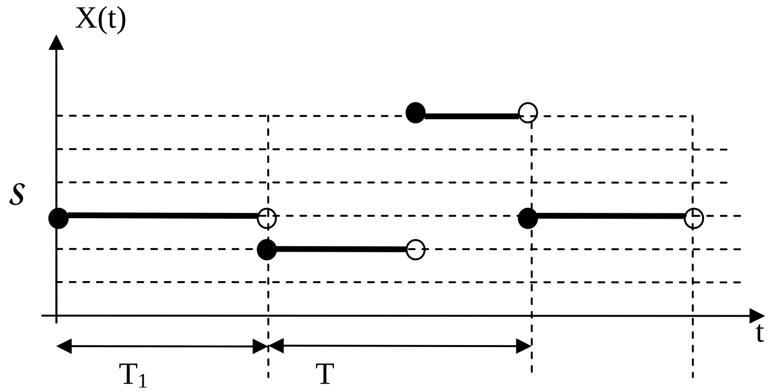
Let $c=\sum_i\pi_i\frac{1}{\alpha_i}$. From $\pi_j=\sum_i\pi_i\tilde{p}_{ij}$, substitute $\pi_i=p_i\alpha_i c$, we then have $p_j\alpha_j\not{c}=\sum_i p_i\alpha_i\not{c}\tilde{p}_{ij}$. Note that $\tilde{p}_{jj}=0$. Hence, $\sum_i p_i\alpha_i\tilde{p}_{ij}=\sum_{i\neq j}p_i\alpha_i\tilde{p}_{ij}$.

Now use $q_{ii}=-\alpha_i$, and $\tilde{p}_{ij}=\frac{q_{ij}}{\alpha_i}$, we then have $-p_j q_{jj}=\sum_{i\neq j}p_i q_{ij}$, the global balance equation. Thus, the p_i's satisfy the global balance equations.

- Now, fix $i\in\mathcal{S}$ and define

  $T_1=\min\{t>0: X(t)\neq i\}$

  $T=\min\{t>0: X(T_1+t)=i\}$ = duration of first sojourn from state i and back.



- Given $X(0) = i$, $T_1 \sim \mathcal{E}(\alpha)$, $\alpha = -q_{ii} > 0 = \sum_{j \neq i} q_{i,j}$. $\mathrm{E}T_1 = \dfrac{1}{-q_{i,i}}$.
- Define $M_i = E\left[T \middle| X(0) = i\right]$

- Independent of $\underline{p}(0)$

$$\lim_{t\to\infty} P_i(t) = \frac{(-q_{ii})^{-1}}{M_i + (-q_{ii})^{-1}} = \begin{cases} \text{equilibrium distribution} & \text{if ergodic} \\ 0 & \text{otherwise} \end{cases}.$$

Example 1:

- For $S = \{1,2\}$, suppose $Q = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}$, where $\alpha > 0$ and $\beta > 0$.

$P'(t) = P(t)Q = P(t)\begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}$; hence,

$$p'_{11}(t) = -\alpha p_{11}(t) + \beta p_{12}(t) = -\alpha p_{11}(t) + \beta\left(1 - p_{11}(t)\right)$$
$$= \beta - (\alpha+\beta) p_{11}(t)$$

With the initial condition $p_{11}(0) = 1$, we have $p_{11}(t) = \dfrac{\beta}{\alpha+\beta} + \dfrac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t}$.

By symmetry, $p_{22}(t) = \dfrac{\alpha}{\alpha+\beta} + \dfrac{\beta}{\alpha+\beta} e^{-(\alpha+\beta)t}$.

$p_{12}(t) = 1 - p_{11}(t)$, $p_{21}(t) = 1 - p_{22}(t)$.

The limiting behavior of $P(t)$ is immediate and should be no surprise. The quantity α and β are the rates of getting out of states 1 and 2 respectively, moving to the other state, so in the long run, the respective chances of being in 1 and 2 will be proportional to α and β. Thus, $p_{i1}(t) \to \dfrac{\beta}{\alpha+\beta}$, $p_{i2}(t) \to \dfrac{\alpha}{\alpha+\beta}$.

