Chapter III. Continuous time Markov process

3.1 Poisson processes

Def: A random point process on $\mathbb{R}_+ = [0,\infty)$ is a sequence of random variables $\{T_n\}_{n\geq 0}$ ~~such that~~ ordered as $0 = T_0 \leq T_1 \leq T_2 \leq \cdots$.

[$\{T_0, T_1, \cdots\}$ gives a countable random set of points, which can be regarded as the sequence of times of occurence of some events.]

Given a point process, we define the interevent sequence: $\tau_n = T_n - T_{n-1}$, $n \geq 1$.

It is convenient to characterize a random point process on $\mathbb{R}_+$ through its counting process.

Def: Given a random point process $\{T_n\}_{n\geq 0}$ on $\mathbb{R}_+$, define its counting process N_t as follows: $N_t = \#\{n : T_n \in (0,t]\} = \sum_{n\geq 1} \mathbb{1}_{(0,t]}(T_n) = \sum_{n\geq 1} \mathbb{1}(T_n \leq t)$.

Note that under this definition, $N_0 = 0$; $\#\{n\geq 1 : T_n \in (a,b]\} = N_b - N_a$, the number of events between a and b; $t \mapsto N_t$ is a right-continuous step function.

One of the most standard random point process to describe the random arrival of customers or random occurence of events is the Poisson Point Process (PPP) or simply the "Poisson Process" (PP).

Def: A counting process N_t is a Poisson Process with intensity $\lambda > 0$ if

(i) for any $k \geq 1$, and $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$, the random variables $N_{t_{i+1}} - N_{t_i}$, $i = 1, 2, \cdots, k-1$, are independent;

(ii) for any $s, t > 0$, $N_{t+s} - N_s$ has distribution Poisson(λt), i.e.,

$$\mathbb{P}(N_{t+s} - N_s = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \cdots;$$

(iii) $N_0 = 0$ a.s.

Lemma: Let $\{N_t\}_{t\geq 0}$ be a PP with intensity λ. Then, $\forall s \geq 0$, $\{\tilde{N}_t\}_{t\geq 0} = \{N_{t+s} - N_s\}_{t\geq 0}$ is also a PP with intensity λ and is independent of $\{N_u\}_{u\leq s}$.

Pf: Check that $\tilde{N}_t$ also satisfies the above three properties. □

Construction of the Poisson process: $\boxed{f(x) = \lambda e^{-\lambda x}}$

Theorem: Consider a sequence of i.i.d. $\operatorname{Exp}(\lambda)$ random variables $\tau_1, \tau_2, \tau_3, \cdots$. Define the random points $T_n = \sum_{i=1}^{n} \tau_i$. Then, the counting process N_t associatad with $\{T_n\}_{n\geq 1}$ is a Poisson Process with intensity λ.

Pf: ~~Since $T_1 = \tau_1 > 0$ a.s., we have~~ We trivially have $N_0 = 0$.

Next, we show that N_t has $\operatorname{Poisson}(\lambda t)$ distribution. First, we claim that T_n has $\operatorname{Gamma}(n, \lambda)$ distribution, i.e., T_n has PDF

$$(*) \quad f_n(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$$

We prove $(*)$ by induction. First, $(*)$ holds for $n=1$. ~~Then~~ Next, suppose T_n has distribution f_n. Then, $T_{n+1} = T_n + \tau_{n+1}$ has PDF $f_n \circ f_1$:

$$f_n \circ f_1 (t) = \int_0^t f_n(s) f_1(t-s)\, ds = \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} \cdot \lambda e^{-\lambda(t-s)}\, ds$$

$$= \lambda e^{-\lambda t} \frac{\lambda^n}{(n-1)!} \int_0^t s^{n-1}\, ds = \lambda e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

This concludes $(*)$ by induction in n.

Now, the distribution of N_t is determined by:

$$\mathbb{P}(N_t \geq n) = \mathbb{P}(T_n \leq t) = \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!}\, ds =: F(t)$$

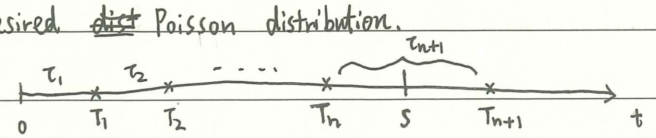
We want it to be equal to $G(t) = \sum_{m=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^m}{m!}$. First, $G(0) = F(0) = 0$.

Second, $G'(t) = \sum_{m=n}^{\infty} (-\lambda) e^{-\lambda t} \frac{(\lambda t)^m}{m!} + \sum_{m=n}^{\infty} e^{-\lambda t} \frac{\lambda^m t^{m-1}}{(m-1)!}$

$$= -\sum_{m=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^m}{m!} + \sum_{m=n-1}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^m}{m!} = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} = \frac{d}{dt} F(t) = f_n(t).$$

Thus, we see that $N_t \sim \operatorname{Poisson}(\lambda t)$.

Finally, we show that N_t has independent ~~and~~ increaments, and the increments satisfy the desired ~~dist~~ Poisson distribution.



$n = n(\omega) = \sup\{n : T_n \leq s\}$.

The key to the proof is the memoryless property of the exponential distribution: $X \sim \operatorname{Exp}(\lambda)$ satisfies $\mathbb{P}(X > t+s \mid X > s) = \mathbb{P}(X > t)$, $\forall\, t \in \mathbb{R}$.

Rmk: Exp distribution is the only continuous distribution satisfying the memoryless property.

Fix $t > 0$, $s \geq r > q$, we want to show that $N_{t+s} - N_s$ is independent of $N_r - N_q$.

(Similar argument can be extended to multiple increments.)

$$\mathbb{P}(N_{t+s} - N_s \leq m,\ N_r - N_q = \ell) = \sum_{n=0}^{\infty} \mathbb{P}(T_n \leq s < T_{n+1},\ T_{n+m} \geq t+s,\ N_r - N_q = \ell)$$

$$= \sum_n \mathbb{P}(T_n \leq s,\ \tau_{n+1} > s - T_n,\ N_r - N_q = \ell,\ T_n + \tau_{n+1} + \cdots + \tau_{n+m} \geq t+s)$$

$$\cancel{= \sum_n \mathbb{P}(\tau_{n+1} + \cdots + \tau_{n+m} \geq t + (s - T_n) \mid T_n \leq s,\ \tau_{n+1} > s - T_n,\ N_r - N_q = \ell)}$$

$$\cancel{\mathbb{P}(T_n \leq s,\ \tau_{n+1} > s - T_n,\ N_r - N_q = \ell)} \qquad \mathcal{F}_n = \sigma(\tau_1, \cdots, \tau_n)$$

$$= \sum_n \mathbb{E}\left[\mathbb{1}(T_n \leq s,\ N_r - N_q = \ell)\, \mathbb{P}(\tau_{n+1} + \cdots + \tau_{n+m} \geq t + (s - T_n),\ \tau_{n+1} > s - T_n \mid \mathcal{F}_n) \right]$$

$$= \sum_n \mathbb{E}\left(\mathbb{1}(T_n \leq s,\ N_r - N_q = \ell)\, \mathbb{P}(\tau_{n+1} + \cdots + \tau_{n+m} \geq t \mid \mathcal{F}_n)\, \mathbb{P}(\tau_{n+1} > s - T_n \mid \mathcal{F}_n) \right]$$

$$= \mathbb{P}(\tau_{n+1} + \cdots + \tau_{n+m} \geq t) \sum_n \mathbb{P}(T_n \leq s,\ N_r - N_q = \ell,\ \tau_{n+1} > s - T_n)$$

$$= \mathbb{P}(\tau_{n+1} + \cdots + \tau_{n+m} \geq t)\, \mathbb{P}(N_r - N_q = \ell)$$

~~This shows the independence between N~~ Summing over ℓ, we get that

$$\mathbb{P}(N_{t+s} - N_s \leq m) = \mathbb{P}(\tau_{n+1} + \cdots + \tau_{n+m} \geq t) = \mathbb{P}(\tau_1 + \cdots + \tau_m \geq t)$$

By the above results, this implies that $N_{t+s} - N_s$ has the same distribution as $N_t \sim \operatorname{Poisson}(\lambda t)$.

In addition, it gives $\mathbb{P}(N_{t+s} - N_s \leq m,\ N_r - N_q = \ell) = \mathbb{P}(N_{t+s} - N_s \leq m)\, \mathbb{P}(N_r - N_q = \ell)$, i.e., $N_{t+s} - N_s$ is independent of $N_r - N_q$. $\square$

The above theorem ~~the Poisson process~~ gives a construction of the Poisson process through i.i.d. $\operatorname{Exp}(\lambda)$ interevent sequence $\tau_1, \tau_2, \cdots$. Conversely, given a Poisson process N_t, define the random points $T_n = \inf\{t \geq 0 : N_t = n\}$. Then, the corresponding interevent times $\tau_n = T_n - T_{n-1}$ are i.i.d. $\operatorname{Exp}(\lambda)$ random variables.

Thm: The ~~intervent~~ interevent sequence $\{\tau_n\}_{n\geq 1}$ of a Poisson process with intensity λ is i.i.d. with $\operatorname{Exp}(\lambda)$ distribution.

