Notes On Markov Process

March 5, 2025

This set of lecture notes deal with discrete time Markov Chains on a countable state space.

Contents

1 DTMC

To start let S be a countable state space, $(\Omega, \mathcal{F})$ a measurable space, $\mathbb{P}$ be a probabily measure.

Definition 1.1 (Markov Chain)**.** A discrete time stochastic process $(X_n)_{n\in\mathbb{N}_0}$ with values in S is a (homogeneous) Markov Chain on S with respect to filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ if

(a) X_n is $\mathcal{F}_n$-measurable
(b) $(X_n)_n$ satisfies the Markov Property: There exist a transition function $p : S \times S \to \mathbb{R}_+$ such that $\mathbb{P}(X_{n+1} = y|\mathcal{F}_n) = p(X_n, y)$, a.s.

Part (b) in the definition is called the markov property. By taking conditional expecation on $\sigma(X_1, \ldots, X_n) \subseteq \mathcal{F}_n$, we see that one can always take $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

There are other ways to define the Markov property. Let

- $F_n = \{X_{n+1} = i_{n+1}, X_{n+2} = i_{n+2}, \ldots\}$ i.e. the future trajectory
- $H_n = \{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\}$ i.e. the past trajectory/history.

Proposition 1.2. *The following are equivalent characterizations of the Markov Property*

(b1) Dependence only on the present.

$$\mathbb{P}\Big(X_{n+1} = i_{n+1}|H_n \cap \{X_n = i_n\}\Big) = \mathbb{P}\Big(X_{n+1} = i_{n+1}|X_n = i_n\Big)$$

(b2) Past and Future are conditionally independent given the present.

$$\mathbb{P}\Big(F \cap H|X_n = i_n\Big) = \mathbb{P}\Big(F|X_n = i_n\Big)\mathbb{P}\Big(H|X_n = i_n\Big)$$

Theorem 1.3. *A markov chain is fully characterized by*

(i) initial distribution $\mu_i = \mathbb{P}(X_0 = i)$, and
(ii) transition probabilities $p(i, j) = \mathbb{P}(X_{n+1} = j|X_n = i)$.

Proof. By fully characterized, we mean all the finite state distribution are determined. Indeed,

$$\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) = \mu_{i_0} p(i_0, i_1) \cdots p(i_{n-1}, i_n).$$

□

Remark 1.4. If the measurable state space $(S, \mathcal{S})$ is nice, knowing the finite state distributions are suffient to know about distribution of sample paths of $(X_n)_{n\in\mathbb{N}}$, which are infinite sequences. This is a consequence of the Komogorov Extension Theorem, which states the existence of a probability measure μ_∞ on the sequence space S^∞, consisting of infinite sequences $(x_0, x_1, \ldots)$, endowed with the product σ-algebra, whose finite dimensional projections agrees with finite dimensional distributions.

The point is from now on, whenver we have a markov chain, we can assume we are working on the probability space $(\Omega, \mathcal{F}) = (S^\infty, \mathcal{S}^\infty)$

1.1 Transition Probabilities

The transition probabilites can be thought of as a stochastic matrix (P_{ij}), satisfying $\sum_j P_{ij} = 1$.

Proposition 1.5 (Chapman-Komogorov Equations)**.**

$$P^{n+m} = P^n P^m$$

Proof. Straightfoward. □

It is useful to think of P as operator.

- When acting on a distribution $\mu^\top$ (it is best to think of distributions as row vectors in this case), we get a new distribution $\mu^\top P$.
- When acting on a bounded measurable function f, we get a new fucntion $Pf : S \to \mathbb{R}$ given by
$$Pf(X_n) = \mathbb{E}[f(X_{n+1})|X_n] = \sum_{y \in S} f(y)p(X_n, y).$$

Notice we have the relationship
$$\langle \mu, Pf \rangle = \langle P^* \mu, f \rangle$$
This is to be interpreted as
$$\mathbb{E}^\mu[Pf(X_n)] = \mathbb{E}^{P^*\mu}[f(X_{n+1})].$$