We can write $Q=\begin{bmatrix}1 & \alpha \\ 1 & -\beta\end{bmatrix}\begin{bmatrix}0 & 0 \\ 0 & -(\alpha+\beta)\end{bmatrix}\begin{bmatrix}1 & \alpha \\ 1 & -\beta\end{bmatrix}^{-1}$.

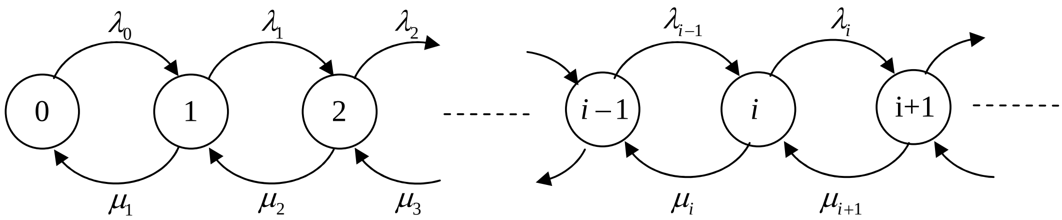
Then, $P(t)=I+Qt+\frac{Q^2t^2}{2!}+\frac{Q^3t^3}{3!}+\cdots$

$$=A\left(I+Dt+\frac{D^2t^2}{2!}+\frac{D^3t^3}{3!}+\cdots\right)A^{-1}$$

$$=A\begin{bmatrix}1 & 0 \\ 0 & \sum_{k=0}^{\infty}\frac{\left(-(\alpha+\beta)t\right)^k}{k!}\end{bmatrix}A^{-1}=A\begin{bmatrix}1 & 0 \\ 0 & e^{-(\alpha+\beta)t}\end{bmatrix}A^{-1}$$

Example 2: Birth-and-death process

- **Def**: **Birth-and-death process** is a Markov chain in which only transitions between adjacent states occur.



- Let $r_j=\frac{\lambda_{j-1}}{\mu_j}$ and $R_j=r_jr_{j-1}\cdots r_1$ for j = 1, 2, …. Also, let $R_0=1$. Then, if $\sum_{j=0}^{\infty}R_j$ converges, then the stationary pmf is given by $p_i=\frac{R_i}{\sum_{j=0}^{\infty}R_j}$. If the series does not converges, then a stationary pmf does not exist, and $\forall i\ \ p_i=0$.

- The global balance equations are

  $\lambda_0p_0=\mu_1p_1$, and $\left(\lambda_j+\mu_j\right)p_j=\lambda_{j-1}p_{j-1}+\mu_{j+1}p_{j+1}$ for j = 1, 2, ….

  We can rewrite the second equation as follows:

  $\lambda_jp_j-\mu_{j+1}p_{j+1}=\lambda_{j-1}p_{j-1}-\mu_jp_j$ for j = 1, 2, …., which implies that

  $\lambda_{j-1}p_{j-1}-\mu_jp_j=k$, a constant for j = 1, 2, ….

  The case when j = 1 gives $\lambda_0p_0-\mu_1p_1=k$. However, we already know that $\lambda_0p_0=\mu_1p_1$. Hence, k = 0, and therefore,

$$\lambda_{j-1} p_{j-1} - \mu_j p_j = 0$$

$$p_j = \left(\frac{\lambda_{j-1}}{\mu_j}\right) p_{j-1} = r_j p_{j-1}$$

, where $r_j = \dfrac{\lambda_{j-1}}{\mu_j}$ for j = 1, 2, ….

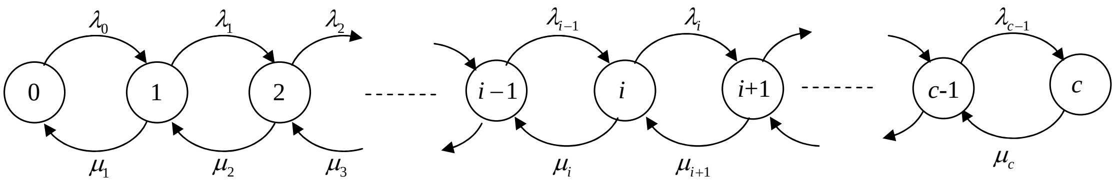
By simple induction argument, $p_j = r_j r_{j-1} \cdots r_1 p_0$.

