Chapter 4: Martingales and Brownian Motions

Lecturer: Kenneth Ng

Preview

This chapter introduces adapted stochastic processes within the framework of evolving information, formalized by filtrations. We begin by defining filtrations and the notion of adaptedness. We then explore two important classes of processes that exhibit special properties: Markov processes and martingales. In the last part of this chapter, we introduce the Brownian motion along with their properties, which is the building block of stochastic calculus and modelling in the subsequent chapters.

Key topics in this chapter:

1. Filtrations and adapted processes;
2. Markov processes;
3. Martingales;
4. Brownian motions and properties.

1 Adapted Stochastic Processes

Generally speaking, a *stochastic process* on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a family of $\mathcal{F}$-measurable random variables $\{X_t\}_{t\in\mathcal{T}}$, indexed by a time parameter $t \in \mathcal{T}$:

1. If $\mathcal{T} = \mathbb{N}_0$, $\{X_t\}_{t\in\mathcal{T}}$ is a **discrete-time process**;
2. If $\mathcal{T} = \mathbb{R}_+ := [0, \infty)$, $\{X_t\}_{t\in\mathcal{T}}$ is a **continuous-time process**;

In this course, we will primarily focus on continuous-time stochastic processes. For convenience, we will use X to denote the process $\{X_t\}_{t\in\mathcal{T}}$ when no confusion is caused.

By definition, each X_t is $\mathcal{F}$-measurable. Yet, this sole requirement does not capture the *non-anticipative* nature of a realistic stochastic process. This property means that the value of the process at time t should depend only on the information available up to time t, and not on any future information.

For example, let $T > 0$ and $\{S_t\}_{t\in[0,T]}$ be the stochastic process such that S_t represents the price of a risky asset at time t. Define $M_t := \max_{u\in[t,T]} S_u$, which represents the maximum price of the asset over the remaining time interval $[t,T]$. The process $\{M_t\}_{t\in[0,T]}$ fails to be non-anticipative, M_t depends on the price of the asset in the future.

The above example highlights the importance of restricting realistic stochastic processes to rely only on the information available up to the current time. We first define a *filtration* as follows.

Definition 1.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. The collection $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ is called a ***filtration*** if

1. for each $t \in \mathcal{T}$, $\mathcal{F}_t$ is a sub-σ-algebra of $\mathcal{F}$;
2. for any $s \leq t$, $s, t \in \mathcal{T}$, $\mathcal{F}_s \subseteq \mathcal{F}_t$.

In that case, we call the tuple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$ a ***filtered probability space***.

Within a filtration $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$, each $\mathcal{F}_t$ represents the collection of events whose outcomes are known by time t. The second defining property of a filtration ensures that information is cumulative: any event observable at an earlier time remains observable at all later times. Given a filtration, we define an adapted stochastic process as follows.

Definition 1.2 A stochastic process is said to be ***adapted*** to the filtration $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ if, for any $t \in \mathcal{T}$, X_t is $\mathcal{F}_t$-measurable.

In other words, the random variable X_t is fully determined by the information up to time t as encoded by $\mathcal{F}_t$.

Definition 1.3 Let $\{X_t\}_{t\in\mathcal{T}}$ be a stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$. For each fixed $\omega \in \Omega$, the ***sample path*** of the process is the function

$$t \mapsto X_t(\omega), \quad t \geq 0.$$

That is, a sample path is a trajectory of the stochastic process as a function of time.

Definition 1.4 Let $X = \{X_t\}_{t\in\mathcal{T}}$ be an adapted process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$. The ***filtration generated by*** X, denoted $\{\mathcal{F}_t^X\}_{t\in\mathcal{T}}$, is defined by

$$\mathcal{F}_t^X := \sigma(X_s : s \leq t \in \mathcal{T}),$$

i.e., $\mathcal{F}_t^X$ is the smallest σ-algebra with respect to which all X_s for $s \leq t \in \mathcal{T}$ are measurable.

2 Markov Processes

Definition 2.1 A $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$-adapted process $X = \{X_t\}_{t\in\mathcal{T}}$ is called a ***Markov process*** if, for any $s \le t \in \mathcal{T}$ and $A \in \mathcal{F}_t$,

$$\mathbb{P}(X_t \in A|\mathcal{F}_s) = \mathbb{P}(X_t \in A|\sigma(X_s)) \text{ a.s.}$$

Equivalent, X is Markov if, for any $s \le t \in \mathcal{T}$ and any bounded, Borel measurable function f, there exists another Borel measurable function g such that

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = g(X_s)$$

The following independence lemma is useful for showing the Markov property of a stochastic process.

