Lecture 14. SDE and Itô's formula

Tiejun Li[1,2]

[1]School of Mathematical Sciences (SMS),
&
[2]Center for Machine Learning Research (CMLR),
Peking University,
Beijing 100871,
P.R. China
tieli@pku.edu.cn

Office: No. 1 Science Building, Room 1376E

Table of Contents

White noise in Physics Literature

- In physics literature, the physicists usually use the stochastic differential equations (SDEs) like

$$\dot{X}_t = b(X_t, t) + \sigma(X_t, t)\dot{W}_t, \quad X|_{t=0} = X_0,$$

where $\dot{W}_t$ is called the temporal Gaussian white noise, which is the formal derivative of the Brownian motion W_t with respect to time.

White noise in Physics Literature

- In physics literature, the physicists usually use the stochastic differential equations (SDEs) like

$$\dot{X}_t = b(X_t, t) + \sigma(X_t, t)\dot{W}_t, \quad X|_{t=0} = X_0,$$

  where $\dot{W}_t$ is called the temporal Gaussian white noise, which is the formal derivative of the Brownian motion W_t with respect to time.
- Its mathematical description is that it is a Gaussian process with mean and covariance functions as

$$m(t) = \mathbb{E}(\dot{W}_t) = 0, \quad K(s,t) = \mathbb{E}(\dot{W}_s\dot{W}_t) = \delta(t-s).$$

White noise in Physics Literature

- In physics literature, the physicists usually use the stochastic differential equations (SDEs) like

$$\dot{X}_t = b(X_t, t) + \sigma(X_t, t)\dot{W}_t, \quad X|_{t=0} = X_0,$$

  where $\dot{W}_t$ is called the temporal Gaussian white noise, which is the formal derivative of the Brownian motion W_t with respect to time.
- Its mathematical description is that it is a Gaussian process with mean and covariance functions as

$$m(t) = \mathbb{E}(\dot{W}_t) = 0, \quad K(s,t) = \mathbb{E}(\dot{W}_s\dot{W}_t) = \delta(t-s).$$

- It can be formally understood as

$$m(t) = \frac{d}{dt}\mathbb{E}(W_t) = 0,$$

$$K(s,t) = \frac{\partial^2}{\partial s \partial t}\mathbb{E}(W_s W_t) = \frac{\partial^2}{\partial s \partial t}(s \wedge t) = \delta(t-s).$$

White noise

- From the regularity theory of the Brownian motion, the function $\dot{W}$ is meaningless since W_t has less than half order smoothness. In fact, it is not a traditional function but a distribution.

White noise

- From the regularity theory of the Brownian motion, the function $\dot{W}$ is meaningless since W_t has less than half order smoothness. In fact, it is not a traditional function but a distribution.
- However, the rigorous mathematical foundation of the white noise calculus can be also established.

White noise

- From the regularity theory of the Brownian motion, the function $\dot{W}$ is meaningless since W_t has less than half order smoothness. In fact, it is not a traditional function but a distribution.
- However, the rigorous mathematical foundation of the white noise calculus can be also established.
- In this Lecture, we will only introduce the Itô's classical way to establish the well-posedness of the stochastic differential equations.

Interpretation of SDEs?

- Mathematically, the SDEs are often denoted as

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t,$$

to avoid the ambiguity of the white noise, where W_t is the standard Wiener process. X_t can be viewed as a process induced by W_t.

Interpretation of SDEs?

- Mathematically, the SDEs are often denoted as

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t,$$

to avoid the ambiguity of the white noise, where W_t is the standard Wiener process. X_t can be viewed as a process induced by W_t.
- The effect of $b(X_t, t)$ is to drive the mean position of the system, while the effect of $\sigma(X_t, t)dW_t$ is to diffuse around the mean position.

Interpretation of SDEs?

- Mathematically, the SDEs are often denoted as

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t,$$

  to avoid the ambiguity of the white noise, where W_t is the standard Wiener process. X_t can be viewed as a process induced by W_t.
- The effect of $b(X_t, t)$ is to drive the mean position of the system, while the effect of $\sigma(X_t, t)dW_t$ is to diffuse around the mean position.
- One natural way is to define X_t through its integral form

$$X_t = X_0 + \int_0^t b(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s.$$

Interpretation of SDEs?

- Mathematically, the SDEs are often denoted as

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t,$$

  to avoid the ambiguity of the white noise, where W_t is the standard Wiener process. X_t can be viewed as a process induced by W_t.
- The effect of $b(X_t, t)$ is to drive the mean position of the system, while the effect of $\sigma(X_t, t)dW_t$ is to diffuse around the mean position.
- One natural way is to define X_t through its integral form

$$X_t = X_0 + \int_0^t b(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s.$$

- The first mathematical issue is how to define the integral $\int_0^t \sigma(X_s, s)dW_s$ involving Brownian motion.

Table of Contents

Stochastic Integral: Necessity

- First suppose X_t is continuous with respect to time t. For a fixed sample ω, we borrow the idea for defining the Riemann-Stieljes integral to make the definition

$$\int_0^t \sigma(X_s, s) dW_s = \lim_{|\Delta| \to 0} \sum_j \sigma(X_j, t_j^*) \Big(W_{t_{j+1}} - W_{t_j} \Big),$$

where Δ is a subdivision of $[0, t]$, X_j is the function value $X_{t_j^*}$ and t_j^* is chosen from the interval $[t_j, t_{j+1}]$.