1.2 A reformulation of Markov Property

We shall assume sample space is the sequence space $(\Omega, \mathcal{F}) = (S^\infty, \mathcal{S}^\infty)$, we shall define a family of shift operators, which are measurable maps $\theta_m : \Omega \to \Omega$ for $m \geq 0$, given by

$$(\theta_m \omega)_n = \omega_{m+n}.$$

Then the Markov Property can be reformulated as the following:

Proposition 1.6. *Let X_n be markov chain, $F : \Omega \to \mathbb{R}$ be bounded measurable function, then we have*
$$\mathbb{E}[F \circ \theta_n | \mathcal{F}_n] = \mathbb{E}^{X_n}[F], \quad a.s.$$

1.3 Strong Markov Property

Definition 1.7. Random variable τ is a stopping time if $\{\tau = n\} \in \mathcal{F}_n$ for all $n < \infty$. The amount of information available at stopping time τ is

$$\mathcal{F}_\tau = \{A \in \bigvee_n \mathcal{F}_n : A \cap \{\tau \leq n\} \in \mathcal{F}_n\}.$$

where $\bigvee_n \mathcal{F}_n := \sigma(\bigcup_n \mathcal{F}_n)$.

It is easy to check $\tau \in \mathcal{F}_\tau$. To see this, take $A = \{\tau = k\}$, clearly $A \cap \{\tau \leq n\} \in \mathcal{F}_n$.

We shall also extend the definition of shift operator to stopping times. Note it is possible that $\tau = \infty$, in which case X_τ is undefined. This motivates the introduction of a "cemetary" state Δ.

$$(\theta_\tau \omega)_n = \begin{cases} \omega_{n+m}, & \tau = m, \\ \Delta, & \tau = \infty. \end{cases}$$

Theorem 1.8 (Strong Markov Property)**.** *Let X_n be a markov chain and F be integrable function on S^∞ with the product σ-algebra. Then on $\{\tau < \infty\}$,*

$$\mathbb{E}[F \circ \theta_\tau | \mathcal{F}_\tau] = \mathbb{E}^{X_\tau}[F].$$

Proof. The proof is a standard technique for results involving stopping times.

$$\begin{aligned}\mathbb{E}[F\circ\theta_\tau; A\cap\{\tau<\infty\}] &= \sum_{n=1}^{\infty}\mathbb{E}[F\circ\theta_n; A\cap\{\tau=n\}]\\ &= \sum_{n=1}^{\infty}\mathbb{E}[\mathbb{E}^{X_n}[F]; A\cap\{\tau=n\}] = \mathbb{E}[\mathbb{E}^{X_n}[F]; A\cap\{\tau<\infty\}].\end{aligned}$$

Where the second = is application of markov inequality since $A\cap\{\tau=n\}\in\mathcal{F}_n$. ☐

Definition 1.9 (Hitting time and visit time)**.** Let $B\subseteq S$,

(a) Hitting time of B: $\tau_B=\min\{n\in\mathbb{N}_0 : X_n\in B\}$
(b) Visit time of B: $\tau_B^+=\min\{n\in\mathbb{N} : X_n\in B\}$

For $i\in S$, consider sequence of stopping times (called successive return times)

$$\tau_i^+(0)=0,\quad \tau_i^+(1)=\tau_i^+,\quad \tau_i^+(2)=\min\{n>\tau_i^+(1) : X_n=i\},\ldots$$

Proposition 1.10. *Assuming $\tau_i^+(n)<\infty$, for all n, then $\{\tau_i^+(n)-\tau_i^+(n-1)\}_{n\in\mathbb{N}}$ are i.i.d.*

