Stochastic Calculus

Author: Kellen Kanarios
Mentor: Ethan Zell

University of Michigan

April 11, 2022

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Note

Intuition for a **measure**:

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Note

Intuition for a **measure**:

1. In $\mathbb{R}$: length,

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Note

Intuition for a **measure**:

1. In $\mathbb{R}$: length,
2. In $\mathbb{R}^2$: area,

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Note

Intuition for a **measure**:

1. In $\mathbb{R}$: length,
2. In $\mathbb{R}^2$: area,
3. In $\mathbb{R}^3$: volume.

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Definition translated into english

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Definition translated into english

1. Unbeknownst to us, someone chooses a random $\omega \in \Omega$. Then we see the $X(\omega) \in \mathbb{R}$.

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Definition translated into english

1. Unbeknownst to us, someone chooses a random $\omega \in \Omega$. Then we see the $X(\omega) \in \mathbb{R}$.
2. We cannot see the corresponding $\omega \in \Omega$, but the $X(\omega) \in \mathbb{R}$ gives us partial information about ω.

Background

Definition

A real-valued random variable X is a mapping $X : \Omega \to \mathbb{R}$. Here, Ω is the **sample space** and $\mathbb{P}$ is the measure of the **probability space**, such that $\mathbb{P}(\Omega) = 1$.

Definition translated into english

1. Unbeknownst to us, someone chooses a random $\omega \in \Omega$. Then we see the $X(\omega) \in \mathbb{R}$.
2. We cannot see the corresponding $\omega \in \Omega$, but the $X(\omega) \in \mathbb{R}$ gives us partial information about ω.
3. $\mathbb{P}$ tells us how likely subsets $A \subseteq \Omega$ are to occur.

Background

Example

Consider the case where you flip a coin. Using our previous definition, this could be described as $\Omega = \{\text{heads}, \text{tails}\}$ and

$$X(\omega) = \begin{cases} 1, & \text{if } \omega = \text{heads} \\ -1, & \text{if } \omega = \text{tails} \end{cases} \qquad \text{where } \omega \in \Omega.$$

This would yield the familiar notation of $\mathbb{P}(X = 1) = .5$ and $\mathbb{P}(X = -1) = .5$ for a fair coin.

Background

Definition

A **stochastic process** is a family of random variables indexed by a time parameter $t \geq 0$.

Background

Definition

A **stochastic process** is a family of random variables indexed by a time parameter $t \geq 0$.

Example

A sequence of coin flips. At each time t your process corresponds to a random variable (aka coinflip) X_t.

Wiener Process

Definition

A stochastic process W is called a **Wiener process** if the follow
conditions hold:

Wiener Process

Definition

A stochastic process W is called a **Wiener process** if the follow conditions hold:

1. $W_0 = 0$,

Wiener Process

Definition

A stochastic process W is called a **Wiener process** if the follow conditions hold:

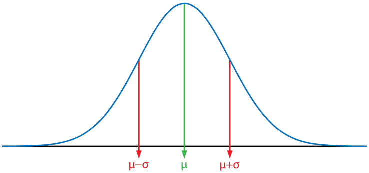
1. $W_0 = 0$,
2. The process W has independent increments,

Wiener Process

Definition

A stochastic process W is called a **Wiener process** if the follow conditions hold:

1. $W_0 = 0$,
2. The process W has independent increments,
3. For $s < t$ the random variable $W_t - W_s$ has the Gaussian distribution $\mathcal{N}(0, t-s)$,

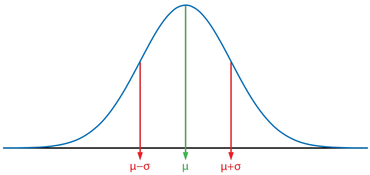


Wiener Process

Definition

A stochastic process W is called a **Wiener process** if the follow conditions hold:

1. $W_0 = 0$,
2. The process W has independent increments,
3. For $s < t$ the random variable $W_t - W_s$ has the Gaussian distribution $\mathcal{N}(0, t - s)$,



4. W has continuous trajectories.

Wiener Process

Theorem

A Wiener trajectory is with probability one, nowhere differentiable, and it has locally infinite total variation.

