Probability and Stochastic Processes I I - Lecture 6c

Michael Evans
University of Toronto
https://utstat.utoronto.ca/mikevans/stac62/staC632024.html

2024

VI.3 Continuous time discrete state space processes

- some of the following material comes from Grimmett and Stirzaker (2001) and Applied Probability by Berestycki and Sousi (2017) (on the web)

Definition VI.6 A process $\{X_t : t \geq 0\}$ is a *Markov process* if for every n and $0 \leq t_1 \leq \cdots \leq t_n$, $B \in \mathcal{B}^1$ then

$$P(X_{t_n} \in B \,|\, X_{t_1}, \ldots, X_{t_{n-1}}) = P(X_{t_n} \in B \,|\, X_{t_{n-1}})$$

and it is *time homogeneous* if this depends only on $t_n - t_{n-1}$.

- generally the time domain T can be any totally ordered set

- when the state space $\mathcal{S}$ is countable we may as well take it to be $\mathbb{Z}$ or $\mathbb{N}_0$ ($\mathcal{S} \subset \mathbb{N}_0$ assumed hereafter in this section) and then for each t we have the (possibly infinitely dimensional) matrix

$$P^{(t)} = (p_{xy}^{(t)}) = (P(X_t = y \,|\, X_0 = x))$$

- then for $u \geq 0$, using the MP and TH

$$p_{xy}^{(t+u)} = P(X_{t+u} = y \,|\, X_0 = x) = \sum_{s \in \mathcal{S}} P(X_{t+u} = y, X_t = s \,|\, X_0 = x)$$

$$= \sum_{s \in \mathcal{S}} P(X_t = s \,|\, X_0 = x) P(X_{t+u} = y \,|\, X_t = s) = \sum_{s \in \mathcal{S}} p_{xs}^{(t)} p_{sy}^{(u)}$$

- so, as with Markov chains these are stochastic matrices (rows nonnegative and sum to 1) and satisfy (i) $P^{(0)} = I$ (ii) $P^{(t+u)} = P^{(t)} P^{(u)}$

- a set of matrices that satisfy these properties is known as a *semigroup*

- so it is immediate that a Poisson process is a continuous time discrete state space MP

Definition VI.7 A Markov process $\{X_t : t \geq 0\}$ is a *standard Markov process* if $\lim_{t\downarrow 0} p_{xy}^{(t)} = \delta_{xy}$ for all $x, y \in \mathcal{S}$ where $\delta_{xy} = 1$ when $x = y$ and is 0 otherwise (Kronecker delta), so $\lim_{t\downarrow 0} P^{(t)} = I$.

- so $p_{xy}^{(t+\varepsilon)} = \sum_{s\in\mathcal{S}} p_{xs}^{(t)} p_{sy}^{(\varepsilon)} \to p_{xy}^{(t)}$ as $\varepsilon \to 0$ (by DCT) and $P^{(t)}$ is continuous in t for a standard MP

- also $p_{xx}^{(0)} = 1$ and so $p_{xx}^{(t)} > 0$ for all t small enough which implies $p_{xx}^{(nt)} \geq (p_{xx}^{(t)})^n > 0$ which implies $p_{xx}^{(t)} > 0$ for all $t \geq 0$ which implies aperiodicity

Example VI.7 *Poisson process*

- then as $t \to 0$, $p_{0y}^{(t)} = P\left(N_t = y \,|\, N_0 = 0\right) = (\lambda t)^y \exp\{-\lambda t\}/y! \to 0$ when $y > 0$ and $p_{00}^{(t)} = \exp\{-\lambda t\} \to 1$ so a Poisson process is standard ■

Definition VI.8 The *generator matrix* G of a standard Markov process $\{X_t : t \geq 0\}$ has (i, j)-th element for $i, j \in \mathcal{S}$

$$g_{ij} = \lim_{t \downarrow 0} \frac{p_{ij}^{(t)} - \delta_{i,j}}{t} = \left. \frac{d^+ p_{ij}^{(t)}}{dt} \right|_{t=0}.$$