Stochastic Integral: Necessity

- First suppose X_t is continuous with respect to time t. For a fixed sample ω, we borrow the idea for defining the Riemann-Stieljes integral to make the definition

$$\int_0^t \sigma(X_s, s)dW_s = \lim_{|\Delta|\to 0} \sum_j \sigma(X_j, t_j^*)\Big(W_{t_{j+1}} - W_{t_j}\Big),$$

  where Δ is a subdivision of $[0, t]$, X_j is the function value $X_{t_j^*}$ and t_j^* is chosen from the interval $[t_j, t_{j+1}]$.
- One critical issue about the above definition is that it depends on the choice of t_j^* when we are handling W_t, which has unbounded variation in any interval almost surely.

Possible Definitions

Consider the Riemann-Stieltjes integral to $\int_a^b f(t)dg(t)$, where f and g are all assumed continuous. So

$$\int_a^b f(t)dg(t) \approx \sum_j f_j\Big(g(t_{j+1}) - g(t_j)\Big).$$

Possible Definitions

Consider the Riemann-Stieltjes integral to $\int_a^b f(t)dg(t)$, where f and g are all assumed continuous. So

$$\int_a^b f(t)dg(t) \approx \sum_j f_j \Big(g(t_{j+1}) - g(t_j) \Big).$$

If one takes another value for f_j in $[t_j, t_{j+1}]$ under the same subdivision, then

$$\int_a^b f(t)dg(t) \approx \sum_j \tilde{f}_j \Big(g(t_{j+1}) - g(t_j) \Big).$$

BV case

- If $g(t)$ has bounded total variation, we subtract the right hand side of the two definitions and obtain

$$\begin{aligned}
&\Big|\sum_j (f_j - \tilde{f}_j)\Big(g(t_{j+1}) - g(t_j)\Big)\Big| \\
\le &\max_j |f_j - \tilde{f}_j| \sum_j \Big|g(t_{j+1}) - g(t_j)\Big| \\
\le &\max_j |f_j - \tilde{f}_j| V(g; [a,b]) \to 0
\end{aligned}$$

as $|\Delta| \to 0$ by the uniform continuity of f on $[a,b]$.

BV case

- If $g(t)$ has bounded total variation, we subtract the right hand side of the two definitions and obtain

$$
\begin{aligned}
&\left|\sum_j (f_j - \tilde{f}_j)\Big(g(t_{j+1}) - g(t_j)\Big)\right| \\
&\le \max_j |f_j - \tilde{f}_j| \sum_j \Big|g(t_{j+1}) - g(t_j)\Big| \\
&\le \max_j |f_j - \tilde{f}_j| V(g; [a,b]) \to 0
\end{aligned}
$$

as $|\Delta| \to 0$ by the uniform continuity of f on $[a, b]$.

- Thus we get a well-defined definition which is independent of the choice of reference point in the approximation.

BV case

- If $g(t)$ has bounded total variation, we subtract the right hand side of the two definitions and obtain

$$\begin{aligned}&\Big|\sum_j (f_j - \tilde{f}_j)\Big(g(t_{j+1}) - g(t_j)\Big)\Big|\\ &\leq \max_j |f_j - \tilde{f}_j| \sum_j \Big|g(t_{j+1}) - g(t_j)\Big|\\ &\leq \max_j |f_j - \tilde{f}_j| V(g;[a,b]) \to 0\end{aligned}$$

  as $|\Delta| \to 0$ by the uniform continuity of f on $[a,b]$.
- Thus we get a well-defined definition which is independent of the choice of reference point in the approximation.
- If $g(t) = W_t(\omega)$, what will happen?

Three Choices

Example

Different choices for the stochastic integral $\int_0^T W_t dW_t$.

Three Choices

Example

Different choices for the stochastic integral $\int_0^T W_t dW_t$.

- *Choice 1:* Leftmost endpoint integral.

$$\int_0^T W_t dW_t \approx \sum_j W_{t_j}(W_{t_{j+1}} - W_{t_j}) := I_N^L.$$

Three Choices

Example

Different choices for the stochastic integral $\int_0^T W_t dW_t$.

- *Choice 1:* Leftmost endpoint integral.

$$\int_0^T W_t dW_t \approx \sum_j W_{t_j}(W_{t_{j+1}} - W_{t_j}) := I_N^L.$$

- *Choice 2:* Rightmost endpoint integral.

$$\int_0^T W_t dW_t \approx \sum_j W_{t_{j+1}}(W_{t_{j+1}} - W_{t_j}) := I_N^R.$$

Three Choices

Example

Different choices for the stochastic integral $\int_0^T W_t dW_t$.