Lemma 2.1 Let $X_1, \ldots, X_n$ be $\mathcal{G}$-measurable random variables, and $Y_1, \ldots, Y_m$ be random variables that are independent of $\mathcal{G}$. Let $f(x_1, \ldots, x_n, y_1, \ldots, y_m)$ be a measurable function and define

$$g(x_1, \ldots, x_n) := \mathbb{E}[f(x_1, \ldots, x_n, Y_1, \ldots, Y_m)].$$

Then,

$$\mathbb{E}[f(X_1, \ldots, X_n, Y_1, \ldots, Y_m)|\mathcal{G}] = g(X_1, \ldots, X_n).$$

Example 2.1 (**Random Walk**) Consider a sequence of i.i.d. random variables $\{\xi_n\}_{n=1}^{\infty}$, where $\mathbb{P}(\xi_n = 1) = p$, and $\mathbb{P}(\xi_n = -1) = 1 - p$, where $p \in [0, 1]$. Define a filtration $\{\mathcal{F}_n\}_{n=0}^{\infty}$ with discrete time step by $\mathcal{F}_n := \sigma(\xi_k : 0 \le k \le n)$, which is generated by the independent trails up to time n. Define a $\{\mathcal{F}_n\}_{n=1}^{\infty}$-adapted process $\{X_n\}_{n=1}^{\infty}$ by $X_1 := \xi_1$, and for $n > 1$,

$$X_n = \sum_{k=1}^{n} \xi_k = X_{n-1} + \xi_n.$$

We show that $\{X_n\}_{n=1}^{\infty}$ is a Markov process: for any $m > n$, $X_m = X_n + \sum_{k=n+1}^{m} \xi_k$. By the independence assumption, for any $k > n$, we have ξ_k being independent of $\mathcal{F}_n$. Now, for any measurable function f, let

$$g(x) := \mathbb{E}\left[f\left(x + \sum_{k=n+1}^{m} \xi_k\right)\right].$$

By Lemma 2.1,

$$\mathbb{E}[f(X_m)|\mathcal{F}_n] = \mathbb{E}\left[f\left(X_n + \sum_{k=n+1}^{m} \xi_m\right)|\mathcal{F}_n\right] = g(X_n).$$

Therefore, X is Markov.

3 Martingales

Definition 3.1 A $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$-adapted process $\{X_t\}_{t\in\mathcal{T}}$ is called a ***martingale*** if it satisfies the following:

1. for any $t \in \mathcal{T}$, $X_t \in L^1$;
2. for any $s \leq t \in \mathcal{T}$,

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \text{ a.s.}$$

If the second property is replaced by $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$ (resp. $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$ a.s.) for any $0 \leq s \leq t \in \mathcal{T}$, the process $\{X_t\}_{t\geq 0}$ is called a ***sub-martingale*** (resp. ***super-martingale***).

In other words, a martingale is a stochastic process whose future expected values given the current information remain constant. A sub-martingale (resp. sub-martingale) is a process whose future expected values given the current information increases (resp. decrease) over time. By definition, a process is a martingale iff it is both a super-martingale and sub-martingale.

Example 3.1 (**Random Walk**) Continuing from Example 2.1, compute $\mathbb{E}[X_m|\mathcal{F}_n]$ for any $1 \leq n < m$. Hence, determine the values of p such that $\{X_n\}_{n=1}^{\infty}$ is a super-martingale/sub-martingale/martingale.

Solution. We first compute $\mathbb{E}[X_{n+1}|\mathcal{F}_n]$ any $n \geq 1$. Note that by the independence of ξ_{n+1} with ξ_k, $1 \leq k \leq n$, and the fact that X_n is $\mathcal{F}_n$-measurable,

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_n + \xi_{n+1}|\mathcal{F}_n] = X_n + \mathbb{E}[\xi_{n+1}] = X_n + 2p - 1.$$

Using this and the tower property of conditional expectations, we further have

$$\mathbb{E}[X_{n+2}|\mathcal{F}_n] = \mathbb{E}\left[\mathbb{E}[X_{n+2}|\mathcal{F}_{n+1}]|\mathcal{F}_n\right] = \mathbb{E}\left[X_{n+1} + 2p - 1|\mathcal{F}_n\right] = X_n + 2(2p-1).$$

Applying this recursively, for any $m > n$, we have

$$\mathbb{E}[X_m|\mathcal{F}_n] = X_n + (m-n)(2p-1).$$

Since $m - n > 0$, $\{X_n\}$ is a martingale if $p = 1/2$, a sub-martingale if $p \geq 1/2$, and a super-martingale if $p \leq 1/2$. $\square$

Example 3.2 Following from Example 3.1, show that $\{X_n - n(2p-1)\}_{n=1}^{\infty}$ is a martingale for any $p \in [0,1]$.

