Lecture 1: Brownian motion, martingales and Markov processes

David Nualart

Department of Mathematics
Kansas University

Gene Golub SIAM Summer School 2016
Drexel University

Outline

1. Stochastic proceses. Brownian motion. Markov processes.
2. Stopping times. Martingales.
3. Stochastic integrals.
4. Itô's formula and applications.
5. Stochastic differential equations.
6. Introduction to Malliavin calculus.

Multivariate normal distribution

- A random vector $X = (X_1, \ldots, X_n)$ has the *multivariate normal distribution* $N(\mu, \Sigma)$, if its characteristic function is

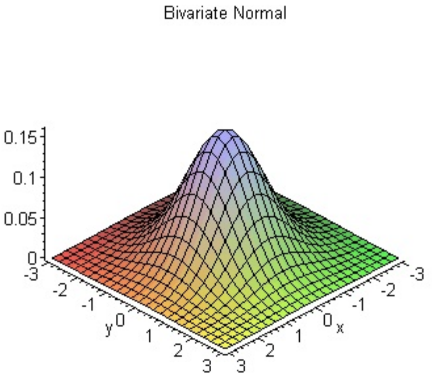
$$E\left(e^{i\langle u,X\rangle}\right) = \exp\left(i\langle u,\mu\rangle - \frac{1}{2}u^T\Sigma u\right),\ u \in \mathbb{R}^n,$$

  where $\mu \in \mathbb{R}^n$ and Σ is an $n \times n$ symmetric and nonnegative definite matrix.

- $\mu = (E(X_1), \ldots, E(X_n))$
- $\Sigma_{ij} = \mathrm{Cov}(X_i, X_j)$
- If X has the $N(\mu, \Sigma)$ distribution, then $Y = AX + b$, where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$, has the $N(A\mu + b, A\Sigma A^T)$ distribution.

- If Σ is nonsingular, then X has a density given by

$$f(x) = (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \right).$$



Stochastic processes

- A stochastic process $X = \{X_t, t \geq 0\}$ is a family of random variables

$$X_t : \Omega \to \mathbb{R}$$

defined on a probability space $(\Omega, \mathcal{F}, P)$.

Stochastic processes

- A stochastic process $X = \{X_t, t \geq 0\}$ is a family of random variables

$$X_t : \Omega \to \mathbb{R}$$

  defined on a probability space $(\Omega, \mathcal{F}, P)$.

- The probabilities on $\mathbb{R}^n$, $n \geq 1$,

$$P_{t_1,\ldots,t_n} = P \circ (X_{t_1}, \ldots, X_{t_n})^{-1}$$

  where $0 \leq t_1 < \cdots < t_n$, are called the *finite-dimensional marginal distributions* of the process X.

Stochastic processes

- A stochastic process $X = \{X_t, t \geq 0\}$ is a family of random variables

$$X_t : \Omega \to \mathbb{R}$$

defined on a probability space $(\Omega, \mathcal{F}, P)$.

- The probabilities on $\mathbb{R}^n$, $n \geq 1$,

$$P_{t_1,\dots,t_n} = P \circ (X_{t_1}, \dots, X_{t_n})^{-1}$$

where $0 \leq t_1 < \cdots < t_n$, are called the *finite-dimensional marginal distributions* of the process X.

- For every $\omega \in \Omega$, the mapping

$$t \to X_t(\omega)$$

is called a *trajectory* of the process X.

Theorem (Kolmogorov's extension theorem)

Consider a family of probability measures

$$\{P_{t_1,\dots,t_n},\ 0 \le t_1 < \cdots < t_n, n \ge 1\}$$

such that :

(i) $P_{t_1,\dots,t_n}$ *is a probability on* $\mathbb{R}^n$.

(ii) *(Consistence condition) : If* $\{t_{k_1} < \cdots < t_{k_m}\} \subset \{t_1 < \cdots < t_n\}$, *then* $P_{t_{k_1},\dots,t_{k_m}}$ *is the marginal of* $P_{t_1,\dots,t_n}$, *corresponding to the indexes* $k_1, \dots, k_m$.

Then, there exists a stochastic process $\{X_t, t \ge 0\}$ *defined in some probability space* $(\Omega, \mathcal{F}, P)$, *which has the family* $\{P_{t_1,\dots,t_n}\}$ *as finite-dimensional marginal distributions.*

- Take Ω as the set of all functions $\omega : [0, \infty) \to \mathbb{R}$, $\mathcal{F}$ the σ-algebra generated by cylindrical sets, extend the probability from cylindrical sets to $\mathcal{F}$, and set $X_t(\omega) = \omega(t)$.

Gaussian processes

- $X = \{X_t, t \geq 0\}$ is called *Gaussian* if all its finite-dimensional marginal distributions are multivariate normal.
- The law of a Gaussian process is determined by the mean function $E(X_t)$ and the covariance function

$$\mathrm{Cov}(X_t, X_s) = E((X_t - E(X_t))(X_s - E(X_s))).$$

- Suppose $\mu : \mathbb{R}_+ \to \mathbb{R}$, and $\Gamma : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ is symmetric and nonnegative definite :

$$\sum_{i,j=1}^{n} \Gamma(t_i, t_j) a_i a_j \geq 0, \quad \forall\, t_i \geq 0,\ a_i \in \mathbb{R}.$$

  Then there exists a Gaussian process with mean μ and covariance function Γ.