Proof. Note $\tau_i^+(n)-\tau_i^+(n-1)\ =\tau_i^+(1)\circ\theta_{\tau_i^+(n-1)}$. The proof uses repeated application of strong markov property. Take any two bounded measurable function f and g on $\mathbb{R}$,

$$\begin{aligned}\mathbb{E}^i[f(\tau_i^+(1))g(\tau_i^+(2)-\tau_i^+(1))] &= \mathbb{E}^i[f(\tau_i^+(1))g(\tau_i^+(1)\circ\theta_{\tau_i^+(1)})]\\ &= \mathbb{E}^i\Big[f(\tau_i^+(1))\mathbb{E}^i[g(\tau_i^+(1)\circ\theta_{\tau_i^+(1)})|\mathcal{F}_{\tau_i^+(1)}]\Big]\\ &= \mathbb{E}^i\Big[f(\tau_i^+(1))\mathbb{E}^{X_{\tau_i^+(1)}}[g(\tau_i^+(1))]\Big]\\ &= \mathbb{E}^i\Big[f(\tau_i^+(1))\Big]\mathbb{E}^i\Big[g(\tau_i^+(1))\Big]\end{aligned}$$

In the above, 3rd line uses strong markov property, last line uses $X_{\tau_i^+(1)}=i$ a.s. Since f and g are arbitrary, the result follows for the case of 2 excursions. The general case follows by doing multiple conditioning. ☐

One can also show that the excursions (the paths between revisits) are i.i.d., but this is more complicated. Need to carefully define the space for excursion paths.

1.4 Transience and Recurrence

Definition 1.11 (Recurrence and Transience)**.** Let $i\in S$, we say

(a) i is transient if $\mathbb{P}^i[\tau_i^+(1)=\infty]>0$
(b) i is null recurrent if $\mathbb{P}^i[\tau_i^+(1)<\infty]=1$ but $\mathbb{E}^i[\tau_i^+(1)]=\infty$
(c) i is positive recurrent if $\mathbb{E}^i[\tau_i^+(1)]<\infty$

1.5 Criteria for Recurrence

We have a family of probability measures $(\mathbb{P}^i)_{i\in S}$, one for each initial state. Can also have a probability measure associated with an initial distribution μ, defined as

$$\mathbb{P}^\mu(A) = \sum_{i\in S} \mu_i \mathbb{P}^i(A)$$

Denote total number of visits to state i:

$$N_i = \sum_{n=1}^{\infty} 1_{\{X_n=i\}}$$

Proposition 1.12. *The following are equivalent*

(i) State i is recurrent

(ii) Infinite expected number of visits to state i: $\mathbb{E}[N_i] = \infty$

(iii) $\sum_n p_{ii}^{(n)} < \infty$

1.6 Recurrence of Random Walks

Example 1.13. The most classical example is that simple symmetric random walk (SSRW) is null recurrent. If asymmetric, still it is recurrent, the simplest proof is via SLLN. To make it positive recurrent, we can either: (1) use state dependent probabilities, or (2) restrict to interval $[a, b]$, i.e. "reflected RW"

1.7 Topology of Markov Chains

Definition 1.14 (Accessibility and Communication)**.**

(i) State j is accessible from i ($j \longrightarrow i$) if exist $m > 0$ such that $p_{ij}^{(m)} > 0$, equivalently if $\mathbb{P}^i[\tau_j < \infty] > 0$.

(ii) State i and j communicate if $i \longleftrightarrow j$.

Proposition 1.15. *Recurrence and transience are class properties.*

Proof. Let $i, j \in S$, $i \leftrightarrow j$, and i be recurrent ($\sum_n p_{ii}^{(n)} = \infty$). It suffice to show $\sum_n p_{jj}^{(n)} = \infty$. We can use the lower bound

$$p_{jj}^{(m+k+n)} \geq p_{ji}^{(m)} p_{ii}^{(n)} p_{ij}^{(k)}$$

Taking the sum on both sides over n yields the result. □

What is important is that "$\longleftrightarrow$" defines an equivalence relationship. So we can partition the state space as a disjoint union of recurrent and transient classes:

$$S = C_1 \cup C_2 \cup \cdots \cup T$$

Ususally we group the tansient classes into one, denoted by T.

Proposition 1.16. *We have*

(a) Recurrent classes are closed

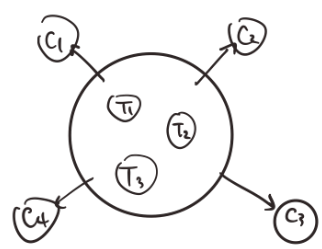
(b) if S finite, then at least one of its classes is recurrent.