- *Choice 1:* Leftmost endpoint integral.

$$\int_0^T W_t dW_t \approx \sum_j W_{t_j}(W_{t_{j+1}} - W_{t_j}) := I_N^L.$$

- *Choice 2:* Rightmost endpoint integral.

$$\int_0^T W_t dW_t \approx \sum_j W_{t_{j+1}}(W_{t_{j+1}} - W_{t_j}) := I_N^R.$$

- *Choice 3:* Midpoint integral.

$$\int_0^T W_t dW_t \approx \sum_j W_{t_{j+\frac{1}{2}}}(W_{t_{j+1}} - W_{t_j}) := I_N^M.$$

Expectation Check

We have the following identities from the statistical average sense.

$$
\begin{aligned}
\mathbb{E}(I_N^L) &= \sum_j \mathbb{E}W_{t_j}\mathbb{E}(W_{t_{j+1}} - W_{t_j}) = 0, \\
\mathbb{E}(I_N^R) &= \sum_j \Big[\mathbb{E}(W_{t_{j+1}} - W_{t_j})^2 + \mathbb{E}W_{t_j}\mathbb{E}(W_{t_{j+1}} - W_{t_j})\Big] \\
&= \sum_j \Delta t_j = T, \\
\mathbb{E}(I_N^M) &= \mathbb{E}\Big[\sum_j W_{t_{j+\frac{1}{2}}}(W_{t_{j+1}} - W_{t_{j+\frac{1}{2}}}) + \sum_j W_{t_{j+\frac{1}{2}}}(W_{t_{j+\frac{1}{2}}} - W_{t_j})\Big] \\
&= \sum_j \mathbb{E}(W_{t_{j+\frac{1}{2}}} - W_{t_j})^2 = \sum_j (t_{j+\frac{1}{2}} - t_j) = \frac{T}{2}.
\end{aligned}
$$

The reason is that the Brownian motion has unbounded variations for any finite interval. Therefore, we should take special attention to stochastic integrals.

Remark on Stochastic Integral

Remark.

- The stochastic integrals can not be defined for arbitrary continuous functions f, otherwise the function g must have bounded variations on compacts (by Banach-Steinhaus Theorem).[1] One rescue is to restrict the integrands to be a special class of functions, the adapted processes. That is the key point of the well-known Itô integral.

[1] P. Protter, Stochastic integration and differential equations, Springer

Remark on Stochastic Integral

Remark.

- The stochastic integrals can not be defined for arbitrary continuous functions f, otherwise the function g must have bounded variations on compacts (by Banach-Steinhaus Theorem).[1] One rescue is to restrict the integrands to be a special class of functions, the adapted processes. That is the key point of the well-known Itô integral.
- Different choices of the reference point correspond to different consistent definitions of stochastic integrals, but they can be connected by some simple transformation rules.

[1]P. Protter, Stochastic integration and differential equations, Springer

Remark on Stochastic Integral

Remark.

- The stochastic integrals can not be defined for arbitrary continuous functions f, otherwise the function g must have bounded variations on compacts (by Banach-Steinhaus Theorem).[1] One rescue is to restrict the integrands to be a special class of functions, the adapted processes. That is the key point of the well-known Itô integral.
- Different choices of the reference point correspond to different consistent definitions of stochastic integrals, but they can be connected by some simple transformation rules.
- Next, we take the filtration generated by standard Wiener process as $\mathcal{F}_t^W$. The construction of Itô integral takes the leftmost endpoint approximation

$$\int_0^T f(t,\omega)dW_t \approx \sum_j f_{t_j}(W_{t_{j+1}} - W_{t_j}).$$

[1] P. Protter, Stochastic integration and differential equations, Springer

Itô integral for Simple Functions

We first establish Itô integral on simple functions.[2]

- $f(t,\omega)$ is called a *simple function* if

$$f(t,\omega) = \sum_{j=1}^{n} e_j(\omega)\chi_{[t_j,t_{j+1})}(t),$$

where $e_j(\omega)$ is $\mathcal{F}_{t_j}^W$-measurable and $\chi_{[t_j,t_{j+1})}(t)$ is the indicator function on $[t_j, t_{j+1})$.

[2] Karatzas and Shreve, Brownian motion and stochastic calculus, Springer.

Itô integral for Simple Functions

We first establish Itô integral on simple functions.[2]

- $f(t,\omega)$ is called a *simple function* if

$$f(t,\omega) = \sum_{j=1}^{n} e_j(\omega)\chi_{[t_j,t_{j+1})}(t),$$

  where $e_j(\omega)$ is $\mathcal{F}_{t_j}^W$-measurable and $\chi_{[t_j,t_{j+1})}(t)$ is the indicator function on $[t_j, t_{j+1})$.
- For simple functions, define

$$\int_0^T f(t,\omega)dW_t = \sum_j e_j(\omega)(W_{t_{j+1}} - W_{t_j})$$

[2] Karatzas and Shreve, Brownian motion and stochastic calculus, Springer.