<u>*Solution*</u>. From Example 3.1, we have shown that, for any $m > n$,

$$\mathbb{E}[X_m|\mathcal{F}_n] = X_n + (m-n)(2p-1).$$

Rearranging yields

$$\mathbb{E}[X_m - m(2p-1)|\mathcal{F}_n] = X_n - n(2p-1).$$

Therefore, the process $\{X_n - n(2p-1)\}_{n=1}^{\infty}$ is a martingale.
Indeed, the term $n(2p-1)$ serves as a *compensator* which adjusts for the drift of the expected value of X_n. By subtracting this drift, the resulting process becomes centered and thus martingale-valued. □

Example 3.3 Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration. Then, $Z_t := \mathbb{E}[X|\mathcal{F}_t]$ is a martingale. To see this, it is clear by definition of conditional expectations that $\{Z_t\}_{t\geq 0}$ is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted. In addition, $X_t \in L^1$, since by triangle inequality,

$$\mathbb{E}[|Z_t|] \leq \mathbb{E}\left[\mathbb{E}[|X||\mathcal{F}_t]\right] = \mathbb{E}[|X|] < \infty,$$

since $X \in L^1$. Finally, for any $0 \leq s \leq t$, using the tower property,

$$\mathbb{E}[Z_t|\mathcal{F}_s] = \mathbb{E}\left[\mathbb{E}[X|\mathcal{F}_t]|\mathcal{F}_s\right] = \mathbb{E}[X|\mathcal{F}_s] = Z_s,$$

which verifies the martingale property.

Proposition 3.1 Let $\{X_t\}_{t\in\mathcal{T}}$ be a martingale. Suppose that $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function and $\varphi(X_t) \in L^1$ for any $t \geq 0$. Then, $\{\varphi(X_t)\}_{t\in\mathcal{T}}$ is a sub-martingale.

As a consequence of Proposition 3.1, if $\{X_t\}_{t\geq 0}$ is a square-integrable martingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, i.e., $\mathbb{E}[X_t^2] < \infty$ for all $t \geq 0$. Then, $\{X_t^2\}_{t\geq 0}$ is a sub-martingale, since the function $\varphi(x) = x^2$ is convex.

Proof. For any $s \leq t \in \mathcal{T}$, using the martingale property of $\{X_t\}$ and Jensen's inequality, we have

$$\mathbb{E}[\varphi(X_t)|\mathcal{F}_s] \geq \varphi\left(\mathbb{E}[X_t|\mathcal{F}_s]\right) = \varphi(X_s) \text{ a.s.}$$

□

4 Brownian Motions

This section introduces the Brownian motion, a fundamental continuous stochastic process and building block of stochastic calculus. It was first observed by Robert Brown in 1828 (and hence the name Brownian motion). The mathematical formulation was later developed by Norbert Wiener, and the process is also known as the Wiener process in his honor.

We first provide the definition of a standard Brownian motion.

Definition 4.1 (Brownian Motion) An adapted process $\{B_t\}_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ is called a ***Brownian motion*** (a.k.a. ***Wiener process***) if it satisfies the following properties:

1. $B_0 = 0$ almost surely;
2. The sample paths of B_t are almost surely continuous;
3. The process has independent increments: for any $0 \leq s \leq t$, $B_t - B_s$ is independent of $\mathcal{F}_s$;
4. The increments are normally distributed: for all $s, t \geq 0$, $B_{t+s} - B_s \sim \mathcal{N}(0, t)$.

Remark 4.1. Since B_s is $\mathcal{F}_s$-measurable, we have $\sigma(B_s) \subset \mathcal{F}_s$. Hence, by Remark 1.1 of Chapter 3, the independent increment implies $B_t - B_s$ and B_s are independent.

There are various ways to construct a continuous-time process satisfying the properties outlined in Definition 4.1. For a comprehensive treatment, readers are referred to the monograph *Brownian Motion and Stochastic Calculus* by Karatzas and Shreve. In this section, we discuss a classical approach by viewing Brownian motion as the (weak) limit of a suitably scaled random walk.

4.1 Limit of Symmetric Random Walk

Consider the symmetric random walk $\{X_n\}_{n=0}^{\infty}$, where

$$X_n = \sum_{k=1}^{n} \xi_k,$$

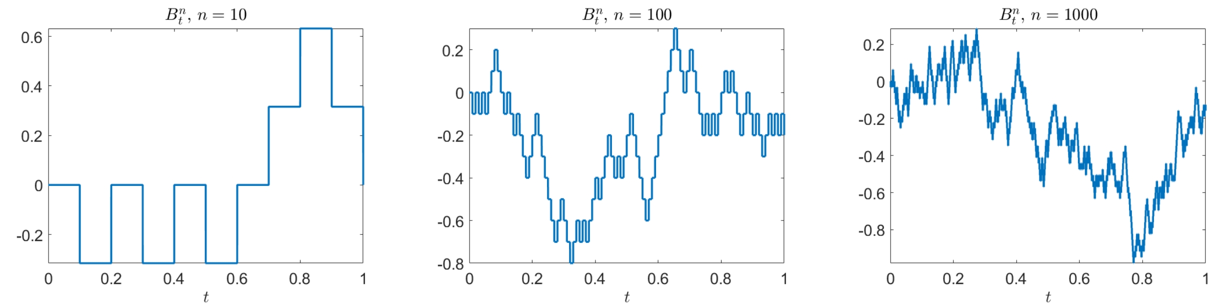
and the random variables $\{\xi_n\}_{n=0}^{\infty}$ are i.i.d. with distribution $\mathbb{P}(\xi_n = 1) = 1/2 = \mathbb{P}(\xi_n = -1)$. For each $n \in \mathbb{N}$, define the scaled process $\{B_t^{(n)}\}_{t\geq 0}$ by

