Theory of Stochastic Processes

Lecture 12: Diffusion processes

Tomonari SEI

July 6, 2017

A diffusion process is a Markov process with continuous trajectories, that is constructed by stochastic integral with respect to a Brownian motion. We briefly study how to manipulate diffusion processes and apply them to specific problems. Most of the proofs are omitted[*1].

1 Brownian motion

Definition 1. A continuous-time process $\{W_t\}_{t\geq 0}$ taking values in $\mathbb{R}$ is called a *standard Brownian motion* (or *standard Wiener process*) if

(i) $W_0 = 0$.
(ii) $W_t - W_s$ is independent of the past history $\{W_r\}_{r\leq s}$ for $0 \leq s < t$,
(iii) $W_t - W_s$ has the normal distribution $N(0, t-s)$ for $0 \leq s < t$, and
(iv) the path $t \mapsto W_t$ is continuous.

A sample path is shown in Figure 1.

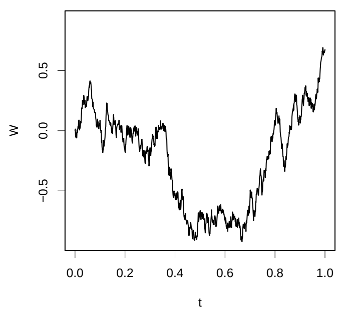


Figure1 A sample path of the standard Brownian motion.

*1 See e.g. B. Øksendal (2003), *Stochastic Differential Equations: An Introduction with Applications*, 6th ed., Springer.

The following theorem is fundamental but the proof is not so obvious (and omitted).

Theorem 1. The Brownian motion exists on a probability space.

We check some properties of the Brownian motion. In general, a stochastic process $\{X_t\}_{t\geq 0}$ is said to be *adapted* if X_t is a function of the past history $\mathcal{F}_t = \{W_s\}_{s\leq t}$[*2]. An adapted process X_t is called a *Markov process* if the conditional distribution of X_t given $\mathcal{F}_s$ depends only on X_s for $s < t$. An adapted process X_t is called *a martingale* if $E[X_t|\mathcal{F}_s] = X_s$ for $s < t$.

Theorem 2. The Brownian motion itself is a Markov process and a martingale.

Proof. The Markov property follows from the condition (ii). We also have

$$E[W_t|\mathcal{F}_s] = E[W_s|\mathcal{F}_s] + E[W_t - W_s|\mathcal{F}_s] \underset{\text{(ii)}}{=} W_s + E[W_t - W_s] \underset{\text{(iii)}}{=} W_s.$$

Hence W_t is a martingale. $\square$

By definition, the density function of $W_t = y$ given $W_s = x$ is

$$p(t, y|s, x) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right). \tag{1}$$

It is easy to see that the density satisfies the *heat equation*

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial y^2}. \tag{2}$$

Therefore the Brownian motion describes a microscopic structure of diffusion phenomena. Historically, this was the motivation for the study of the Brownian motion. Nowadays the Brownian motion is applied to various fields of science and technology. We give an example.

Example 1 (Brownian bridge). Let W_t be a standard Brownian motion and define

$$B_t = W_t - tW_1, \quad 0 \leq t \leq 1. \tag{3}$$

This process is called a *Brownian bridge*. It is easy to see that (Problem 1)

$$E[B_s B_t] = s(1-t) \quad \text{for} \quad s \leq t. \tag{4}$$

In statistics, for given data $X_1, \ldots, X_n \in \mathbb{R}$, the empirical distribution is defined by

$$\hat{F}(t) = \frac{1}{N}\sum_{i=1}^{N} 1_{\{X_i \leq t\}}, \quad t \in \mathbb{R},$$

If X_i's are independent and uniformly distributed on $[0, 1]$, then $\hat{F}(t)$ has mean t and covariance $E[\hat{F}(s)\hat{F}(t)] = s(1-t)/N$ for $s \leq t$. Furthermore, it is known that $\sqrt{N}(\hat{F}(t) - t)$ converges to the Brownian bridge as $N \to \infty$. This fact is applied to a goodness-of-fit test such as the Kolmogorov-Smirnov test[*3].

