Introduction to Stochastic Calculus

Yuewu Xu, Ph.D.

School of Business

Fordham University

Fall, 2022

Overview

- Continuous-time finance is one of the most exciting and most important developments of modern finance. It has fundamentally changed how risks are priced and managed in the real world.
- Main asset pricing ideas from single-period models:
  - Mean-variance portfolio theory: Markowitz (1952).
  - Capital asset pricing model: Sharpe (1964) and Lintner (1965).
  - Arbitrage pricing theory: Ross (1976, 1977).
- Continuous-time finance models extend the insights of single-period models and provide more realistic modeling of financial markets.
- Discrete vs. continuous-time models
  - Multi-period models can be studied in discrete or continuous time.
  - Each approach has its own dis/advantages.
  - Discrete-time models are easier to understand, but sometimes are more difficult to solve.
  - Continuous-time models rely on more complicated mathematics, but can lead to more elegant and powerful solutions.
- Potential topics covered
  - Introduction to stochastic calculus.
  - Option pricing: Black-Scholes model, risk-neutral pricing.
  - Term structure of interest rates (time permits)
  - Credit risk models (time permits).
- Teaching philosophy
  - Focusing on main ideas and sketch of proofs.
  - Focusing on understanding the results: why it is true, and how to use it in practice.
  - Measure-theoretic type of proofs and conditions kept at minimum (if you are interested, read the textbook).

Part 1. Introduction to Stochastic Calculus

Main topics

- Brownian motion
- Stochastic integration
- Ito's formula
- Applications of Ito's formula

1.2 Brownian Motion

Brownian motion is one of the most widely studied continuous-time stochastic processes and is a major building block for continuous-time asset pricing models

A little history: Scatter a few grains of pollen on the surface of an apparently still beaker of liquid. Under a microscope you will observe that each grain is not still but jitters about on the liquid surface. This was first noticed by a botanist called Robert Brown in 1827. He was looking for microscopic life in a drop of water when he noticed that small grains in the water were jiggling around in a strange way - almost as if they were alive! This type of random motion is called Brownian motion after him.

Einstein showed that continual collision with water molecules causes Brownian motion. When he investigated Brownian motion at the beginning of the 20th century not all scientists believed in molecules, and Einstein was trying to demonstrate that they really did exist.

Einstein (1905) listed the following three properties of BM:

- (i) The sample paths must be continuous (based on physics);
- (ii) The increments follow a normal distribution with a variance that is proportion to the time elapsed (based on CLT);

- (iii) The increments of BM are independent, i.e., pollen grain has no memory.

Einstein could not prove that such process exists. In 1920s, Nobert Wiener proved BM exists and hence it is also called Wiener process.

Definition of Brownian Motion. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A family of random variables W_t indexed by time t (assume that $W_0 \equiv 0$) is called a Brownian Motion if it satisfies

- Continuous sample path: For each $\omega \in \Omega$, the function $W_t(\omega)$ is a continuous function of $t \geq 0$
- Independent increments: for all $0 = t_0 < t_1 < ... < t_m$ the incrementsis

  $$W_{t_1} - W_0, W_{t_2} - W_{t_1}, ..., W_{t_m} - W_{t_{m-1}}$$

  are independent of each other.
- Normality: each of the increments is normally distributed

  $$\mathbb{E}\left[W_{t_{i+1}} - W_{t_i}\right] = 0 \text{ and } \mathbb{E}\left[W_{t_{i+1}} - W_{t_i}\right]^2 = t_{i+1} - t_i.$$

A filtration for Brownian motion is defined as $\mathcal{F}_t = \sigma\left(W_s, s \leq t\right)$. That is, all observable event is based on observing the BM before t. Adaptivity means that $W\left(t\right)$ is $\mathcal{F}_t$-measurable. Independence of future increments means that for $0 \leq t < u$, $W\left(u\right) - W\left(t\right)$ is independent of $\mathcal{F}_t$.

A few important properties of Brownian motion

(1). Brownian motion is a martingale.

Proof. Let $0 \leq s < t$ be given. Then

$$\begin{aligned}
& \mathbb{E}\left[W_t|\mathcal{F}_s\right] \\
= & \mathbb{E}\left[W_t - W_s + W_s|\mathcal{F}_s\right] \\
= & \mathbb{E}\left[W_t - W_s|\mathcal{F}_s\right] + \mathbb{E}\left[W_s|\mathcal{F}_s\right] \text{ (linearity property)} \\
= & \mathbb{E}\left[W_t - W_s|\mathcal{F}_s\right] + W_s \text{ (} W_s \text{ is known given } \mathcal{F}_s\text{)} \\
= & \mathbb{E}\left[W_t - W_s\right] + W_s \text{ (independence of increments)} = W_s.
\end{aligned}$$

This property means that the best forecast of tomorrow's value is today's value. This is the idea behind the random walk model

(2). $W_t^2 - t$ is a martingale.