Properties of Itô integral

Lemma

For any $S \leq T$, the stochastic integral for the simple functions satisfies

$$(1) \qquad \mathbb{E}\left(\int_S^T f(t,\omega)dW_t\right) = 0,$$

$$(2) \qquad (\textit{Itô isometry}) \;\; \mathbb{E}\Big(\int_S^T f(t,\omega)dW_t\Big)^2 = \mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right).$$

Properties of Itô integral

Lemma

For any $S \leq T$, the stochastic integral for the simple functions satisfies

$$(1) \qquad \mathbb{E}\left(\int_S^T f(t,\omega)dW_t\right) = 0,$$

$$(2) \qquad (\textit{Itô isometry}) \;\; \mathbb{E}\Big(\int_S^T f(t,\omega)dW_t\Big)^2 = \mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right).$$

Proof. The first property is straightforward by the independence between $\Delta W_j := W_{t_{j+1}} - W_{t_j}$ and $e_j(\omega)$ and $\Delta W_j \sim N(0, t_{j+1} - t_j)$.

Properties of Itô integral

For the second property, we have

$$\begin{aligned}
\mathbb{E}\left(\int_S^T f(t,\omega)dW_t\right)^2 &= \mathbb{E}\Big(\sum_j e_j \Delta W_j\Big)^2 = \mathbb{E}\Big(\sum_{j,k} e_j e_k \Delta W_j \Delta W_k\Big) \\
&= \mathbb{E}\Big(\sum_j e_j^2 \Delta W_j^2 + 2\sum_{j<k} e_j e_k \Delta W_j \Delta W_k\Big) \\
&= \sum_j \mathbb{E}e_j^2 \cdot \mathbb{E}\Delta W_j^2 + \sum_{j<k} \mathbb{E}(e_j e_k \Delta W_j) \cdot \mathbb{E}(\Delta W_k) \\
&= \sum_j \mathbb{E}e_j^2 \Delta t_j = \mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right).
\end{aligned}$$

The last third identity holds because of the independence between ΔW_k and $e_j e_k \Delta W_j$ for $j < k$. □

Itô integral: Definition

Now we $f(t,\omega)$ belongs to the class of functions $\mathcal{V}[S,T]$ which defined as

(i) f is $\mathcal{B}([0,\infty)) \times \mathcal{F}$-measurable as a function from (t,ω) to $\mathbb{R}$,

(ii) $f(t,\omega)$ is $\mathcal{F}_t^W$-adapted,

(iii) $f \in L_P^2 L_t^2$, that is $\mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right) < \infty$.

Itô integral: Definition

Now we $f(t,\omega)$ belongs to the class of functions $\mathcal{V}[S,T]$ which defined as

(i) f is $\mathcal{B}([0,\infty)) \times \mathcal{F}$-measurable as a function from (t,ω) to $\mathbb{R}$,

(ii) $f(t,\omega)$ is $\mathcal{F}_t^W$-adapted,

(iii) $f \in L_P^2 L_t^2$, that is $\mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right) < \infty$.

- Recall the approximation property through simple functions $\phi_n(t,\omega)$

$$\mathbb{E}\left(\int_S^T (f(t,\omega) - \phi_n(t,\omega))^2 dt\right) \to 0,$$

i.e. $\phi_n \to f$ in $L_P^2 L_t^2$.

Itô integral: Definition

Now we $f(t,\omega)$ belongs to the class of functions $\mathcal{V}[S,T]$ which defined as

(i) f is $\mathcal{B}([0,\infty)) \times \mathcal{F}$-measurable as a function from (t,ω) to $\mathbb{R}$,

(ii) $f(t,\omega)$ is $\mathcal{F}_t^W$-adapted,

(iii) $f \in L_P^2 L_t^2$, that is $\mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right) < \infty$.

- Recall the approximation property through simple functions $\phi_n(t,\omega)$

$$\mathbb{E}\left(\int_S^T (f(t,\omega) - \phi_n(t,\omega))^2 dt\right) \to 0,$$

  i.e. $\phi_n \to f$ in $L_P^2 L_t^2$.

- We define the Itô integral as

$$\int_S^T f(t,\omega)dW_t = \lim_{n\to\infty} \int_S^T \phi_n(t,\omega)dW_t \qquad \text{in } \; L_P^2.$$

Itô integral: Definition

- From the Itô isometry, $\int_S^T \phi_n(t,\omega)dW_t$ is in L_P^2 for any simple function $\phi_n(t,\omega)$ and

$$\mathbb{E}\Big(\int_S^T \phi_n dW_t - \int_S^T \phi_m dW_t\Big)^2 = \mathbb{E}\left(\int_S^T (\phi_n - \phi_m)^2 dt\right).$$

Itô integral: Definition

- From the Itô isometry, $\int_S^T \phi_n(t,\omega)dW_t$ is in L_P^2 for any simple function $\phi_n(t,\omega)$ and

$$\mathbb{E}\Big(\int_S^T \phi_n dW_t - \int_S^T \phi_m dW_t\Big)^2 = \mathbb{E}\left(\int_S^T (\phi_n - \phi_m)^2 dt\right).$$

- The approximation sequence $\{\phi_n\}$ is a Cauchy sequence in $L_P^2(\Omega; L_t^2[S,T])$. This implies $\{\int_S^T \phi_n dW_t\}$ is also a Cauchy sequence in L_P^2.