$$B_t^{(n)} := \frac{1}{\sqrt{n}} X_{\lfloor nt \rfloor}, \ t \geq 0.$$

Figure 1 shows a simulated path of $\{B_t^{(n)}\}_{t\in[0,1]}$ for different values of n. One can observe that as n increases, the path becomes more spiky with an increasing frequency of oscillations since more ξ_k's are included in $X_{\lfloor nt \rfloor}$. In addition, it is easy to check that

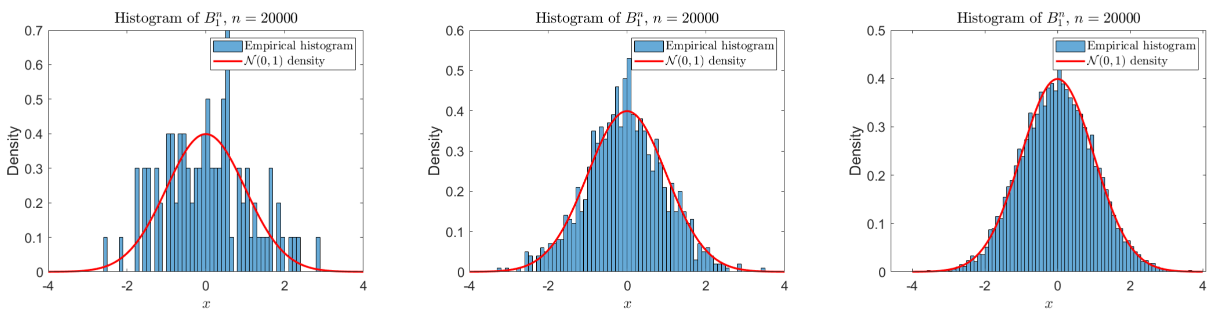
$$\mathbb{E}[B_t^{(n)}] = 0 \text{ and } \mathrm{Var}[B_t^{(n)}] = \frac{\lfloor nt \rfloor}{n}.$$

Figure 2 depicts the distribution of N random samples drawn from B_1^n for $n = 20,000$. As N increases, we see that the histogram converges to the pdf of the standard normal variable.



(a) $n = 10$ (b) $n = 100$ (c) $n = 1000$

Figure 1: Simulations of sample paths of B_t^n for different n



(a) $N = 100$ (b) $N = 1000$ (c) $N = 10000$

Figure 2: Distribution of B_1^n for $n = 20000$ and different number of simulations N. The red curve depicts the density function of $\mathcal{N}(0, 1)$

The following shows the convergence of the finite-dimensional distributions of the scaled random walk to a standard Brownian motion.

Theorem 4.2 Let $0 \le t_1 < \cdots < t_k < \infty$, we have $(B_{t_1}^{(n)}, B_{t_2}^{(n)} - B_{t_1}^{(n)}, \ldots, B_{t_k}^{(n)} - B_{t_{k-1}}^{(n)})$ converges in distribution to $(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}})$.

Proof. For simplicity, we consider $k = 2$, and let $s = t_1$, $t = t_2$. Consider the mgf of $(B_s^{(n)}, B_t^{(n)} - B_s^{(n)})$: for $u, v \in \mathbb{R}$,

$$\begin{aligned} M_{B_s^{(n)}, B_t^{(n)} - B_s^{(n)}}(u, v) &= \mathbb{E}\left[e^{uB_s^{(n)} + v(B_t^{(n)} - B_s^{(n)})}\right] \\ &= \mathbb{E}\left[\exp\left(\frac{u}{\sqrt{n}} \sum_{j=1}^{\lfloor ns \rfloor} \xi_j + \frac{v}{\sqrt{n}} \sum_{j=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \xi_j\right)\right] \end{aligned}$$

$$= \mathbb{E}\left[\exp\left(\frac{u}{\sqrt{n}}\sum_{j=1}^{\lfloor ns\rfloor}\xi_j\right)\right]\mathbb{E}\left[\exp\left(\frac{v}{\sqrt{n}}\sum_{j=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}\xi_j\right)\right].$$

where we have used the i.i.d. property of $\{\xi_j\}_{j=1}^{\infty}$.

