

Stochastic Modelling and Random Processes

Yiming MA

December 14, 2020

Contents

1	Discrete-Time Markov Chains	2
1.1	Countable Discrete-Time Markov Chains	2
2	Continuous-Time Markov Chains	4
2.1	Continuous-Time Markov Chains	4
2.1.1	The Rate Matrix	5
2.1.2	Stationarity and Reversibility	7
2.1.3	The Jump Chain	7
2.2	Countable Continuous-Time Markov Chains	8
2.3	Semi-Markov Chains	9
2.4	Gaussian Processes	9
2.5	Uncountable Markov Processes	10
2.5.1	Jump Processes	10
2.5.2	Diffusion Processes	11
2.6	Generators as Operators	12
2.6.1	Generators of Discrete Continuous-Time Markov Chains	12
2.6.2	Generators of Continuous Continuous-Time Markov Chains	12

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 $\text{Var}[Y_n]$.

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 infinite.

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*

$$\pi_x = \frac{1}{\mathbb{E}[T_x]}.$$

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**

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

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

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*

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

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

$$G := \left. \frac{dP_t}{dt} \right|_{t=0}$$

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

$$\begin{aligned}
P_{t+u} &= P_t P_u \\
P_{t+u} - P_t &= P_t (P_u - I) \\
\frac{P_{t+u} - P_t}{u} &= P_t \cdot \frac{P_u - I}{u} \\
\lim_{u \rightarrow 0} \frac{P_{t+u} - P_t}{u} &= \lim_{u \rightarrow 0} P_t \cdot \frac{P_u - I}{u} \\
\lim_{u \rightarrow 0} \frac{P_{t+u} - P_t}{u} &= P_t \cdot \lim_{u \rightarrow 0} \frac{P_u - I}{u} \\
\frac{dP_t}{dt} &= P_t G,
\end{aligned}$$

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

Proposition 2.1.2. The generator G also satisfies

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

Proof. For any probability distribution $\pi_t = \pi_0 P_t$ with initial distribution π_0 , evolves by

$$\frac{d\pi_t}{dt} = \pi_0 \frac{dP_t}{dt} = \pi_0 P_t G = \pi_t G.$$

And by conservation of probability, we have $\pi_t \vec{1} = \vec{1}$, which implies $\pi_t G \vec{1} = \frac{d\pi_t \vec{1}}{dt} = 0$. Since π_t is arbitrary, we have $G\vec{1} = 0$. \square

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

$$\frac{d\pi_t}{dt} = \pi_t G$$

can be written into

$$\frac{d(\pi_t)_i}{dt} = \underbrace{\sum_{j \neq i} (\pi_t)_j G_{j,i}}_{\text{"gain"}} - \underbrace{\sum_{j \neq i} (\pi_t)_i G_{i,j}}_{\text{"loss"}},$$

*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

$$\begin{aligned} \frac{d(\pi_t)_i}{dt} &= \pi_t G_{:,i} \\ &= \sum_{j \in S} (\pi_t)_j G_{j,i} \\ &= \sum_{j \neq i} (\pi_t)_j G_{j,i} - \sum_{j \neq i} (\pi_t)_i G_{i,j}. \end{aligned}$$

□

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 matrix G such that

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

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 **$M/M/\infty$ 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 $\pi \in \Delta$ is **stationary** if $\pi G = 0$.

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

$$\pi_i G_{i,j} = \pi_j G_{j,i}, \quad \forall i, j \in S.$$

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

Proposition 2.1.4. *S is finite $\implies \exists$ stationary π .*

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 π , and the space of stationary π 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*

$$\lim_{t \rightarrow \infty} \pi_t = \pi_A,$$

where π_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

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

$\frac{d}{dt} \mathbb{P}[W_x > t] \Big|_{t=0} = G_{x,x}$ 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 satisfy 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*

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

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 π , and*

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

In continuous time, the process can get “explosion”.

Definition 2.2.2 (Explosion). Let $J_\infty = \lim_{n \rightarrow \infty} J_n$. 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 explosive.*

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 stochastic process. $X(t)$ is called a **Gaussian process** if $\forall t_1, \dots, t_n \in T$, $(X(t_1), \dots, X(t_n))$ is a multivariate Gaussian random vector, i.e. it has the probability density function

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

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

Remark. A Gaussian process is not necessarily Markov.

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 $\forall t_1, \dots, t_n \in T$.*

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, e.g. Ornstein-Uhlenbeck with specified initial condition $X(0) = 0$, then $f(x_0) = \delta_0(x_0)$, which is not a Gaussian probability density function but can be viewed as the limit of a Gaussian density.

The best way to generate a Gaussian distribution is to use its characteristic function instead of its PDF.