[*2] More precisely, $\mathcal{F}_t$ is the σ-field generated by $\{W_s\}_{s\leq t}$ and X_t is called adapted if X_t is $\mathcal{F}_t$-measurable.

[*3] See Chapter 19 of A. W. van der Vaart (1998), *Asymptotic Statistics*, Cambridge University Press.

2 Itô calculus

Let $\{W_t\}$ be a standard Brownian motion. Suppose that b_t is an adapted process and $\int_0^T E[b_t^2]dt < \infty$. Then the *stochastic integral* (or Itô integral) of b_t with respect to W_t is defined by[*4]

$$\int_0^T b_t dW_t = \lim_{n\to\infty} \sum_{i=1}^{2^n} b_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}), \quad t_i = \frac{iT}{2^n},$$

where the limit is interpreted as the mean-square convergence.

Example 2 (continuous-time betting game). Imagine a stock price W_t generated by the standard Brownian motion. Your initial capital is $X_0 > 0$. If you have b_t tickets during the time $[t, t+\Delta t]$, you will get $b_t(W_{t+\Delta t} - W_t)$ as a profit. Then your capital process X_t will be described as $X_t = X_0 + \int_0^t b_s dW_s$.

It is shown that (c.f. Problem 2)

$$E\left[\int_0^T b_t dW_t\right] = 0, \quad E\left[\left(\int_0^T b_t dW_t\right)^2\right] = \int_0^T E[b_t^2]dt.$$

The latter identity is called Itô isometry. The following theorem also holds.

Theorem 3. For any adapted b_t, the process $\int_0^t b_s dW_s$ is a martingale.

Example 3. We can show that

$$\int_0^t W_s dW_s = \frac{W_t^2 - t}{2}.$$

Indeed, by definition, the left hand side is the limit of

$$\begin{aligned}\sum_{i=1}^{2^n} W_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) &= \frac{1}{2}\sum_i \left(W_{t_i}^2 - W_{t_{i-1}}^2\right) - \frac{1}{2}\sum_i (W_{t_i} - W_{t_{i-1}})^2 \\ &= \frac{1}{2}W_t^2 - \frac{1}{2}\underbrace{\sum_i (W_{t_i} - W_{t_{i-1}})^2}_{Z},\end{aligned}$$

where $t_i = it/2^n$. The result follows from $E[Z] = t$ and $V[Z] = (2t^2)\cdot 2^{-n} \to 0$ as $n \to \infty$. The process $W_t^2 - t$ is directly shown to be a martingale (see Problem 3).

A process X_t is called an *Itô process* if there exist adapted processes μ_t and σ_t such that

$$dX_t = \mu_t dt + \sigma_t dW_t, \tag{5}$$

[*4] The definition is valid only for a reasonable class of b_t. In general, we have to use a limiting argument.

meaning that

$$X_t - X_0 = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s.$$

Theorem 4 (Itô's formula)**.** Let X_t be an Itô process and $f(t,x)$ be a smooth function. Then the process $Y_t = f(t, X_t)$ is an Itô process with

$$dY_t = (\partial_t f)dt + (\partial_x f)dX_t + \frac{1}{2}(\partial_x^2 f)(dX_t)^2,$$

where dX_t and $(dX_t)^2$ are interpreted as $\mu_t dt + \sigma_t dW_t$ and $\sigma_t^2 dt$, respectively.

Example 4. The process W_t^2 satisfies

$$d(W_t^2) = 2W_t dW_t + dt.$$

Therefore we have $W_t^2 = 2\int_0^t W_s dW_s + t$. This result is consistent with Example 3.