Proof. Let $0 \leq s < t$ be given. Then

$$\begin{aligned}
\mathbb{E}\left[W_t^2 - t|\mathcal{F}_s\right] &= \mathbb{E}\left[(W_t - W_s + W_s)^2 - t|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[(W_t - W_s)^2 + 2(W_t - W_s)W_s + W_s^2 - t|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[(W_t - W_s)^2|\mathcal{F}_s\right] + 2\mathbb{E}\left[(W_t - W_s)W_s|\mathcal{F}_s\right] \\
&\quad +\mathbb{E}\left[W_s^2 - t|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[(W_t - W_s)^2\right] + 2\mathbb{E}\left[W_t - W_s|\mathcal{F}_s\right]W_s + W_s^2 - t \\
&= (t - s) + W_s^2 - t = W_s^2 - s.
\end{aligned}$$

(3). For $0 \leq s < t$, $cov(W_s, W_t) = s$.

Proof. The covariance of W_s and W_t is

$$\begin{aligned}
\mathbb{E}[W_s W_t] &= \mathbb{E}[\mathbb{E}[W_s W_t|\mathcal{F}_s]] = \mathbb{E}[W_s \mathbb{E}[W_t|\mathcal{F}_s]] \\
&= \mathbb{E}\left[W_s^2\right] = s.
\end{aligned}$$

Or

$$\begin{aligned}
\mathbb{E}[W_s W_t] &= \mathbb{E}[W_s[W_t - W_s + W_s]] \\
&= \mathbb{E}\left[W_s[W_t - W_s] + W_s^2\right] \\
&= \mathbb{E}[W_s]\,\mathbb{E}[W_t - W_s] + \mathbb{E}\left[W_s^2\right] \\
&= 0 + s = s.
\end{aligned}$$

Nondifferentiability of Brownian path

One important property of BM is that is its sample path is not a differentiable as a function of t.

- This can be understand intuitively from the fact $W(t) \sim N(0,t)$.
- The nondifferentiability result is see from the fact the fact that BM has unbounded total variation.

Total Variation [First Variation]. Let $f(t)$ be a function defined for $0 \le t \le T$. If

$$\sup_{\text{all } \Pi} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)| < \infty,$$

then f is of finite total variation, where $\Pi = \{t_0, t_1, ..., t_n\}$ is a grid: $0 = t_0 < t_1 < ... < t_n = T$.

We can define the size of the grid by

$$\|\Pi\| = \max_{j=0,\ldots,n-1} (t_{j+1} - t_j).$$

Quadratic Variation. The quadratic variation of f up to time T is

$$[f,f](T) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2.$$

If f has a continuous derivative, then its quadratic variation is zero.

$$\begin{aligned}
\sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2 &= \sum_{j=0}^{n-1} \left|f'\left(t_j^*\right)\right|^2 (t_{j+1} - t_j)^2 \\
&\le \|\Pi\| \cdot \sum_{j=0}^{n-1} \left|f'\left(t_j^*\right)\right|^2 (t_{j+1} - t_j), \quad t_j^* \in [t_j, t_{j+1}],
\end{aligned}$$

$$\begin{aligned}
[f,f](T) &\le \lim_{\|\Pi\| \to 0} \left[\|\Pi\| \cdot \sum_{j=0}^{n-1} \left|f'\left(t_j^*\right)\right|^2 (t_{j+1} - t_j)\right] \\
&= \lim_{\|\Pi\| \to 0} \|\Pi\| \cdot \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \left|f'\left(t_j^*\right)\right|^2 (t_{j+1} - t_j) \\
&= \lim_{\|\Pi\| \to 0} \|\Pi\| \cdot \int_0^T |f'(t)|^2 \, dt = 0.
\end{aligned}$$

Because $f'(t)$ is continuous, $\int_0^T |f'(t)|^2 \, dt$ is finite.

Theorem [Quadratic Variation of BM]. Let W be a BM, then $[W,W](T) = T$ for all $T \ge 0$ a.s.