- **fact** - these right derivatives always exist

- note - (conditions for G to be a generator matrix) $g_{ii} \leq 0$ and $g_{ij} \geq 0$ when $i \neq j$ and

$$\sum_{j \in \mathcal{S}} g_{ij} = \sum_{j \in \mathcal{S}} \lim_{t \downarrow 0} \frac{p_{ij}^{(t)} - \delta_{i,j}}{t} \overset{*}{=} \lim_{t \downarrow 0} \frac{\sum_{j \in \mathcal{S}} p_{ij}^{(t)} - \sum_{j \in \mathcal{S}} \delta_{ij}}{t}$$

$$= \lim_{t \downarrow 0} \frac{1 - 1}{t} = 0$$ where * holds under conditions, e.g. $\mathcal{S}$ finite

- so $G\mathbf{1} = \mathbf{0}$

- why is G called the generator of the process?

Lemma VI.19 If $A \in R^{m\times m}$, then
$\lim_{n\to\infty}(I+\frac{A}{n})^n = \sum_{k=0}^{\infty}\frac{A^k}{k!} \stackrel{def}{=} \exp(A)$.

Proof: Let $||A||$ be any matrix norm (e.g. Frobenius norm $||A||^2 = \sum_{i,j} a_{ij}^2$) so it satisfies

$$\begin{aligned} \text{(i) } ||A|| &\geq 0 \text{ with } ||A|| = 0 \text{ iff } A = 0 \text{ (ii) } ||aA|| = |a|||A|| \\ \text{(iii) } ||A+B|| &\leq ||A|| + ||B|| \text{ (iv) } ||AB|| \leq ||A||\,||B|| \end{aligned}$$

By the binomial theorem

$$\begin{aligned} \left(I+\frac{A}{n}\right)^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{A^k}{n^k}I^{n-k} = \sum_{k=0}^{n}\frac{n(n-1)\cdots(n-k+1)}{n^k}\frac{A^k}{k!} \\ &= \sum_{k=0}^{n}(1-1/n)\cdots(1-(k-1)/n)\frac{A^k}{k!}. \end{aligned}$$

Note that $||\sum_{k=0}^{\infty}\frac{A^k}{k!}|| \leq \sum_{k=0}^{\infty}\frac{||A||^k}{k!} = \exp\{||A||\} < \infty$ so $\sum_{k=0}^{\infty}\frac{A^k}{k!}$ exists finitely.

For given N and $n > N$

$$||(I + \frac{A}{n})^n - \sum_{k=0}^{N} \frac{A^k}{k!}||$$

$$= \quad ||\textstyle\sum_{k=0}^{N}[(1-1/n)\cdots(1-(k-1)/n)-1]\frac{A^k}{k!} - \sum_{k=N+1}^{n}(1-1/n)\cdots(1-(k-1)/n)\frac{A^k}{k!}||$$

$$\leq \quad ||\textstyle\sum_{k=0}^{N}[(1-1/n)\cdots(1-(k-1)/n)-1]\frac{A^k}{k!}|| + ||\sum_{k=N+1}^{n}(1-1/n)\cdots(1-(k-1)/n)\frac{A^k}{k!}||$$

and the first term goes to 0 as $n \to \infty$ while

$$|| \sum_{k=N+1}^{n} (1-1/n)\cdots(1-(k-1)/n)\frac{A^k}{k!}||$$

$$\leq \quad \sum_{k=N+1}^{n} \frac{||A||^k}{k!} \leq \sum_{k=N+1}^{\infty} \frac{||A||^k}{k!} \to 0 \text{ as } N \to \infty.$$

So choose N large to make the second term less than $\epsilon/2$ and $n > N$ to make the first term less than $\epsilon/2$. ■

- note - $||\exp(A)|| \leq \sum_{k=0}^{\infty} \frac{||A||^k}{k!} = \exp(||A||) < \infty$