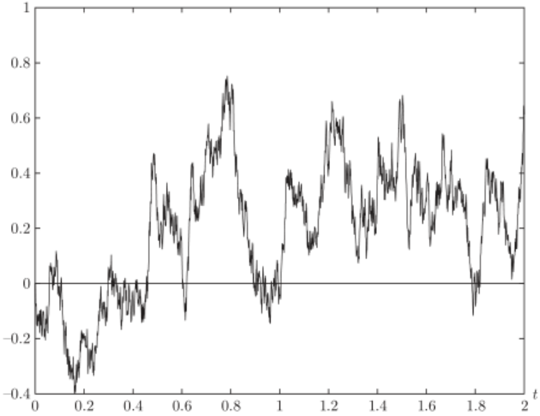


Figure: Wiener trajectory

Quadratic Variation

Definition

Let X be a stochastic process. Suppose P is a partition of $[0, t]$ denoted t_k and let $\|P\|$ be the mesh of the partition then the **quadratic variation** of X is defined to be:

$$[X]_t = \lim_{\|P\| \to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2.$$

Quadratic Variation

Definition

Let X be a stochastic process. Suppose P is a partition of $[0, t]$ denoted t_k and let $\|P\|$ be the mesh of the partition then the **quadratic variation** of X is defined to be:

$$[X]_t = \lim_{\|P\| \to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2.$$

Comments

Note that Quadratic Variation itself is a stochastic process. An intuitive way to think about quadratic variation is the internal clock of a process, describing how randomness accumulates over time.

Quadratic Variation of Wiener Process

Theorem

Quadratic variation of a Wiener Process on the interval $[0, t]$ is t.

Quadratic Variation of Wiener Process

Theorem

Quadratic variation of a Wiener Process on the interval $[0, t]$ *is* t.

Proof.

Let us fix a point t and subdivide the interval $[0, t]$ into n equally large subintervals of the form $[k\frac{t}{n}, (k+1)\frac{t}{n}]$, where $k = 0, 1, \ldots, n-1$. Then the quadratic variation on P is:

$$[W]_t^P = \sum_{i=1}^{n} (W_{i\frac{t}{n}} - W_{(i-1)(\frac{t}{n})})^2.$$

Quadratic Variation of Wiener Process

Theorem

Quadratic variation of a Wiener Process on the interval $[0, t]$ is t.

Proof.

Let us fix a point t and subdivide the interval $[0, t]$ into n equally large subintervals of the form $[k\frac{t}{n}, (k+1)\frac{t}{n}]$, where $k = 0, 1, \ldots, n-1$. Then the quadratic variation on P is:

$$[W]_t^P = \sum_{i=1}^{n} (W_{i\frac{t}{n}} - W_{(i-1)(\frac{t}{n})})^2.$$

Therefore,

$$\mathbb{E}\left[[W]_t^P\right] = \sum_{i=1}^{n} \mathbb{E}\left[(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})})^2\right].$$

Quadratic Variation of Wiener Process

Theorem

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$

Quadratic Variation of Wiener Process

Proof cont.

Since $W_{(i)(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})} \in \mathcal{N}(0, i(\frac{t}{n}) - (i-1)(\frac{t}{n}))$,

$$\mathbb{E}(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}) = 0.$$

Quadratic Variation of Wiener Process

Proof cont.

Since $W_{(i)(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})} \in \mathcal{N}(0, i(\frac{t}{n}) - (i-1)(\frac{t}{n}))$,

$$\mathbb{E}(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}) = 0.$$

It follows that:

$$(\mathbb{E}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right])^2 = 0.$$

Quadratic Variation of Wiener Process

Proof cont.

Since $W_{(i)(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})} \in \mathcal{N}(0, i(\frac{t}{n}) - (i-1)(\frac{t}{n}))$,

$$\mathbb{E}(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}) = 0.$$

It follows that:

$$(\mathbb{E}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right])^2 = 0.$$

Thus, $Var(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}) = E(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2)$.

Quadratic Variation of Wiener Process

Proof cont.

Since $W_{(i)(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})} \in \mathcal{N}(0, i(\frac{t}{n}) - (i-1)(\frac{t}{n}))$,

$$\mathbb{E}(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}) = 0.$$

It follows that:

$$(\mathbb{E}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right])^2 = 0.$$

Thus, $Var(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}) = E(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2)$.