If we define $R_j = r_j r_{j-1} \cdots r_1$, and $R_0 = 1$, then $p_j = R_j p_0$ for j = 0, 1, 2, ….

From $\sum_{j=0}^{\infty} p_j = 1$, we require $p_0 = \dfrac{1}{\sum_{j=0}^{\infty} R_j}$.

- Sufficient condition for ergodicity is $\sum_{j=0}^{\infty} R_j < \infty$.

Example 3: Truncated birth-and-death process



- Let $r_j = \dfrac{\lambda_{j-1}}{\mu_j}$ and $R_j = r_j r_{j-1} \cdots r_1$ for j = 1, 2, …, c. Also, let $R_0 = 1$. Then, the stationary pmf is given by $p_i = \dfrac{R_i}{\sum_{j=0}^{c} R_j}$.

- The argument follows exactly the argument used in the birth-and-death process except the last step which requires $\sum_{j=0}^{c} p_j = 1$ instead of $\sum_{j=0}^{\infty} p_j = 1$. Note that the sum is finite, and hence, always converges.

- $\boxed{P_b = \dfrac{\lambda_c R_c}{\sum_{j=0}^{c} \lambda_i R_j}}$. If $\forall i \ \lambda_i = \lambda$, then 1) $\forall i \ f_i = p_i$ 2) $P_b = f_c = p_c$.

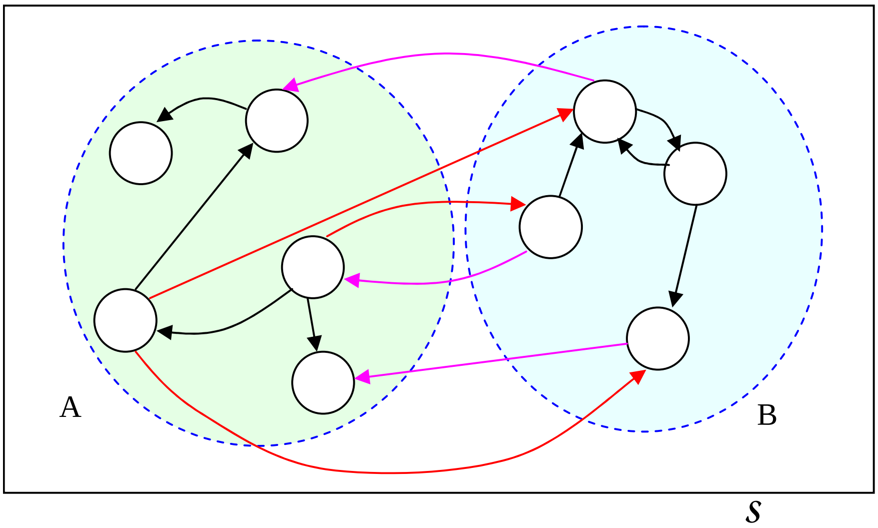
Proof. In the steady state, the value of p_i represents the fraction of the time axis during which the system is in state i, or equivalently the probability that the system is in state i at a “randomly chosen instant.” However, the density of calling attempts varies with the state of the system. When in state i, call attempts occur at rate λ_i. Hence, The fraction of all call attempts that occur when the system is in state i is not p_i but rather $f_i = \dfrac{\lambda_i p_i}{\sum_{j=0}^{c} \lambda_i p_j}$. Call attempts get blocked if and only if they occur when the system is in state c, the blocking probability is

$$P_b = f_c = \frac{\lambda_c p_c}{\sum_{j=0}^{c} \lambda_i p_j} = \frac{\lambda_c \frac{R_c}{\cancel{\sum_{k'=0}^{c} R_{k'}}}}{\sum_{j=0}^{c} \lambda_i \frac{R_j}{\cancel{\sum_{k'=0}^{c} R_{k'}}}} = \frac{\lambda_c R_c}{\sum_{j=0}^{c} \lambda_i R_j}$$