Itô integral: Definition

- From the Itô isometry, $\int_S^T \phi_n(t,\omega)dW_t$ is in L_P^2 for any simple function $\phi_n(t,\omega)$ and

$$\mathbb{E}\Big(\int_S^T \phi_n dW_t - \int_S^T \phi_m dW_t\Big)^2 = \mathbb{E}\left(\int_S^T (\phi_n - \phi_m)^2 dt\right).$$

- The approximation sequence $\{\phi_n\}$ is a Cauchy sequence in $L_P^2(\Omega; L_t^2[S,T])$. This implies $\{\int_S^T \phi_n dW_t\}$ is also a Cauchy sequence in L_P^2.
- From the completeness of $L_P^2(\Omega)$, it has a unique limit and we define it as

$$\int_S^T f(t,\omega)dW_t$$

Itô integral: Definition

- From the Itô isometry, $\int_S^T \phi_n(t,\omega)dW_t$ is in L_P^2 for any simple function $\phi_n(t,\omega)$ and

$$\mathbb{E}\Big(\int_S^T \phi_n dW_t - \int_S^T \phi_m dW_t\Big)^2 = \mathbb{E}\left(\int_S^T (\phi_n - \phi_m)^2 dt\right).$$

- The approximation sequence $\{\phi_n\}$ is a Cauchy sequence in $L_P^2(\Omega; L_t^2[S,T])$. This implies $\{\int_S^T \phi_n dW_t\}$ is also a Cauchy sequence in L_P^2.
- From the completeness of $L_P^2(\Omega)$, it has a unique limit and we define it as

$$\int_S^T f(t,\omega)dW_t$$

- The independence on the choice of the approximating sequence $\{\phi_n\}$ is left as an exercise.

Itô Isometry

Theorem

For $f \in \mathcal{V}[S,T]$, the Itô integral satisfies

$$(1) \qquad \mathbb{E}\left(\int_S^T f(t,\omega)dW_t\right) = 0,$$

$$(2) \qquad (\textit{Itô isometry}) \ \mathbb{E}\Big(\int_S^T f(t,\omega)dW_t\Big)^2 = \mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right).$$

Proof. Firstly,

$$\left|\mathbb{E}\left(\int_S^T f(t,\omega dW_t\right)\right| = \left|\mathbb{E}\left(\int_S^T f(t,\omega)dW_t - \int_S^T \phi_n(t,\omega)dW_t\right)\right|$$

$$\leq \left(\mathbb{E}\Big(\int_S^T f(t,\omega)dW_t - \int_S^T \phi_n(t,\omega)dW_t\Big)^2\right)^{\frac{1}{2}} \to 0$$

by Hölder's inequality.

Ito Isometry

It is a standard result that if $X_n \to X$ in a Hilbert space H, then $|X_n| \to |X|$ and thus $|X_n|^2 \to |X|^2$, where $|\cdot|$ is the corresponding norm in Hilbert space H.

Ito Isometry

It is a standard result that if $X_n \to X$ in a Hilbert space H, then $|X_n| \to |X|$ and thus $|X_n|^2 \to |X|^2$, where $|\cdot|$ is the corresponding norm in Hilbert space H.

So we have

$$\mathbb{E}\left(\int_S^T \phi_n(t,\omega)dW_t\right)^2 \to \mathbb{E}\left(\int_S^T f(t,\omega)dW_t\right)^2 \qquad \text{in } L_P^2(\Omega)$$

and

$$\mathbb{E}\left(\int_S^T \phi_n^2(t,\omega)dt\right) \to \mathbb{E}\left(\int_S^T f^2(t,\omega)dt\right) \qquad \text{in } L_P^2(\Omega; L_t^2[S,T])$$

From the Itô isometry for simple functions, we obtain Itô isometry for $f \in \mathcal{V}[S,T]$.

□

Itô integral: Properties

The properties of the Itô integral

Proposition

For $f, g \in \mathcal{V}[S,T]$ and $U \in [S,T]$, we have

(i) $\int_S^T f dW_t = \int_S^U f dW_t + \int_U^T f dW_t$ *a.s..*

(ii) $\int_S^T (f + cg) dW_t = \int_S^T f dW_t + c \int_S^T g dW_t$ *(c is a constant) a.s..*

(iii) $\int_S^T f dW_t$ *is $\mathcal{F}_t^W$-measurable.*

Itô integral: Properties

The properties of the Itô integral

Proposition

For $f, g \in \mathcal{V}[S,T]$ and $U \in [S,T]$, we have

(i) $\int_S^T f dW_t = \int_S^U f dW_t + \int_U^T f dW_t$ *a.s..*

(ii) $\int_S^T (f + cg) dW_t = \int_S^T f dW_t + c \int_S^T g dW_t$ *(c is a constant) a.s..*

(iii) $\int_S^T f dW_t$ *is $\mathcal{F}_t^W$-measurable.*

Lemma

For $f \in \mathcal{V}[0,T]$, $X_t := \int_0^t f(s,\omega) dW_s$ has continuous trajectories in the almost sure sense.