Proposition VI.20 If standard Markov process $\{X_t : t \geq 0\}$ has generator G and $||\frac{P^{(t)}-I}{t} - G|| \to 0$ as $t \to 0$, then

$$P^{(t)} = \exp(tG) = I + \frac{t}{1!}G + \frac{t^2}{2!}G^2 + \frac{t^3}{3!}G^3 + \cdots$$

Proof: We have

$$\begin{aligned} P^{(t)} &= (P^{(t/m)})^m = (P^{(t/m)})^m = (I + P^{(t/m)} - I)^m \\ &= \left(I + \frac{t}{m} \frac{P^{(t/m)} - I}{t/m} \right)^m = \left(I + \frac{t}{m} G_m \right)^m \end{aligned}$$

and note $||G_m - G|| \to 0$ as $m \to \infty$. Therefore,

$$
\begin{aligned}
&||P^{(t)} - \exp(tG)|| \\
&= ||\left(I + \tfrac{t}{m} G_m\right)^m - \exp(tG)|| \\
&= ||\left(I + \tfrac{t}{m} G + \tfrac{t}{m}(G_m - G)\right)^m - \exp(tG)|| \\
&= ||(I + \tfrac{t}{m} G)^m - \exp(tG) + \\
&\qquad \textstyle\sum_{k=1}^m (1 - 1/m) \cdots (1 - (k-1)/m) \frac{t^k}{k!} (G_m - G)^k (I + \frac{t}{m} G)^{m-k}|| \\
&\leq ||(I + \tfrac{t}{m} G)^m - \exp(tG)|| + \textstyle\sum_{k=1}^m \frac{t^k}{k!} ||G_m - G||^k ||I + \frac{t}{m} G||^{m-k} \\
&= ||(I + \tfrac{t}{m} G)^m - \exp(tG)|| + ||I + \tfrac{t}{m} G||^m \textstyle\sum_{k=1}^m \frac{t^k}{k!} \left(\frac{||G_m - G||}{||I + \frac{t}{m} G||}\right)^k \to 0
\end{aligned}
$$

since $(I + \frac{t}{m} G)^m \to \exp(tG)$ by Lemma VI.19 and for any $\varepsilon > 0$ there exists m_0 such that for all $m \geq m_0$, then $||G_m - G|| / ||I + \frac{t}{m} G|| \leq \varepsilon$ so

$$
\sum_{k=1}^m \frac{t^k}{k!} \left(\frac{||G_m - G||}{||I + \frac{t}{m} G||}\right)^k \leq \sum_{k=1}^\infty \frac{t^k}{k!} \varepsilon^k = e^\varepsilon - 1 \to 0 \text{ as } \varepsilon \to 0.
$$

This proves the result. ■

Example VI.8 *Poisson process with intensity λ*

- recall (see Prop. VI.17), as $t \to 0$

$$p_{ii}^{(t)} = e^{-\lambda t} \to 1, p_{ii+1}^{(t)} = \lambda t + o(t), p_{ij}^{(t)} = o(t) \text{ otherwise}$$

which implies $g_{i,i} = -\lambda, g_{i,i+1} = \lambda, g_{i,j} = 0$ otherwise so G is bidiagonal with $-\lambda$ on the main diagonal and λ on the first upper diagonal ■

- note $G = \lim_{t\to 0} \frac{P^{(t)}-I}{t}$ so G can be obtained from knowledge of the transitions $P^{(t)}$

- given G, how to compute $P^{(t)}$?