Since $\mathbb{E}[(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})})^2] = Var\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]$, we can write:

$$\mathbb{E}\left[[W]_t^P\right] = \mathbb{E}\left[\sum_{k=1}^{n}(W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})})^2\right] = Var\left[\sum_{k=1}^{n}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]\right].$$

Quadratic Variation of Wiener Process

Stats fact

$$Var\left[\sum_{k=1}^{n}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]\right] = \sum_{k=1}^{n} Var\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]$$
$$+ \sum_{i\neq\ell}^{n} Cov\left(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right], \left[W_{\ell(\frac{t}{n})} - W_{(\ell-1)(\frac{t}{n})}\right]\right)$$

Quadratic Variation of Wiener Process

Proof cont.

Since Wiener increments are independent of eachother,

$$Cov\left(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right], \left[W_{\ell(\frac{t}{n})} - W_{(\ell-1)(\frac{t}{n})}\right]\right) = 0.$$

Quadratic Variation of Wiener Process

Proof cont.

Since Wiener increments are independent of eachother,

$$Cov\left(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right], \left[W_{\ell(\frac{t}{n})} - W_{(\ell-1)(\frac{t}{n})}\right]\right) = 0.$$

Therefore,

$$E\left[[W]_t^P\right] = \sum_{k=1}^{n} Var\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]$$

Quadratic Variation of Wiener Process

Proof cont.

Since Wiener increments are independent of eachother,

$$Cov\left(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right], \left[W_{\ell(\frac{t}{n})} - W_{(\ell-1)(\frac{t}{n})}\right]\right) = 0.$$

Therefore,

$$E\left[[W]_t^P\right] = \sum_{k=1}^{n} Var\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]$$

Using $W_{(i)(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})} \in \mathcal{N}(0, i(\frac{t}{n}) - (i-1)(\frac{t}{n}))$:

$$\begin{aligned}\mathbb{E}\left[[W]_t^P\right] &= \sum_{k=1}^{n}(i(\frac{t}{n}) - (i-1)(\frac{t}{n})) \\ &= t\end{aligned}$$

Quadratic Variation of Wiener Process

Proof cont.

Since Wiener increments are independent of eachother,

$$Cov\left(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right], \left[W_{\ell(\frac{t}{n})} - W_{(\ell-1)(\frac{t}{n})}\right]\right) = 0.$$

Therefore,

$$E\left[[W]_t^P\right] = \sum_{k=1}^{n} Var\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]$$

Using $W_{(i)(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})} \in \mathcal{N}(0, i(\frac{t}{n}) - (i-1)(\frac{t}{n}))$:

$$\begin{aligned}\mathbb{E}\left[[W]_t^P\right] &= \sum_{k=1}^{n}(i(\frac{t}{n}) - (i-1)(\frac{t}{n})) \\ &= t\end{aligned}$$

Now we must show $\mathbb{E}([W]_t^P) = [W]_t = t$.

Quadratic Variation of Wiener Process

Stats fact

If $Var(X) = 0$ then $X = \mathbb{E}[X]$.

Quadratic Variation of Wiener Process

Proof.

Therefore, we must show that :

$$Var(\lim_{\|P\|\to 0}[W]_t^P) = 0.$$

Or equivalently,

$$Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2) = 0,$$

Quadratic Variation of Wiener Process

Proof.

Therefore, we must show that :

$$Var(\lim_{\|P\|\to 0}[W]_t^P) = 0.$$

Or equivalently,

$$Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2) = 0,$$

and we showed earlier:

$$Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2) = \lim_{n\to\infty}\sum_{i=1}^{n}Var(\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2).$$

Quadratic Variation of Wiener Process

Proof.

Also from earlier:

$$Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})}-W_{(i-1)(\frac{t}{n})}\right]^2) = \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}(\left[W_{i(\frac{t}{n})}-W_{(i-1)(\frac{t}{n})}\right]^4) - \mathbb{E}\left(\left[W_{i(\frac{t}{n})}-W_{(i-1)(\frac{t}{n})}\right]^2\right)^2.$$

Quadratic Variation of Wiener Process

Stats fact

$$\mathrm{E}\big[(X-\mu)^p\big] = \begin{cases} 0, & \text{if } p \text{ is odd} \\ \sigma^p (p-1)!!, & \text{if } p \text{ is even.} \end{cases}$$

Quadratic Variation of Wiener Process

Proof.