3 Diffusion processes

A process X_t is called a *diffusion process* if there exist functions $\mu(t,x)$ and $\sigma(t,x)$ such that

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t. \tag{6}$$

The coefficients $\mu(t,x)$ and $\sigma(t,x)$ are called the *drift* and *diffusion* coefficients. The equation (6) is called a *stochastic differential equation* (SDE). We do not discuss the existence and uniqueness of solutions. The following theorem is quite natural.

Theorem 5. A diffusion process defined by (6) is a Markov process.

We give two examples of diffusion processes.

Example 5 (Ornstein-Uhlenbeck process). Let $\theta > 0$ and $\sigma > 0$. Define a diffusion process $\{X_t\}_{t\geq 0}$ by

$$dX_t = -\theta X_t dt + \sigma dW_t. \tag{7}$$

Applying Itô's formula to $f(t, X_t) = e^{\theta t} X_t$, we have

$$d(e^{\theta t} X_t) = e^{\theta t}(dX_t + \theta X_t dt) = \sigma e^{\theta t} dW_t.$$

By integrating out, we have

$$X_t = e^{-\theta t}\left(X_0 + \sigma \int_0^t e^{\theta s} dW_s\right).$$

The process X_t is called the *Ornstein-Uhlenbeck process.* Letting $\sigma \to 0$, we have the deterministic motion $X_t = X_0 e^{-\theta t}$. The Ornstein-Uhlenbeck process has a stationary distribution $N(0, \sigma^2/(2\theta))$. See Problem 4.

Example 6. Define a diffusion process X_t by $X_0 = 0$ and

$$dX_t = -\frac{X_t}{1-t}dt + dW_t, \quad t \in [0, 1).$$

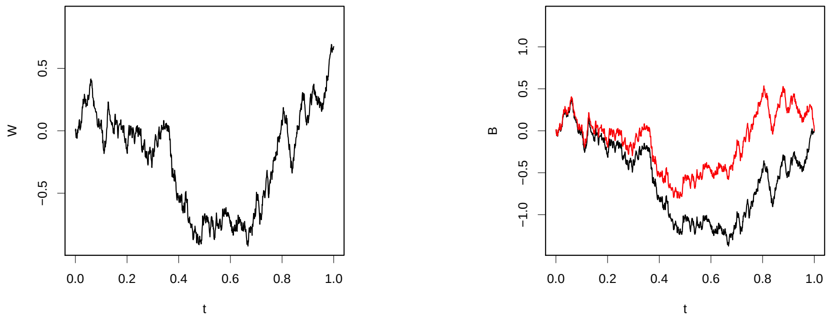
Applying Itô's formula to $f(t, X_t) = X_t/(1-t)$ and integrating out, we have

$$X_t = (1-t)\int_0^t \frac{1}{1-s}dW_s. \tag{8}$$

The process X_t is a mean-zero Gaussian process with the autocovariance function

$$E[X_s X_t] = (1-s)(1-t)\int_0^s \frac{1}{(1-u)^2}du = (1-s)(1-t)\left(\frac{1}{1-s} - 1\right) = s(1-t)$$

for $s < t$. Hence X_t has the same distribution as the Brownian bridge B_t defined in (3). Note that the two processes B_t and X_t are different as a functional of W_t, even though they have the same distribution. See Figure 2.



Brownian motion W_t Brownian bridges B_t (black) and X_t (red)

Figure2 Two ways of constructing a Brownian bridge. Refer to Equations (3) and (8).

Let $p(t, y|s, x)$ be the conditional density function of $X_t = y$ given $X_s = x$ for $s < t$. We call it the *transition density function* of the diffusion process.

Theorem 6. The transition density function $p(t, y|s, x)$ of a diffusion process (6) satisfies

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial y}\{\mu(t, y)p\} + \frac{1}{2}\frac{\partial^2}{\partial y^2}\{\sigma(t, y)^2 p\}. \tag{9}$$

This is called the *forward equation* or *Fokker-Planck equation*.