Itô integral: Properties

One can define the multi-dimensional Itô integral $\int_0^T \boldsymbol{\sigma}(t,\omega)\cdot d\boldsymbol{W}_t$. To compute their expectation, we have the similar property as the Ito isometry

$$\mathbb{E}\left(\int_S^T \sigma(t,\omega)dW_t^j\right) = 0,$$

$$\mathbb{E}\Big(\int_S^T \sigma(t,\omega)dW_t^j\Big)^2 = \mathbb{E}\left(\int_S^T \sigma^2(t,\omega)dt\right), \quad \forall j.$$

and especially the cross product expectation

$$\mathbb{E}\left(\int_S^T \sigma_1(t,\omega)dW_t^i \cdot \int_S^T \sigma_2(t,\omega)dW_t^j\right) = 0, \qquad \forall i \neq j,$$

$$\mathbb{E}\Big(\int_S^T \sigma_1(t,\omega)dW_t^j \int_S^T \sigma_2(t,\omega)dW_t^j\Big) = \mathbb{E}\left(\int_S^T \sigma_1(t,\omega)\sigma_2(t,\omega)dt\right).$$

Itô integral: Example

Example

With Itô integral we have

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

Itô integral: Example

Example

With Itô integral we have

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

Proof. From the definition of Itô integral

$$\begin{aligned}\int_0^t W_s dW_s &\approx \sum_j W_{t_j}(W_{t_{j+1}} - W_{t_j}) = \sum_j \frac{2W_{t_j}W_{t_{j+1}} - 2W_{t_j}^2}{2} \\ &= \sum_j \frac{W_{t_{j+1}}^2 - W_{t_j}^2}{2} - \sum_j \frac{W_{t_{j+1}}^2 - 2W_{t_{j+1}}W_{t_j} + W_{t_j}^2}{2} \\ &= \frac{W_t^2}{2} - \frac{1}{2}\sum_j (W_{t_{j+1}} - W_{t_j})^2 \to \frac{W_t^2}{2} - \frac{t}{2},\end{aligned}$$

where the last limit is due to the fact $\langle W, W\rangle_t = t$. □

Table of Contents

Itô's formula: Simplest Case

- Let's take the differential form of $\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}$, then we have
$$dW_t^2 = 2W_t dW_t + dt.$$

Itô's formula: Simplest Case

- Let's take the differential form of $\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}$, then we have
$$dW_t^2 = 2W_t dW_t + dt.$$

- It is different from the traditional Newton-Leibnitz calculus which suggests $dW_t^2 = 2W_t dW_t$ with chain rule.

Itô's formula: Simplest Case

- Let's take the differential form of $\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}$, then we have
$$dW_t^2 = 2W_t dW_t + dt.$$

- It is different from the traditional Newton-Leibnitz calculus which suggests $dW_t^2 = 2W_t dW_t$ with chain rule.

Proposition

For any bounded and continuous function $f(t, \omega)$ in t,

$$\sum_j f(t_j^*, \omega)(W_{t_{j+1}} - W_{t_j})^2 \to \int_0^t f(s, \omega) ds, \quad \forall t_j^* \in [t_j, t_{j+1}]$$

in probability when the subdivision size goes to zero.

We simply denoted it as $(dW_t)^2 = dt$.

Proof. Using the uniform continuity of f on $[0, t]$, we have

$$\begin{aligned}
&\mathbb{E}\left(\sum_j f(t_j)\Delta W_{t_j}^2 - \sum_j f(t_j)\Delta t_j\right)^2 \\
=&\mathbb{E}\left(\sum_{j,k} f(t_j)f(t_k)(\Delta W_{t_j}^2 - \Delta t_j)(\Delta W_{t_k}^2 - \Delta t_k)\right) \\
=&\mathbb{E}\left(\sum_j f^2(t_j)\cdot\mathbb{E}\left((\Delta W_{t_j}^2 - \Delta t_j)^2|\mathcal{F}_{t_j}\right)\right) \\
=&2\sum_j \mathbb{E}f^2(t_j)\Delta t_j^2 \to 0.
\end{aligned}$$

At the same time, we have

$$|\sum_j (f(t_j^*) - f(t_j))\Delta W_{t_j}^2| \le \sup_j |f(t_j^*) - f(t_j)| \cdot \sum_j \Delta W_{t_j}^2.$$

The second term of the RHS converges to the quadratic variation of W_t in probability. Combining the results above leads to the desired conclusion.

Itô process

Now let us consider the Itô process defined as

$$X_t = X_0 + \int_0^t b(s,\omega)ds + \int_0^t \sigma(s,\omega)dW_s,$$

which is usually denoted as

$$dX_t = b(t,\omega)dt + \sigma(t,\omega)dW_t, \quad X_t|_{t=0} = X_0$$

for $\sigma \in \mathcal{W}[0,T]$, b is $\mathcal{F}_t$-adapted and $\int_0^T |b(t,\omega)|dt < \infty$ a.s.

1D Itô's formula

Theorem (1D Itô's formula)

If X_t *is an Itô process,* $Y_t = f(X_t)$ *where* f *is a twice differentiable function. Then* Y_t *is also an Itô process and*

$$dY_t = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2,$$

where the rule of simplification is $dt^2 = 0,\ dtdW_t = dW_tdt = 0$ *and* $(dW_t)^2 = dt$*, i.e.*

$$(dX_t)^2 = (bdt + \sigma dW_t)^2 = b^2dt^2 + 2b\sigma dtdW_t + \sigma^2(dW_t)^2 = \sigma^2 dt.$$

Thus finally

$$dY_t = \left(b(t,\omega)f'(X_t) + \frac{1}{2}\sigma^2(t,\omega)f''(X_t)\right)dt + \sigma(t,\omega)f'(X_t)dW_t.$$

1D Itô's formula

Sketch of Proof. We will only consider the situation that f, f' and f'' are bounded and continuous here.