Proposition 1.17 (Canonical Form of P)**.** *The transition matrix P can be writen as*

$$P = \begin{array}{c|ccc|c} & C_1 & C_2 & \cdots & T \\ \hline C_1 & P_1 & 0 & \cdots & \\ C_2 & 0 & P_2 & 0 & 0 \\ \vdots & 0 & 0 & \ddots & \\ \hline T & & R & & Q \end{array}$$

1.8 Method of First Step Decomposition and the Fundamental Equation

For simplicity, assume there is only one recurrent class: $S = C \cup T$. We should think of the recurrent class C as boundary points and the transient states T as the interior. Also let $f : C \to \mathbb{R}$ be a function. Can also think of f as a vector $(f_i)_{i \in C}$.



Define $u : T \to \mathbb{R}$, and as a vector $u = (u_i)_{i \in T}$ by

$$u_i = u(i) = \mathbb{E}^i[f(X_{\tau_C})]$$

where τ_C is the hitting time of the set C. Observe that for $i \in T$

$$u_i = \sum_{j \in T} P_{ij} u_j + \sum_{j \in C} P_{ij} u_j = \sum_{j \in T} Q_{ij} u_j + \sum_{j \in C} R_{ij} f_j$$

In matrix form: $u = Qu + Rf$. Then we can solve for u which by considering $(I - Q)u = Rf$.

Definition 1.18 (Fundamental Equation for Markov Chains)**.** The Fundamental Equation is given by

$$-\mathcal{A}u = Rf$$

where $\mathcal{A} = (Q - I)$ is called the generator for Markov Chain.

Remark 1.19. There is a deep connection between the generator and elliptic operators. Just take the SSRW in 1D, we calculate

$$[(P - I)u]_i = \frac{1}{2}(u(i-1) - 2u(i) + u(i+1))$$

This should remind us of the discrete 2nd order derivative. And for the SSRW in 2D, we have

$$[(P - I)u]_{i,j} = \frac{u(i+1, j) + u(i-1, j) + u(i, j+1) + u(i, j-1)}{4} - u(i, j)$$

This should remind us of the discrete Laplacian operator.

There is a result for solving the fundamental equation:

Theorem 1.20. *Assume for all i, $\mathbb{E}^i[\tau_C] < \infty$, and $f : C \to \mathbb{R}$ is bounded. Then*

$$u = FRf$$

is the unique bounded solution to the fundamental equation, where F is a portion of the potential matrix, given by

$$F = \sum_{n=0}^{\infty} Q^n$$

Proof. Claim 1: $F_{ij} < \infty$ for all $i, j \in T$. To see this

$$F_{ij} = \sum_n (Q^n)_{ij} = \sum_n (P^n)_{ij} = \mathbb{E}^i[\sum_n 1_{\{X_n=j\}}] \leq \mathbb{E}^j[\sum_n 1_{\{X_n=j\}}] = \sum_n p_{jj}^{(n)} < \infty.$$

The inequality is because starting from transient state i, there is a probabiliy of escaping into C, in which case, expected number of returns to state j is 0.

Claim 2: $\sum_j F_{ij} = \mathbb{E}^i[\tau_C] < \infty$. We can interpret F_{ij} as expected number of visits to j starting from i. Summing across j, we get the expected time until absorption, which is finite by assumption.