Also from earlier:

$$Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})}-W_{(i-1)(\frac{t}{n})}\right]^2)=\lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}(\left[W_{i(\frac{t}{n})}-W_{(i-1)(\frac{t}{n})}\right]^4)$$
$$-\mathbb{E}\left(\left[W_{i(\frac{t}{n})}-W_{(i-1)(\frac{t}{n})}\right]^2\right)^2.$$

Therefore,

$$Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})}-W_{(i-1)(\frac{t}{n})}\right]^2)=\lim_{n\to\infty}\sum_{i=1}^{n}3\left(i(\frac{t}{n})-(i-1)(\frac{t}{n})\right)^2$$
$$-\left(i(\frac{t}{n})-(i-1)(\frac{t}{n})\right)^2.$$

Quadratic Variation of Wiener Process

Proof.

Simplifying,

$$\begin{aligned} Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2) &= \lim_{n\to\infty}\sum_{i=1}^{n} 2\frac{t^2}{n^2}, \\ &= \lim_{n\to\infty} 2\frac{t^2}{n}, \\ &= 0. \end{aligned}$$

Quadratic Variation of Wiener Process

Proof.

Simplifying,

$$Var(\lim_{n\to\infty}\sum_{i=1}^{n}\left[W_{i(\frac{t}{n})} - W_{(i-1)(\frac{t}{n})}\right]^2) = \lim_{n\to\infty}\sum_{i=1}^{n} 2\frac{t^2}{n^2},$$
$$= \lim_{n\to\infty} 2\frac{t^2}{n},$$
$$= 0.$$

Thus,

$$[W]_t = \mathbb{E}\left([W]_t\right) = \mathbb{E}\left(\lim_{\|P\|\to 0}[W]_t^P\right).$$

Quadratic Variation of Wiener Process

Proof.

By dominating convergence theorem (outside scope),

$$\begin{aligned}
\mathbb{E}\left(\lim_{\|P\|\to 0}[W]_t^P\right) &= \lim_{\|P\|\to 0}, \mathbb{E}\left([W]_t^P\right), \\
&= \lim_{\|P\|\to 0} t, \\
&= t.
\end{aligned}$$

$\square$

Quadratic Variation of Wiener Process

Proof.

By dominating convergence theorem (outside scope),

$$\begin{aligned}
\mathbb{E}\left(\lim_{\|P\|\to 0} [W]_t^P\right) &= \lim_{\|P\|\to 0}, \mathbb{E}\left([W]_t^P\right), \\
&= \lim_{\|P\|\to 0} t, \\
&= t.
\end{aligned}$$

Thus, $[W]_t = t$. $\square$

Quadratic Variation of Wiener Process

Implications

This motivates us to write:

$$\int_0^t (dW_t)^2 = t.$$

Or equivalently,

$$(dW_t)^2 = dt.$$

This will come up frequently later as we transitition into Ito calculus.

The Stochastic Integral

Motivation

Let h_t be a stochastic process that represents our trading strategy and let W_t be the price of the stock at the given time. Then

$$\int_0^t h_t dW_t$$

would be our gains or losses from this strategy.

The Stochastic Integral

The Problem

Integrals of the form $\int_0^t g_s dW_s$ for some stochastic process g_s.

The Stochastic Integral

The Problem

Integrals of the form $\int_0^t g_s dW_s$ for some stochastic process g_s.

Possible Solution 1

Try to define like the Riemman Integral:

1. $\sum_{k=1}^{n} g_s(W_{t_{k+1}} - W_{t_k})$,

The Stochastic Integral

The Problem

Integrals of the form $\int_0^t g_s dW_s$ for some stochastic process g_s.

Possible Solution 1

Try to define like the Riemman Integral:

1. $\sum_{k=1}^{n} g_s(W_{t_{k+1}} - W_{t_k})$,
2. Not possible due to locally unbounded variation of a Wiener Process.

The Stochastic Integral

Definition

A stochastic process is **simple** on $[a, b]$ when there exists deterministic points in time $a = t_0 < t_1 < \cdots < t_n = b$ such that $g_s = g_{t_k}$ for all $s \in [t_k, t_{k+1}]$.