- **method 1**: truncate the series in Prop. VI.20 after finitely many terms to approximate $P^{(t)}$

- **method 2:** (finite case) let $G \in R^{k\times k}$ have left eigenvalues and vectors $(\lambda_i, \mathbf{w}_i)$ so $\mathbf{w}_i' G = \lambda_i \mathbf{w}_i'$ and suppose the initial distribution can be expressed as $\upsilon = \sum_{i=1}^k a_i \mathbf{w}_i'$, then the distribution of X_t is given by

$$
\begin{aligned}
\upsilon P^{(t)} &= \upsilon \exp(tG) = \sum_{i=1}^k a_i \mathbf{w}_i' \left(I + \frac{t}{1!}G + \frac{t^2}{2!}G^2 + \frac{t^3}{3!}G^3 + \cdots\right) \\
&= \sum_{i=1}^k a_i \left(\mathbf{w}_i' + \frac{t}{1!}\mathbf{w}_i' G + \frac{t^2}{2!}\mathbf{w}_i' G^2 + \frac{t^3}{3!}\mathbf{w}_i' G^3 + \cdots\right) \\
&= \sum_{i=1}^k a_i \mathbf{w}_i' \left(1 + \frac{\lambda_i t}{1!} + \frac{(\lambda_i t)^2}{2!} + \frac{(\lambda_i t)^3}{3!} + \cdots\right) = \sum_{i=1}^k a_i e^{\lambda_i t} \mathbf{w}_i'
\end{aligned}
$$

- **method 3:** *Kolmogorov's forward equations* (under conditions)

$$\frac{d^+ p_{i,j}^{(t)}}{dt} = \lim_{h \downarrow 0} \frac{p_{i,j}^{(t+h)} - p_{i,j}^{(t)}}{h} = \lim_{h \downarrow 0} \frac{\left(\sum_{k \in \mathcal{S}} p_{i,k}^{(t)} p_{k,j}^{(h)} \right) - p_{i,j}^{(t)}}{h}$$

$$= \lim_{h \downarrow 0} \frac{\left(\sum_{k \in \mathcal{S}} p_{i,k}^{(t)} (\delta_{kj} + g_{kj} h + o(h)) \right) - p_{i,j}^{(t)}}{h} = \sum_{k \in \mathcal{S}} p_{i,k}^{(t)} g_{kj}$$

$$\text{so } \frac{d^+ P^{(t)}}{dt} = P^{(t)} G$$

which is a system of linear differential equations which can be solved subject to boundary condition $P^{(0)} = I$

- **method 4:** *Kolmogorov's backward equations* (under conditions)

$$\frac{d^+ p_{i,j}^{(t)}}{dt} = \lim_{h\downarrow 0} \frac{p_{i,j}^{(t+h)} - p_{i,j}^{(t)}}{h} = \lim_{h\downarrow 0} \frac{\left(\sum_{k\in\mathcal{S}} p_{i,k}^{(h)} p_{k,j}^{(t)}\right) - p_{i,j}^{(t)}}{h}$$

$$= \lim_{h\downarrow 0} \frac{\left(\sum_{k\in\mathcal{S}} (\delta_{ik} + g_{ik}h + o(h)) p_{k,j}^{(t)}\right) - p_{i,j}^{(t)}}{h} = \sum_{k\in\mathcal{S}} g_{ik} p_{k,j}^{(t)}$$

$$\text{so } \frac{d^+ P^{(t)}}{dt} = GP^{(t)}$$

which is a system of linear differential equations which can be solved subject to boundary condition $P^{(0)} = I$

- suppose $X_0 = x$ then we can consider $H_x = \sup\{t : X_u = x$ for $u \in [0, t)\}$ = the *holding time at state* x, so it is the time of the next transition to a new state

Proposition VI.20 If $\{X_t : t \geq 0\}$ is a time homogeneous MP then

$$P(H_x > t + u \,|\, X_0 = x, H_x > t) = P(H_x > u \,|\, X_0 = x)$$

Proof: We have

$$\begin{aligned} & P(H_x > t + u \,|\, X_0 = x, H_x > t) \\ = \; & P(X_v = x \text{ for all } v \in [0, t + u] \,|\, X_v = x \text{ for all } v \in [0, t]) \\ = \; & P(X_v = x \text{ for all } v \in [t, t + u] \,|\, X_t = x \text{ for all } v \in [0, t]) \\ = \; & P(X_v = x \text{ for all } v \in [t, t + u] \,|\, X_t = x) \text{ by MP} \\ = \; & P(X_v = x \text{ for all } v \in [0, u] \,|\, X_0 = x) \text{ by TH} \\ = \; & P(H_x > u \,|\, X_0 = x). \end{aligned}$$