Equivalent processes

- Two processes, X, Y are *equivalent* (or X is a version of Y) if for all $t \geq 0$,
$$P\{X_t = Y_t\} = 1.$$

Equivalent processes

- Two processes, X, Y are *equivalent* (or X is a version of Y) if for all $t \geq 0$,
$$P\{X_t = Y_t\} = 1.$$

- Two equivalent processes may have quite different trajectories. For example, the processes $X_t = 0$ for all $t \geq 0$ and
$$Y_t = \begin{cases} 0 & \text{if} \quad \xi \neq t \\ 1 & \text{if} \quad \xi = t \end{cases}$$
where $\xi \geq 0$ is a continuous random variable, are equivalent, because $P(\xi = t) = 0$, but their trajectories are different.

Equivalent processes

- Two processes, X, Y are *equivalent* (or X is a version of Y) if for all $t \geq 0$,
$$P\{X_t = Y_t\} = 1.$$

- Two equivalent processes may have quite different trajectories. For example, the processes $X_t = 0$ for all $t \geq 0$ and
$$Y_t = \begin{cases} 0 & \text{if} \quad \xi \neq t \\ 1 & \text{if} \quad \xi = t \end{cases}$$
where $\xi \geq 0$ is a continuous random variable, are equivalent, because $P(\xi = t) = 0$, but their trajectories are different.
- Two processes X and Y are said to be *indistinguishable* if
$$X_t(\omega) = Y_t(\omega)$$
for all $t \geq 0$ and for all $\omega \in \Omega^*$, with $P(\Omega^*) = 1$.

Exercise : Two equivalent processes with right-continuous trajectories are indistinguishable.

Regularity of trajectories

Theorem (Kolmogorov's continuity theorem)

Suppose that $X = \{X_t, t \in [0, T]\}$ *satisfies*

$$E(|X_t - X_s|^\beta) \le K|t - s|^{1+\alpha},$$

for all $s, t \in [0, T]$*, and for some constants* $\beta, \alpha > 0$*. Then, there exists a version* $\widetilde{X}$ *of* X *such that, if* $\gamma < \alpha/\beta$*,*

$$|\widetilde{X}_t - \widetilde{X}_s| \le G_\gamma |t - s|^\gamma$$

for all $s, t \in [0, T]$*, where* G_γ *is a random variable.*

- The trajectories of $\widetilde{X}$ are Hölder continuous of order γ for any $\gamma < \alpha/\beta$.

Sketch of the proof :

(i) Suppose $T = 1$. Take $\gamma < \alpha/\beta$ and set $\mathcal{D}_n = \{\frac{k}{2^n}, 0 \leq k \leq 2^n\}$ and $\mathcal{D} = \cup_{n\geq 1}\mathcal{D}_n$. From Chebychev's inequality,

$$\begin{aligned} P(\max_{1\leq k\leq 2^n} |X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| \geq 2^{-\gamma n}) &\leq \sum_{k=1}^{2^n} P(|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| \geq 2^{-\gamma n}) \\ &\leq \sum_{k=1}^{2^n} 2^{\gamma\beta n} E[|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}|^{\beta}) \\ &\leq K 2^{-n(\alpha-\gamma\beta)}. \end{aligned}$$

Because this series of probabilities is convergent, from the Borel-Cantelli lemma, there is a set $\Omega^* \in \mathcal{F}$ with $P(\Omega^*) = 1$ such that for all $\omega \in \Omega^*$, there exists $N(\omega)$ with

$$|X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)| < 2^{-\gamma n}, \quad \forall n \geq N(\omega), \quad \forall 1 \leq k \leq 2^n.$$

(ii) Suppose that $s, t \in \mathcal{D}$ are such that

$$|s - t| \leq 2^{-n}, \quad n \geq N.$$

Then , there exists two increasing sequences $s_k \in \mathcal{D}_k$ and $t_k \in \mathcal{D}_k$, $k \geq n$, converging to s and t respectively, and such that $|s_{k+1} - s_k| \leq 2^{-(k+1)}$, $|t_{k+1} - t_k| \leq 2^{-(k+1)}$ and $|s_n - t_n| \leq 2^{-n}$. Then, from the decomposition

$$X_s - X_t = \sum_{i=n}^{\infty} (X_{s_{i+1}} - X_{s_i}) + (X_{s_n} - X_{t_n}) + \sum_{i=n}^{\infty} (X_{t_i} - X_{t_{i+1}})$$

we obtain

$$|X_t - X_s| \leq \frac{2}{1 - 2^{-\gamma}} 2^{-\gamma n}.$$

This implies that the paths $t \to X_t(\omega)$ are γ-Hölder on $\mathcal{D}$ for all $\omega \in \Omega^*$, which allows us to conclude the proof. $\square$

Brownian motion

A stochastic process $B = \{B_t, t \geq 0\}$ is called a *Brownian motion* if :

i) $B_0 = 0$ almost surely.

ii) *Independent increments :* For all $0 \leq t_1 < \cdots < t_n$ the increments $B_{t_n} - B_{t_{n-1}}, \ldots, B_{t_2} - B_{t_1}$, are independent random variables.

iii) If $0 \leq s < t$, the increment $B_t - B_s$ has the normal distribution $N(0, t-s)$.

iv) With probability one, $t \to B_t(\omega)$ is continuous.

Proposition

Properties i), ii), iii) are equivalent to :

($\star$) *B is a Gaussian process with mean zero and covariance*

$$\Gamma(s,t) = \min(s,t).$$

Proposition

Properties i), ii), iii) are equivalent to :

($\star$) *B is a Gaussian process with mean zero and covariance*

$$\Gamma(s,t) = \min(s,t).$$

Proof :

a) Suppose i), i) and iii). The distribution of $(B_{t_1}, \ldots, B_{t_n})$, for $0 < t_1 < \cdots < t_n$, is normal, because this vector is a linear transformation of $\left(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}\right)$ which has independent and normal components.