The Stochastic Integral

Definition

A stochastic process is **simple** on $[a, b]$ when there exists deterministic points in time $a = t_0 < t_1 < \cdots < t_n = b$ such that $g_s = g_{t_k}$ for all $s \in [t_k, t_{k+1}]$.

Real Solution

Now we can define the stochastic integral by

$$\int_a^b g_s dW_s = \sum_{k=0}^{n-1} g_{t_k} [W_{t_{k+1}} - W_{t_k}],$$

for some simple stochastic process g_s.

The Stochastic Integral

Generalized Version

By the **simple approximation theorem** (outside of scope), for some stochastic process g_s there exists a sequence of simple functions g_s^n, such that $g_s^n \to g_s$. Therefore, we define our integral for non-simple functions as:

$$\int_a^b g_s dW_t = \lim_{n \to \infty} \int_a^b g_s^n dW_t.$$

The Stochastic Integral

Generalized Version

By the **simple approximation theorem** (outside of scope), for some stochastic process g_s there exists a sequence of simple functions g_s^n, such that $g_s^n \to g_s$. Therefore, we define our integral for non-simple functions as:

$$\int_a^b g_s dW_t = \lim_{n\to\infty} \int_a^b g_s^n dW_t.$$

Something to think about

How might we evaluate $\int_0^t W_s dW_s$?

Stochastic Differential Equations

Definition

An ito process, X_t, is a process that can be represented as:

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

Stochastic Differential Equations

Definition

An ito process, X_t, is a process that can be represented as:

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

We often use the notation:

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

Stochastic Differential Equations

Definition

An ito process, X_t, is a process that can be represented as:

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

We often use the notation:

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

This is known as a stochastic differential equation, and a useful tool for solving these is Ito's lemma, which we will see shortly.

Ito's Multiplication Table

Definition

Ito's multiplication table is:

	dt	dW_t
dt	0	0
dW_t	0	dt

We have already shown the only interesting result: $(dW_t)^2 = dt$.

Ito's Multiplication Table

Definition

Ito's multiplication table is:

	dt	dW_t
dt	0	0
dW_t	0	dt

We have already shown the only interesting result: $(dW_t)^2 = dt$.

Quick sketch.

Suppose you are integrating on the interval $[a, b]$. Partition [a,b] into N increments. Then $dt \sim \frac{1}{N}$ and $(dt)^2 \sim \frac{1}{N^2}$. It follows that:

$$\sum_{i=1}^{N} \frac{1}{N^2} = \frac{1}{N},$$

$$\lim_{N\to\infty} \sum_{i=1}^{N} \frac{1}{N^2} = \lim_{N\to\infty} \frac{1}{N} = \int_a^b (dt)^2 = 0.$$

Ito's Multiplication Table

Quick sketch.

Similarly, suppose you are integrating on the interval $[a, b]$. Partition [a,b] into N increments. Since $dW_t = \sqrt{dt}$, $dt \sim \frac{1}{N}$ and $(dt)(dW_t) \sim \frac{1}{N^{3/2}}$. It follows that:

$$\sum_{i=1}^{N} \frac{1}{N^{3/2}} = \frac{1}{N^{1/2}},$$

$$\lim_{N\to\infty} \sum_{i=1}^{N} \frac{1}{N^{3/2}} = \lim_{N\to\infty} \frac{1}{N^{1/2}} = \int_a^b (dt)(dW_t) = 0.$$

$\square$

Ito's Multiplication Table

Quick sketch.

Similarly, suppose you are integrating on the interval $[a, b]$. Partition [a,b] into N increments. Since $dW_t = \sqrt{dt}$, $dt \sim \frac{1}{N}$ and $(dt)(dW_t) \sim \frac{1}{N^{3/2}}$. It follows that:

$$\sum_{i=1}^{N} \frac{1}{N^{3/2}} = \frac{1}{N^{1/2}},$$

$$\lim_{N\to\infty} \sum_{i=1}^{N} \frac{1}{N^{3/2}} = \lim_{N\to\infty} \frac{1}{N^{1/2}} = \int_a^b (dt)(dW_t) = 0.$$

□

This completes the multiplication table and provides the foundation for introducing Ito's lemma.