■

- this result implies that H_x has a fixed distribution, given by $P(H_x \leq u \,|\, X_0 = x)$ and this distribution has the *memoryless property*

- also if we define $H_{t,x}$ to be the holding time when $X_t = x$ then

$$\begin{aligned} & P(H_{t,x} > u \,|\, X_t = x) = P(X_{t+v} = x \text{ for every } v \in [0, u] \,|\, X_t = x) \\ = \ & P(X_v = x \text{ for every } v \in [0, u] \,|\, X_0 = x) \text{ by TH} \\ = \ & P(H_x > u \,|\, X_0 = x) \end{aligned}$$

and so the distribution of the holding time only depends on the state x

- for a Poisson process with intensity λ the distribution of H_x is exponential(λ) because of the memoryless feature of the exponential distribution so it is independent of x

Proposition VI.21 A positive r.v. X has the memoryless property iff it has an exponential distribution.
Proof: We already proved in Example VI.6 that the exponential distribution has the memoryless property. So consider the converse which implies for $x, y \geq 0$

$$\begin{aligned} G(x+y) &= P(X > x+y) = P(X > x+y \,|\, X > x)P(X > x) \\ &= P(X > y)P(X > x) = G(x)G(y). \end{aligned}$$

As $P(X > 0) = 1$, $\exists\ n$ s.t. $P(X > 1/n) > 0$ so with $\lambda = -\log G(1)$

$$e^{-\lambda} = G(1) = G\left(\frac{1}{n} + \frac{n-1}{n}\right) = G\left(\frac{1}{n}\right) G\left(\frac{n-1}{n}\right) = \cdots = G^n\left(\frac{1}{n}\right)$$

so $G\left(\frac{1}{n}\right) = e^{-\lambda/n}$. Similarly, $G(k) = e^{-k\lambda}$ for any $k \in \mathbb{N}$ and for any positive rational p/q then $G(p/q) = G^p(1/q) = e^{-(p/q)\lambda}$. Since the cdf of X is right continuous this implies $G(x_n) \downarrow G(x)$ for any rational sequence $x_n \downarrow x$ which implies $G(x) = e^{-\lambda x}$. ■

- note - Prop VI.21 implies that $H_x \sim \text{exponential}(\lambda_x)$ but it does not say that these have constant rate λ

- for a standard Markov chain $\{X_t : t \geq 0\}$ define jump times $J_0, J_1, J_2, \ldots$ by $J_0 = 0$ and for $n \geq 1$, $J_n = \inf\{t \geq J_{n-1} : X_t \neq X_{J_{n-1}}\}$

Proposition VI.22 Under conditions, a standard Markov chain $\{X_t : t \geq 0\}$ is equivalent to an initial distribution, as given by the distribution of X_0, a set of independent *holding times* $H_0, H_1, H_2, \ldots$ where $H_i \sim \text{exponential}(-g_{ii})$ and a *jump Markov chain* $\{\hat{X}_n : n \in \mathbb{N}_0\}$ with $\hat{X}_0 = X_0$, transition probabilities

$$\hat{p}_{ij} = P(\hat{X}_1 = j \,|\, \hat{X}_0 = i) = -g_{ij}/g_{ii}$$

and $X_t = \hat{X}_n$ when $J_n \leq t < J_{n+1}$.

- this gives a method for simulating the process

1. generate X_0 according to the initial distribution and suppose $X_0 = i_1$
2. generate $H_{i_1} \sim \text{exponential}(-g_{i_1 i_1})$ and $\hat{X}_1 \sim \hat{p}_{i_1 j}$ for $j \in \mathcal{S}$ and suppose $\hat{X}_1 = i_2$
3. generate $H_{i_2} \sim \text{exponential}(-g_{i_1 i_1})$ and $\hat{X}_2 \sim \hat{p}_{i_2 j}$ for $j \in \mathcal{S}$ and suppose $\hat{X}_1 = i_3$

etc.