The mean is zero, and for $s < t$, the covariance is

$$E(B_s B_t) = E(B_s(B_t - B_s + B_s)) = E(B_s(B_t - B_s)) + E(B_s^2) = s.$$

b) The converse is also easy to show. $\square$

First construction of the Brownian motion

1. The function $\Gamma(s,t) = \min(s,t)$ is symmetric and nonnegative definite because it can be written as

$$\min(s,t) = \int_0^\infty \mathbf{1}_{[0,s]}(r)\mathbf{1}_{[0,t]}(r)dr,$$

so

$$\begin{aligned}\sum_{i,j=1}^n a_i a_j \min(t_i,t_j) &= \sum_{i,j=1}^n a_i a_j \int_0^\infty \mathbf{1}_{[0,t_i]}(r)\mathbf{1}_{[0,t_j]}(r)dr \\ &= \int_0^\infty \left[\sum_{i=1}^n a_i \mathbf{1}_{[0,t_i]}(r)\right]^2 dr \geq 0.\end{aligned}$$

Therefore, by Kolmogorov's extension theorem there exists a Gaussian process B with zero mean and covariance function $\min(s,t)$.

2. The process B satisfies

$$E\left[(B_t - B_s)^{2k}\right] = \frac{(2k)!}{2^k k!}(t-s)^k, \quad s \le t$$

for any $k \ge 1$, because the distribution of $B_t - B_s$ is $N(0, t-s)$.

3. Therefore, by the Kolmogorov's continuity theorem, there exist a version $\widetilde{B}$ of B, such that $\widetilde{B}$ has Hölder continuous trajectories of order γ for any $\gamma < \frac{k-1}{2k}$ on any interval $[0, T]$. This implies that the paths are γ-Hölder on $[0, T]$ for any $\gamma < \frac{1}{2}$ and for any $T > 0$.

Second construction of Brownian motion

Fix $T > 0$.

(i) $\{e_n, n \geq 0\}$ is an orthonormal basis of $L^2([0, T])$.

(ii) $\{Z_n, n \geq 0\}$ are independent $N(0, 1)$ random variables.

Then, as $N \to \infty$,

$$\sup_{0 \leq t \leq T} \left| \sum_{n=0}^{N} Z_n \int_0^t e_n(s) ds - B_t \right| \xrightarrow{\text{a.s.}, L^2} 0.$$

Notice that

$$\begin{aligned}
& E\left[\left(\sum_{n=0}^{N} Z_n \int_0^t e_n(r) dr \right) \left(\sum_{n=0}^{N} Z_n \int_0^s e_n(r) dr \right) \right] \\
= & \sum_{n=0}^{N} \left(\int_0^t e_n(r) dr \right) \left(\int_0^s e_n(r) dr \right) \\
= & \sum_{n=0}^{N} \left\langle \mathbf{1}_{[0,t]}, e_n \right\rangle_{L^2([0,T])} \left\langle \mathbf{1}_{[0,s]}, e_n \right\rangle_{L^2([0,T])} \overset{N \to \infty}{\to} \left\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \right\rangle_{L^2([0,T])} = s \wedge t.
\end{aligned}$$

- In particular, if $T = 2\pi$, $e_0(t) = \frac{1}{\sqrt{2\pi}}$ and $e_n(t) = \frac{1}{\sqrt{\pi}} \cos(nt/2)$, for $n \geq 1$, we obtain the Paley-Wiener representation of Brownian motion :

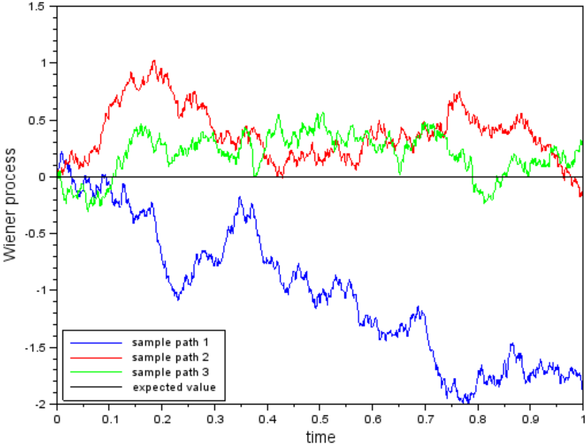
$$B_t = Z_0 \frac{t}{\sqrt{2\pi}} + \frac{2}{\sqrt{\pi}} \sum_{n=1}^{\infty} Z_n \frac{\sin(nt/2)}{n}, \; t \in [0, 2\pi].$$

- In order to use this formula to get a simulation of Brownian motion, we have to choose some number M of trigonometric functions and a number N of discretization points.

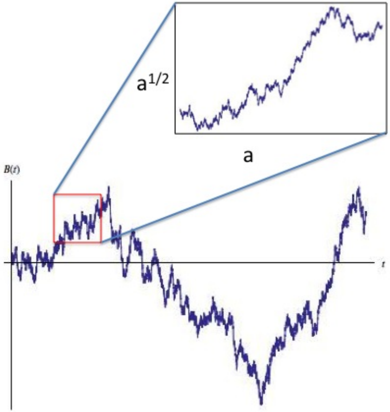
Third construction of Brownian motion

- Let $\{\xi_k, 1 \leq k \leq n\}$ be independent and identically distributed random variables with zero mean and variance one.
- Define $S_n(0) = 0$,
$$S_n(\frac{kT}{n}) = \sqrt{T}\frac{\xi_1 + \cdots + \xi_k}{\sqrt{n}}, \quad k = 1, \ldots, n$$
and extend $S_n(t)$ to $t \in [0, T]$ by linear interpolation.
- *Donsker Invariance Principle* : The law of the random walk S_n on $C([0, T])$ converges to the *Wiener measure*, which is the law of the Brownian motion. That is, that for any continuous and bounded function $\varphi : C([0, T]) \to \mathbb{R}$,
$$E(\varphi(S_n)) \stackrel{n\to\infty}{\longrightarrow} E(\varphi(B)),$$

Simulations of Brownian motion



1. *Selfsimilarity :*
   For any $a > 0$, the process $\{a^{-\frac{1}{2}} B_{at}, t \geq 0\}$ is also a Brownian motion.