We consider the first expectation on the right, which is essentially the mgf of $\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor ns\rfloor}\xi_j$. The second expectation can be handled in a similar fashion. Let $m_n := \lfloor ns\rfloor$. Then

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor ns\rfloor}\xi_j = \frac{1}{\sqrt{n}}\sum_{j=1}^{m_n}\xi_j = \frac{\sqrt{m_n}}{\sqrt{n}}\cdot\frac{1}{\sqrt{m_n}}\sum_{j=1}^{m_n}\xi_j.$$

Since $m_n/n \to s$, we have $\sqrt{m_n}/\sqrt{n} \to \sqrt{s}$ deterministically. By the central limit theorem,

$$\frac{1}{\sqrt{m_n}}\sum_{j=1}^{m_n}\xi_j \xrightarrow{d} \mathcal{N}(0,1).$$

Consequently [1]

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor ns\rfloor}\xi_j \to \mathcal{N}(0,s)$$

in distribution, whence

$$\lim_{n\to\infty}\mathbb{E}\left[\exp\left(\frac{u}{\sqrt{n}}\sum_{j=1}^{\lfloor ns\rfloor}\xi_j\right)\right] = M_{\mathcal{N}(0,s)}(u) = e^{\frac{su^2}{2}}.$$

Likewise,

$$\lim_{n\to\infty}\mathbb{E}\left[\exp\left(\frac{v}{\sqrt{n}}\sum_{j=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}\xi_j\right)\right] = e^{\frac{v^2(t-s)}{2}}.$$

Therefore, we have shown that

$$\lim_{n\to\infty} M_{B_s^{(n)},B_t^{(n)}-B_s^{(n)}}(u,v) = e^{\frac{u^2 s}{2}+\frac{v^2(t-s)}{2}} = M_{B_s,B_t-B_s}(u,v).$$

$\square$

[1] This is a result of Slutsky's theorem

4.2 Properties of Brownian Motion

The first property is concerned with the behavior of the sample paths of Brownian motions.

Theorem 4.3 The sample paths of a Brownian motion are almost surely continuous, but almost surely nowhere differentiable.

By definition, sample paths of Brownian motion are almost surely continuous. When viewing Brownian motion as the weak limit of a scaled symmetric random walk, each sample path $t \mapsto B_t^{(n)}(\omega)$ remains continuous for any fixed n, despite exhibiting a spiky appearance; see Figure 1. However, as n increases, the frequency of oscillations grows, resulting in increasingly irregular paths. This limiting behavior ultimately leads to a function that is continuous everywhere but differentiable nowhere. Consequently, the derivative $\frac{dB_t(\omega)}{dt}$ is ill-defined for a.a. $\omega \in \Omega$.

Theorem 4.4 The Brownian motion $\{B_t\}_{t\geq 0}$ is a Markov process and a martingale.

Proof. We first show that $B = \{B_t\}_{t\geq 0}$ is Markov. For any $0 \leq s \leq t$, we can write $B_t = (B_t - B_s) + B_s$, where $B_t - B_s \perp\!\!\!\perp \mathcal{F}_s$, thanks to the independent increment property of Brownian motions. Let f be a bounded measurable function, and define g by

$$g(x) := \mathbb{E}[f(x + B_t - B_s)].$$

Then, by the independent increment and Lemma 2.1,

$$\mathbb{E}[f(B_t)|\mathcal{F}_s] = \mathbb{E}[f(B_s + (B_t - B_s))|\mathcal{F}_s] = g(B_s).$$

Hence, B is a Markov process.

Next, we prove that the Brownian motion is a martingale. The adaptedness of $\{B_t\}_{t\geq 0}$ follows from the definition. In addition, $B_t \sim \mathcal{N}(0, t)$ and thus $B_t \in L^1$ (Indeed, $\mathbb{E}[|B_t|] = \sqrt{2/\pi}$). Finally, for any $0 \leq s \leq t$, using the independent increment,

$$\mathbb{E}[B_t|\mathcal{F}_s] = \mathbb{E}[B_t - B_s + B_s|\mathcal{F}_s] = \mathbb{E}[B_t - B_s|\mathcal{F}_s] + B_s = \mathbb{E}[B_t - B_s] + B_s = B_s.$$

Hence, $\{B_t\}_{t\geq 0}$ is a martingale.

□

The correlation structure of Brownian motion across different time points offers another distinctive characterization of the process.