Exercise VI.11 Suppose $\mathcal{S} = \{1, 2, 3\}, \lambda_1 = 3, \lambda_2 = 2, \lambda_3 = 4$ and

$$G = \begin{pmatrix} -3 & 3 & 0 \\ 1 & -2 & 1 \\ 0 & 4 & -4 \end{pmatrix}.$$

Generate this process for 20 steps.

Example VI.9 *Simple Birth and Death process*

- suppose $X_0 = 1$

- for a *pure birth process* suppose holding times $B_1, B_2, \ldots$ are stat. ind. and $B_i \sim \text{exponential}(\lambda_i)$ is the holding time when in state i a transition is to state $i+1$ (so $\hat{p}_{ii+1} = 1$ when $i \geq 1$) so $J_i = B_1 + \cdots + B_i$ and note this is a Poisson process of intensity λ when $\lambda_i = \lambda$ for every i and note a natural choice might be $\lambda_i = i\lambda$ for some $\lambda > 0$

- a process is called *explosive* if there can be infinitely many jumps (arrivals) in a finite interval which occurs for a birth process whenever the *explosion time* $\zeta = \sum_{i=0}^{\infty} B_i < \infty$

Lemma VI.23 (i) If $\sum_{i=0}^{\infty} 1/\lambda_i < \infty$, then $P(\zeta < \infty) = 1$. (ii) If $\sum_{i=0}^{\infty} 1/\lambda_i = \infty$, then $P(\zeta < \infty) = 0$.

Proof: (i) $E(\sum_{i=0}^{\infty} B_i) \overset{MCT}{=} \sum_{i=0}^{\infty} E(B_i) = \sum_{i=0}^{\infty} 1/\lambda_i < \infty$ which implies $P(\sum_{i=0}^{\infty} B_i < \infty) = 1$.

(ii)

$$E\left(\exp\left\{-\sum_{i=0}^{\infty} B_i\right\}\right) = E\left(\prod_{i=0}^{\infty}\exp\{-B_i\}\right) \overset{MCT}{=} \prod_{i=0}^{\infty} E(\exp\{-B_i\})$$

$$= \prod_{i=1}^{\infty} m_{B_i}(-1) = \prod_{i=0}^{\infty} \frac{\lambda_i}{1+\lambda_i} = 0 \text{ with } m_{B_i} = \text{ mgf of } B_i$$

which implies $P\left(\sum_{i=0}^{\infty} B_i = \infty\right) = 1$.

- a simple birth process only makes transitions $i \longrightarrow i+1$ but a birth and death process can also make the transition $i \longrightarrow i-1$ where a death occurs according to stat. ind. holding times $D_1, D_2, \ldots$ with $D_i \sim$ exponential(μ_i) and the B_i and D_i processes are independent

- $P(H_i > t) = P(\min(B_i, D_i) > t) = P(B_i > t)P(D_i > t) =$ $\exp(-(\lambda_i + \mu_i))$ so $H_i \sim$ exponential$(\lambda_i + \mu_i)$

- $p_{i,i+1} = P(B_i < D_i) = \int_0^\infty (1 - e^{-\lambda_i x -})\mu_i e^{-\mu_i x}\, dx =$ $1 - \mu_i/(\lambda_i + \mu_i) = \lambda_i/(\lambda_i + \mu_i)$ so $p_{i,i-1} = \mu_i/(\lambda_i + \mu_i)$

- when $\lambda_0 = 0$ concern is with the probability of extinction

■

Proposition VI.24 Suppose $\hat{P}$ is a transition matrix for a Markov chain $\{\hat{X}_n : n \in \mathbb{N}_0\}$ and $\{N_t : t \geq 0\}$ is an independent Poisson process of intensity λ. The process $\{X_t : t \geq 0\}$ given by $X_t = \hat{X}_{N_t}$ has generator $G = \lambda(\hat{P} - I)$.