Proof: This amounts to showing that $\forall\, n \geq 1$, the density function for $(\tau_1, \cdots, \tau_n)$ is

$$f_{\tau}(\tau_1, \cdots, \tau_n) = \prod_{i=1}^{n} \lambda e^{-\lambda \tau_i} = \lambda^n e^{-\lambda T_n}.$$

It is equivalent to showing that the density function of $(T_1, \cdots, T_n)$ is

$$f_{\vec{T}}(t_1, \cdots, t_n) = \lambda^n e^{-\lambda t_n} \mathbb{1}\{0 < t_1 < t_2 < \cdots < t_n\}.$$

We can calculate directly that for $0 < t_1 < t_2 < \cdots < t_n$,

$$f_{\vec{T}}(t_1, \cdots, t_n) = \lim_{\varepsilon \to 0} \varepsilon^{-n} \mathbb{P}(T_1 \in (t_1, t_1+\varepsilon], \cdots, T_n \in (t_n, t_n+\varepsilon]).$$

For small enough $\varepsilon > 0$, we have

$$\mathbb{P}(T_i \in (t_i, t_i+\varepsilon], 1 \le i \le n) = \mathbb{P}(N_{t_1} = 0, N_{t_1+\varepsilon} - N_{t_1} = 1, N_{t_2} - N_{t_1+\varepsilon} = 0,$$

$$N_{t_2+\varepsilon} - N_{t_2} = 1, \cdots, N_{t_n+\varepsilon} - N_{t_n} \ge 1)$$

$$= e^{-\lambda t_1} \prod_{i=1}^{n-1} \underbrace{\lambda\varepsilon e^{-\lambda\varepsilon}}_{\mathbb{P}(N_{t_i+\varepsilon} - N_{t_i} = 1)} \cdot \underbrace{e^{-\lambda(t_{i+1} - t_i - \varepsilon)}}_{\mathbb{P}(N_{t_{i+1}} - N_{t_i+\varepsilon} = 0)} \times (1 - e^{-\lambda\varepsilon})$$

$$= \lambda^{n-1} \varepsilon^{n-1} e^{-\lambda t_n} (1 - e^{-\lambda\varepsilon})$$

Hence, $f_{\vec{T}}(t_1, \cdots, t_n) = \lim_{\varepsilon \to 0} \lambda^{n-1} e^{-\lambda t_n} \frac{1 - e^{-\lambda\varepsilon}}{\varepsilon} = \lambda^n e^{-\lambda t_n}$. □

Thm: Almost surely, $N_t - N_{t-} \in \{0, 1\}$ $\forall t \in [0, \infty)$. In other words, almost surely, Poisson process has no double jump (or multi-jump) at any time.

Pf: Fix an arbitrary $T > 0$. Consider a partition $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ with $t_i = \frac{i}{n}T$.

If $N_t - N_{t-} \ge 2$, then there exists t_i s.t. $N_{t_{i+1}} - N_{t_i} \ge 2$. In other words,

$$\mathbb{P}(\exists t \in [0, T] \text{ s.t. } N_t - N_{t-} \ge 2) \le \mathbb{P}(\exists\, 0 \le i \le n-1 \text{ s.t. } N_{t_{i+1}} - N_{t_i} \ge 2)$$

$$\le \sum_{i=0}^{n-1} \mathbb{P}(N_{t_{i+1}} - N_{t_i} \ge 2) = \sum_{i=0}^{n-1} \left(1 - e^{-\lambda T/n} - \frac{\lambda T}{n} e^{-\lambda T/n}\right) = O\left(n \left(\frac{\lambda T}{n}\right)^2\right) = O(\lambda^2 T^2 / n).$$

Taking $n \to \infty$, we get $\mathbb{P}(\exists t \in [0, T] \text{ s.t. } N_t - N_{t-} \ge 2) = 0$. □

Thm (Superposition of Poisson processes) Let $\{N_t^i\}$ be a family of independent processes with intensities $\{\lambda_i\}$. ~~Then~~ If $\sum_i \lambda_i = \lambda$, then $N_t = \sum_i N_t^i$ is a Poisson process with intensity λ.

Pf: First, it is trivial to see that N_t has independent increments since each N_t^i has independent increments. Second, $N_{t+s} - N_s = \sum_i (N_{t+s}^i - N_s^i)$ is a sum of independent Poisson$(\lambda_i t)$ random variables. By the property of Poisson distribution, $N_{t+s} - N_s$ has Poisson$(\sum_i \lambda_i t)$ = Poisson(λt) distribution. □

Thm (Thinning of Poisson processes) Given a Poisson process N_t with intensity λ. We associate to ~~each~~ the i-th arrival / point / event a random variable Y_i. The random variables Y_i are i.i.d with distribution $\mathbb{P}(Y_i = C_j) = p_j$, $1 \le j \le m$. (You can consider C_j's as colors of each event.) We define the Poisson processes N_t^j of colors C_j as:

$$N_t^j = \sum_{n \ge t} \mathbb{1}(T_n \le t, Y_n = C_j).$$

Then, $\{N_t^j\}$ are independent Poisson processes with rates λp_j.

Pf: First, by the independent increments property for N_t, the increments

$$(N_{t_{i+1}}^1 - N_{t_i}^1, \cdots, N_{t_{i+1}}^m - N_{t_i}^m), \quad 0 < t_1 < t_2 < \cdots < t_k, \text{ are independent random } m\text{-tuples}.$$

It remains to show that the random variables $N_{t+s}^1 - N_s^1, \cdots, N_{t+s}^m - N_s^m$ are independent random variables with the desired distribution Poisson$(\lambda p_j t)$, $1 \le j \le m$.

For simplicity, ~~dentte~~ denote $X_j = N_{t+s}^j - N_s^j$, $1 \le j \le m$.

We need to calculate $(*) = \mathbb{P}(X_1 = k_1, \cdots, X_m = k_m)$ for $k_1, \cdots, k_m \in \mathbb{N}$.

Recall ~~Consi~~ ~~Using~~ the multi-nomial distribution describes the ~~property~~ probability that k_j of the $k = \sum_{j=1}^m k_j$ balls are of color j if the k balls are assigned colors according $\mathbb{P}(Y_i = C_j) = p_j$:

$$\frac{k!}{k_1! \cdots k_m!} p_1^{k_1} \cdots p_m^{k_m}.$$

Hence,

$$\mathbb{P}(X_1 = k_1, \cdots, X_m = k_m) = e^{-\lambda t} \frac{(\lambda t)^k}{k!} \frac{k!}{k_1! \cdots k_m!} p_1^{k_1} \cdots p_m^{k_m}$$

$$= \prod_{j=1}^m e^{-\lambda p_j t} \frac{(p_j \lambda t)^{k_j}}{k_j!} = \prod_{j=1}^m \mathbb{P}(\text{Poisson}(p_j \lambda t) = k_j).$$

This shows both independence and the distribution of $N_{t+s}^j - N_s^j$. □

~~We can define uniform distribution on [0,~~ Poisson process ~~on [0,∞)~~ can be regarded as "uniform point process" on $[0, \infty)$. We will show that conditioning on $N_t = n$, i.e., there are n events inside $[0, t]$, then the unordered times of these events are uniformly distributed on $[0, t]$.

To see how uniform distribution arises, suppose $N_t = 1$:

$$\mathbb{P}(T_1 \le s \mid N_t = 1) = \frac{\mathbb{P}(T_1 \le s, N_t = 1)}{\mathbb{P}(N_t = 1)} = \frac{\mathbb{P}(N_s = 1, N_t - N_s = 0)}{\mathbb{P}(N_t = 1)} = \frac{\lambda s e^{-\lambda s} \cdot e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} = \frac{s}{t}.$$

More generally, let $u_1, u_2, \cdots, u_n$ be an independent sequence of random variables distributed on $[0,t]$. Arrange them in increasing order $u_{(1)} \le u_{(2)} \le \cdots \le u_{(n)}$. This is called the order statistics of $u_1, \cdots, u_n$. The joint density of $(u_{(1)}, \cdots, u_{(n)})$ is

$$f(u_1, \cdots, u_n) = \frac{n!}{t^n}, \quad \text{for } 0 \le u_1 \le u_2 \le \cdots \le u_n \le t.$$

Thm: Conditioning on $N_t = n$, the vector $(T_1, T_2, \cdots, T_n)$ has the same distribution as $(u_{(1)}, \cdots, u_{(n)})$.

Pf: For $t_0 = 0 < t_1 < t_2 < \cdots < t_n < t = t_{n+1}$, the joint density of $(T_1, T_2, \cdots, T_n)$ conditioning on $N_t = n$ at $(t_1, \cdots, t_n)$ is

$$\lim_{\varepsilon \to 0} \frac{\mathbb{P}(T_1 \in (t_1, t_1+\varepsilon], \cdots, T_n \in (t_n, t_n+\varepsilon] \mid N_t = n)}{\varepsilon^n}$$

$$\mathbb{P}(T_1 \in (t_1, t_1+\varepsilon], \cdots, T_n \in (t_n, t_n+\varepsilon] \mid N_t = n) = \frac{\mathbb{P}(T_i \in (t_i, t_i+\varepsilon], 1 \le i \le n, N_t = n)}{\mathbb{P}(N_t = n)}$$

For small enough $\varepsilon > 0$, we have

$$\mathbb{P}(T_i \in (t_i, t_i+\varepsilon], 1 \le i \le n, N_t = n) = \mathbb{P}(N_{t_1} = 0, N_{t_1+\varepsilon} - N_{t_1} = 1, \cdots, N_{t_n} - N_{t_{n-1}+\varepsilon} = 0,$$
$$N_{t_n+\varepsilon} - N_{t_n} = 1, N_t - N_{t_n+\varepsilon} = 0) = e^{-\lambda t_1} \prod_{i=1}^{n} \lambda \varepsilon e^{-\lambda \varepsilon} \cdot e^{-\lambda (t_{i+1} - t_i - \varepsilon)}$$
$$= \lambda^n \varepsilon^n e^{-\lambda t}$$

$$\Rightarrow \mathbb{P}(T_i \in (t_i, t_i+\varepsilon], 1 \le i \le n \mid N_t = n) = \frac{\lambda^n \varepsilon^n e^{-\lambda t}}{\frac{(\lambda t)^n}{n!} e^{-\lambda t}} = \frac{n!}{t^n} \varepsilon^n.$$

$\Rightarrow$ Hence, the conditional density is $f(t_1, \cdots, t_n) = \frac{n!}{t^n} \mathbb{1}_{0 < t_1 < \cdots < t_n < t}$. □

Corollary: $\mathbb{P}(N_s = m \mid N_t = n) = \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m}$.