State diagram and Balance

- Arrows represent probability transition **rates**, not transition probabilities.
- Global balance

  Let S = A∪B and A∩B = ∅

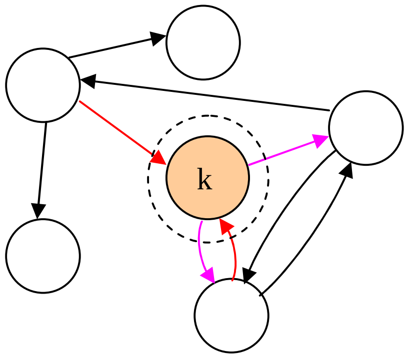


Can get global balance equation by equating flow from A to B to flow from B to A:

$$\sum_{\substack{i\in A \\ j\in B}} p_i q_{i,j} = \sum_{\substack{i\in B \\ j\in A}} p_i q_{i,j}$$

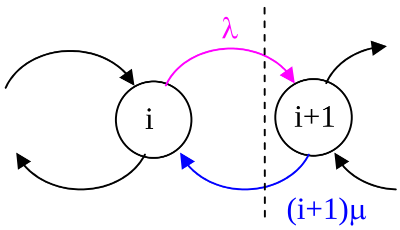
- Suppose A = $\{k\}$

  B = S-$\{k\}$ = $\{j: j \neq k\}$



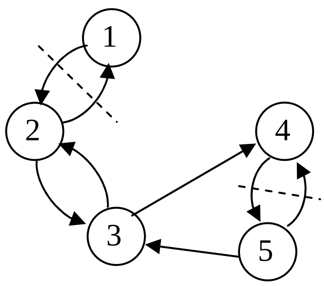
Then, global balance equation becomes $\sum_{j \neq k} p_k q_{k,j} = \sum_{i \neq k} p_i q_{i,k}$ which is precisely the k^{th} equation in $\underline{p}Q = \underline{0}$.

- Some "global" balance looks very "local"



- Local Balance

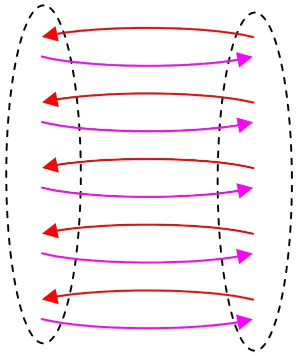
Consider:



Cutting both links between states 1 and 2 partitions the state spaces in to {1} and {2,3,4,5}. Hence, the local balance equations $p_1q_{1,2} = p_2q_{2,1}$ must hold.

But, cutting both links between 4 and 5 doesn't partition, so $p_4q_{4,5}$ may not necessarily equal to $p_5q_{5,4}$.

- If a solution $\{p_i, i \in S\} = \underline{p}$ can be found that satisfies the local balance equations for all pairs of states, then this $\underline{p}$ satisfies the global balance equations and hence is an equilibrium distribution.

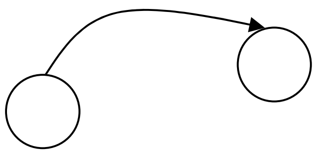


If each red arrow flow is equal to the corresponding magenta flow, then the sum of red flow is equal to the sum of magenta flow.

- In such instances, need $\{\{q_{i,j} > 0 \text{ and } q_{j,i} > 0\} \text{ or } q_{i,j} = q_{j,i} = 0\}\ \forall(i,j)$

$\equiv q_{i,j} > 0 \leftrightarrow q_{j,i} > 0$

thus, can't have



- When there exists a solution to the set of all local balance equations, the Markov chain is reversible.

Time-Reversible Continuous-Time Markov Chains

- Consider a stationary, continuous-time Markov chain.
- The reverse process also spends an exponentially distributed amount of time with rate α_i in state *i*.

Proof.