2. For any $h > 0$, the process $\{B_{t+h} - B_t, t \geq 0\}$ is a Brownian motion.
3. The process $\{-B_t, t \geq 0\}$ is a Brownian motion.
4. Almost surely $\lim_{t\to\infty} \frac{B_t}{t} = 0$ and the process

$$X_t = \begin{cases} tB_{1/t}, & t > 0 \\ 0, & t = 0 \end{cases}$$

   is a Brownian motion.
5. Almost surely the paths of B are not differentiable at any point $t \geq 0$.
6. $\limsup_{t\to\infty} \frac{B_t}{\sqrt{t}} = \infty$, $\liminf_{t\to\infty} \frac{B_t}{\sqrt{t}} = -\infty$,

Quadratic variation

Fix a time interval $[0, t]$ and consider a partition

$$\pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}.$$

Define $\Delta t_k = t_k - t_{k-1}$, $\Delta B_k = B_{t_k} - B_{t_{k-1}}$ and $|\pi| = \max_{1 \leq k \leq n} \Delta t_k$.

Proposition

The following convergence holds in L^2 :

$$\lim_{|\pi| \to 0} \sum_{k=1}^{n} (\Delta B_k)^2 = t.$$

- We can say that $(\Delta B_t)^2 \sim \Delta t$

Proof : Set $\xi_k = (\Delta B_k)^2 - \Delta t_k$. The random variables ξ_k are independent and centered. Thus,

$$
\begin{aligned}
E\left[\left(\sum_{k=1}^{n} (\Delta B_k)^2 - t\right)^2\right] &= E\left[\left(\sum_{k=1}^{n} \xi_k\right)^2\right] = \sum_{k=1}^{n} E\left[\xi_k^2\right] \\
&= \sum_{k=1}^{n} \left[3\left(\Delta t_k\right)^2 - 2\left(\Delta t_k\right)^2 + \left(\Delta t_k\right)^2\right] \\
&= 2\sum_{k=1}^{n} \left(\Delta t_k\right)^2 \le 2t|\pi| \overset{|\pi|\to 0}{\longrightarrow} 0.
\end{aligned}
$$

□

Exercise : Using the Borel-Cantelli lemma, show that if $\{\pi^n\}$ is a sequence of partitions of $[0, t]$ such that $\sum_n |\pi^n| < \infty$, then $\sum_{k=1}^{n} \left(\Delta B_k\right)^2$ converges almost surely to t.

Infinite total variation

- Define
$$V_t = \sup_{\pi} \sum_{k=1}^{n} |\Delta B_k|$$
- Then,
$$P(V_t = \infty) = 1.$$
In fact, using the continuity of the trajectories of the Brownian motion, we have, on the set $V < \infty$,
$$\sum_{k=1}^{n} (\Delta B_k)^2 \le \sup_k |\Delta B_k| \left(\sum_{k=1}^{n} |\Delta B_k| \right) \le V \sup_k |\Delta B_k| \overset{|\pi| \to 0}{\longrightarrow} 0.$$
Then, $V < \infty$ contradicts the fact that $\sum_{k=1}^{n} (\Delta B_k)^2$ converges in L^2 to t as $|\pi| \to 0$.

Fine properties of the trajectories

- *Lévy's modulus of continuity :*

$$\limsup_{\delta \downarrow 0} \sup_{s,t \in [0,1], |t-s|<\delta} \frac{|B_t - B_s|}{\sqrt{2|t-s| \log |t-s|}} = 1, \quad \text{a.s.}$$

- In contrast, the behavior at at single point is given by the *law of iterated logarithm :*

$$\limsup_{t \downarrow s} \frac{|B_t - B_s|}{\sqrt{2|t-s| \log \log |t-s|}} = 1, \quad \text{a.s.}$$

for any $s \geq 0$.

Conditional expectation

Let X be an integrable random variable on a probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subset \mathcal{F}$ a σ-algebra.

Definition

The conditional expectation $E(X|\mathcal{G})$ is a random variable Y safisfying :

(i) Y is $\mathcal{G}$-measurable.

(ii) For all $A \in \mathcal{G}$,

$$\int_A XdP = \int_A YdP.$$

- if $X \geq 0$, $E(X|\mathcal{G})$ is the density of the measure $\mu(A) = \int_A XdP$, restricted to $\mathcal{G}$, with respect to P.
- By the Radon-Nikodym theorem, $E(X|\mathcal{G})$ exists and it is unique almost surely.

Properties of the conditional expectation

1. *Linearity :*
$$E(aX+bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G}).$$
2. $E\left(E(X|\mathcal{G})\right) = E(X)$.
3. If X and $\mathcal{G}$ are independent, then $E(X|\mathcal{G}) = E(X)$.
4. If X is $\mathcal{G}$-measurable, then $E(X|\mathcal{G}) = X$.
5. If Y is bounded and $\mathcal{G}$-measurable, then
$$E(YX|\mathcal{G}) = YE(X|\mathcal{G}).$$
6. Given two σ-fields $\mathcal{B} \subset \mathcal{G}$, then
$$E\left(E(X|\mathcal{B})|\mathcal{G}\right) = E\left(E(X|\mathcal{G})|\mathcal{B}\right) = E(X|\mathcal{B}).$$

7. Let X and Z be such that :

   (i) Z is $\mathcal{G}$-measurable.
   (ii) X is independent of $\mathcal{G}$.