To be continued.. □

1.9 Exercises

Exercise 1.21 (Dur Ex 5.2.1)**.** Denote $H = \sigma(X_0, \ldots, X_n)$, $F = \sigma(X_n, X_n + 1, \ldots)$, then

$$\begin{aligned} P_\mu(H \cap F|X_n) = E_\mu[E_\mu(1_H 1_F|\mathcal{F}_n)|X_n] &= E_\mu[1_H E_\mu[1_F|\mathcal{F}_n]|X_n] \\ &= E_\mu[1_H E_\mu[1_F|X_n]|X_n] = P_\mu(H|X_n)P_\mu(F|X_n). \end{aligned}$$

2 CTMC

2.1 Basic Definitions

For DTMC, we saw we could assume the probability space is the sequence space S^∞. For continuous time, here is our set up for the probabilty space

- ❖ The canonical space

$$\Omega = \{\omega : [0,\infty) \to S| \quad \omega \text{ right continuous} \\ \text{with finitely many jumps on any finite time interval } [0,T]\}$$

- ❖ A trajectory of the markov chain is then

$$X(\cdot,\omega) = \omega(\cdot),$$

with the time shit operator $(\theta_s\omega)(t) = \omega(s+t)$. The σ-algebra on Ω is the smallest σ-algebra such that $\omega \to \omega(t)$ is measurable for each $t \geq 0$.

- ❖ We also need a collection of probability measures $\{\mathbb{P}^x\}_{x\in\Omega}$ on Ω and a right continuous filtration $\{\mathcal{F}_t, t\geq 0\}$ on $(\Omega,\mathcal{F})$, with respect to which X is adapted.

Definition 2.1 (Continuous Time Markov Chain)**.** We say X is a CTMC on S if X has the markov property

$$\mathbb{P}^x[Y\circ\theta_s|\mathcal{F}_s] = \mathbb{E}^{X(s)}Y \quad \mathbb{P}^x\text{-a.s.}$$

for all $x\in S$ and all bounded measurable Y on Ω.

Definition 2.2 (Transition Function)**.** A transition function $p_t(x,y)$, $t\in[0,\infty)$, $x,y\in S$ satisfies

(i) (Stochastic) $p_t(x,y)\geq 0$, $\sum_y p_t(x,y) = 1$

(ii) (Regularity) $p_t(x,x)\to 1$ and $p_t(x,y)\to 0$ as $t\downarrow 0$

(iii) (Chapman-Kolmogorov Eqn) $p_{s+t}(x,y) = \sum_z p_s(x,z)p_t(z,y)$.

Remark 2.3.

- ❖ On regularity conditions (ii): consider Cauchy Equation $\varphi : [0,\infty)\to\mathbb{R}$ satisfies $\varphi(t+s) = \varphi(t)+\varphi(s)$. There could be many solutions, but if we impose some sort of regularity(measurability or continuity), then we have a unique class of solution $\varphi(t) = at$, $a\in\mathbb{R}$.
- ❖ Notation: can use the operator notation to denote $(p_t(x,y))_{x,y}$ as P_t, then condition (ii) becomes $P_t\to I$ as $t\downarrow 0$. Chapman-Kolmogorov becomes $P_tP_s = P_{t+s}$ (the familiar semigroup property). Also necessarily, we have $P_0 = I$. This means the operator P_t is continuous at 0, some call this the C^0 semigroup.

Definition 2.4 (Q-Matrix)**.** A Q-matrix is $S\times S$ matrix such that

(i) Off-diagonal entries $q(x,y)\geq 0$ for $x\neq y$

(ii) Diagonal entries $q(x,x)\leq 0$ for all $x\in S$. (denote $c(x) = -q(x,x)$)

(iii) $\sum_{y\in S} q(x,y) = 0$.

As we shall see, Q-matrix is the generator.

2.2 Connections between CTMC, Transition Function and Q-matrix

Markov Chain to Transition Function

Theorem 2.5. *Given a markov chain, define*

$$p_t(x, y) = \mathbb{P}^x(X(t) = y)$$

then $p_t(x, y)$ is a transition function and it determines $\mathbb{P}^x$ uniquely.

Sketch proof. Just need to check continuity at 0. Right continuity of path implies first jumping time $\tau = \inf\{t > 0 : X(t) \neq X(0)\} > 0$ a.s. for any $\mathbb{P}^x$. Since $p_t(x, x) \geq \mathbb{P}^x(\tau > t)$, take $t \to 0$ gives the regularity condition.