Ito's Lemma

Theorem (Ito's formula)

Assume that process X has a stochastic differential given by

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

Define the process Z by $Z(t) = f(t, X_t)$. Then Z has a stochastic differential given by

$$df(t, X_t) = \left\{ \frac{\partial f}{\partial t}(t, X_t) + \mu_t \frac{\partial f}{\partial x}(t, X_t) + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}(t, X_t) \right\} dt + \sigma \frac{\partial f}{\partial x}(t, X_t) dW_t.$$

Ito's Lemma

Heuristic proof.

Using f from the theorem, we will consider the second order Taylor expansion:

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dt dX_t.$$

Ito's Lemma

Heuristic proof.

Using f from the theorem, we will consider the second order Taylor expansion:

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2 + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dt dX_t.$$

By definition we have:

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

Ito's Lemma

Heuristic proof.

Using f from the theorem, we will consider the second order Taylor expansion:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x}dt dX_t.$$

By definition we have:

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

so, we obtain

$$(dX_t)^2 = \mu_t^2 (dt)^2 + 2\mu_t \sigma_t (dt)(dW_t) + \sigma_t^2 (dW_t)^2.$$

Ito's Lemma

Heuristic proof.

Using f from the theorem, we will consider the second order Taylor expansion:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x}dtdX_t.$$

By definition we have:

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

so, we obtain

$$(dX_t)^2 = \mu_t^2(dt)^2 + 2\mu_t\sigma_t(dt)(dW_t) + \sigma_t^2(dW_t)^2.$$

From Ito's multiplication table we have:

$$(dX_t)^2 = \sigma_t^2 dt.$$

Ito's Lemma

Heuristic proof cont.

Substituting back for $(dX_t)^2$,

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\sigma_t^2)(dt) + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x}dtdX_t.$$

$\square$

Ito's Lemma

Heuristic proof cont.

Substituting back for $(dX_t)^2$,

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\sigma_t^2)(dt) + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x}dtdX_t.$$

Applying Ito's multiplication table one more time,

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\sigma_t^2)(dt) + \frac{\partial^2 f}{\partial t \partial x}dtdX_t.$$

□

Ito's Lemma

Heuristic proof cont.

Substituting back for $(dX_t)^2$,

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\sigma_t^2)(dt) + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x}dtdX_t.$$

Applying Ito's multiplication table one more time,

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\sigma_t^2)(dt) + \frac{\partial^2 f}{\partial t \partial x}dtdX_t.$$

Now substituing in (dX_t),

$$\begin{aligned} df &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}(\mu_t dt + \sigma_t dW_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\sigma_t^2)(dt) + \frac{\partial^2 f}{\partial t \partial x}(\mu_t dt + \sigma_t dW_t)dt, \\ &= \left\{ \frac{\partial f}{\partial t}(t, X_t) + \mu_t \frac{\partial f}{\partial x}(t, X_t) + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}(t, X_t) \right\} dt + \sigma \frac{\partial f}{\partial x}(t, X_t)dW_t. \end{aligned}$$

$\square$

Ito's Lemma

Example

Find $d(W_t^2)$.

Ito's Lemma

Example

Find $d(W_t^2)$.

Solution

Define X by $X_t = W_t$ and set $f(t,x) = x^2$. In terms of Ito's formula we have

$$\frac{\partial f}{\partial t}(t,x) = 0,\ \frac{\partial f}{\partial x}(t,x) = 2x,\ \frac{\partial^2 f}{\partial t^2}(t,x) = 2.$$

Ito's Lemma

Example

Find $d(W_t^2)$.

Solution

Define X by $X_t = W_t$ and set $f(t,x) = x^2$. In terms of Ito's formula we have

$$\frac{\partial f}{\partial t}(t,x) = 0, \ \frac{\partial f}{\partial x}(t,x) = 2x, \ \frac{\partial^2 f}{\partial t^2}(t,x) = 2.$$

Substituting,

$$d(W_t^2) = \left\{0 + 0 \cdot 2X_t + \frac{1}{2} \cdot 1 \cdot 2\right\} dt + 1 \cdot 2X_t dW_t.$$

Ito's Lemma

Example

Find $d(W_t^2)$.