Definition 2.4.2 (Characteristic Functions). Let \vec{X} be a random vector, then its characteristic function is

$$\phi(\vec{\theta}) := \mathbb{E}[e^{i\vec{\theta}^T \vec{X}}].$$

Remark. For a multivariate Gaussian distribution with the mean vector $\vec{\mu}$ and covariance matrix Σ (which is allowed to be positive semi-definite), its characteristic function is

$$\phi(\vec{\theta}) = e^{i\vec{\theta}^T \vec{\mu} - \frac{1}{2} \vec{\theta}^T \Sigma \vec{\theta}}.$$

We can include vector-valued Gaussian processes.

Definition 2.4.3 (Multivariate Gaussian Processes). $\vec{X} : T \mapsto \mathbb{R}^n$ is a **multivariate Gaussian process** if $X : T \times \{1, \dots, n\} \mapsto \mathbb{R}$ is a Gaussian process.

Definition 2.4.4 (Stationary Gaussian Processes). Suppose $T = \mathbb{R} \times \mathbb{K}$. The Gaussian process is **stationary**, if its mean function $m(t, k)$ is independent of t , and its covariance function $c(t, k; t', k')$ is dependent only on $t - t'$ and $k - k'$.

Remark. Gaussian processes are great for inference, because $\mathbb{P}[\text{parameters}|\text{data}]$ reduces to linear algebra.

2.5 Markov Processes with $S = \mathbb{R}$

Suppose $X : T \mapsto \mathbb{R}$, where T can be \mathbb{Z} or \mathbb{R} .

Definition 2.5.1 (Markov Processes). $\{X(t) : t \in T\}$ is a **Markov Processes** if it satisfies the Markov property

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

where $A \subset \mathbb{R}$ and $t_{n+1} > t_n > \dots > t_1$.

Remark. There is a technical problem in the definition. The conditional probability is not well defined, since random variables $X_{t_n}, \dots, X(t_1)$ now take values in \mathbb{R} , and the probability that they take particular values is 0. This will not be a problem if we restrict to any choice of interpretation of conditional probability such that

$$\mathbb{P}[X(t) \in A] = \int \mathbb{P}[X(t) \in A | X(0) = x] d\mathbb{P}[X(0) \leq x] \quad (\text{a Stieltjes integral}).$$

Definition 2.5.2 (Homogeneity). A Markov process is **homogeneous** if

$$\mathbb{P}[X(t) \in A | X(t') = s] = \mathbb{P}[X(t - t') \in A | X(0) = s].$$

It is unlikely that $\mathbb{P}[X(t) = y | X(0) = x] > 0$, so instead we specify $\mathbb{P}[X(t) \in A | X(0) = x]$ for any measurable set $A \subset \mathbb{R}$ as

$$\int_A p_t(x, y) dy$$

for a transition density $p_t(\cdot, \cdot)$.

Definition 2.5.3 (Transition Densities). A **transition probability** is a function $p_t(\cdot, \cdot) : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ such that

$$\mathbb{P}[X(t) \in A | X(0) = x] = \int_A p_t(x, y) dy$$

Theorem 2.5.1 (The Chapman-Kolmogorov Equation). *The Markov property and homogeneity implies the **Chapman-Kolmogorov equation***

$$p_{t+u}(x, y) = \int_{\mathbb{R}} p_t(x, z) p_u(z, y) dz.$$

2.5.1 Jump Processes

Definition 2.5.4 (Jump Processes). $\{X(t) : t \in \mathbb{R}\}$ is a **jump process** if

- there is a **jump rate density** $r(x, y)$ with the **exit rate**

$$R(x) = \int_{\mathbb{R}} r(x, y) dy \leq M < \infty, \quad \forall x \in \mathbb{R},$$

where $M \in \mathbb{R}$ is a constant;

- its transition density satisfies

$$p_{\Delta t}(x, y) = r(x, y)\Delta t + (1 - R(x)\Delta t) \delta(y - x) + o(\Delta t), \quad \text{as } \Delta t \rightarrow 0.$$

Theorem 2.5.2 (The Kolmogorov-Feller Equation). *The Chapman-Kolmogorov equation of a jump process turns into the **Kolmogorov-Feller equation** for initial condition $x \in \mathbb{R}$*

$$\frac{\partial}{\partial t} p_t(x, y) = \int_{\mathbb{R}} p_t(x, z) r(z, y) - p_t(x, y) r(y, z) dz$$

2.5.2 Diffusion Processes

Definition 2.5.5 (The Brownian Motion). The **Brownian motion** is a Gaussian process $B : \mathbb{R}_+ \mapsto \mathbb{R}$ with $m(t) = 0$ and $c(t, t') = \min(t, t')$ and almost surely continuous paths.