For uniqueness, the $p_t(x, y)$ defined determines the finite-dimensional distribution of $\mathbb{P}^x$. By $\pi - \lambda$ theorem, it determines the full distribution. □

Proposition 2.6 (Properties of the Transition Function)**.**

(i) $p_t(x, x) > 0$, $\forall t \geq 0$, $x \in S$.

(ii) If $p_t(x, x) = 1$ *for some* t, x, *then* $p_t(x, x) = 1$ *for all* t, x.

(iii) For all $x, y \in S$, $p_t(x, y)$ *is uniformly continuous in* t. *In fact,*

$$|p_t(x, y) - p_s(x, y)| \leq 1 - p_{|t-s|}(x, x)$$

Proof. Basic all 3 comes from using the C-K equations. See [Lig Thm 2.13]. Potential Error: (iii) should be $|p_t(x, y) - p_s(x, y)| \leq 2(1 - p_{|t-s|}(x, x))$, the rest of proof still go through. □

Remark 2.7. Property (i) implies there is no periodicity problems in CTMC.

Theorem 2.8 (Differentiability of Transition Function)**.**

Let $p_t(x, y)$ be a transition function. Then

(i) For all x, the RHS derivative

$$c(x) = -q(x, x) = -\lim_{t \to 0} \frac{1}{t}(p_t(x, x) - 1) \in [0, \infty] \text{ exists.}$$

And it satisfies $p_t(x, x) \geq e^{-c(x)t}$.

(ii) If $c(x) < \infty$, *then*

$$q(x, y) = \lim_{t \to 0} \frac{1}{t} p_t(x, y) \in [0, \infty) \text{ exists}$$

(iii) If $c(x) < \infty$ *and* $\sum_y q(x, y) = 0$, *then* $p_t(x, y)$ *satisfies the Backward Kolmogorov Equation*

$$\frac{d}{dt} p_t(x, y) = \sum_z q(x, z) p_t(z, y)$$

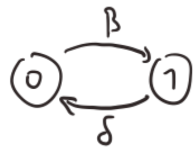
Or more compactly in operator notation, $\frac{d}{dt} P_t = Q P_t$.

Proof. Define $f(t) = -\log p_t(x, x)$. Note f is subadditive ($f(t + s) \leq f(t) + f(s)$). Then using Feteke's subadditive lemma, $c(x) = -\frac{d}{dt} e^{-f(t)}\Big|_{t=0} = e^{-f(t)} \frac{df(t)}{dt}\Big|_{t=0} = \frac{df(t)}{dt}\Big|_{t=0} \in [0, \infty]$ □

2.2.1 Blackwell's Example

This example constuct a transition function for which all states are instantaneous.

First consider a two-state chain with rate diagram:



So

$$Q = \begin{pmatrix} -\beta & +\beta \\ +\delta & -\delta \end{pmatrix},$$

By computing $P_t = e^{tQ}$, we can verify

$$P_t = \begin{pmatrix} \frac{\delta}{\beta+\delta} + \frac{\beta}{\beta+\delta} e^{-t(\beta+\delta)}, & \frac{\beta}{\beta+\delta}\left[1 - e^{-t(\beta+\delta)}\right] \\ \frac{\delta}{\beta+\delta}\left[1 - e^{-t(\beta+\delta)}\right], & \frac{\beta}{\beta+\delta} + \frac{\delta}{\beta+\delta} e^{-t(\beta+\delta)} \end{pmatrix}$$

Now consider a sequence of such chains X_i, $i \in \mathbb{N}$, with transition rates β_i and δ_i, and independent of each other. Let

$$X(t) = (X_1(t), X_2(t), \ldots) \in \{0,1\}^{\mathbb{N}}$$

Now consider the state space

$$S = \left\{ x \in \{0,1\}^{\mathbb{N}} : \sum_i x_i < \infty \right\}$$

Note that if $x \in S$, then there exist n such that for all $i > n$, $x_i = 0$.

Let $x, y \in S$, define

$$p_t(x,y) = \mathbb{P}^x(X(t) = y) = \prod_{i \geq 1} \mathbb{P}^{x_i}(X_i(t) = y_i)$$

Proposition 2.9.

