Stochastic Modelling and Random Processes

Yiming MA

December 13, 2020

Contents

Chapter 1

Discrete-Time Markov Chains

1.1 Countable Discrete-Time Markov Chains

One can extend much of what we have done for finite discrete-time Markov chains to the countably infinite case, e.g. the **simple random walk** on $\mathbb{Z}$, but some results become more subtle. For example, the simple random walk is *not SP-ergodic*, despite being *irreducible*. Actually, it even *fails to have a stationary probability*; also it is *not aperiodic*, and it has a *period* 2.

Example 1.1.1. Using definition of the simple random walk:

$$Y_n = \sum_{i=0}^{n-1} X_i,$$

where X_i's are independent and identically distributed, with

$$X_i = \begin{cases} +1 & \text{with probability}\, p \\ -1 & \text{with probability}\, 1-p \end{cases},$$

Compute the $\mathbb{E}[Y_n]$ and $\mathrm{Var}[Y_n]$.

One has to refine various concepts.

Definition 1.1.1 (The First Return Time)**.** The **first return time** to state x is defined as

$$T_x = \inf\{n \geq 1 : X_n = x | X_0 = x\}.$$

Remark. Notice that when the state space is finite and x is recurrent, T_x is finite. Since the state space here is countably infinite, T_x is allowed to be infinte.

Definition 1.1.2 (Transience)**.** Say $x \in S$ is **transient** if

$$\mathbb{P}[T_x = \infty] > 0.$$

Remark. If $x \in S$ is transient, then with probability 1 X_n comes back to x only finitely many times.

Definition 1.1.3 (Null Recurrence)**.** Say $x \in S$ is **null recurrent** if

$$\mathbb{P}[T_x < \infty] = 1 \quad \text{and} \quad \mathbb{E}[T_x] = \infty.$$

Definition 1.1.4 (Positive Recurrence)**.** Say $x \in S$ is **positive recurrent** if

$$\mathbb{P}[T_x < \infty] = 1 \quad \text{and} \quad \mathbb{E}[T_x] < \infty.$$

Remark. A communicating class is either **null recurrent**, which means every member is null recurrent, or **positive recurrent** which means every member is positive recurrent.

Theorem 1.1.1 (Stationarity $\iff$ Positive Recurrence)**.** *An absorbing class has a stationary probability if and only if it is positive recurrent. Furthermore, if the class has one stationary probability, then it is uniquely determined by*

$$\boldsymbol{\pi}_x = \frac{1}{\mathbb{E}[T_x]}.$$

Chapter 2

Continuous-Time Markov Chains

2.1 Continuous-Time Markov Chains

We are now considering a continuous-time markov chain with a countable state space S and the domain $T \in \mathbb{R}$ (or $T \in \mathbb{R}_+$), and we restrict $X : \mathbb{R} \mapsto S$ to those which are *piecewise constant* and *right-continuous*, meaning

$$X(t) = \begin{cases} \vdots & \vdots \\ s & t \in [J_s, J_{s'}) \\ s' & t \in [J_{s'}, J_{s''}) \\ \vdots & \vdots \end{cases}$$

Definition 2.1.1 (Continuous-Time Markov Chains)**.** $X(t) : \mathbb{R} \mapsto S$ is a **continuous-time Markov chain**, if it satisfies the **Markov property**

$$\mathbb{P}[X(t_{n+1}) \in A | X(t_n) = s_n, \cdots, X(t_1) = s_1] = \mathbb{P}[X(t_{n+1}) \in A | X(t_n) = s_n],$$

where $A \subset S$ and $t_1 < \cdots t_n < t_{n+1}$.

Definition 2.1.2 (Homogeneity)**.** A continuous-time Markov chain is **homogeneous** if

$$\mathbb{P}[X(t+u) \in A | X(u) = s] = \mathbb{P}[X(t) \in A | X(0) = s].$$

Remark. Homogeneity means time translation invariance.

Definition 2.1.3 (Transition Matrices)**.** Let $(P_t)_{i,j} := \mathbb{P}[X(t) = j | X(0) = i]$, then P_t is the transition matrix with time step t.

Remark. The (i, j) element of the transition matrix P_t can also be expressed as $P_t(i, j)$.