The Poisson process on $[0,\infty)$ can be extended to Poisson Point Processes (PPP) on arbitrary measure space.

Def: Let $(S, \mathcal{A}, \mu)$ be a σ-finite measure space. A process $\{N(A) : A \in \mathcal{A}\}$ indexed by the measurable sets is a PPP with mean measure μ if

(i) Almost surely, $N(\cdot)$ is a $\mathbb{N} \cup \{\infty\}$-valued measure on $(S, \mathcal{A})$.

(ii) $N(A)$ is Poisson distributed with parameter $\mu(A)$.

(iii) For any disjoint sets $A_1, \cdots, A_n \in \mathcal{A}$, the random variables $N(A_1), \cdots, N(A_n)$ are independent.

The Poisson process $[0,\infty)$ is a special case with $S = [0,\infty)$, $\mathcal{A}$ the Borel/Lebesgue σ-field, and μ the Lebesgue measure.

When people refer to the d-dim PPP, they refer to the case where $S = \mathbb{R}^d$, $\mathcal{A}$ the Borel σ-field, and $\mu = \mathcal{L}$ the Lebesgue measure.

Theorem: Let $(S, \mathcal{A}, \mu)$ be a σ-finite measure space. Then, a PPP $\{N(A) : A \in \mathcal{A}\}$ with mean measure μ exists.

Its proof is left as an exercise.

3.2 Continuous Time Markov Process: discrete state space S

Consider a Markov process $\{X_t : t \in [0,\infty)\}$ on a finite/countable state space S.

We first give a more elementary definition.

$\boxed{\text{Def I}}$: $\{X_t\}_{t \ge 0}$ is a continuous time Markov process if $\forall\, t_0 = 0 \le t_1 \le t_2 \le \cdots \le t_n \le t_{n+1}$ and $x_0, x_1, x_2, \cdots, x_n, x_{n+1} \in S$, we have

$$\mathbb{P}[X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n, \cdots, X_{t_1} = x_1, X_{t_0} = x_0] = \mathbb{P}[X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n]$$
$$= p_{t_{n+1} - t_n}(x_n, x_{n+1}).$$

Here, $p_t(x,y)$ is called the transition function satisfying the following properties for all $t \ge 0$ and $x, y \in S$:

$$p_t(x,y) \ge 0, \quad \sum_{y \in S} p_t(x,y) = 1, \quad p_0(x,x) = \lim_{t \downarrow 0} p_t(x,x) = 1 \quad \text{and}$$

$$p_{s+t}(x,y) = \sum_{z \in S} p_s(x,z) p_t(z,y). \quad \text{(Chapman-Kolmogorov equation)}$$

$p_{s+t}(x,y) =$

Rmk: $\mathbb{P}(X_{t+s} = y \mid X_0 = x) = \sum_z \mathbb{P}(X_{t+s} = y \mid X_s = z, X_0 = x)\, \mathbb{P}(X_s = z \mid X_0 = x)$
$$= \sum_z \mathbb{P}(X_{t+s} = y \mid X_s = z)\, p_s(x,z) = \sum_z p_s(x,z)\, p_t(z,y).$$

Example: Poisson process is a Markov process:

$$\mathbb{P}[N_{t_{n+1}} = k_{n+1} \mid N_{t_n} = k_n, \cdots, N_{t_1} = k_1] = \mathbb{P}[N_{t_{n+1}} - N_{t_n} = k_{n+1} - k_n \mid N_{t_n} = k_n]$$
$$= \mathbb{P}[N_{t_{n+1}} - N_{t_n} = k_{n+1} - k_n].$$

Its transition function is $p_t(x,y) = e^{-\lambda t} \frac{(\lambda t)^{y-x}}{(y-x)!} \mathbb{1}_{y \ge x}$.

The CK equation guarantees that these finite-dimensional distributions are consistent.

From the transition function, we can define the finite-dimensional distributions of the chain as follows:

$$\mathbb{P}_x[X_{t_1}=x_1, \cdots, X_{t_n}=x_n] = p_{t_1}(x, x_1)\, p_{t_2-t_1}(x_1, x_2) \cdots p_{t_n-t_{n-1}}(x_{n-1}, x_n).$$

Troughout this course, we assume the following regularity condition: the process is right-continuous.

Here is another definition in terms of the Markov property.

Def II: Let Ω = the set of right continuous functions $w: [0,\infty) \to S$ with finitely many jumps in any $[0,T]$, $\forall\, T \geq 0$.

Define $X(t,w) = w(t)$ and $\theta_s : \Omega \to \Omega$ as $(\theta_s w)(t) = w(t+s)$. The σ-algebra $\mathcal{F}$ on Ω is generated by the coordinate maps $X_t : w \mapsto w(t)$. ~~and~~ Then, ~~the~~ a $\underbrace{\text{Markov chain}}_{\text{continuous-time}}$ on S consists of

(a) a collection of probability measures $\{P^x : x \in S\}$ on Ω,

(b) a right-continuous filtration $\{\mathcal{F}_t : t \geq 0\}$ on $(\Omega, \mathcal{F})$ with respect to which the process X_t is adapted (i.e., $X_t \in \mathcal{F}_t$),

such that the followings hold $\forall\, x \in S$:

(1) $\mathbb{P}_x(X_0 = x) = 1$ ~~and~~

(2) For any bdd measurable function Y on Ω,

$$\mathbb{E}_x(Y \circ \theta_s \mid \mathcal{F}_s) = \mathbb{E}_{X_s} Y \quad \text{a.s.} \quad \mathbb{P}_x\text{-a.s.} \qquad (\text{Markov property})$$

(Strong Markov Property)

Thm: Let X_t be a continuous time Markov chain defined as in the above Def II. Let $Y_s(w)$ be a bounded and jointly measurable function on $[0,\infty) \times \Omega$. Let τ be a stopping time w.r.t. to the filtration $\{\mathcal{F}_t\}$. Then,

$$\mathbb{E}_x(Y_\tau \circ \theta_\tau \mid \mathcal{F}_\tau) = \mathbb{E}_{X_\tau} Y_\tau \quad \mathbb{P}_x\text{-a.s.} \quad \text{on } \{\tau < \infty\}.$$

Pf: The proof of this SMP is similar to that of the BM. In that proof, we used the MP (which is contained in the definition), the right continuity of the paths, and the Feller property ← This is trivial since all functions on S are continuous. □

The Markov chain is also described infinitesimally in time in terms of the Q-matrix defined as: $q(x,y) = \frac{d}{dt} p_t(x,y)\big|_{t=0}$

Def (Q-matrix) A Q-matrix $Q = (q(x,y) : x, y \in S)$ is a real matrix whose entries satisfy that $q(x,y) \geq 0$ for $x \neq y$ and $\sum_y q(x,y) = 0$. In particular, the diagonal terms are ≤ 0 and we denote them as: $\boxed{c(x) := -q(x,x).}$

When S is finite, there are always 1-1 correspondances among the MP, transition functions, and Q-matrices: MP $\rightleftarrows$ ~~transi~~ $p_t(x,y) \rightleftarrows q(x,y)$

However, this is not necessarily the case if $|S| = \infty$.

We will see that there is no problem in: MP $\to p_t(x,y) \to q(x,y)$. However, for the reverse direction to hold, we need some extra "mild" conditions on the Q-matrix.

Example Given a discrete time MC with transition probability $p(x,y)$. We can get a continuous time MC by assuming that transitions between different states occur at the event times of a Poisson process. In other words, the ~~waiter~~ time between transitions are i.i.d. Exp(λ) random variables. The transition function is

$$p_t(x,y) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} p^k(x,y).$$

The transition matrix is given by: $q(x,x) = -\lambda(1 - p(x,x))$,

$q(x,y) = \lambda p(x,y)$ if $x \neq y$.

3.3 MP $\to$ transition function $\to$ Q matrix (The easier direction)

Theorem (Markov property $\to$ transition function)

Given a continuous time MC satisfying the Markov property. Let

$$p_t(x,y) = \mathbb{P}_x(X_t = y). \quad \text{Then}$$

(a) p_t is a transition function, and

(b) p_t determines the measure $\mathbb{P}_x$ uniquely.

Pf: (a) It is trivial to see that $p_t(x,y) \geq 0$ and $\sum_y p_t(x,y) = 1$.

By the right continuity of paths and DCT, $\lim_{t\downarrow 0} p_t(x,x) = \lim_{t\downarrow 0} \mathbb{P}_x(X_t = x)$

$= \mathbb{P}_x(\lim_{t\downarrow 0} X_t = x) = 1$. For the CK equation, using the MP, we get

$$\mathbb{P}_x(X_{t+s} = y \mid \mathcal{F}_s) = \mathbb{P}_{X_s}(X_t = y) = p_t(X_s, y) \quad \text{a.s.}$$

Taking expectation gives the CK equation.

(b) The transition function gives the finite-dimensional distributions of $\mathbb{P}_x$.

By π-λ thm, the prob. measure on $(\Omega, \mathcal{F})$ are determined by them. □

Thm (Some properties of the transition function) Suppose p_t is a transition function.

(a) $p_t(x,x) > 0$ for all $t \geq 0$ and $x \in S$.

(b) If $p_t(x,x) = 1$ for some $t > 0$ and $x \in S$, then $p_t(x,x) = 1$ for all $t > 0$ and that x.