(i) If $\sum_i \frac{\beta_i}{\beta_i+\delta_i} < \infty$, *then* $\mathbb{P}^x[X(t) \in S] = 1$, *for all* $t > 0$.

(ii) $p_t(x,y)$ *is a transition function on* S.

Proof. See [Lig Thm 2.17], Basically, (i) is an application of Borel-Cantelli Lemma. For (ii), we need to check the 3 properties. The book left checking the Chapman-Kolmogorov as an exercise [Lig Ex 2.18]. Here is solution sketch: First let $p_t^{(n)}(x,y) = \prod_{i=1}^n P^{x_i}(X_i(t) = y_i)$, we fist show $p_t^{(n)}(x,y)$ is a transition function. Property (1) and (2) are obvious, for C-K equation, write $a_{z_i} = P^{x_i}(X_i(s) = z_i) P^{z_i}(X_i(t) = y_i)$, and noting $\sum_{z_i} a_{z_i} = P^{x_i}(X_i(s+t) = y_i)$, we have

$$\begin{aligned} p_{t+s}^n(x,y) &= \prod_{i=1}^n P^{x_i}(X_i(t+s) = y_i) \\ &= \sum_{z_n} \cdots \sum_{z_1} a_{z_1} a_{z_2} \cdots a_{z_n} = \sum_z p_t^n(x,z) p_t^n(z,y) \end{aligned}$$

Now we establish pointwise convergence. Observe for any $x, y \in S$, for all $t > 0$,

$$
\begin{aligned}
p_t^n(x,y) - p_t(x,y) &= \prod_{i=1}^{n} P^{x_i}(X_i(t) = y_i) \left[1 - \prod_{i=n+1}^{\infty} P^{x_i}(X_i(t) = y_i)\right] \\
&\leq 1 - \prod_{i=n+1}^{\infty} P^{x_i}(X_i(t) = y_i) \to 0 \text{ as } n \to \infty
\end{aligned}
$$

Then

$$
p_{t+s}(x,y) = \lim_{n\to\infty} p_{t+s}^n(x,y) = \lim_{n\to\infty} \sum_z p_s^n(x,z) p_t^n(z,y) = \sum_z p_s(x,z) p_t(z,y)
$$

- we used DCT. Note $p_s^n(x,z)p_t^n(z,y) \leq p_s^n(x,z)$ and $\sum_z p_s^n(x,z) = 1$.

$\square$

Proposition 2.10. *Suppose* $\sum_i \frac{\beta_i}{\beta_i + \delta_i} < \infty$ *and* $\sum_i \beta_i = \infty$ *,then*

(i) all states are instantaneous ($c(x) = \infty$)

(ii) For all $x \in S$, $\mathbb{P}^x(X(t) \in S$ for all $t \geq 0) = 0$,

Proof. For $x \in S$, choose m so that $x_i = 0$ for $i \geq m$ (can do this by definiton of S). Then for $n \geq m$,

$$
\begin{aligned}
1 - p_t(x,x) &= 1 - \prod_i P^{x_i}(X_i(t) = x_i) \geq 1 - \prod_{i=m}^{n} P^0(X_i(t) = 0) \\
&= 1 - \prod_{i=m}^{n} \left[\frac{\delta_i}{\beta_i + \delta_i} + \frac{\beta_i}{\beta_i + \delta_i} e^{-t(\beta_i + \delta_i)}\right] \\
&\sim 1 - \prod_{i=m}^{n} (1 - \beta_i t) \geq 1 - \exp\left(-t \sum_{i=m}^{n} \beta_i\right),
\end{aligned}
$$

Then $c(x) \geq \sum_{i=m}^n \beta_i$ (this uses the fact that for any transition function: $p_t(x,x) \geq e^{-c(x)t}$), then take $n \to \infty$.