Theorem 2.1.1 (Chapman-Kolmogorov Equation)**.** *The transition matrix P of a homogeneous Markov chain satisfies*

$$P_{t+u} = P_t P_u,\ P_0 = I.$$

Proof. Notice that

$$
\begin{aligned}
(P_{t+u})_{i,j} =&\mathbb{P}[X(t+u)=j|X(0)=i]\\
=&\sum_{k\in S}\mathbb{P}[X(t+u)=j|X(t)=k,\ X(0)=i]\mathbb{P}[X(t)=k|X(0)=i]\\
=&\sum_{k\in S}\mathbb{P}[X(t+u)=j|X(t)=k]\mathbb{P}[X(t)=k|X(0)=i]\\
=&\sum_{k\in S}\mathbb{P}[X(u)=j|X(0)=k]\mathbb{P}[X(t)=k|X(0)=i]\\
=&\sum_{k\in S}(P_u)_{k,j}(P_t)_{i,k}\\
=&(P_t)_{i,:}\ (P_u)_{:,j},
\end{aligned}
$$

where $(P_t)_{i,:}$ is the i-th row of P_t and $(P_u)_{:,j}$ is the j-th column of P_u. Thus, $P_{t+u}=P_tP_u$. And by definition, $(P_0)_{i,j}=\mathbb{P}[X_0=j|X_0=i]=\delta_{i,j}$, so $P_0=I$. $\square$

2.1.1 The Rate Matrix

Definition 2.1.4 (Rate Matrix)**.** Suppose P_t is differentiable with respect to t at $t=0$, then

$$
G := \left.\frac{\mathrm{d}P_t}{\mathrm{d}t}\right|_{t=0}
$$

is called the **generator** or the **rate matrix** of the process.

Proposition 2.1.1. $P_t=\exp(tG)$ *in the sense of power series.*

Proof. By the Chapman-Kolmogorov equation, we have

$$
\begin{aligned}
P_{t+u} =&P_tP_u\\
P_{t+u}-P_t =&P_t(P_u-I)\\
\frac{P_{t+u}-P_t}{u} =&P_t\cdot\frac{P_u-I}{u}\\
\lim_{u\to 0}\frac{P_{t+u}-P_t}{u} =&\lim_{u\to 0}P_t\cdot\frac{P_u-I}{u}\\
\lim_{u\to 0}\frac{P_{t+u}-P_t}{u} =&P_t\cdot\lim_{u\to 0}\frac{P_u-I}{u}\\
\frac{\mathrm{d}P_t}{\mathrm{d}t} =&P_tG,
\end{aligned}
$$

So $P_t=C\cdot\exp(tG)$, where C is a constant diagonal matrix with diagonal elements being equal. By $P_0=I$, we konw $C=I$. $\square$

Proposition 2.1.2. *The generator G also satisfies*

$$
G\vec{1}=\vec{0}.
$$

Proof. For any probability distribution $\boldsymbol{\pi}_t=\boldsymbol{\pi}_0P_t$ with initial distribution $\boldsymbol{\pi}_0$, evolves by

$$
\frac{\mathrm{d}\boldsymbol{\pi}_t}{\mathrm{d}t}=\boldsymbol{\pi}_0\frac{\mathrm{d}P_t}{\mathrm{d}t}=\boldsymbol{\pi}_0P_tG=\boldsymbol{\pi}_tG.
$$

And by conservation of probability, we have $\boldsymbol{\pi}_t\vec{1}=\vec{1}$, which implies $\boldsymbol{\pi}_tG\vec{1}=\frac{\mathrm{d}\boldsymbol{\pi}_t\vec{1}}{\mathrm{d}t}=0$. Since $\boldsymbol{\pi}_t$ is arbitrary, we have $G\vec{1}=0$. $\square$

Theorem 2.1.2 (The Master Equation)**.** *The equation*

$$\frac{\mathrm{d}\boldsymbol{\pi}_t}{\mathrm{d}t} = \boldsymbol{\pi}_t G$$

can be written into

$$\frac{\mathrm{d}(\boldsymbol{\pi}_t)_i}{\mathrm{d}t} = \underbrace{\sum_{j\neq i}(\boldsymbol{\pi}_t)_j G_{j,i}}_{\text{"gain"}} - \underbrace{\sum_{j\neq i}(\boldsymbol{\pi}_t)_i G_{i,j}}_{\text{"loss"}},$$

which is called the ***master equation****.*

Proof. For $i \neq j$, since $G_{i,j}$ is the rate at which the process goes from state i to j, we have $G_{i,j} \geq 0$. By $G\vec{1} = \vec{0}$, we have

$$G_{i,i} = -\sum_{j\neq i} G_{i,j}.$$

So

$$\begin{aligned}\frac{\mathrm{d}(\boldsymbol{\pi}_t)_i}{\mathrm{d}t} =& \boldsymbol{\pi}_t G_{:,i} \\ =& \sum_{j\in S}(\boldsymbol{\pi}_t)_j G_{j,i} \\ =& \sum_{j\neq i}(\boldsymbol{\pi}_t)_j G_{j,i} - \sum_{j\neq i}(\boldsymbol{\pi}_t)_i G_{i,j}.\end{aligned}$$

□

Remark. The name "master equation" is exaggerated; it does not tell everything about the process, such as the correlations between states at different times.