   Suppose that $E(|h(X,Z)|) < \infty$. Then,

$$\boxed{E(h(X,Z)|\mathcal{G}) = E(h(X,z))\,|_{z=Z}.}$$

Markov processes

- A *filtration* $\{\mathcal{F}_t \subset \mathcal{F}, t \geq 0\}$ is an increasing family of σ-fields.
- A process $\{X_t, t \geq 0\}$ is $\mathcal{F}_t$-adapted if X_t is $\mathcal{F}_t$-measurable for all $t \geq 0$.

Definition

An adapted process X_t is a Markov process with respect to $\mathcal{F}_t$ if for any $s \geq 0$, $t > 0$ and any $f \in C_b(\mathbb{R})$,

$$E[f(X_{s+t})|\mathcal{F}_s] = E[f(X_{s+t})|X_s], \quad \text{a.s.}$$

- This implies that X_t is also an $\mathcal{F}_t^X$-Markov process, where $\mathcal{F}_t^X = \sigma\{X_u, 0 \leq u \leq t\}$.
- The finite-dimensional marginal distributions of a Markov process are characterized by the transition probabilities

$$p(s, x, s+t, B) = P(X_{s+t} \in B | X_s = x).$$

Markov property of Brownian motion

Theorem

The Brownian motion B_t is an $\mathcal{F}_t^B$-Markov process such that, for any $f \in C_b(\mathbb{R})$, $s \geq 0$ and $t > 0$,

$$E[f(X_{s+t})|\mathcal{F}_s^B] = (P_t f)(B_s),$$

where $(P_t f)(x) = \int_{\mathbb{R}} f(y) \frac{1}{\sqrt{2\pi t}} e^{-\frac{|x-y|^2}{2t}} dy$.

- $\{P_t, t \geq 0\}$ is the the semigroup of operators associated with the Brownian motion :

$$\begin{aligned} P_t \circ P_s &= P_{t+s} \\ P_0 &= Id \end{aligned}$$

Proof :

We have

$$E[f(B_{s+t})|\mathcal{F}_s^B] = E[f(B_{s+t} - B_s + B_s)|\mathcal{F}_s^B].$$

Since $B_{s+t} - B_s$ is independent of $\mathcal{F}_s^B$, we obtain

$$\begin{aligned} E[f(B_{s+t})|\mathcal{F}_s^B] &= E[f(B_{s+t} - B_s + x)]|_{x=B_s} \\ &= \int_{\mathbb{R}} f(y + B_s) \frac{1}{\sqrt{2\pi t}} e^{-\frac{|y|^2}{2t}} dy \\ &= \int_{\mathbb{R}} f(y) \frac{1}{\sqrt{2\pi t}} e^{-\frac{|B_s - y|^2}{2t}} dy = (P_t f)(B_s). \end{aligned}$$

□

Multidimensional Brownian motion

- $B_t = (B_t^1, \dots, B_t^d)$ is called a *d-dimensional Brownian motion* if its components are independent Brownian motions.
- It is a Markov process with semigroup

$$(P_t f)(x) = \int_{\mathbb{R}^n} f(y)(2\pi t)^{-\frac{n}{2}} \exp\left(-\frac{|x-y|^2}{2t}\right).$$

- The transition density $p_t(x,y) = (2\pi t)^{-\frac{n}{2}} \exp\left(-\frac{|x-y|^2}{2t}\right)$ satisfies the heat equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\Delta p, \quad t > 0,$$

with initial condition $p_0(x,y) = \delta_x(y)$.

Stopping times

- Consider a filtration $\{\mathcal{F}_t, t \geq 0\}$ in a probability space $(\Omega, \mathcal{F}, P)$, that satisfies the following conditions :
  (i) If $A \in \mathcal{F}$ is such that $P(A) = 0$, then $A \in \mathcal{F}_0$.
  (ii) The filtration is *right-continuous*, that is, for every $t \geq 0$,

$$\mathcal{F}_t = \cap_{n \geq 1} \mathcal{F}_{t+\frac{1}{n}}.$$

Definition

A random variable $T : \Omega \to [0, \infty]$ is a *stopping time* with respect to a filtration $\{\mathcal{F}_t, t \geq 0\}$ if

$$\{T \leq t\} \in \mathcal{F}_t, \quad \forall t \geq 0.$$

Properties of stopping times

1. T is a stopping time if and only if $\{T < t\} \in \mathcal{F}_t$ for all $t \geq 0$.

   Proof :

   $$\{T < t\} = \cup_n \{T \leq t - \frac{1}{n}\} \in \mathcal{F}_t.$$

   Conversely,

   $$\{T \leq t\} = \cap_n \{T < t + \frac{1}{n}\} \in \cap \mathcal{F}_{t+\frac{1}{n}} = \mathcal{F}_t. \quad \square$$

2. $S \vee T$ and $S \wedge T$ are stopping times.
3. Given a stopping time T,

   $$\mathcal{F}_T = \{A : A \cap \{T \leq t\} \in \mathcal{F}_t, \text{ for all } t \geq 0\}.$$

   is a σ-field.
4. $S \leq T \Rightarrow \mathcal{F}_S \subset \mathcal{F}_T$.