Proposition 4.5 Let $\{B_t\}_{t\geq 0}$ be a Brownian motion. Then, for any $s, t \geq 0$,

$$\mathrm{Cov}(B_s, B_t) = \min\{s, t\}.$$

Proof. Without loss of generality, we assume that $s \leq t$ and show that $\mathrm{Cov}(B_s, B_t) = s$. Indeed, using the fact that $B_t \sim \mathcal{N}(0,t)$, $B_s \sim \mathcal{N}(0,s)$, and the independent increment of Brownian motions,

$$\begin{aligned}\mathrm{Cov}(B_s, B_t) &= \mathbb{E}[B_s B_t] - \mathbb{E}[B_s]\mathbb{E}[B_t] \\ &= \mathbb{E}[B_s B_t] - 0 \\ &= \mathbb{E}[B_s(B_t - B_s + B_s)] \\ &= \mathbb{E}[B_s(B_t - B_s)] + \mathbb{E}[B_s^2] \\ &= \mathbb{E}[B_s]\mathbb{E}[B_s - B_t] + \mathbb{E}[B_s^2] \quad \text{(independent increment)} \\ &= s.\end{aligned}$$

□

The following theorem shows that Brownian motions are scale-invariant.

Theorem 4.6 Let $\{B_t\}_{t\geq 0}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. Then, for any $c > 0$, $\{\frac{1}{c}B_{c^2 t}\}_{t\geq 0}$ is a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_{c^2 t}\}_{t\geq 0}, \mathbb{P})$.

Proof. Let $\tilde{B}_t := \frac{1}{c}B_{c^2 t}$ and $\mathcal{G}_t := \mathcal{F}_{c^2 t}$. It is clear that $\tilde{B}_t$ has continuous sample paths, and is adapted to $\mathcal{G}_t$ since $B_{c^2 t}$ is $\mathcal{F}_{c^2 t}$-measurable. Therefore, it suffices to show that $\tilde{B}_t$ has independent and Gaussian increment.

Since B is a standard Brownian motion, for any $0 \leq s < t$ we have $B_{c^2 t} - B_{c^2 s} \perp\!\!\!\perp \mathcal{F}_{c^2 s}$, which implies $\tilde{B}_t - \tilde{B}_s \perp\!\!\!\perp \mathcal{G}_s$. This shows that $\tilde{B}_t$ has independent increment. Finally, using the fact that $B_{c^2 t} - B_{c^2 s} \sim \mathcal{N}(0, c^2(t-s))$ for any $0 \leq s < t$, we have

$$\tilde{B}_t - \tilde{B}_s = \frac{1}{c}(B_{c^2 t} - B_{c^2 s}) \sim \mathcal{N}(0, t-s).$$

Therefore, $\tilde{B}_t$ has Gaussian increment. □

Example 4.1 (Brownian motions with drift) Let $\{B_t\}_{t\geq 0}$ be a standard Brownian motion. Then, for any $\mu \in \mathbb{R}$, the process $B_t^\mu := \mu t + B_t$ is called a ***Brownian motion with drift*** μ. Note that $\mathbb{E}[B_t^\mu] = \mu t$, and B^μ is a super-martingale (resp. sub-martingale) if $\mu \leq 0$ (resp. $\mu \geq 0$).

4.3 Quadratic Variations

Let $\{X_t\}_{t\geq 0}$ be a square-integrable martingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. Recall from Proposition 3.1 and the discussion following it, $\{X_t^2\}_{t\geq 0}$ is a sub-martingale. Under mild conditions, this sub-martingale has the following (unique) representation:

$$X_t^2 = M_t + A_t,$$

where

1. $\{M_t\}_{t\geq 0}$ is a martingale;
2. $\{A_t\}_{t\geq 0}$ is an increasing process, i.e., $A_t \geq A_s$ a.s. for any $t \geq s$.

The representation is called the ***Doob-Meyer decomposition***. It says that a sub-martingale can be written as a martingale part, and an increasing part which drives up the conditional expectations. In particular, the process $\{A_t\}_{t\geq 0}$ is called the ***quadratic variation*** of the martingale $\{X_t\}_{t\geq 0}$:

Definition 4.2 Let $\{X_t\}_{t\geq 0}$ be a square-integrable martingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. The ***quadratic variation process*** of X_t, denoted by $\langle X\rangle_t$, is the unique process such that $\langle X\rangle_0 = 0$ and $X_t^2 - \langle X\rangle_t$ is a martingale.

Example 4.2 Following from Example 2.1, when $p = 1/2$, $\{X_n\}_{n=1}^\infty$ is a martingale. Show that $\{X_n^2 - n\}_{n=1}^\infty$ is a martingale, and thus the quadratic variation of X_n is n.