Proof. Let $f(x)$ be any smooth function. By Itô's formula, we see that, given $X_s = x$,

$$\begin{aligned} E[f(X_{t+dt}) - f(X_t)] &= E\left[(\partial_x f)(\mu_t dt + \sigma_t dW_t) + \frac{1}{2}(\partial_x^2 f)\sigma_t^2 dt\right] \\ &= E\left[(\partial_x f)\mu_t + \frac{1}{2}(\partial_x^2 f)\sigma_t^2\right] dt, \end{aligned}$$

where μ_t denotes $\mu(t, X_t)$ and so on. Dividing both sides by dt and substituting $E[f(X_t)] = \int f(y)p(t, y|s, x)dy$, we obtain

$$\begin{aligned}\int f(y)(\partial_t p)dy &= \int \{(\partial_y f(y))\mu_t + \frac{1}{2}(\partial_y^2 f(y))\sigma_t^2\} p dy \\ &= \int f(y)\{-\partial_y(\mu_t p) + \frac{1}{2}\partial_y^2(\sigma_t^2 p)\}dy,\end{aligned}$$

where the integration-by-parts formula is used. Since f is arbitrary, we obtain the result. $\square$

4 Application: Langevin Monte Carlo

Suppose that we want to generate samples from a density function $\pi(x) = \pi_*(x)/Z$ on $\mathbb{R}$, where $Z = \int_{-\infty}^{\infty} \pi_*(x)dx$ is the normalizing constant. Consider a diffusion process

$$dX_t = \frac{1}{2}(\log \pi_*)'(X_t)dt + dW_t. \tag{10}$$

The process has the stationary distribution $\pi(x)$. Indeed, the forward equation (9) is satisfied by $p(t, y|s, x) = \pi(y)$ (Problem 8).

Hence we can generate a sample by simulating (10). This sampling algorithm is called the *Langevin Monte Carlo*. In practice, the exact solution of (10) is not necessary since the truncation errors are adjusted by the Metropolis-Hastings algorithm[*5].

Figure 3 compares the Langevin MC with the random walk MC ($X_t = W_t$). The target distribution is a skewed normal distribution $\pi(x) = 2\phi(x)\Phi(x)$. The equation (10) is approximated by $\Delta X_t = (1/2)(\log \pi_*)'(X_t)\Delta t + \Delta W_t$, where $\Delta W_t \sim N(0, \Delta t)$ and $\Delta t = 0.3$.

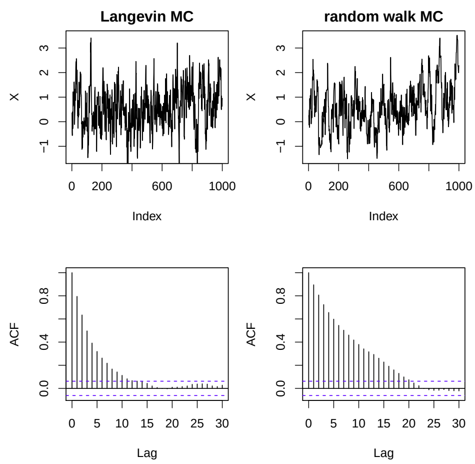


Figure3 Comparison of the two methods: Langevin MC (left) and random walk MC (right).

[*5] See e.g. C. P. Robert and G. Casella (2004), *Monte Carlo Statistical Methods*, 2nd ed., Springer.

5 Exercises

In the following, W_t denotes the standard Brownian motion.

Problem 1. Let B_t be the Brownian bridge defined by (3). Show that B_t has the mean zero and autocovariance function (4).

Problem 2. Let $Z = \sum_{i=1}^{N} b_i(W_{t_i} - W_{t_{i-1}})$ for $0 = t_0 < t_1 < \cdots < t_N$, where b_i is a function of $\{W_s\}_{s \le t_i}$ with $E[b_i^2] < \infty$. Show that $E[Z] = 0$ and $V[Z] = \sum_{i=1}^{N} E[b_{t_i}^2](t_i - t_{i-1})$.