- At first, if b and σ are simple functions, we have

$$
\begin{aligned}
Y_t - Y_0 &= \sum_j (f(X_{t_{j+1}}) - f(X_{t_j})) \\
&= \sum_j \left(f'(X_{t_j})\Delta X_{t_j} + \frac{1}{2} f''(X_{t_j})\Delta X_{t_j}^2 + R_j \right),
\end{aligned}
$$

where $\Delta X_{t_j} = X_{t_{j+1}} - X_{t_j}$ and $R_j = o(|\Delta X_{t_j}|^2)$.

1D Itô's formula

Sketch of Proof. We will only consider the situation that f, f' and f'' are bounded and continuous here.

- At first, if b and σ are simple functions, we have

$$Y_t - Y_0 = \sum_j (f(X_{t_{j+1}}) - f(X_{t_j}))$$

$$= \sum_j \left(f'(X_{t_j})\Delta X_{t_j} + \frac{1}{2} f''(X_{t_j})\Delta X_{t_j}^2 + R_j \right),$$

  where $\Delta X_{t_j} = X_{t_{j+1}} - X_{t_j}$ and $R_j = o(|\Delta X_{t_j}|^2)$.
- Without loss of generality we assume the discontinuity of the step functions are embedded in the current subdivision grid points.

1D Itô's formula

Sketch of Proof. We will only consider the situation that f, f' and f'' are bounded and continuous here.

- At first, if b and σ are simple functions, we have

$$Y_t - Y_0 = \sum_j (f(X_{t_{j+1}}) - f(X_{t_j}))$$
$$= \sum_j \left(f'(X_{t_j})\Delta X_{t_j} + \frac{1}{2} f''(X_{t_j})\Delta X_{t_j}^2 + R_j \right),$$

  where $\Delta X_{t_j} = X_{t_{j+1}} - X_{t_j}$ and $R_j = o(|\Delta X_{t_j}|^2)$.
- Without loss of generality we assume the discontinuity of the step functions are embedded in the current subdivision grid points.
- We obtain

$$\sum_j f'(X_{t_j})\Delta X_{t_j} = \sum_j f'(X_{t_j}) b(t_j)\Delta t_j + \sum_j f'(X_{t_j})\sigma(t_j)\Delta W_{t_j}$$
$$\to \int_0^t b(s) f'(X_s) ds + \int_0^t \sigma(s) f'(X_s) dW_s$$

- And

$$
\begin{aligned}
&\sum_j f''(X_{t_j})\Delta X_{t_j}^2 \\
=&\sum_j f''(X_{t_j})\left(b^2(t_j)\Delta t_j^2 + 2b(t_j)\sigma(t_j)\Delta t_j \Delta W_{t_j} + \sigma^2(t_j)\Delta W_{t_j}^2\right).
\end{aligned}
$$

- And

$$\sum_j f''(X_{t_j})\Delta X_{t_j}^2$$
$$= \sum_j f''(X_{t_j})\left(b^2(t_j)\Delta t_j^2 + 2b(t_j)\sigma(t_j)\Delta t_j \Delta W_{t_j} + \sigma^2(t_j)\Delta W_{t_j}^2\right).$$

- Set K be the bound of b, σ and f'', we have

$$|\sum_j f''(X_{t_j})b^2(t_j)\Delta t_j^2| \le K\sum_j \Delta t_j^2 \le KT\sup_j \Delta t_j \to 0,$$

$$|\sum_j f''(X_{t_j})b(t_j)\sigma(t_j)\Delta t_j \Delta W_{t_j}| \le K\sum_j |\Delta t_j \Delta W_{t_j}|$$
$$\le KT\sup_j |\Delta W_{t_j}| \to 0$$

$$\sum_j f''(X_{t_j})\sigma^2(t_j)\Delta W_{t_j}^2 \to \int_0^t \sigma^2(s)f''(X_s)ds \qquad \text{in} \quad L_P^2.$$

- And

$$\sum_j f''(X_{t_j})\Delta X_{t_j}^2$$

$$= \sum_j f''(X_{t_j}) \left(b^2(t_j)\Delta t_j^2 + 2b(t_j)\sigma(t_j)\Delta t_j \Delta W_{t_j} + \sigma^2(t_j)\Delta W_{t_j}^2 \right).$$

- Set K be the bound of b, σ and f'', we have

$$|\sum_j f''(X_{t_j})b^2(t_j)\Delta t_j^2| \leq K \sum_j \Delta t_j^2 \leq KT \sup_j \Delta t_j \to 0,$$

$$|\sum_j f''(X_{t_j})b(t_j)\sigma(t_j)\Delta t_j \Delta W_{t_j}| \leq K \sum_j |\Delta t_j \Delta W_{t_j}| \leq KT \sup_j |\Delta W_{t_j}| \to 0$$

$$\sum_j f''(X_{t_j})\sigma^2(t_j)\Delta W_{t_j}^2 \to \int_0^t \sigma^2(s) f''(X_s) ds \quad \text{in} \quad L_P^2.$$

- The general situation can be done by taking approximation through simple functions.