(c) For every $x, y \in S$, $p_t(x,y)$ is uniformly continuous in t. In fact,

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

Proof: (a) We know that $p_t(x,x) > 0$ for small $t > 0$. By the CK equation,

$$(*) \quad p_{s+t}(x,x) \geq p_s(x,x)\, p_t(x,x), \text{ so we can extend the positivity to all } t > 0.$$

(b) $p_{s+t}(x,x) = \sum_z p_s(x,z) p_t(z,y) \leq p_s(x,x) p_t(x,x) + \sum_{z \neq x} p_s(x,z) = p_s(x,x) p_t(x,x) + [1 - p_s(x,x)]$

$= 1 - p_s(x,x)[1 - p_t(x,x)]$.

Thus, if $p_{s+t}(x,x) = 1$, then $p_t(x,x) = 1$ since $p_t(x,x) > 0$. This implies that the set

$I = \{t \geq 0 : p_t(x,x) = 1\}$ is an interval starting at 1. By $(*)$, I must be $[0,\infty)$.

(c) Using the CK equation, $p_{t+s}(x,y) - p_t(x,y) = \underbrace{p_t(x,y)[p_s(x,x) - 1]}_{\leq 0} + \underbrace{\sum_{z \neq x} p_s(x,z) p_t(z,y)}_{\geq 0}$

Each term has absolute value $\leq 1 - p_s(x,x)$.

Hence, $|p_{t+s}(x,y) - p_t(x,y)| \leq 1 - p_s(x,x)$. Since $\lim_{s\downarrow 0} p_s(x,x) = 1$, we get uniform continuity. □

Thm (Differentials of transition functions)

(a) $\forall x \in S$, the right derivative

$$c(x) = -q(x,x) = -\frac{d}{dt} p_t(x,x)\Big|_{t=0} \in [0,\infty] \text{ exists, and we have}$$

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

(b) If $c(x) < \infty$, then $\forall y \neq x$, the right derivative

$$q(x,y) = \frac{d}{dt} p_t(x,y)\Big|_{t=0} \in [0,\infty) \text{ exists, and then } \sum_y q(x,y) \leq 0.$$

(c) If for some x, $c(x) < \infty$ and $\sum_y q(x,y) = 0$, then $p_t(x,y)$ is continuously differentiable in t $\forall y \in S$. Furthermore, the Kolmogorov backward equation holds:

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

Proof: (a) Define $f(t) = -\log p_t(x,x)$. It is well-defined and continuous by the previous lemma.

Moreover, $f(t)$ is subadditive: $f(t+s) = -\log p_{t+s}(x,x) \leq -\log p_t(x,x) - \log p_s(x,x)$

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

Then, $c(x) = \lim_{t\downarrow 0} \frac{f(t)}{t} \in [0,\infty]$ exists and satisfies

$f(x) \leq c(x)t$. This concludes (a).

Proof of the claim: $f : [0,\infty) \to \mathbb{R}$ is right continuous at 0 and $f(0) = 0$.

If f is subadditive, then $c = \lim_{t\downarrow 0} \frac{f(t)}{t} = \sup_{t>0} \frac{f(t)}{t} \in (-\infty, \infty]$ exists.

Define $c = \sup_{t>0} \frac{f(t)}{t}$. Fix $s > 0$ and $\forall\, 0 < t \leq s$, we can choose $n \in \mathbb{N}$ and $0 \leq \varepsilon < t$ s.t.

$s = nt + \varepsilon$. By subadditivity,

$$\frac{f(s)}{s} \leq \frac{n f(t) + f(\varepsilon)}{nt + \varepsilon} = \frac{nt}{nt+\varepsilon} \frac{f(t)}{t} + \frac{f(\varepsilon)}{s}$$

Taking $\liminf_{t\downarrow 0}$ we get, $\frac{f(s)}{s} \leq \liminf_{t\downarrow 0} \frac{f(t)}{t} \Rightarrow c \leq \liminf_{t\downarrow 0} \frac{f(t)}{t}$.

(b) Suppose $c(x) < \infty$. By (a), we have

$$1 - p_t(x,x) \leq 1 - e^{-c(x)t} \leq c(x)t, \text{ so that}$$

$$(\dagger) \quad \sum_{y : y \neq x} \frac{1}{t} p_t(x,y) = \frac{1}{t}(1 - p_t(x,x)) \leq c(x). \quad \text{Thus, } \limsup_{t\downarrow 0} \frac{p_t(x,y)}{t} < +\infty, \ \forall y \neq x.$$

Let $q(x,y)$ denote this limsup. Next, we show that the $\lim_{t\downarrow 0} \frac{p_t(x,y)}{t}$ exists and is $= q(x,y)$.

Take any (small) $\delta > 0$ and $n \in \mathbb{N}$. We have

$$p_{n\delta}(x,y) \geq \sum_{k=0}^{n-1} p_{k\delta}(x,x)\, p_\delta(x,y)\, p_{(n-k-1)\delta}(y,y).$$

$$\Rightarrow \frac{p_{n\delta}(x,y)}{n\delta} \geq \frac{p_\delta(x,y)}{\delta} e^{-c(x)n\delta} \inf_{0 \leq s \leq n\delta} p_s(y,y).$$

Take $\delta \downarrow 0$ along a subsequence that $\frac{p_\delta(x,y)}{\delta} \to q(x,y)$ and $n \to \infty$ so that $n\delta \to t$, we ~~const~~ conclude that $\frac{p_t(x,y)}{t} \ge q(x,y) e^{-c(x)t} \inf_{0 \le s \le t} p_s(y,y)$. Taking $t \downarrow 0$, we get

$\liminf_{t \downarrow 0} \frac{p_t(x,y)}{t} \ge q(x,y)$. This gives $\lim_{t \downarrow 0} \frac{p_t(x,y)}{t} = q(x,y)$.

Now taking $t \downarrow 0$ in the equation $\sum_{y \ne x} \frac{p_t(x,y)}{t} \le c(x)$, by Fatou's lemma, we have

$$\sum_{y : y \ne x} q(x,y) \le c(x) = -q(x,x). \quad \Rightarrow \quad \sum_y q(x,y) \le 0.$$

(c) $$\frac{p_{t+s}(x,y) - p_t(x,y)}{s} - \sum_z q(x,z) p_t(z,y) = \sum_z \left[\frac{p_s(x,z) - p_0(x,z)}{s} - q(x,z) \right] p_t(z,y).$$

By (a) and (b), each term in the sum $\to 0$ as $s \downarrow 0$. We need to control the tails of the sum. Take a finite $T \subset S$ containing x, we note that

$$\sum_{z \notin T} \left| \frac{p_s(x,z)}{s} - q(x,z) \right| p_t(z,y) \le \sum_{z \notin T} \frac{p_s(x,z)}{s} + \sum_{z \notin T} q(x,z)$$

$$= s^{-1} \left[1 - \sum_{z \in T} p_s(x,z) \right] - \sum_{z \in T} q(x,z) \to -2 \sum_{z \in T} q(x,z) \quad \text{as } s \downarrow 0.$$

Note $\sum_{z \in T} q(x,z)$ can be made arbitrarily small by taking T large since $\sum_z q(x,z) = 0$.

This shows that the right derivative of $p_t(x,y)$ exists and ~~sta~~ satisfies the Kolmogorov backward equation. Similar argument gives

$$\lim_{s \downarrow 0} \frac{p_t(x,y) - p_{t-s}(x,y)}{s} = \lim_{s \downarrow 0} \sum_z q(x,z) p_{t-s}(z,y) = \sum_z q(x,z) p_t(z,y),$$

where we use the continuity of $p_{t-s}(z,y)$ and $\sum_{z \ne x} q(x,z) = c(x) < +\infty$. □

Rmk: Taking derivative ~~over s~~ of the CK equation over s at $s=0$:

$p_{t+s}(x,y) = \sum_z p_s(x,z) p_t(z,y)$ gives the Kolmogorov backward equation.

Instead, we can also formally differentiate this equation with respect to t at $t=0$, and get:

$$\frac{d}{ds} p_s(x,y) = \sum_z p_s(x,z) q(z,y)$$ — Kolmogorov forward equation

The backward is usually "more useful" theoretically.

The backward and forward equations hold under different conditions. We will return to this issue later.

Rmk: If $c(x) = 0$, then $p_t(x,x) \equiv 1 \ \forall t$. Then, we say state x is "absorbing".

If $c(x) = \infty$, we say the state is "instantaneous".

States with $c(x) < \infty$ are called "stable".

3.4 Blackwell's example (Optional)

Before talking about the harder direction from Q matrix $\to p_t(x,y) \to$ MP, we first look at a classical example where every state is instantaneous. Our construction of the MC from Q matrices is heavily inspired by this example.

First, consider a 2-state MC with $S = \{0,1\}$. The most general Q-matrix is

$Q = \begin{pmatrix} -\beta & \beta \\ \delta & -\delta \end{pmatrix}$. We can calculate ~~that~~ the transition function from the equation

$$\frac{dp_t}{dt} = Q p_t \Rightarrow p_t = e^{tQ}$$

We can calculate e^{tQ} to get that

$$p_t(0,0) = \frac{\delta}{\beta+\delta} + \frac{\beta}{\beta+\delta} e^{-t(\beta+\delta)}, \quad p_t(0,1) = \frac{\beta}{\beta+\delta} [1 - e^{-t(\beta+\delta)}]$$

$$p_t(1,0) = \frac{\delta}{\beta+\delta} [1 - e^{-t(\beta+\delta)}], \quad p_t(1,1) = \frac{\beta}{\beta+\delta} + \frac{\delta}{\beta+\delta} e^{-t(\beta+\delta)}$$

Now, we take positive sequence (β_i, δ_i), and let $X_i(t)$ be independent 2-state chains as above. Then, we let $X(t) = (X_1(t), X_2(t), \cdots) \in \{0,1\}^\infty$. This is well-defined by the Kolmogorov extension thm. Now, we choose the state space

$$S = \{ x = (x_1, x_2, \cdots) : x_i \in \{0,1\} \ \forall i \text{ and } \sum_i x_i < \infty \}.$$

This is countable. For $x, y \in S$, define the transition function

$$(\dagger) \quad p_t(x,y) = \mathbb{P}(X(t) = y \mid X(0) = x) = \prod_i \mathbb{P}_{x_i}(X_i(t) = y_i).$$

Thm: Suppose that $\sum_i \frac{\beta_i}{\beta_i + \delta_i} < \infty$. Then, $\mathbb{P}(X(t) \in S \mid X(0) \in S) = 1 \quad \forall t \ge 0$, and $p_t(x,y)$ is a transition function on S.