Problem 3. Show that $W_t^2 - t$ is a martingale.

Problem 4. For the Ornstein-Uhlenbeck process (7), show that the conditional density function of X_t given $X_0 = x$ is $N(xe^{-\theta t}, \sigma^2(1 - e^{-2\theta t})/2\theta)$. You may use the fact that $\int_0^t b_s dW_s$ is a Gaussian random variable if b_t is deterministic (non-random). Deduce that $N(0, \sigma^2/(2\theta))$ is a stationary distribution.

Problem 5. Use Itô's formula to show the following identities.

(a) $\int_0^t W_s^2 dW_s = W_t^3/3 - \int_0^t W_s ds$.

(b) $\int_0^t f(W_s)dW_s = F(W_t) - (1/2)\int_0^t f'(W_s)ds$, where $f(x)$ is a smooth function and $F(x) = \int_0^x f(y)dy$ is the Riemann integral.

Problem 6. Let f and g be smooth functions and let $Y_t = f(g(W_t))$. Obtain the expression of dY_t in the following two ways, and confirm that the results are the same.

(a) First apply Itô's formula to $X_t = g(W_t)$. Then apply the formula to $Y_t = f(X_t)$.

(b) Apply Itô's formula to $Y_t = h(W_t)$, where $h(x) = (f \circ g)(x) = f(g(x))$.

Problem 7 (Multi-dimensional diffusion)**.** Let $W_1(t), \ldots, W_d(t)$ be independent standard Brownian motions. A d-dimensional diffusion process $X(t) = (X_1(t), \ldots, X_d(t))$ is defined by

$$dX_i(t) = \mu_i(t, X(t))dt + \sum_{a=1}^{d} \sigma_{ia}(t, X(t))dW_a(t),$$

where μ_i and σ_{ia} are smooth functions. Itô's formula for $Y(t) = f(X(t))$ is known to be

$$dY(t) = \sum_{j=1}^{d} \frac{\partial f}{\partial x_i} dX_j(t) + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{d} \frac{\partial^2 f}{\partial x_j \partial x_k} dX_j(t) \cdot dX_k(t),$$

where the second-order term is interpreted as

$$(dt)^2 = 0, \quad (dt)(dW_a(t)) = 0, \quad \text{and} \quad dW_a(t) \cdot dW_b(t) = \delta_{ab} dt$$

with Kronecker's delta δ_{ab}. Show that the forward equation is

$$\frac{\partial p}{\partial t} = -\sum_i \frac{\partial}{\partial y_i}(\mu_i p) + \frac{1}{2}\sum_i \sum_j \frac{\partial^2}{\partial y_i \partial y_j}\left(\sum_a \sigma_{ia}\sigma_{ja} p\right),$$

in the same way as the proof of Theorem 4.

Problem 8. For a diffusion process (10), confirm that the forward equation (9) is satisfied by $p(t, y|s, x) = \pi(y) = \pi_*(y)/Z$.

Problem 9. Let $\pi_*(x)$ and $g(x)$ be positive functions on $\mathbb{R}$. Define a diffusion process X_t by

$$dX_t = \frac{1}{2g}\left(\log \frac{\pi_*}{\sqrt{g}}\right)' dt + \frac{1}{2\sqrt{g}}\left(\frac{1}{\sqrt{g}}\right)' dt + \frac{1}{\sqrt{g}} dW_t,$$

where $g = g(X_t)$ and so on. Let $\pi(x) = \pi_*(x)/Z$, where $Z = \int \pi_*(x)dx$ is the normalizing constant. Then show that $\pi(x)$ is a stationary solution of the forward equation. This result is related to the Riemannian manifold Langevin Monte Carlo[*6]. Note that (10) is a special case, $g = 1$.

[*6] M. Girolami and B. Calderhead (2011), Riemann manifold Langevin and Hamitonian Monte Carlo methods, *J. Roy. Statist. Soc.*, **73** (2), 123–214.