Solution. For any $n \geq 1$,

$$X_{n+1}^2 = (X_n + \xi_{n+1})^2 = X_n^2 + 2\xi_{n+1}X_n + \xi_{n+1}^2.$$

Using this, the i.i.d. property of $\{\xi_n\}_{n=1}^\infty$, and the $\mathcal{F}_n$-measurability of X_n, we have

$$\begin{aligned}
\mathbb{E}[X_{n+1}^2 - (n+1)|\mathcal{F}_n] &= \mathbb{E}[X_n^2|\mathcal{F}_n] + 2\mathbb{E}[\xi_{n+1}X_n|\mathcal{F}_n] + \mathbb{E}[\xi_{n+1}^2|\mathcal{F}_n] - n - 1 \\
&= X_n^2 + 2X_n\mathbb{E}[\xi_n] + \mathbb{E}[\xi_{n+1}^2] - n - 1 \\
&= X_n^2 + \frac{1^2 + (-1)^2}{2} - n - 1 \\
&= X_n^2 - n.
\end{aligned}$$

Applying this recursively using the tower property of conditional expectations, we have that

$$\mathbb{E}[X_m^2 - m|\mathcal{F}_n] = X_n^2 - n$$

for any $m \geq n \geq 1$. □

Theorem 4.7 The quadratic variation of the standard Brownian motion $\{B_t\}_{t\geq 0}$ is given by $\langle B\rangle_t = t$, i.e., $B_t^2 - t$ is a martingale.

Proof. It is clear that $\{Y_t := B_t^2 - t\}_{t\geq 0}$ is $\{\mathcal{F}_t\}$-adapted, and $Y_t \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ for all $t \geq 0$.

It remains to verify that, for any $0 \le s \le t$, $\mathbb{E}[Y_t|\mathcal{F}_s] = Y_s$ a.s. Indeed,

$$\begin{aligned}
\mathbb{E}[Y_t|\mathcal{F}_s] &= \mathbb{E}[B_t^2 - t|\mathcal{F}_s] \\
&= \mathbb{E}[(B_t - B_s + B_s)^2|\mathcal{F}_s] - t \\
&= \mathbb{E}[(B_t - B_s)^2] + 2\mathbb{E}[B_s(B_t - B_s)|\mathcal{F}_s] + \mathbb{E}[B_s^2|\mathcal{F}_s] - t \\
&= \mathbb{E}[(B_t - B_s)^2] + 2B_s\mathbb{E}[B_t - B_s] + B_s^2 - t \\
&= t - s + B_s^2 - t \\
&= B_s^2 - s = Y_s,
\end{aligned}$$

where we have used the independent increment of Brownian motions, and the fact that B_s is $\mathcal{F}_s$-measurable. □

The term *variation* often refers to partial sums of increments. An alternative definition of quadratic variation considers the limit of squared increments over increasingly finer partitions as below. Quadratic variation plays a central role in stochastic calculus. As we will see in the next chapter, it gives rise to the additional term in Itô's lemma that distinguishes stochastic calculus from classical calculus.

Definition 4.3 Let $\{X_t\}_{t\ge0}$ be an adapted process. Fix $t > 0$ and let $\Pi = \{t_0, t_1, \dots, t_m\}$ with $0 \le t_0 \le t_1 \le \cdots \le t_m = t$ be a partition of $[0, t]$. The ***p-th variation of X over the partition*** Π is defined as

$$V_t^{(p)}(\Pi) := \sum_{k=1}^{m} \left|X_{t_k} - X_{t_{k-1}}\right|^p .$$

In particular, if $V_t^{(2)}(\Pi)$ is convergent in some sense as $\|\Pi\| := \max_{1\le k\le m} |t_k - t_{k-1}| \to 0$, the limit is referred to as the ***quadratic variation***.

The following theorem shows the equivalence of Definitions 4.2 and 4.3 concerning the quadratic variation of a continuous, square-integrable martingale.

Theorem 4.8 Let $\{X_t\}_{t\ge0}$ be a square-integrable martingale with continuous sample paths, and Π be a partition of $[0, t]$. Then,

$$\lim_{\|\Pi\|\to0} V_t^{(2)}(\Pi) = \langle X\rangle_t$$

in probability.

Example 4.3 For any $t > 0$, and any partition Π of $[0, t]$, show that the 2nd-variation of a standard Brownian motion $V_t^{(2)}(\Pi)$ satisfies

$$V_t^{(2)}(\Pi) \xrightarrow{L^2} t.$$

<u>*Solution*</u>. Fix $t > 0$ and let $\Pi = \{t_0, t_1, \ldots, t_m\}$ be a partition of $[0, t]$. Using the fact that $B_{t_k} - B_{t_{k-1}} \sim \mathcal{N}(0, t_k - t_{k-1})$, we have

$$\mathbb{E}\left[V_t^{(2)}(\Pi)\right] = \sum_{k=1}^{m} \mathbb{E}\left[(B_{t_k} - B_{t_{k-1}})^2\right] = \sum_{k=1}^{m} t_k - t_{k-1} = t.$$