Pf: Note $\forall t \ge 0$, $\mathbb{P}_0(X_i(t) = 1) = \frac{\beta_i}{\beta_i + \delta_i} [1 - e^{-t(\beta_i + \delta_i)}] \le \frac{\beta_i}{\beta_i + \delta_i}$.

By the Borel-Cantelli lemma, $\forall\, x \in S$, $\mathbb{P}_x(\|X(t)\|_1 < \infty)$ a.s. This shows that $\mathbb{P}(X(t) \in S \mid X(0) \in S) = 1$. Thus,

$$\sum_y p_t(x,y) = \mathbb{P}(X(t) \in S \mid X(0) = x) = 1, \quad \forall\, x \in S.$$

To show $\lim_{t \downarrow 0} p_t(x,x) = 1$, suppose $n \in \mathbb{N}$ is sufficiently large such that $x_i = 0$ $\forall\, i \geq n$. Then, $p_t(x,x) \geq \prod_{i \leq n} \mathbb{P}_{x_i}(X_i(t) = x_i) \prod_{i > n} \frac{\delta_i}{\beta_i + \delta_i}$. Note the product $\prod_{i>n} \frac{\delta_i}{\beta_i+\delta_i}$ converges to 1:

$$\prod_{i>n}\left(1 - \frac{\beta_i}{\beta_i+\delta_i}\right) = \exp\left(\sum_{i>n} \log\left(1 - \frac{\beta_i}{\beta_i+\delta_i}\right)\right) \leq \exp\left(-\sum_{i>n} \frac{\beta_i}{\beta_i+\delta_i}\right) \to 1 \text{ as } n \to +\infty.$$

First, taking $t \downarrow 0$, we get $\lim_{t \downarrow 0} p_t(x,x) \geq \prod_{i>n} \frac{\delta_i}{\beta_i+\delta_i}$. Then, taking $n \to \infty$, we get $\lim_{t\downarrow 0} p_t(x,x) \geq 1$.

A similar argument gives the CK equation for p_t. □

Thm: Suppose the sequence (β_i, δ_i) satisfies $\sum_i \frac{\beta_i}{\beta_i+\delta_i} < \infty$ and $\sum_i \beta_i = \infty$. Then, all states are instantaneous. Moreover,

$$(*) \quad \mathbb{P}(X(t) \in S \ \forall\, t \geq 0 \mid X(0) = x) = 0, \quad x \in S.$$

Proof: $\forall\, x \in S$, choose m s.t. $x_i = 0$ for $i \geq m$. Then, for $n \geq m$,

$$\begin{aligned}
1 - p_t(x,x) &= 1 - \prod_i \mathbb{P}_{x_i}(X_i(t) = x_i) \geq 1 - \prod_{i=m}^n \mathbb{P}_0(X_i(t) = 0) \\
&= 1 - \prod_{i=m}^n \left[\frac{\delta_i}{\beta_i+\delta_i} + \frac{\beta_i}{\beta_i+\delta_i} e^{-t(\beta_i+\delta_i)}\right] \geq 1 - \prod_{i=m}^n \left(1 - \beta_i t + O(\beta_i(\beta_i+\delta_i)^2 t^2)\right) \\
&\geq 1 - \exp\left(-t \sum_{i=m}^n \beta_i\right) + O(t^2)
\end{aligned}$$

$$\Rightarrow \lim_{t \downarrow 0} \frac{1 - p_t(x,x)}{t} \geq \sum_{i=m}^n \beta_i, \quad \forall\, n. \quad \text{Taking } n \to \infty \text{ gives that } c(x) = \infty.$$

The proof ~~the~~ of $(*)$ uses the Baire category theorem: A complete metric space is not the union of countably many nowhere dense sets.

Note that the event in $(*)$ is a tail event, so its probability is either 0 or 1 by Kolmogorov's 0-1 law. Suppose it has probability 1. Then, letting

$$A(y,\omega) := \{t \geq 0 : X(t,\omega) = y\},$$

it follows that $\bigcup_{y \in S} A(y,\omega) = [0,\infty)$. Then, the conclusion follows by contradiction if we can show that $A(y,\omega)$ is nowhere dense in $[0,\infty)$ with prob. 1 $\forall\, y \in S$.

For any rationals $0 \leq a < b \in \mathbb{Q}$,

$$\{\omega : X(t,\omega) = y \text{ for a dense set of } t \in (a,b)\} \subset \{\omega : X(t,\omega) = y \ \forall\, t \in [a,b)\}$$

This follows from the right continuity of the paths of X_i. On the other hand, suppose $y_i = 0$ $\forall\, i \geq m$. By the MP, we have

$$\begin{aligned}
\mathbb{P}(X(t) = y \ \forall\, t \in [a,b) \mid X(0) = x) &= \prod_i \mathbb{P}(X_i(a) = y_i \mid X_i(0) = x_i)\, \mathbb{P}(X_i(t) = y_i \ \forall\, t \in [a,b) \mid X_i(a) = y_i) \\
&\cancel{\leq \prod_{i=m}^\infty \left[\frac{\delta_i}{\beta_i+\delta_i} + \frac{\beta_i}{\beta_i+\delta_i} e^{-(b-a)(\beta_i+\delta_i)}\right] = \prod_{i=m}^\infty \left[1 - \frac{\beta_i}{\beta_i+\delta_i}\left(1 - e^{(b-a)(\beta_i+\delta_i)}\right)\right]} \\
&\leq \prod_{i=m}^\infty e^{-\beta_i(b-a)} = 0.
\end{aligned}$$

The conclusion then follows.

The inequality follows from the following description of the 2-state chain: $X(t)$ on $S = \{0,1\}$. If it starts at 0, then the chain stays at 0 for an exponential time with parameter β and then moves to 1. If it starts at 1, then the chain stays at 1 for an exponential time with parameter δ and then moves to 0. The chain repeats this role. □

From the above result, we see that there is no MC on S with transition function $p_t(x,y)$ as in (+). (If there were, it would have the same finite dimensional distribution as $X(t)$ defined above.) It would then follow that $\forall\, \varepsilon > 0$, $x \in S$,

$$\mathbb{P}_x(X(t) = x \text{ for all rational } t < \varepsilon) = 0.$$

This would violate the right continuity of paths in our def.

Rmk: While the above example behaves badly on a countable state space, it is well-behaved if regarded as a Markov process on the uncountable state space $\{0,1\}^\infty$.

3.5 From infinitesimal generator to Markov chain

We would like to determine when a given Q-matrix determines a unique transition function $p_t(x,y)$ that satisfies the Kolmogorov backward equation. When it does, we will then construct the corresponding MC. First, we will construct $p_t(x,y)$ analytically. Then, we will provide the probabilistic interpretation of the construction.

When S is finite, the construction of p_t is immediate: $p_t = e^{tQ}$, which is the unique solution to the system of linear differential equations: $\frac{d}{dt}p_t = Qp_t$. When S is infinite, ~~we can still construct~~ the existence and uniqueness theory for ∞-systems of differential equations is more involved.

Analytical construction

We first turn the KBE into an equivalent form of integral equations.

Prop: Suppose $p_t(x,y)$ is a uniformly bounded function in t, x, y. Then, TFAE:

(a) $p_t(x,y)$ is continuously differentiable in t and satisfies the KBE: $\frac{dp_t(x,y)}{dt} = \sum_z q(x,z)\,p_t(z,y)$ and the initial cond. $p_0(x,y) = \delta(x,y)$ $\forall\, x,y \in S$.

(b) $p_t(x,y)$ is continuous in t and satisfies

(*) $p_t(x,y) = \delta(x,y)e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)} \sum_{z:z\neq x} q(x,z)\,p_s(z,y)\,ds \quad \forall\, x,y\in S,\ t\geq 0.$

Proof: From (b) to (a) is trivial. Now, suppose (a) holds. We rewrite the KBE as

$$\frac{d}{dt}p_t(x,y) + c(x)\,p_t(x,y) = \sum_{z:z\neq x} q(x,z)\,p_t(z,y)$$

$$\Rightarrow \frac{d}{dt}\left[e^{c(x)t}p_t(x,y)\right] = e^{c(x)t}\sum_{z:z\neq x} q(x,z)\,p_t(z,y).$$

Integrating over t and using the initial condition, we get

$$e^{c(x)t}p_t(x,y) = \delta(x,y) + \int_0^t e^{c(x)s}\sum_{z:z\neq x} q(x,z)\,p_s(z,y)\,ds. \quad \square$$

One ~~essential~~ key feature of the equation (*) is that all quantities on the RHS are non-negative. This is ~~key~~ essential for the monotonicity argument below to work. We can solve (*) by interations. Let $p_t^{(0)}(x,y) \equiv 0$ and

(**) $p_t^{(n+1)}(x,y) = \delta(x,y)e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)} \sum_{z:z\neq x} q(x,z)\,p_s^{(n)}(z,y)\,ds, \quad n\geq 0.$

By induction, it is trivial to verify the following properties for $p_t^{(n)}(x,y)$:

$p_t^{(n)}(x,y) \geq 0$, $\quad \sum_y p_t^{(n)}(x,y) \leq 1$, $\quad$ & $\quad p_t^{(n+1)}(x,y) \geq p_t^{(n)}(x,y)$.