Example 2.1.1 (Poisson Processes)**.** The **Poisson process** with rate $\lambda > 0$ has the state space $S = \mathbb{N}$, $X(0) = 0$, and the transition matirx G such that

$$G_{i,j} = \begin{cases} \lambda & j = i+1 \\ -\lambda & j = 1 \end{cases}.$$

It has $\mathbb{P}[X(t+u) = n+k | X(u) = n] = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$, $\forall n, k \in \mathbb{N}$, $\forall t, u \in \mathbb{R}_+$.

Example 2.1.2 (Birth and Death Processes)**.** Suppose we have the birth rates α_i and the death rates β_i ($\beta_0 = 0$), for $i \in S = \mathbb{N}$. The rate matrix G is defined by

$$G_{i,j} = \begin{cases} \alpha_i & j = i+1 \\ \beta_i & j = i-1 \\ -(\alpha_i + \beta_i) & j = i \end{cases}.$$

Then the process is called the **Birth and Death Process**.

Example 2.1.3 ($M/M/1$ queue)**.** The birth and death process has a special case - the ***M/M/*1 queue**, in which $\alpha_i = \alpha$, $\beta_i = \beta$ for $i \neq 0$ and $\beta_0 = 0$. M means "memoryless", and 1 means there is only one cashier to serve customers.

Example 2.1.4 ($M/M/\infty$ queue)**.** Another example is the ***M/M/∞* queue**, in which there are infinitely many servers so that customers do not have to wait for people in front of them. In this model $\alpha_i = \alpha$ and $\beta = i\beta$.

Example 2.1.5 (Population Growth)**.** Population growth can be modelled by the birth and death process with $\alpha_i = i\alpha$ and $\beta_i = i\beta$, where i is the size of population.

2.1.2 Stationarity and Reversibility

Definition 2.1.5 (Stationarity)**.** Say $\boldsymbol{\pi} \in \Delta$ is **stationary** if $\boldsymbol{\pi} G = 0$.

Definition 2.1.6 (Reversibility)**.** Say $\boldsymbol{\pi} \in \Delta$ is **reversible** if

$$\boldsymbol{\pi}_i G_{i,j} = \boldsymbol{\pi}_j G_{j,i}, \ \forall i, j \in S.$$

Proposition 2.1.3 (Reversibility $\implies$ Stationarity)**.** *If $\boldsymbol{\pi} \in \Delta$ is reversibile, then it is also stationary.*

Proposition 2.1.4. *S is fintie $\implies \exists$ stationary $\boldsymbol{\pi}$.*

There is an analogous decomposition of the state space S into transient and recurrent states, and of the set of recurrent states into communicating components. And we have the same definition of an absorbing component.

Proposition 2.1.5. *If S is finite, then each absorbing component has a unique stationary probability $\boldsymbol{\pi}$, and the space of starionary $\boldsymbol{\pi}$ for the whole continuous-time Markov chain (up to normalisation) is the span of those for its absorbing components. Furthermore, 0 is a semisimple eigenvalue of G.*

Theorem 2.1.3. *Suppose S is finite and G has a unique absorbing component, then the process is SP-ergodic, which means*

$$\lim_{t\to\infty} \boldsymbol{\pi}_t = \boldsymbol{\pi}_A,$$

where $\boldsymbol{\pi}_A$ is the stationary distribution of the absorbing component.

Remark. Aperiodicity is automatic in continuous time.

2.1.3 The Jump Chain

Definition 2.1.7 (Waiting Times)**.** The **waiting time** or the **holding time** W_x is defined as

$$W_x = \inf\{t > 0 : X(t) \neq x | X(0) = x\}.$$

Proposition 2.1.6. *The waiting time W_x is exponentially distributed with mean $\frac{1}{|G_{x,x}|}$.*

Proof.

$$\begin{aligned}\mathbb{P}[W_x > t+u | W_x > t] =& \mathbb{P}[W_x > t+u | X(s) = x, \forall s \leq t] \\ =& \mathbb{P}[W_x > t+u | X(t) = x] \\ =& \mathbb{P}[W_x > u | X(0) = x] \\ =& \mathbb{P}[W_x > u].\end{aligned}$$

So $\mathbb{P}[W_x > t+u] = \mathbb{P}[W_x > u]\mathbb{P}[W_x > t]$. So $\exists \gamma \in \mathbb{R}$, such that

$$\mathbb{P}[W_x > t] = e^{-\gamma t}.$$

$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}[W_x > t]\Big|_{t=0} = G_{x,x}$ shows $\gamma = -G_{x,x}$. □

Definition 2.1.8 (Jump Times)**.** Define **jump times** $J_{n+1} = \inf\{t > J_n : X(t) \neq X(J_n)\}$, with $J_0 = 0$.