Hence, to show that $V_t^{(2)}(\Pi) \xrightarrow{L^2} t$, it suffices to show that

$$\operatorname{Var}\left[V_t^{(2)}(\Pi)\right] = \mathbb{E}\left[\left(V_t^{(2)}(\Pi) - t\right)^2\right] \to 0$$

as $\|\Pi\| \to 0$.
Using the independent increment, we have

$$\begin{aligned}
\operatorname{Var}\left[V_t^{(2)}(\Pi)\right] &= \sum_{k=1}^{m} \operatorname{Var}\left[|B_{t_k} - B_{t_{k-1}}|^2\right] \\
&= \sum_{k=1}^{m} \left(\mathbb{E}\left[|B_{t_k} - B_{t_{k-1}}|^4\right] - \mathbb{E}^2\left[|B_{t_k} - B_{t_{k-1}}|^2\right]\right)
\end{aligned}$$

Recall that for $X \sim \mathcal{N}(0, \sigma^2)$,

$$\mathbb{E}\left[X^2\right] = \sigma^2 \text{ and } \mathbb{E}[X^4] = 3\sigma^4.$$

Hence,

$$\begin{aligned}
\operatorname{Var}\left[V_t^{(2)}(\Pi)\right] &= \sum_{k=1}^{m} \left(\mathbb{E}\left[|B_{t_k} - B_{t_{k-1}}|^4\right] - \mathbb{E}^2\left[|B_{t_k} - B_{t_{k-1}}|^2\right]\right) \\
&= \sum_{k=1}^{m} \left(3(t_k - t_{k-1})^2 - (t_k - t_{k-1})^2\right) \\
&= 2\sum_{k=1}^{m} (t_k - t_{k-1})^2 \\
&\leq 2\|\Pi\| \sum_{k=1}^{m} (t_k - t_{k-1}) \\
&= 2t\|\Pi\| \to 0
\end{aligned}$$

as $\|\Pi\| \to 0$. Therefore, we conclude that $V_t^{(2)}(\Pi) \xrightarrow{L^2} t$. □

The above example shows that the quadratic variation of B is finite, yet non-zero. Indeed, as shown below, any measurable and differentiable function has a zero quadratic variation. This echos the nowhere differentiability of the sample paths of Brownian motions.

Theorem 4.9 Let $f : [0, \infty) \to \mathbb{R}$ be a differentiable function such that for every $t \geq 0$,

$$\int_0^t |f'(s)|^2 \, ds < \infty,$$

where the integral is understood in the Riemann sense. Then, the quadratic variation of f over $[0, t]$ is zero for all $t \geq 0$.

Proof. Fix $t > 0$ and let $\Pi = \{t_0, \ldots, t_m\}$ be a partition of the interval $[0, t]$. Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a differentiable function, and define

$$V_t^{(2)}(\Pi) := \sum_{k=1}^{m} |f(t_k) - f(t_{k-1})|^2 .$$

By the mean value theorem, for any $k = 1, \ldots, m$, there exists $\xi_k \in [t_{k-1}, t_k]$ such that $f(t_k) - f(t_{k-1}) = f'(\xi_k)(t_k - t_{k-1})$. Hence,

$$V_t^{(2)}(\Pi) = \sum_{k=1}^{m} |f'(\xi_k)|^2 (t_k - t_{k-1})^2 \leq \|\Pi\| \sum_{k=1}^{m} |f'(\xi_k)|^2 (t_k - t_{k-1})$$

As $\|\Pi\| \to 0$, we have

$$\sum_{k=1}^{m} |f'(\xi_k)|^2 (t_k - t_{k-1}) \to \int_0^t |f'(s)|^2 ds.$$

Therefore, given that $\int_0^t |f'(s)|^2 ds < \infty$, we have

$$\lim_{\|\Pi\| \to 0} V_t^{(2)}(\Pi) \leq \lim_{\|\Pi\| \to 0} \|\Pi\| \sum_{k=1}^{m} |f'(\xi_k)|^2 (t_k - t_{k-1}) = 0.$$

$\square$

4.4 Multivariate Brownian Motions

We define a standard d-dimensional Brownian motion as follows:

Definition 4.4 Let d be a positive integer. An adapted process $\{B_t = (B_t^1, \ldots, B_t^d)\}_{t\geq 0}$ on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ taking values in $\mathbb{R}^d$ is said to be a *standard d-**dimensional Brownian motion*** if it satisfies the following:

1. $B_0 = \mathbf{0}_d$ almost surely;
2. The sample paths of B_t are almost surely continuous;
3. The process has independent increments: for any $0 \leq s \leq t$, $B_t - B_s$ is independent of $\mathcal{F}_s$;
4. The increments follow multivariate normal distributions: for all $s, t \geq 0$, $B_{t+s} - B_s \sim \mathcal{N}(\mathbf{0}_d, (t-s)I_d)$, where I_d is the $d \times d$ identity matrix.

It is easy to verify the following properties for a standard d-dimensional Brownian motion B:

1. B is a Markov process and a martingale;
2. Write $B_t = (B_t^1, \ldots, B_t^d)$. Then, for each $i = 1, \ldots, d$, B_t^i is a standard one-dimensional Brownian motion.