Proof: We have

$$
\begin{aligned}
p_{i,j}^{(t)} &= P(\hat{X}_{N_t} = j \,|\, \hat{X}_0 = i) = \sum_{k=0}^{\infty} P(N_t = k, \hat{X}_k = j \,|\, \hat{X}_0 = i) \\
&= \sum_{k=0}^{\infty} P(N_t = k) P(\hat{X}_k = j \,|\, \hat{X}_0 = i) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t}\, \hat{p}_{ij}^{(k)} \\
&= \sum_{k=0}^{1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}\, \hat{p}_{ij}^{(k)} + o(t) = e^{-\lambda t}\delta_{ij} + \lambda t e^{-\lambda t} \hat{p}_{ij} + o(t) \\
&= (1 - \lambda t)\delta_{ij} + \lambda t \hat{p}_{ij} + o(t) = \delta_{ij} + \lambda t(\hat{p}_{ij} - \delta_{ij}) + o(t)
\end{aligned}
$$

Therefore,

$$
g_{i,j} = \lim_{t \downarrow 0} \frac{p_{i,j}^{(t)} - \delta_{i,j}}{t} = \lim_{t \downarrow 0} \frac{\lambda t(\hat{p}_{ij} - \delta_{ij}) + o(t)}{t} = \lambda(\hat{p}_{ij} - \delta_{ij}).
$$

■

- what about stationary distributions and convergence?

- for a stationary distribution π we want $\pi P^{(t)} = \pi$ for all t and so we must have $\pi P^{(t+h)} - \pi P^{(t)} = 0$ and so by Kolmogorov's forward equations

$$0 = \lim_{h \downarrow 0} \frac{\pi(P^{(t+h)} - P^{(t)})}{h} = \pi P^{(t)} G = \pi G$$

and so π is a solution of the linear equations $\pi G = 0$

Definition VI.9 The probability distribution π is *stationary* for the MP $\{X_t : t \geq 0\}$ with generator G whenever $\pi G = 0$.

Definition VI.10 The MP $\{X_t : t \geq 0\}$ with generator G is *reversible* wrt probability distribution π if $\pi_i g_{ij} = \pi_j g_{ji}$ for all $i, j \in \mathcal{S}$.

Proposition VI.25 If the MP $\{X_t : t \geq 0\}$ with generator G is reversible wrt probability distribution π, then π is stationary.

Proof: We have

$$\sum_{k \in \mathcal{S}} \pi_k g_{kj} = \sum_{k \in \mathcal{S}} \pi_j g_{jk} = \pi_j \sum_{k \in \mathcal{S}} g_{jk} = 0$$

since the rows of G sum to 0. ■

- the MP $\{X_t : t \geq 0\}$ is irreducible if the jump chain is irreducible ($i \to j$ for all $i, j \in \mathcal{S}$) and recall it is always aperiodic

Proposition VI.26 If the standard MP $\{X_t : t \geq 0\}$ is irreducible with stationary distribution π, then (under conditions) $\lim_{t\to\infty} p_{i,j}^{(t)} = \pi_j$.

Proof: Since π is stationary $\pi G = 0$ which implies $\pi G P^{(t)} = 0$ and by Kolmogorov's backward equations this implies

$$\frac{d^+ \pi P^{(t)}}{dt} = \pi \frac{d^+ P^{(t)}}{dt} = \pi G P^{(t)} = 0$$

so $\pi P^{(t)}$ is constant in t. Since $\lim_{t\downarrow 0} P^{(t)} = I$ (standard Markov process) this implies that the constant value of $\pi P^{(t)}$ is $\lim_{t\downarrow 0} \pi P^{(t)} = \pi \lim_{t\downarrow 0} P^{(t)} = \pi I = \pi$.
■

Exercise VI.12 Text 4.4.4

Exercise VI.13 Text 4.4.14