For (ii), Let $A = \{X(t) \in S \text{ for all } t \geq 0\}$,

- X_i's are independent of each other. A is a tail event. Can apply Kolmogorov 0-1 Law, which says $P(A) = 0$ or 1.
- For the sake of contradiction, say $P(A) = 1$. Let $A(y) = \{t \geq 0 : X(t) = y\}$. Note w.p 1, $\cup_{y \in S} A(y) = [0, \infty)$.
- Claim that w.p 1, $A(y;\omega)$ is no-where dense in $[0,\infty)$ for each $y \in S$.
  - Suppose it is dense w.p.1 in open subset (a,b) for some y. From right-continuity of paths, $X(t,\omega) = y$ for a dense subset of (a,b) implies $X(t,\omega) = y$ for all $t \in [a,b)$.
  - Suppose $y_i = 0$ for all $i \geq m$, since holding times are exponential

$$
\begin{aligned}
P^x(X(t) = y &\text{ for dense subset of } [a,b)) \leq P^x(X(t) = y \text{ for all } t \in [a,b)) \\
&= \prod_i [P^{x_i}(X_i(a) = y_i) \cdot P(X_i(t) = y_i \text{ for all } t \in [a,b) \mid X_i(a) = y_i)] \\
&\leq \prod_{i=m}^{\infty} e^{-\beta_i(b-a)} = e^{-(b-a)\sum \beta_i} = 0.
\end{aligned}
$$

which is a contradiction.

- By Baire's Category Theorem, $[0, \infty)$ cannot be countable union of no-where dense sets, so this leads to contradiction.

$\square$

3 Feller Process

3.1 Operator Semigroup

- Let E be a locally compact Hausdorff space. (that is $\forall x \in E$, x has a compact neighborhood).
- Let $(\mathcal{X}, \|\cdot\|)$ be a Banach Space. And denote $\mathcal{L}(\mathcal{X})$ to be the space of bounded linear operators on $\mathcal{X}$.

Topologies on $\mathcal{L}(\mathcal{X})$

(1) Weak Operator Topology (WOT)
  - Equip $\mathcal{X}$ with the weak topolgy. (recall weak convergence: $f^* \in \mathcal{X}^*$, $\langle f^*, T_n f\rangle \to \langle f^*, Tf\rangle$)
  - The coarsest topology on $\mathcal{L}(\mathcal{X})$ such that all maps: $T \mapsto Tf$ are continuous, where $T \in \mathcal{L}(\mathcal{X})$ and $f \in \mathcal{X}$.

(2) Strong Operator Topology (SOT)
  - Equip $\mathcal{X}$ with the strong topolgy.
  - The coarsest topology on $\mathcal{L}(\mathcal{X})$ such that all maps: $T \mapsto Tf$ are continuous, where $T \in \mathcal{L}(\mathcal{X})$ and $f \in \mathcal{X}$.
  - In other words, this is the topology with pointwise convergence: $T_n \to T$ if $T_n f \to Tf$ in $\mathcal{X}$, for all $f \in \mathcal{X}$.

(3) Norm Topology (Uniform Topology)
  - Topology induced by the operator norm ($\|T\| = \sup\{\|Tf\| : \|f\| \leq 1\}$).
  - $T_n \to T$ if $\|T_n - T\| \to 0$.

Definition 3.1 (Monoid)**.**

- An algebraic semigroup is a pairing $(M, \circ)$, where M is an non-empty set and $\circ$ is an associative binary operation $M \times M \to M$.
- M is a monoid if it is an algebraic semigroup and have an unit element. i.e. $\exists e \in M$ s.t. $e \circ a = a \circ e = a$, $\forall a \in M$.
- A topological monoid is a monoid with a topology, in which $\circ$ is continuous.

Definition 3.2 (Algebraic Representation)**.** let M be a monoid. A map $T : M \to \mathcal{L}(\mathcal{X})$ is called an algebraic representation if

(1) $T(e) = Id$

(2) $T(a \circ b) = T(a)T(b)$, for all $a, b \in M$.

If in addition, M is a topological monoid and $a \mapsto T(a)$ is continuous when $\mathcal{L}(\mathcal{X})$ is given the strong operator topology, then we say T is strongly continuous represenatiton.