Proof. For any finite partition Π of $[0,T]$, denote by $Q=\sum_{j=0}^{n-1}\left(W_{t_{j+1}}-W_{t_j}\right)^2$.

We must show that $Q\to T$ as $\|\Pi\|\to 0$. We can show that

$$\mathbb{E}\left[Q\right]=T \text{ and } \lim_{\|\Pi\|\to 0}\mathbb{E}\left[Q-T\right]^2=0.$$

This means that Q converges to T in $\mathcal{L}^2$-norm. We have

$$\begin{aligned} &\mathbb{E}\left[\left(W_{t_{j+1}}-W_{t_j}\right)^2\right]=Var\left[W_{t_{j+1}}-W_{t_j}\right]=t_{j+1}-t_j\\ \Rightarrow\quad &\mathbb{E}\left[Q\right]=\mathbb{E}\left[\sum_{j=0}^{n-1}\left(W_{t_{j+1}}-W_{t_j}\right)^2\right]=\sum_{j=0}^{n-1}\left(t_{j+1}-t_j\right)=T.\end{aligned}$$

Moreover,

$$Var\left[\left(W_{t_{j+1}}-W_{t_j}\right)^2\right]=\mathbb{E}\left[\left(\left(W_{t_{j+1}}-W_{t_j}\right)^2-\left(t_{j+1}-t_j\right)\right)^2\right]$$

$$\begin{aligned}=\quad&\mathbb{E}\left[\left(W_{t_{j+1}}-W_{t_j}\right)^4\right]-2\left(t_{j+1}-t_j\right)\mathbb{E}\left[\left(W_{t_{j+1}}-W_{t_j}\right)^2\right]\\ &+\left(t_{j+1}-t_j\right)^2\quad=2\left(t_{j+1}-t_j\right)^2.\end{aligned}$$

Note $\mathbb{E}\left[\left(W_{t_{j+1}}-W_{t_j}\right)^4\right]=3\left(t_{j+1}-t_j\right)^2$. Therefore,

$$\begin{aligned} Var\left(Q\right)\quad&=\quad\sum_{j=0}^{n-1}Var\left[\left(W_{t_{j+1}}-W_{t_j}\right)^2\right]=\sum_{j=0}^{n-1}2\left(t_{j+1}-t_j\right)^2\\ &\leq\quad\sum_{j=0}^{n-1}\|\Pi\|\,2\left(t_{j+1}-t_j\right)=2\|\Pi\|\,T.\end{aligned}$$

Therefore, $\lim_{\|\Pi\|\to 0}Var\left(Q\right)=0$ and $\lim_{\|\Pi\|\to 0}Q=\mathbb{E}\left(Q\right)=T$. These two properties imply the result in the Theorem■

Corollary. $[W,W]\left(T\right)=T\Rightarrow TV\left(W\right)=\infty$ almost surely.

Proof. We prove by contradiction. Suppose $TV\left(W\right)<\infty$, then

$$\lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\left(W_{t_{j+1}}-W_{t_j}\right)^2\leq\lim_{\|\Pi\|\to 0}\max_{t_j\in\Pi}\left|W_{t_{j+1}}-W_{t_j}\right|\sum_{j=0}^{n-1}\left|W_{t_{j+1}}-W_{t_j}\right|$$

$$\begin{aligned}\leq\quad&\lim_{\|\Pi\|\to 0}\max_{t_j\in\Pi}\left|W_{t_{j+1}}-W_{t_j}\right|\sum_{j=0}^{n-1}\left|W_{t_{j+1}}-W_{t_j}\right|\\ \leq\quad&\lim_{\|\Pi\|\to 0}\max_{t_j\in\Pi}\left|W_{t_{j+1}}-W_{t_j}\right|TV\left(W\right)=0.\end{aligned}$$

This contradicts with $[W, W](T) = T$.■

In the above derivation, we have used

$$\begin{aligned} \mathbb{E}\left[\left(W_{t_{j+1}} - W_{t_j}\right)^2\right] &= t_{j+1} - t_j, \\ Var\left[\left(W_{t_{j+1}} - W_{t_j}\right)^2\right] &= 2\left(t_{j+1} - t_j\right)^2. \end{aligned}$$

It is tempting to argue that when $t_{j+1} - t_j$ is small, $(t_{j+1} - t_j)^2$ is very small. Therefore, $\left(W_{t_{j+1}} - W_{t_j}\right)^2$, although random, is with high probability near its mean $t_{j+1} - t_j$.

Review: L^2 convergence, R^n, L^2, Second Moment, etc. Convergence in L^2.