By monotonicity, we can define the limit $p_t^*(x,y) = \lim_{n\to\infty} p_t^{(n)}(x,y)$. We show that p_t^* satisfies the backward equation and is "almost" a transition function.

Thm: The function $p_t^*(x,y)$ satisfies that:

(a) $p_t^*(x,y) \geq 0$ $\forall\, t\geq 0,\ x,y\in S$.

(b) $\sum_y p_t^*(x,y) \leq 1$ $\forall\, t\geq 0,\ x\in S$.

(c) $p_t^*(x,y)$ satisfies the equation (*) and hence the KBE.

(d) $p_t^*(x,y)$ satisfies the Chapman-Kolmogorov equation.

Pf: (a) and (b) follow directly from the properties of $p_t^{(n)}$. To show (*), we pass to the $n\to\infty$ limit in (**) and use MCT. As a consequence, we get that $p_t(x,y)$ is continuous in t. To prove (d), we use the following equation in "Prop. 2.24 of Liggett": $\Delta_t^{(n)}(x,y) = p_t^{(n+1)}(x,y) - p_t^{(n)}(x,y) \geq 0$,

$$\Delta_{t+s}^{(n)}(x,y) = \sum_z \sum_{k=0}^{n} \Delta_s^{(k)}(x,z)\,\Delta_t^{(n-k)}(z,y)$$

Summing over n and using $\sum_{n=0}^{\infty} \Delta_t^{(n)}(x,y) = p_t^*(x,y)$, we get from Fubini's theorem:

$$p_{t+s}^*(x,y) = \sum_z p_s^*(x,z)\,p_t^*(z,y). \quad \square$$

The solution p_t^* satisfies all the desired properties, except that it may be substochastic: $\sum_y p_t^*(x,y)$ may <1. We also need to deal with the uniqueness of the solution.

Thm: (a) If $p_t(x,y)$ is any non-negative solution to be backward equation satisfying $p_0(x,y) = \delta(x,y)$, then we have $p_t(x,y) \geq p_t^*(x,y)$, $\forall\, t\geq 0,\ x,y \in S$.

[As a cosequence, we also call p_t^* the minimal solution.]

(b) If $\sum_y p_t^*(x,y) = 1$ $\forall\, t\geq 0,\ x\in S$, then $p_t^*(x,y)$ is the unique transition function satisfying the backward equation.

Pf: (a) It suffices to prove that $p_t(x,y) \geq p_t^{(n)}(x,y)$ $\forall\, n$.

This inequality holds when $n=0$. Suppose it is true for n. Then, by def.,

$$p_t^{(n+1)}(x,y) = \delta(x,y)e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)} \sum_{z:z\neq x} q(x,z)\, p_s^{(n)}(x,y)\,ds$$

$$\le \delta(x,y)e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)} \sum_{z:z\neq x} q(x,y)\, p_s(x,y)\,ds = p_t(x,y).$$

(b) If $p_t(x,y)$ is a transition function satisfying KBE. Then, $p_t(x,y) \ge p_t^*(x,y)$. Summing on y shows that we indeed have "=". □

Probabilistic construction of the MC

We now provide probabilistic interpretations for $p_t^{(n)}(x,y)$ and $p_t^*(x,y)$, and give a construction of the MC when $p_t^*(x,y)$ is stochastic. We first show that for Markov processes, the inter-event times should be exponential RVs.

Prop: Suppose $X(t)$ is a Markov chain, and let $\tau = \inf\{t\ge 0: X_t \ne X_0\}$. Then,

$$\mathbb{P}_x(\tau>t) = e^{-c(x)t} \text{ for some } 0 \le c(x) \le \infty. \quad (+)$$

Pf: For $0\le s\le t$ and $x\in S$, applying the MP to $Y = \mathbb{1}_{\{\omega:\, \omega(r)=x \text{ for } 0\le r\le t-s\}}$,

we get $\mathbb{P}_x(X_r = x\ \forall s\le r\le t \mid \mathcal{F}_s) = \mathbb{P}_{X_s}(\tau > t-s)$ $\mathbb{P}_x$-a.s.

$\Rightarrow \mathbb{E}_x[\mathbb{1}_{\tau>s}\, \mathbb{P}_x[X_r = x\ \forall s\le r\le t \mid \mathcal{F}_s]] = \mathbb{E}_x[\mathbb{1}_{\tau>s}\, \mathbb{P}_{X_s}(\tau>t-s)]$

$\Rightarrow \underbrace{\mathbb{P}_x(\tau>s, \text{ and } X_r = x\ \forall s\le r\le t)}_{=\,\mathbb{P}_x(\tau>t)} = \mathbb{P}_x(\tau>s)\,\mathbb{P}_x(\tau>t-s)$

This equation gives the memoryless property for τ: either $\mathbb{P}_x(\tau>t)=0$ $\forall t>0$, or it is strictly positive $\forall t>0$. In the former case, (+) holds with $c(x)=\infty$. In the latter case, we have proved that (+) must hold for some $c(x)\in[0,\infty)$ in the HW. □

Now, given the Q-matrix $q(x,y)$ with $c(x) := -q(x,x) < \infty$, we define transition probabilities $p(x,y)$ for a "discrete time MC" on S as: If $c(x)=0$, we let $p(x,x)=1$ and $p(x,y)=0$ for $y\ne x$; if $c(x)>0$, then we let

$$p(x,y) = \begin{cases} \dfrac{q(x,y)}{c(x)} & \text{if } y \ne x, \\ 0 & \text{if } y = x. \end{cases}$$

Note $\sum_y p(x,y) = 1$ since $\sum_y q(x,y) = 0$. Now, let $\pi(x)$ be a probability measure on S, and let Z_n be a discrete time MC on S with initial distribution π and transition probabilities $p(x,y)$. Z_n is called the "embedded discrete time chain". It provides information about the sequence of sites that $X(t)$ will visit, but no info. about the time spent at a site between jumps.

Motivated by the above prop., let $\tau_0, \tau_1, \tau_2, \cdots$ be sequence of RVs whose distribution is determined as follows: when conditioned on the chain $Z_0, Z_1, Z_2, \cdots$, the τ_k's are independent $\mathrm{Exp}(c(Z_k))$ RVs. (If $c(Z_k)=0$, then $\tau_k=\infty$.) More precisely, the finite-dim distributions of $(Z_1,\tau_1), \cdots, (Z_m,\tau_m)$ are given by

$$\mathbb{P}_\pi(Z_0 = x_0, \tau_0>t_0, Z_1 = x_1, \tau_1 > t_1, \cdots, Z_m = x_m, \tau_m > t_m)$$
$$= \pi(x_0)\,p(x_0,x_1)\,p(x_1,x_2)\cdots p(x_{m-1},x_m)\, e^{-c(x_0)t_0} e^{-c(x_1)t_1} \cdots e^{-c(x_m)t_m}.$$

By Kolmogorov extension thm, we can define the prob. measure for the ∞-sequence $\{(Z_k,\tau_k)\}_{k=1}^\infty$ on some prob. space. (Note that unconditionally, the τ_k's are neither independent nor exponentially distributed.)

Next, define

$$N_t = \begin{cases} \inf\{n\ge 0: \tau_0+\tau_1+\cdots+\tau_n > t\} & \text{if } \sum_{k=0}^\infty \tau_k > t \\ \infty & \text{otherwise} \end{cases}$$

Finally, we set $X_t = Z_{N_t}$ on $\{N_t<\infty\}$. Note that this process has right continuous paths where they are defined.

In sum, X_t evolves as follows: If $X_t = x$, then the process stays at x for an $\mathrm{Exp}(c(x))$ time. At the end of that time, it moves to y with prob. $p(x,y)$, stays there ~~with~~ for an ind. $\mathrm{Exp}(c(y))$ time, etc. This is well-behaved provided that the process does not jump ∞-many times within a finite time interval. This leads to substochastic solutions of the backward equation.

Now, we establish the connection between $p_t^*(x,y)$ and the probabilistic construction X_t.

☆ Thm (a) $p_t^{(n)}(x,y) = \mathbb{P}_x(X_t = y,\ N_t < n)$

(b) $p_t^*(x,y) = \mathbb{P}_x(X_t = y,\ N_t < \infty)$

(c) $\sum_y p_t^*(x,y) = \mathbb{P}_x(N_t < \infty)$

Pf: (b) follows from (a) by taking $n \to \infty$ and using MCT. (c) follows by summing (b) over $y \in S$. For (a), ~~we have~~ suppose it holds for n. Then, we have

$$\mathbb{P}_x(X_t = y,\ N_t < n+1 \mid Z_1 = z,\ \tau_0 = s) = \delta(x,y) \quad \text{if } s > t,$$

or $= \mathbb{P}_z(X_{t-s} = y,\ N_{t-s} < n)$ if $s < t$.

Taking the expectation of both sides, we get

$$\mathbb{P}_x(X_t = y,\ N_t < n+1) = \delta(x,y)e^{-c(x)t} + \int_0^t c(x)e^{-c(x)s}\sum_z p(x,z)\cdot \mathbb{P}_z(X_{t-s} = y,\ N_{t-s} < n)\,ds \quad s \to t-s$$

$$= \delta(x,y)e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)}\sum_{z: z\neq x} q(x,z)\,p_s^{(n)}(z,y)\,ds = p_t^{(n+1)}(x,y).$$

This proves (a) by induction noticing that $p_t^{(0)}(x,y) = \mathbb{P}_x(X_t = y, N_t < 0) = 0$. □

With this thm, we get:

Thm: TFAE: (a) The minimal solution $p_t^*(x,y)$ to KBE is stochastic.

(b) $\mathbb{P}(N(t) < \infty) = 1 \quad \forall t \geq 0$.

(c) $\sum_n \tau_n = \infty$ a.s. (d) $\sum_n \frac{1}{c(Z_n)} = \infty$ a.s.