5. Let $\{X_t, t \geq 0\}$ be a continuous and adapted process. The *hitting time* of a set $A \subset \mathbb{R}$ is defined by

$$T_A = \inf\{t \geq 0 : X_t \in A\}.$$

Then, if A is open or closed, T_A is a stopping time.

6. Let X_t be an adapted stochastic process with right-continuous paths and $T < \infty$ a stopping time. Then the random variable

$$X_T(\omega) = X_{T(\omega)}(\omega)$$

is $\mathcal{F}_T$-measurable.

Martingales

- We assume that $\{\mathcal{F}_t, t \geq 0\}$ is a filtration.

Definition

An adapted process $M = \{M_t, t \geq 0\}$ is called a *martingale* with respect to $\mathcal{F}_t$ if

(i) For all $t \geq 0$, $E(|M_t|) < \infty$.

(ii) For each $s \leq t$, $E(M_t | \mathcal{F}_s) = M_s$.

Martingales

- We assume that $\{\mathcal{F}_t, t \geq 0\}$ is a filtration.

Definition

An adapted process $M = \{M_t, t \geq 0\}$ is called a *martingale* with respect to $\mathcal{F}_t$ if

(i) For all $t \geq 0$, $E(|M_t|) < \infty$.

(ii) For each $s \leq t$, $E(M_t|\mathcal{F}_s) = M_s$.

- Property (ii) can also be written as :

$$E(M_t - M_s|\mathcal{F}_s) = 0$$

Martingales

- We assume that $\{\mathcal{F}_t, t \geq 0\}$ is a filtration.

Definition

An adapted process $M = \{M_t, t \geq 0\}$ is called a *martingale* with respect to $\mathcal{F}_t$ if

(i) For all $t \geq 0$, $E(|M_t|) < \infty$.

(ii) For each $s \leq t$, $E(M_t|\mathcal{F}_s) = M_s$.

- Property (ii) can also be written as :

$$E(M_t - M_s|\mathcal{F}_s) = 0$$

- M_t is a *supermartingale* (or *submartingale*) if property (ii) is replaced by $E(M_t|\mathcal{F}_s) \leq M_s$ (or $E(M_t|\mathcal{F}_s) \geq M_s$).

Basic properties

1. For any integrable random variable X, $\{E(X|\mathcal{F}_t), t \geq 0\}$ is a martingale.
2. If M_t is a submartingale, then $t \to E[M_t]$ is nondecreasing.
3. If M_t is a martingale and φ is a convex function such that $E(|\varphi(M_t)|) < \infty$ for all $t \geq 0$, then $\varphi(M_t)$ is a submartingale.

   Proof : By Jensen's inequality, if $s \leq t$,

$$E\left(\varphi(M_t)|\mathcal{F}_s\right) \geq \varphi\left(E(M_t|\mathcal{F}_s)\right) = \varphi(M_s). \quad \square$$

   In particular, if M_t is a martingale such that $E(|M_t|^p) < \infty$ for all $t \geq 0$ and for some $p \geq 1$, then $|M_t|^p$ is a submartingale.

Examples :

Let B_t be a Brownian motion $\mathcal{F}_t$ the filtration generated by B_t :

$$\mathcal{F}_t = \sigma\{B_s, 0 \leq s \leq t\}.$$

Then, the processes

$$M_t^{(1)} = B_t$$
$$M_t^{(2)} = B_t^2 - t$$
$$M_t^{(3)} = \exp(aB_t - \frac{a^2 t}{2})$$

where $a \in \mathbb{R}$, are martingales.

1. B_t is a martingale because

$$E(B_t - B_s|\mathcal{F}_s) = E(B_t - B_s) = 0.$$

1. B_t is a martingale because

$$E(B_t - B_s|\mathcal{F}_s) = E(B_t - B_s) = 0.$$

2. For $B_t^2 - t$, we can write, using the properties of the conditional expectation, for $s < t$

$$\begin{aligned}
E(B_t^2|\mathcal{F}_s) &= E((B_t - B_s + B_s)^2 |\mathcal{F}_s) \\
&= E((B_t - B_s\ )^2 |\mathcal{F}_s) + 2E((B_t - B_s\ ) B_s|\mathcal{F}_s) \\
&\quad + E(B_s^2|\mathcal{F}_s) \\
&= E\,(B_t - B_s\ )^2 + 2B_s E((B_t - B_s\ ) |\mathcal{F}_s) + B_s^2 \\
&= t - s + B_s^2.
\end{aligned}$$

1. B_t is a martingale because

$$E(B_t - B_s|\mathcal{F}_s) = E(B_t - B_s) = 0.$$

2. For $B_t^2 - t$, we can write, using the properties of the conditional expectation, for $s < t$

$$\begin{aligned} E(B_t^2|\mathcal{F}_s) &= E((B_t - B_s + B_s)^2 |\mathcal{F}_s) \\ &= E((B_t - B_s)^2 |\mathcal{F}_s) + 2E((B_t - B_s) B_s|\mathcal{F}_s) \\ &\quad + E(B_s^2|\mathcal{F}_s) \\ &= E(B_t - B_s)^2 + 2B_s E((B_t - B_s)|\mathcal{F}_s) + B_s^2 \\ &= t - s + B_s^2. \end{aligned}$$

3. Finally, for $\exp(aB_t - \frac{a^2 t}{2})$ we have

$$\begin{aligned} E(e^{aB_t - \frac{a^2 t}{2}}|\mathcal{F}_s) &= e^{aB_s} E(e^{a(B_t - B_s) - \frac{a^2 t}{2}}|\mathcal{F}_s) \\ &= e^{aB_s} E(e^{a(B_t - B_s) - \frac{a^2 t}{2}}) \\ &= e^{aB_s} e^{\frac{a^2(t-s)}{2} - \frac{a^2 t}{2}} = e^{aB_s - \frac{a^2 s}{2}}. \end{aligned}$$