Proposition 2.5.1 (Brownian Motions are Markov). *A Brownian motion is Markov, and it has independent increments: $\forall t_1 < \dots < t_n$, $(X(t_{k+1}) - X(t_k))_{k=1, \dots, n-1}$ are independent variables.*

Proposition 2.5.2 (Brownian Motions are Homeogeneous). *Furthermore, the increments are stationary: $X(t) - X(s)$ and $X(t - s) - X(0) = X(t - s)$ have the same distribution, for $t \geq s$. So $B(t)$ is homoeogeneous.*

Remark. $B(t)$ is not stationary.

Proposition 2.5.3. *The transition density $p_t(x, y)$ of a Brownian motion is a Gaussian PDF with mean $y - x$ and variance t , which satisfies the heat equation (or diffusion equation):*

$$\frac{\partial p_t}{\partial t} = \frac{1}{2} \frac{\partial^2 p_t}{\partial y^2}$$

with the initial condition $p_0(x, y) = \delta(y - x)$.

Proposition 2.5.4. *Brownian motions are normally distributed: $B(t) \sim \mathcal{N}(0, t)$.*

Proposition 2.5.5. *$B(t)$ is scale-invariant: $B(\lambda t)$ and $\sqrt{\lambda} B(t)$ have the same distribution.*

Proposition 2.5.6. *$B(t)$ is almost surely continuous, but it is also almost surely nowhere differentiable. Actually,*

$$\xi_{t,h} := \frac{B(t+h) - B(t)}{h} \sim \mathcal{N}\left(0, \frac{1}{h}\right).$$

Although Brownian motions are almost surely nowhere differentiable, we can still informally talk about the limit process $\xi_t := \lim_{h \rightarrow 0} \xi_{t,h}$.

Definition 2.5.6 (Gaussian White Noises). $\xi_t := \lim_{h \rightarrow 0} \xi_{t,h}$ is called the **Gaussian white noise**.

Remark. The Gaussian white noise can be considered as a limiting case of a Gaussian process with mean $m(t) = 0$ and $c(t, t') = \delta(t - t')$.

Proposition 2.5.7. *$B(t) = \int_0^t \xi_{t'} dt'$, or we can write it as a stochastic differential equation*

$$\frac{dB}{dt} = \xi,$$

with $B(0) = 0$.

2.6 Generators as Operators

2.6.1 Generators of Discrete Continuous-Time Markov Chains

For a continuous-time Markov chain with a countable state space S , for any function $f : S \mapsto \mathbb{R}$, we have

$$\mathbb{E}[f(X(t))] = \sum_{x \in S} \pi_t(x) f(x) = \pi_t \vec{f},$$

where \vec{f} is a column vector of values of f at all the state $x \in S$.

We may be interested in how fast $\mathbb{E}[f(X(t))]$ varies with time t , so

$$\frac{d}{dt} \mathbb{E}[f(X(t))] = \frac{d}{dt} \pi_t \vec{f} = \pi_t G \vec{f}.$$

Thus, we can think of the generator G as acting on the function f by

$$(Gf)(x) = \sum_{y \in S} G_{x,y} f(y) = \sum_{\substack{y \neq x \\ y \in S}} G_{x,y} (f(y) - f(x)).$$

2.6.2 Generators of Continuous Continuous-Time Markov Chains

The idea of generators as operators can be extended to $S = \mathbb{R}$ by replacing matrices and vectors with operators and functions.

Generators of Brownian Motions

For a Brownian motion,

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[f(X(t))] &= \frac{\partial}{\partial t} \int_{\mathbb{R}} p_t(x, y) f(y) dy \\ &= \int_{\mathbb{R}} \frac{\partial}{\partial t} p_t(x, y) f(y) dy \\ &= \frac{1}{2} \int_{\mathbb{R}} \frac{\partial^2}{\partial y^2} p_t(x, y) f(y) dy \\ &= \mathbb{E}[(\mathcal{L}f)](X(t)) \end{aligned}$$

with $(\mathcal{L}f)(x) := \frac{1}{2} f''(x)$, assuming f is twice differentiable and $f(x) \& f'(x) \rightarrow 0$ as $x \rightarrow \pm\infty$ (integration by parts). \mathcal{L} is the generator but now a linear operator on functions.

Generators of Jump Processes

For a jump process on \mathbb{R} ,

$$(\mathcal{L}f)(x) = \int_{\mathbb{R}} r(x, y) [f(y) - f(x)] dy.$$

We can obtain the Brownian motion as a scaling limit of a jump process. Take a jump process $X(t)$ with $r(x, y) = q(y - x)$ such that $\int_{\mathbb{R}} z q(z) dz = 0$ and $\int_{\mathbb{R}} z^2 q(z) dz = \sigma^2 \in (0, \infty)$. Then $\forall T > 0$, with $X(0) = 0$,

$$\frac{\epsilon}{\sigma} X\left(\frac{t}{\epsilon^2}\right) \Big|_{t \in [0, T]} \xrightarrow{d} B(t) \Big|_{t \in [0, T]}, \text{ as } \epsilon \rightarrow 0.$$