Pf: (a) $\Leftrightarrow$ (b) by the above thm. By def, $\{N_t < \infty\} = \{\sum_k \tau_k > t\}$, so (b) $\Leftrightarrow$ (c).

To see (c) $\Leftrightarrow$ (d), we take $\lambda > 0$ and use the fact that τ_n's are ind. Exp random variables conditioning on Z_n's:

$$\mathbb{E}\left(\exp\left(-\lambda\sum_{k=0}^n \tau_k\right) \Big| Z_0, Z_1, Z_2, \cdots\right) = \prod_{k=0}^n \frac{c(Z_k)}{c(Z_k)+\lambda}.$$

$$\Rightarrow \mathbb{E}\left(\exp\left(-\lambda\sum_{k=0}^n \tau_k\right)\right) = \mathbb{E}\prod_{k=0}^n \frac{c(Z_k)}{c(Z_k)+\lambda} \xRightarrow[\text{DCT}]{n\to+\infty} \mathbb{E}\exp\left(-\lambda\sum_{k=0}^\infty \tau_k\right) = \mathbb{E}\prod_{k=0}^\infty \frac{c(Z_k)}{c(Z_k)+\lambda}$$

Finally, taking $\lambda \downarrow 0$, we get $\mathbb{P}\left(\sum_k \tau_k < \infty\right) = \lim_{\lambda\downarrow 0}\mathbb{E}\left[\prod_{k=0}^\infty \left(1 + \frac{\lambda}{c(Z_k)}\right)^{-1}\right] = \mathbb{P}\left(\sum_k \frac{1}{c(Z_k)} < \infty\right)$ □

Cor: The minimal ~~soluti~~ solution p_t^* is stochastic if one of the followings holds.

(a) $\sup_x c(x) < \infty$.

or (b) The discrete MC Z_n is irreducible and recurrent.

Pf: If (a) holds, then $\frac{1}{c(x)}$ is bounded from below, so the series $\sum_n \frac{1}{c(Z_n)}$ diverges.

Under (b), $Z_n = x$ i.o. almost surely for any x, so a summand $\frac{1}{c(x)}$ occurs i.o. in the series $\sum_n \frac{1}{c(Z_n)}$. □

Thm Suppose $p_t^*(x,y)$ is stochastic. Then, p_t^* is a transition function, and there is a unique MC (in the sense of equal distribution) such that

$$\mathbb{P}_x(X_t = y) = p_t^*(x,y)$$

Pf: We have $p_t^*(x,x) \geq \mathbb{P}_x(N_t = 0) = e^{-c(x)t}$, so $\lim_{t\downarrow 0} p_t(x,x) = 1$. All the other properties of a transition have been established in a previous thm.

To construct the prob. measures $\mathbb{P}_x$ for the MC for any starting points $x \in S$, we let $\mathbb{P}_x(\cdot)$ be the distribution of $(X(\cdot) \mid X(0) = x)$, where X is the process defined before.

~~The~~ By $p_t^*(x,y) = \mathbb{P}_x(X_t = y,\ N_t < \infty)$ and the fact that $N_t < \infty$ a.s., p_t^* is its transition fcn.

The MP of X_t can be proved in the same way as that for BM, where only the following facts are used: [(a) right continuity of paths; (b) the joint continuity of $\phi(y,h) = \mathbb{E}_y f_1(X_{t_1-h}) \cdots f_n(X_{t_n-h})$, $\forall h < t_1 < t_2 < \cdots < t_n$, bdd measurable $f_1, \cdots, f_n$; (c) $\phi(y,0) = \mathbb{E}_y Y$; (d) the CK equation.]

The uniqueness follows from that the measure for the MC is uniquely determined by the transition functions. □

The Kolmogorov Forward equation

Thm: Suppose that $\sum_y p_t^*(x,y)c(y) < \infty \quad \forall t \geq 0$. Then, we have

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

Pf. For any $h > 0$, we have

$$\frac{p_{t+h}^*(x,y) - p_t^*(x,y)}{h} = \sum_z p_t^*(x,z)\frac{p_h^*(z,y) - \delta(z,y)}{h}.$$

Since $\sum_{y:y\neq z} p_h^*(z,y) \leq 1 - p_h^*(z,z) \leq 1 - e^{-c(z)h} \leq hc(z)$. Thus, by the DCT, letting

$h \downarrow 0$ in the above equation gives that the right derivative of $p_t^*(x,y)$ satisfies the forward equation. Since the RHS is continuous in t, the two-sided derivative of $p_t^*(x,y)$ exists and satisfies the forward equation. □

3.6 Stationary measures, recurrence, and transience

By the results in the previous section, we can ~~construct~~ get a MC ~~using~~ by simply describing the Q-matrix. ~~In reality, only~~ Often, only the Q-matrix (instead of the transition function) is given explicitly. Hence, it is desirable to deduce properties of the MC from the Q-matrix directly.

In the following, we always assume that the minimal solution p_t^* is stochastic, and denote it by $p_t(x,y)$. We also assume that $0 < c(x) < \infty$ $\forall x \in S$, and $\sup_x c(x) < \infty$.

Stationary and reversible measures

Def: A measure π on S is said to be stationary w.r.t. $p_t(x,y)$ if

$$\pi(y) = \sum_x \pi(x) p_t(x,y) \quad \forall\ y \in S,\ t > 0.$$

Thm: Suppose that $\sum_x c(x)\pi(x) < \infty$. ~~Then~~ Then, π is stationary iff $\sum_x \pi(x) q(x,y) = 0$, $\forall y \in S$.

Pf: ~~If $\sum_x \pi(x) q(x,y) = 0$, then~~ By DCT,

$$\frac{d}{dt}\sum_x \pi(x) p_t(x,y) = \sum_x \pi(x) q(x,y) = 0 \iff \sum_x \pi(x) p_t(x,y) = \sum_x \pi(x) p_0(x,y) = \pi(y). \quad □$$

Def: A measure π on S is said to be reversible w.r.t. $p_t(x,y)$ if

$$\pi(x) p_t(x,y) = \pi(y) p_t(y,x), \quad \forall\ x,y \in S,\ t > 0.$$

Thm: Suppose $\pi(x) > 0$ $\forall x \in S$. Then, $\pi(x)$ is reversible iff $\pi(x) q(x,y) = \pi(y) q(y,x)$ $\forall x,y \in S$.

Def: If a stationary (resp. reversible) measure is a probability distribution, then we call it a stationary (resp. reversible) distribution.

Recurrent and transient states

Def: The MC $X(t)$ is said to be irreducible if $p_t(x,y) > 0$ $\forall x,y \in S$ and $t > 0$.

Rmk: ~~By the CK equation, we have $p_{t+s}(x,$~~ By the equation

$$p_t(x,y) = \delta(x,y) e^{-c(x)t} + \int_0^t e^{-c(x)(t-s)} \sum_{z: z \neq x} q(x,z) p_s(z,y)\, ds,$$

若 $p_t(x,y) > 0$ for some $t > 0$ iff $p_t(x,y) > 0$ for all $t > 0$. This is one feature that makes continuous time MC simpler than the discrete time counterpart. [In fact, if $p_t(x,y) = 0$ for some t, then we have $p_s(z,y) = 0$ $\forall z \in S$ s.t. $q(x,z) > 0$ and $s \in [0,t)$. Then, using the CK equation, we get that $p_t(x,y) \equiv 0$ $\forall t$.]

Def: A state $x \in S$ is said to be recurrent if $\forall c > 0$,

$$\mathbb{P}_x(\exists\, t > c \text{ s.t. } X_t = x) = 1.$$

Otherwise, it is said to be transient.

The Green's function of a MC is defined by $G(x,y) = \int_0^\infty p_t(x,y)\,dt$.

Note $\int_0^\infty \mathbb{P}_x(X_t = y)\,dt = \mathbb{E}_x \int_0^\infty \mathbb{1}(X_t = y)\,dt$, i.e., $G(x,y)$ is the expected amount of time spent at y by the chain started at x.

Thm: The state x is transient iff $G(x,x) < \infty$.

Pf: By the construction of $X_t = Z_{N_t}$, the times spent at x are i.i.d. $\mathrm{Exp}(c(x))$ RVs. Let N_x be the # of visits to x for the embedded chain Z_n: $N_x = \sum_{n=1}^\infty \mathbb{1}(Z_n = x)$.

By Wald's identity, we have $G(x,x) = \frac{1}{c(x)} \mathbb{E}_x N_x$ ~~If $\mathbb{E}N_x = \infty$, then $G(x,x) = \infty$, then $\mathbb{E}N_x = \infty$~~

$$= \frac{1}{c(x)} \frac{\rho_{xx}}{1-\rho_{xx}},$$ where $\rho_{xx} = \mathbb{P}_x(T_x < \infty)$ is defined for the chain Z_n.

If $G(x,x) = \infty$, then $\rho_{xx} = 1$ and Z_n visits x i.o. In this case, the chain X_t also visits x at arbitrarily large t. If $G(x,x) < \infty$, then $\int_0^\infty \mathbb{1}(X_t = x)\,dt < \infty$ $\mathbb{P}_x$-a.s., and the state is transient. □

Thm: For an irreducible MC, either all states are recurrent, or all states are transient. In the transient case, $G(x,y) < \infty$ $\forall x,y \in S$.

Pf: By the CK equation, $p_{t+s}(x,x) \ge p_t(x,y) p_s(y,y) p_t(y,x)$. Taking integral over s,

$$G(x,x) \ge p_t(x,y)\, G(y,y)\, p_t(y,x).$$

Thus, for irreducible chains, $G(x,x) < \infty \iff G(y,y) < \infty$ $\forall x,y \in S$.