Optional Stopping Theorem

Theorem (Optional Stopping Theorem)

Suppose that M_t is a continuous martingale and let $S \leq T \leq K$ two bounded stopping times. Then

$$E(M_T|\mathcal{F}_S) = M_S.$$

- This theorem implies that $E(M_T) = E(M_S)$.
- In the submartingale case we have $E(M_T|\mathcal{F}_S) \geq M_S$.
- As a consequence, if T is a bounded stopping time,

$$M_t \quad \text{(sub)martingale} \Rightarrow M_{t \wedge T} \quad \text{(sub)martingale}$$

Proof :

- We will show that $E(M_T) = E(M_0)$.
- Assume first that T takes value in a finite set :

$$0 \le t_1 \le \cdots \le t_n \le K.$$

Then, by the martingale property

$$\begin{aligned} E(M_T) &= \sum_{i=1}^{n} E(M_T \mathbf{1}_{\{T=t_i\}}) = \sum_{i=1}^{n} E(M_{t_i} \mathbf{1}_{\{T=t_i\}}) \\ &= \sum_{i=1}^{n} E(M_{t_n} \mathbf{1}_{\{T=t_i\}}) = E(M_{t_n}) = E(M_0). \end{aligned}$$

- In the general case we approximate T by the following nonincreasing sequence of stopping times

$$\tau_n = \sum_{k=1}^{2^n} \frac{kK}{2^n} \mathbf{1}_{\{\frac{(k-1)K}{2^n} \le T < \frac{kK}{2^n}\}}.$$

- By continuity
$$M_{\tau_n} \stackrel{\text{a.s.}}{\to} M_T.$$
- To show that $E(M_0) = E(M_{\tau_n}) \to E(M_T)$, it suffices to check that the sequence M_{τ_n} is uniformly integrable. This follows from :

$$\begin{aligned}
E(|M_{\tau_n}|\mathbf{1}_{\{|M_{\tau_n}|\geq A\}}) &= \sum_{k=1}^{2^n} E(|M_{\frac{kK}{2^n}}|\mathbf{1}_{\{|M_{\frac{kK}{2^n}}|\geq A, \tau_n = \frac{kK}{2^n}\}}) \\
&\leq \sum_{k=1}^{2^n} E(|M_K|\mathbf{1}_{\{|M_{\frac{kK}{2^n}}|\geq A, \tau_n = \frac{kK}{2^n}\}}) \\
&= E(|M_K|\mathbf{1}_{\{|M_{\tau_n}|\geq A\}}) \\
&\leq E(|M_K|\mathbf{1}_{\{\sup_{0\leq s\leq K}|M_s|\geq A\}}),
\end{aligned}$$

which converges to zero as $A \uparrow \infty$, uniformly in n. $\square$

Doob's maximal inequalities

Theorem

Let $\{M_t, t \in [0,T]\}$ *be a continuous martingale such that* $E(|M_T|^p) < \infty$ *for some* $p \geq 1$*. Then, for all* $\lambda > 0$ *we have*

$$P\left(\sup_{0\leq t\leq T} |M_t| > \lambda\right) \leq \frac{1}{\lambda^p} E(|M_T|^p). \tag{1}$$

If $p > 1$*, then*

$$E\left(\sup_{0\leq t\leq T} |M_t|^p\right) \leq \left(\frac{p}{p-1}\right)^p E(|M_T|^p). \tag{2}$$

Proof of (1) :

- Set
$$\tau = \inf\{s \geq 0 : |M_s| \geq \lambda\} \wedge T.$$
Because τ is a bounded stopping time and $|M_t|^p$ is a submartingale,
$$E(|M_\tau|^p) \leq E(|M_T|^p).$$

- From the definition of τ,
$$|M_\tau|^p \geq \mathbf{1}_{\{\sup_{0\leq t\leq T}|M_t|\geq\lambda\}}\lambda^p + \mathbf{1}_{\{\sup_{0\leq t\leq T}|M_t|<\lambda\}}|M_T|^p.$$
Therefore,
$$P\left(\sup_{0\leq t\leq T}|M_t| > \lambda\right) \leq \frac{1}{\lambda^p}E(\mathbf{1}_{\{\sup_{0\leq t\leq T}|M_t|<\lambda\}}|M_T|^p) \leq \frac{1}{\lambda^p}E(|M_T|^p).$$

Application to Brownian hitting times

Let B_t be a Brownian motion. Fix $a \in \mathbb{R}$ and consider the *hitting time*

$$\tau_a = \inf\{t \geq 0 : B_t = a\}$$

Proposition

If $a < 0 < b$, then

$$P(\tau_a < \tau_b) = \frac{b}{b-a}.$$

Proof : By the optional stopping theorem

$$E(B_{t \wedge \tau_a}) = E(B_0) = 0.$$

Letting $t \to \infty$ and using the dominated convergence theorem, it follows that

$$0 = aP(\tau_a < \tau_b) + b(1 - P(\tau_a < \tau_b)). \quad \square$$

Proposition

Let $T = \inf\{t \geq 0 : B_t \notin (a,b)\}$*, where* $a < 0 < b$*. Then*

$$E(T) = -ab.$$

Proof : Using that $B_t^2 - t$ is a martingale, we get

$$E(B_{T\wedge t}^2) = E(T \wedge t).$$

Therefore,

$$E(T) = \lim_{t\to\infty} E(B_{T\wedge t}^2) = E(B_T^2) = -ab. \quad \square$$

Proposition

Fix $a > 0$. The hitting time

$$\tau_a = \inf\{t \geq 0 : B_t = a\},$$

satisfies

$$E\left[\exp\left(-\alpha\tau_a\right)\right] = e^{-\sqrt{2\alpha}a}. \quad \alpha > 0 \tag{3}$$