Let T_i be the forward process's state occupancy time for state *i*.

$$\Pr\left[\begin{matrix} X(t') = i \text{ at least for } \tau \text{ more time unit} \\ \text{in the reversed direction } (t-\tau \le t' \le t) \end{matrix} \middle| X(t) = i \right]$$

$$= \frac{\Pr\left[X(t-\tau) = i, X(t') = i \text{ for at least } \tau \text{ more time unit}\right]}{\Pr\left[X(t) = i\right]}$$

$$= \frac{\Pr\left[X(t-\tau) = i\right]\Pr\left[T_i > \tau\right]}{\Pr\left[X(t) = i\right]} = \Pr\left[T_i > \tau\right]$$

Note that $\Pr\left[X(t-\tau) = i\right] = \Pr\left[X(t) = i\right]$ by the stationaryness of the process.

- The jumps in the forward process $X(t)$ are determined by the embedded Markov chain $\tilde{p}_{ij}$, so the jumps in the reverse process are determined by the discrete-time Markov chain corresponding to the time-reversed embedded Markov chain given by $\hat{\tilde{p}}_{ij} = \frac{\pi_j \tilde{p}_{ji}}{\pi_i}$.
- The transition rates for the time-reversed continuous-time process are given by $\hat{q}_{ij} = \frac{p_j}{p_i} q_{ji}$

  Proof. $\hat{q}_{ij} = \alpha_i \hat{\tilde{p}}_{ij} = \alpha_i \frac{\pi_j \tilde{p}_{ji}}{\pi_i} = \alpha_i \frac{\pi_j}{\pi_i} \frac{q_{ji}}{\alpha_j} = \frac{\pi_j / \alpha_j}{\pi_i / \alpha_i} q_{ji} = \frac{\frac{1}{\sum_k \pi_k / \alpha_k} \pi_j / \alpha_j}{\frac{1}{\sum_k \pi_k / \alpha_k} \pi_i / \alpha_i} q_{ji} = \frac{p_j}{p_i} q_{ji}$.
- If we can guess a set of transition rates $\{\hat{q}_{ij}\}$ and a pmf $\{p_j\}$ such that $\forall i, j \;\; p_i \hat{q}_{ij} = p_j q_{ji}$ and $\forall i \;\; \sum_{j \neq i} \hat{q}_{ij} = \sum_{j \neq i} q_{ij} \left(= \alpha_i\right)$, then $\{p_j\}$ is the stationary pmf for $X(t)$ and $\{\hat{q}_{ij}\}$ are the transition rates for the reverse process.

- The continuous-time Markov chain $X(t)$ is reversible

  ≡ its embedded Markov chain is reversible

  ≡ $\forall i, j \;\; p_i q_{ij} = p_j q_{ji}$

  (The rate at which $X(t)$ goes from state *i* to state *j* is equal to the rate at which $X(t)$ goes from state *j* to state *i*).

  Proof. Since the state occupancy times in the forward and reverse processes are exponential random variables with the same mean,

  the continuous-time Markov chain $X(t)$ is reversible if and only if its embedded Markov chain is reversible, i.e., $\forall i, j \;\; \pi_i \tilde{p}_{ij} = \pi_j \tilde{p}_{ji}$ where $\{\pi_i\}$ is the stationary pmf of the embedded Markov chain. Now, recall that the stationary pmf $X(t)$, $p_i = \frac{1}{c} \frac{\pi_i}{\alpha_i}$ where $c = \sum_i \frac{\pi_i}{\alpha_i}$ is just a constant. So, $\pi_i \tilde{p}_{ij} = \pi_j \tilde{p}_{ji}$ is equivalent to $p_i \alpha_i \tilde{p}_{ij} = p_j \alpha_j \tilde{p}_{ji}$. Now, use $q_{ij} = \alpha_i \tilde{p}_{ij}$. We then have $p_i q_{ij} = p_j q_{ji}$.
- All continuous-time birth-and-death processes are time-reversible.

  Proof. The embedded Markov chain is a discrete-time birth-and-death process which is time-reversible.