Solution

Define X by $X_t = W_t$ and set $f(t,x) = x^2$. In terms of Ito's formula we have

$$\frac{\partial f}{\partial t}(t,x) = 0,\ \frac{\partial f}{\partial x}(t,x) = 2x,\ \frac{\partial^2 f}{\partial t^2}(t,x) = 2.$$

Substituting,

$$d(W_t^2) = \left\{0 + 0 \cdot 2X_t + \frac{1}{2} \cdot 1 \cdot 2\right\} dt + 1 \cdot 2X_t dW_t.$$

Thus,

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

Ito's Lemma

Revisiting

Now if we integrate both side such that:

$$\int_0^t d(W_t^2) = \int_0^t dt + \int_0^t 2W_t dW_t.$$

Ito's Lemma

Revisiting

Now if we integrate both side such that:

$$\int_0^t d(W_t^2) = \int_0^t dt + \int_0^t 2W_t dW_t.$$

This gives us:

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

Real world application

Definition

Geometric Brownian Motion is a stochastic process whose dynamics follow the stochastic differential equation

$$dX_t = \alpha X_t dt + \sigma X_t dW_t \text{ for some } \alpha, \beta \in \mathbb{R},$$

where dW_t is the infinitesimal increment of the Wiener process.

Real world application

Definition

Geometric Brownian Motion is a stochastic process whose dynamics follow the stochastic differential equation

$$dX_t = \alpha X_t dt + \sigma X_t dW_t \text{ for some } \alpha, \beta \in \mathbb{R},$$

where dW_t is the infinitesimal increment of the Wiener process.

Geometric Brownian Motion is one of the building blocks for the modeling of asset prices, and turns up naturally in many other places.

Real world application

Closed form

With our toolbox of Ito calculus and the ability to solve linear ODE's, it can be easily shown that the closed form of Geometric Brownian Motion is:

$$dX_t = X_0 e^{(\mu - \frac{\sigma^2}{2})dt + \sigma dW_t}.$$

Real world application

Closed form

With our toolbox of Ito calculus and the ability to solve linear ODE's, it can be easily shown that the closed form of Geometric Brownian Motion is:

$$dX_t = X_0 e^{(\mu - \frac{\sigma^2}{2})dt + \sigma dW_t}.$$

Now let's see how accurate this elementary model is when trying to predict the price of real stocks.

Real world application

Use in Finance

Black Scholes Model is one of the most widely used models of stock price behaviour and is built on Geometric Brownian motion.

Why GBM?

Some of the arguments for GBM are:

Real world application

Use in Finance

Black Scholes Model is one of the most widely used models of stock price behaviour and is built on Geometric Brownian motion.

Why GBM?

Some of the arguments for GBM are:

- The expected returns are independent of the value of the process, which is how stocks behave in reality,

Real world application

Use in Finance

Black Scholes Model is one of the most widely used models of stock price behaviour and is built on Geometric Brownian motion.

Why GBM?

Some of the arguments for GBM are:

- The expected returns are independent of the value of the process, which is how stocks behave in reality,
- A GBM process only assumes positive values, just like real stock prices.

Real world application

Use in Finance

Black Scholes Model is one of the most widely used models of stock price behaviour and is built on Geometric Brownian motion.

Why GBM?

Some of the arguments for GBM are:

- The expected returns are independent of the value of the process, which is how stocks behave in reality,
- A GBM process only assumes positive values, just like real stock prices.
- Calculations with GBM processes are relatively easy.

Real world application

Drawbacks

Real world application

Drawbacks

- In real stock prices, volatility changes over time (possibly stochastically), but in GBM, volatility is assumed constant.

Real world application

Drawbacks

- In real stock prices, volatility changes over time (possibly stochastically), but in GBM, volatility is assumed constant.
- In real life, stock prices often show jumps caused by unpredictable events or news, but in GBM, the path is continuous (no discontinuity).

Real world application

Drawbacks

- In real stock prices, volatility changes over time (possibly stochastically), but in GBM, volatility is assumed constant.
- In real life, stock prices often show jumps caused by unpredictable events or news, but in GBM, the path is continuous (no discontinuity).

This is why GBM is just a building block for more advanced models. For example, one that models volatility stochastically is a stochastic volatility model, which is built off GBM.