Proof :

- For any $\lambda > 0$, the process $M_t = e^{\lambda B_t - \frac{\lambda^2 t}{2}}$ is a martingale such that

$$E(M_t) = E(M_0) = 1.$$

- By the optional stopping theorem we obtain, for all $N \geq 1$.

$$E(M_{\tau_a \wedge N}) = 1.$$

- Notice that $M_{\tau_a \wedge N} = \exp\left(\lambda B_{\tau_a \wedge N} - \frac{\lambda^2(\tau_a \wedge N)}{2}\right) \leq e^{a\lambda}$. So, by the dominated convergence theorem we obtain

$$E\left(M_{\tau_a}\right) = 1,$$

that is,

$$E\left(\exp\left(-\frac{\lambda^2 \tau_a}{2}\right)\right) = e^{-\lambda a}.$$

With the change of variables $\frac{\lambda^2}{2} = \alpha$, we get

$$E\left(\ \exp\left(-\alpha \tau_a\right)\right) = e^{-\sqrt{2\alpha} a}. \quad \square \tag{4}$$

- We can compute the distribution function of τ_a :

$$P(\tau_a \le t) = \int_0^t \frac{ae^{-a^2/2s}}{\sqrt{2\pi s^3}} ds.$$

- The expectation of τ_a can be obtained by computing the derivative of (4) with respect to the variable a :

$$E(\tau_a \exp(-\alpha\tau_a)) = \frac{ae^{-\sqrt{2\alpha}a}}{\sqrt{2\alpha}},$$

and letting $\alpha \downarrow 0$ we obtain $E(\tau_a) = +\infty$.

Strong Markov property

Theorem

Let B be a Brownian motion and let T be a finite stopping time with respect to the filtration $\mathcal{F}_t^B$ generated by B. Then the process

$$\{B_{T+t} - B_T, t \geq 0\}$$

is a Brownian motion independent of B_T.

- As a consequence, for any $f \in C_b(\mathbb{R})$ and any finite stopping time T for the filtration $\mathcal{F}_t^B$, we have

$$E[f(B_{T+t})|\mathcal{F}_T^B] = (P_t f)(B_T),$$

where P_t is the semigroup associated with the Brownian motion B.

Proof :

- Consider the process $\tilde{B}_t = B_{T+t} - B_T$ and suppose first that T is bounded. Let $\lambda \in \mathbb{R}$ and $0 \leq s \leq t$. Applying the optional stopping theorem to the martingale

$$\exp\left(i\lambda\tilde{B}_t + \frac{\lambda^2 t}{2}\right),$$

  yields

$$E\left[e^{i\lambda B_{T+t}+\frac{\lambda^2}{2}(T+t)}|\mathcal{F}_{T+s}\right] = e^{i\lambda B_{T+s}+\frac{\lambda^2}{2}(T+s)}.$$

  Therefore,

$$E\left[e^{i\lambda(B_{T+t}-B_{T+s})}|\mathcal{F}_{T+s}\right] = e^{-\frac{\lambda^2}{2}(t-s)}.$$

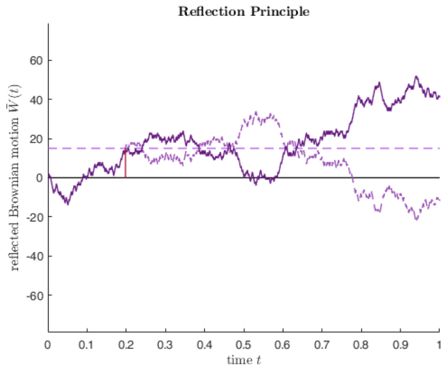
- This implies that the increments of $\tilde{B}$ are independent, stationary and normally distributed.
- If T is not bounded, then we can consider the stopping time $T \wedge N$ and let $N \to \infty$.

Reflection principle

Theorem

Let $M_t = \sup_{0 \le s \le t} B_s$. Then

$$P(M_t \geq a) = 2P(B_t > a) = 2\frac{1}{\sqrt{2\pi t}} \int_a^\infty e^{-\frac{x^2}{2}} dx.$$



Proof :

- We have

$$\begin{aligned} P(B_t \geq a) &= P(B_t \geq a, M_t \geq a) = P(B_t \geq a | M_t \geq a) P(M_t \geq a) \\ &= P(B_t \geq a | \tau_a \leq t) P(M_t \geq a). \end{aligned}$$

- We know that $\{B_{\tau_a + s} - a, s \geq 0\}$ is a Brownian motion independent of $\mathcal{F}_{\tau_a}$. Therefore,

$$P(B_t \geq a | \tau_a \leq t) = E[P(B_{\tau_a + (t - \tau_a)} - a \geq 0 | \mathcal{F}_{\tau_a}) | \tau_a \leq t] = \frac{1}{2}.$$

Brownian filtration

- Define
$$\mathcal{F}_t^B = \sigma\left\{B_s, 0 \le s \le t\right\}.$$
Denote by $\mathcal{N}$ the family of sets in $\mathcal{F}$ of probability zero (null sets).

Proposition

The filtration
$$\mathcal{F}_t = \sigma\left\{\mathcal{F}_t^B, \mathcal{N}\right\}.$$
is right-continuous. Therefore, it satisfies conditions (i) and (ii).