Remark. The jump times are an example of "stopping times", i.e. random variables such that $\{J_n \leq t\}$ is independent of $\{X(s) : s > t\}$ given $\{X(s) : s \leq t\}$.

Theorem 2.1.4. *Markov chains satisfiy the* ***strong Markov property****: let T be a stopping time conditional on $X_T = i$, then X_{T+t} $(t \geq 0)$ is Markov and independent of $\{X(s) : s \leq T\}$.*

Definition 2.1.9 (The Jump Chain)**.** Let $Y_n = X(J_n)$, then $\{Y_n : n \in \mathbb{N}\}$ is called the **jump chain** of $\{X_t : t \in \mathbb{R}\}$.

Remark. The jump chain $\{Y_n : n \in \mathbb{N}\}$ is a discrete-time Markov chain.

Proposition 2.1.7. *The one-step transition matrix of the jump chain $\{Y_n : n \in \mathbb{N}\}$ is*

$$P_{i,j} = \begin{cases} 0 & j = i \\ \frac{G_{i,j}}{|G_{i,i}|} & j \neq i \,\&\, G_{i,i} = 0 \\ \delta_{i,\cdot} & G_{i,i} = 0 \end{cases} .$$

Remark. We can make sample paths for the continuous-time Markov chain by making paths for the associated jump chain and choosing independent waiting times W_{Y_n} with mean $1/|G_{Y_n,Y_n}|$, and let

$$J_n = \sum_{0 \leq k < n} W_{Y_k}.$$

2.2 Countable Continuous-Time Markov Chains

Now suppose the state space S of a continuous-time Markov chains is countable. We can define the null and positive recurrence as in the discrete-time case, but we have to find the return time differently.

Definition 2.2.1 (First Return Time)**.** The **first return time** to state $x \in S$ is defined as

$$\inf\{t > J_1 : X(t) = x\},$$

for $X(0) = x$.

Proposition 2.2.1. *Each positive recurrent absorbing component has a unique stationary probability distribution $\boldsymbol{\pi}$, and*

$$\boldsymbol{\pi} = \frac{\mathbb{E}[W_x]}{\mathbb{E}[T_x]}.$$

In continuous time, the process can get "explosion".

Definition 2.2.2 (Explosion)**.** Let $J_\infty = \lim_{n\to\infty} J_n$. If $\mathbb{P}[J_\infty = \infty] < 1$, then the continuous-time Markov chain is called **explosive**, which means there is a positive probability for infinitely many events in a bounded time.

Proposition 2.2.2. *If $\sup_{i \in S} |G_{i,i}| < \infty$, then the continuous-time Markov chain is not eplosive.*

Example 2.2.1 (Explosion)**.** Consider a birth and death process with $X(0) = 1$, $\alpha_i = i^2$ and $\beta_i = 0$. Then

$$\mathbb{E}[J_\infty] = \sum_{i=2}^{\infty} \mathbb{E}[W_i] = \sum_{i=2}^{\infty} \frac{1}{\alpha_i} = \sum_{i=2}^{\infty} \frac{1}{i^2} < \infty,$$

which means with probability 1 J_∞ is finite.

2.3 Semi-Markov Chains

Definition 2.3.1 (Semi-Markov Chains)**.** Take a discrete-time Markov chain and make a continuous-time process by waiting a time W_x in each state $x \in S$ independently of previous and future states but not necessarily exponentially distributed.

Remark. Semi-Markov chains allow for latent periods and variations of infectivity with time from infection.

2.4 Gaussian Processes

Definition 2.4.1 (Gaussian Processes)**.** Let $X : T \mapsto \mathbb{R}$ be a stochasti process. $X(t)$ is called a **Gaussian process** if $\forall t_1, \cdots, t_n \in T$, $(X(t_1), \cdots, X(t_2))$ is a multivariate Gaussian random vector, i.e. it has the probability density function

$$f(x_1, \cdots x_n) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu}^T)\Sigma^{-1}(\vec{x} - \vec{\mu})\right),$$

for some $\vec{\mu} = [\mu_1, \cdots, \mu_n]^T$ and some positive definite symmetric $n \times n$ matrix Σ.

Proposition 2.4.1. *There exist functions $m : T \mapsto \mathbb{R}$ and $c : T \times T \mapsto \mathbb{R}$ such that $\mu_i = m(t_i)$ and $\Sigma_{i,j} = c(t_i, t_j)$ with c being "positive definite" i.e. such that Σ is positive definite $\forall t_1, \cdots, t_n \in T$.*

Example 2.4.1 (Stationary Ornstein-Uhlenbeck Processes)**.** Let $T = \mathbb{R}$, $m(t) = 0$ and $c(t, t') = e^{-|t'-t|}$, then the process is called a **stationary Ornstein-Uhlenbeck process**.

One can allow degenerate Gaussians.