Let T_y be the hitting time of y. Then, by the SMP,

$$G(x,y) = \mathbb{E}_x\left[\int_{T_y}^{\infty} \mathbb{1}(X_t=y)\,dt,\ T_y<\infty\right]$$

$$= \mathbb{E}_y\left[\int_0^{\infty} \mathbb{1}(X_t=y)\,dt\right]\mathbb{E}_x(\mathbb{1}_{T_y<\infty}) = \mathbb{P}_x(T_y<\infty)\,G(y,y) \le G(y,y) < \infty. \quad \square$$

Since the MC X_t visits the same states as the embedded discrete time chain Z_n, either both are recurrent or both are transient. Thus, we can determine whether a continuous time chain is recurrent / transient from discrete time results, and vice versa.

Def: A function f on S is <u>superharmonic</u> for the chain $X(t)$ if for each t and x, $\mathbb{E}_x|f(X(t))|<\infty$ and $\mathbb{E}_x f(X(t)) \le f(x)$. (I.e. f is superharmonic iff $X(t)$ is a supermartingale w.r.t. $\mathbb{P}_x$ for each x.)

Thm: An irreducible MC is transient iff it has a <u>non-constant</u> <u>bounded</u> superharmonic function.

Pf: If the chain is transient, then we take $f(x)=G(x,y)$ for some fixed y. Then, $f(x)$ is bounded (by $G(y,y)$), and is superharmonic:

$$\mathbb{E}_x f(X(t)) = \sum_z p_t(x,z)\,G(z,y) = \int_0^\infty \sum_z p_t(x,z)\,p_s(z,y)\,ds$$

$$= \int_0^\infty p_{t+s}(x,y)\,ds = \int_t^\infty p_s(x,y)\,ds < G(x,y) = f(x).$$

The strict ineq. implies that f is non-constant.

Conversely, suppose X_t is recurrent and f is bdd superharmonic. Then, $f(X(t))$ is a bdd supermartingale, which converges a.s. by the martingale convergence thm. On the other hand, X_t visits every site i.o., so f must be constant. $\square$

Thm: If an irreducible chain has a stationary distribution, then it is recurrent.

Pf: Suppose the chain has a stationary distribution π. Then,

$$\pi(y) = \sum_x \pi(x)\,p_s(x,y) \Rightarrow \pi(y) = \sum_x \pi(x)\,\frac{1}{t}\int_0^t p_s(x,y)\,ds$$

If the chain is transient, then $\frac{1}{t}\int_0^t p_s(x,y)\,ds \to 0$. Using the DCT, we get $\sum_x \pi(x)\,\frac{1}{t}\int_0^t p_s(x,y)\,ds \to 0 \quad \forall x\in S$. This gives a contradiction. $\square$

Thm: Every irreducible recurrent MC has a non-zero stationary measure.

Pf: Fix a $z\in S$, and let ~~$\tau = \inf\{t>0: X_t = z\}$~~ τ be the first return time to z, i.e., the first time the chain hits z after visiting some other state:

$$\tau = \inf\{t>0:\ X_t = z \text{ and } X_s \ne z \text{ for some } s\in[0,t)\}.$$

Define $\pi(x) = \mathbb{E}_z\int_0^\tau \mathbb{1}(X_t=x)\,dt$. This is $\not\equiv\infty$ or $\not\equiv 0$ since $\pi(z)=\frac{1}{c(z)}$.

Now, we check that π is stationary. By the SMP, we have

$$\mathbb{E}_z\int_0^s \mathbb{1}(X_t=x)\,dt = \mathbb{E}_z\int_\tau^{\tau+s}\mathbb{1}(X_t=x)\,dt, \quad \forall s>0, \quad \text{on } \{\tau<\infty\}.$$

Then, we have

$$\pi(x) = \mathbb{E}_z\left[\int_0^{\tau+s}\mathbb{1}(X_t=x)\,dt - \int_\tau^{\tau+s}\mathbb{1}(X_t=x)\,dt\right] = \mathbb{E}_z\left[\int_0^{\tau+s}\mathbb{1}(X_t=x)\,dt - \int_0^s \mathbb{1}(X_t=x)\,dt\right]$$

$$= \mathbb{E}_z\left[\int_s^{\tau+s}\mathbb{1}(X_t=x)\,dt\right] = \int_0^\infty \mathbb{E}_z[X_{t+s}=x,\ \tau>t]$$

$$= \int_0^\infty \sum_y \mathbb{P}_z(X_t=y,\ \tau>t)\,p_s(y,x)\,dt = \sum_y \pi(y)\,p_s(y,x). \quad \square$$

Rmk: Similar to discrete time MC, we can check that $\sum_x \pi(x) = \mathbb{E}_z\tau$. Hence, π can be normalized to a stationary distribution iff $E_z\tau<\infty$.

Thm: For an irreducible recurrent MC, the stationary measure is unique up to constant multiples.

~~Pf: Suppose π_1 and π_2 are nonzero stationary measures. By irreducibility, $p_t(x,y)>0$ $\forall t>0$ and $x,y\in S$. Hence, from the eq. $\pi(y)=\sum_x \pi(x)p_t(x,y)$, we can derive that $\pi_1(x)>0$ and $\pi_2(x)>0$ $\forall x\in S$. Then, we can define a new transition function~~

~~$$\bar{p}_t(x,y) = \frac{\pi_1(y)\,p_t(y,x)}{\pi_1(x)}.$$ We have that for the new chain,~~

~~$$\bar{G}(x,x) = \int_0^\infty \bar{p}_t(x,x)\,dt = G(x,x) = \infty,$$ so the new chain is also recurrent.~~

Pf: π is a stationary measure iff $\sum_x \pi(x)\,q(x,y) = 0 \Rightarrow$

$$\sum_{x \neq y} \pi(x) q(x,y) = \pi(y) c(y) \Rightarrow \sum_x \pi(x) p(x,y) = \pi(y)$$, i.e., π is a stationary measure for the discrete time embedded chain Z_n. The chain Z_n is also irreducible and recurrent. Hence, π is unique (up to constant multiples). □

Def: A state x with $\begin{cases} \mathbb{E}_x T_x < \infty \\ \mathbb{E}_x T_x = \infty \end{cases}$ is said to be $\begin{cases} \text{positive recurrent.} \\ \text{null} \end{cases}$ ~~if $\mathbb{E}_x T_x < \infty$~~

The above two theorems imply that:

Thm: If an irreducible MC has a stationary distribution π, then $\forall z \in S$,

$$\cancel{\mathbb{E}_x T_x = \frac{1}{\pi(x)}} \qquad \cancel{\pi(z) = \frac{1}{c(z)\mathbb{E}_z T_z}} \qquad \mathbb{E}_z T_z = \frac{1}{c(z)\pi(z)}.$$

Pf: By irreducibility, $p_t(x,y) > 0 \ \forall t > 0$ and $x, y \in S$. From the equation $\pi(y) = \sum_x \pi(x) p_t(x,y)$, we can deduce that $\pi(x) > 0 \ \forall x \in S$. ~~Then taking integral over the equation~~ $\cancel{\sum_x \pi(x) p_s(x,y) = \pi(y)}$, ~~we get~~ $\cancel{\sum_x \pi(x) G(x,y) = \infty}$.

~~Thus,~~ $\cancel{G(y,y) \geq \sum_x \pi(x) G(x,y) = \infty \Rightarrow y}$ ~~is recurrent~~ $\cancel{\forall y \in S}$.

Now, by the above two theorems, we have $\forall z \in S$,

$$(*) \quad \pi(x) = \frac{1}{\mathbb{E}_z T_z} \mathbb{E}_z \int_0^{\tau} \mathbb{1}(X_t = x)\, dt.$$ Taking $x = z$, we conclude the pf. □

Thm: Given an irreducible MC, TFAE:

(i) Some z is positive recurrent. (ii) There is a stationary distribution.

(iii) All states are positive recurrent.

Pf: (i) ⇒ (ii) If z is positive recurrent, then $(*)$ defines a stationary distribution.

(ii) ⇒ (iii) follows from the above thm. (iii) ⇒ (i) Trivial. □

Convergence of MC

Thm: Suppose the irreducible MC X_t has stationary distribution π. Then, $\forall x$,

$$\|p_t(x,\cdot) - \pi\|_{TV} = \frac{1}{2}\sum_y |p_t(x,y) - \pi(y)| \to 0 \quad \text{as } t \to \infty.$$

Rmk: We do not need "aperiodicity condition" for this thm. This is another feature that makes continuous time MC simpler than its discrete time counterpart.

Pf: We again use the coupling method. Let X_t and Y_t be independent copies of the chain, and let $Z_t = (X_t, Y_t)$. Then, Z_t is an irreducible MC on $S \times S$ with transition function $p'_t((x_1,y_1),(x_2,y_2)) = p_t(x_1,x_2)\, p_t(y_1,y_2) > 0$. It has stationary distribution $\pi'(x,y) = \pi(x)\pi(y)$, so it is recurrent. In particular, we know

$$\mathbb{P}_x(\tau < \infty) = 1 \quad \text{for} \quad \tau := \inf\{t > 0 : X(t) = Y(t)\}.$$

Then, we define a coupling between $X(t)$ and $Y(t)$ as $W(t)$ on $S \times S$ with

$$W(t) = (X'_t, Y'_t) = \begin{cases} (X_t, Y_t) & \text{for } t \leq \tau, \\ (X_t, X_t) & \text{for } t > \tau. \end{cases}$$

Note W_t is also a MC with transition function p'. Taking ~~X(0)~~ X_0 with distribution π and Y_0 with distribution δ_x, we get

$$\sum_y |p_t(x,y) - \pi(y)| = \sum_y |\mathbb{P}_\pi(X'_t = y) - \mathbb{P}_x(Y'_t = y)| \leq 2\mathbb{P}(X'_t \neq Y'_t)$$
$$= 2\mathbb{P}(\tau > t) \to 0 \quad \text{as} \quad t \to \infty. \qquad □$$