Theorem (Multidimensional Itô formula)

If $d\boldsymbol{X}_t = \boldsymbol{b}(t,\omega)dt + \boldsymbol{\sigma}(t,\omega)\cdot d\boldsymbol{W}_t$, *where* $\boldsymbol{X}_t \in \mathbb{R}^n$, $\boldsymbol{\sigma} \in \mathbb{R}^{n\times m}$, $\boldsymbol{W} \in \mathbb{R}^m$. *Define* $Y_t = f(\boldsymbol{X}_t)$, *where* $f \in C^2(\mathbb{R}^n)$. *Then*

$$dY_t = \nabla f(\boldsymbol{X}_t)\cdot d\boldsymbol{X}_t + \frac{1}{2}(d\boldsymbol{X}_t)^T \cdot \nabla^2 f(\boldsymbol{X}_t)\cdot(d\boldsymbol{X}_t),$$

where the rule of simplification is $dt^2 = 0$, $(dW_t^i)^2 = dt$, $dtdW_t^i = dW_t^i dt = dW_t^i dW_t^j = 0 \ (i \neq j)$. *That is*

$$(d\boldsymbol{X}_t)^T \cdot \nabla^2 f(\boldsymbol{X}_t)\cdot(d\boldsymbol{X}_t) = \sum_{l,k,i,j} dW_t^l \sigma_{il}\partial_{ij}^2 f \sigma_{jk} dW_t^k$$
$$= \sum_{k,i,j} \sigma_{ik}\sigma_{jk}\partial_{ij}^2 f dt = \boldsymbol{\sigma}\boldsymbol{\sigma}^T : \nabla^2 f dt,$$

where $\boldsymbol{A} : \boldsymbol{B} = \sum_{ij} a_{ij} b_{ji}$ *is the twice contraction for second order tensors. Finally*

$$dY_t = (\boldsymbol{b}\cdot\nabla f + \frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}^T : \nabla^2 f)dt + \nabla f \cdot \boldsymbol{\sigma}\cdot d\boldsymbol{W}_t.$$

Itô's formula: Applications

Example

Integration by part

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds.$$

Itô's formula: Applications

Example

Integration by part

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds.$$

Proof. Define $f(x, y) = xy$, $X_t = t$, $Y_t = W_t$, then from multidimensional Itô's formula

$$df(X_t, Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t.$$

With the rule $dtdW_t = 0$, we obtain $d(tW_t) = tdW_t + W_t dt$ and the result follows.

Iterated Itô integrals

Example

Iterated Itô integrals

$$\int_0^t dW_{t_1} \int_0^{t_1} dW_{t_2} \dots \int_0^{t_{n-1}} dW_{t_n} = \frac{1}{n!} t^{\frac{n}{2}} h_n \left(\frac{W_t}{\sqrt{t}} \right),$$

where $h_n(x)$ *is the* n*-th order Hermite polynomial*

$$h_n(x) = (-1)^n e^{\frac{1}{2}x^2} \frac{d^n}{dx^n} \left(e^{-\frac{1}{2}x^2} \right).$$

Iterated Itô integrals

Example

Iterated Itô integrals

$$\int_0^t dW_{t_1} \int_0^{t_1} dW_{t_2} \dots \int_0^{t_{n-1}} dW_{t_n} = \frac{1}{n!} t^{\frac{n}{2}} h_n \left(\frac{W_t}{\sqrt{t}} \right),$$

where $h_n(x)$ is the n-th order Hermite polynomial

$$h_n(x) = (-1)^n e^{\frac{1}{2}x^2} \frac{d^n}{dx^n} \left(e^{-\frac{1}{2}x^2} \right).$$

Proof. It is easy to verify that

$$\int_0^t W_s dW_s = \frac{t}{2!} h_2 \left(\frac{W_t}{\sqrt{t}} \right),$$

where $h_2(x) = x^2 - 1$ is the second order Hermite polynomial.

Iterated Itô integrals

In the same fashion, we have

$$\int_0^t \Big(\int_0^s W_u dW_u \Big) dW_s = \frac{1}{2} \int_0^t (W_s^2 - s) dW_s.$$

Iterated Itô integrals

In the same fashion, we have

$$\int_0^t \Big(\int_0^s W_u dW_u \Big) dW_s = \frac{1}{2} \int_0^t (W_s^2 - s) dW_s.$$

Using Itô's formula, we have

$$\int_0^t W_s^2 dW_s = \frac{1}{3} W_t^3 - \int_0^t W_s ds.$$

Iterated Itô integrals

In the same fashion, we have

$$\int_0^t \Big(\int_0^s W_u dW_u\Big) dW_s = \frac{1}{2}\int_0^t (W_s^2 - s) dW_s.$$

Using Itô's formula, we have

$$\int_0^t W_s^2 dW_s = \frac{1}{3}W_t^3 - \int_0^t W_s ds.$$

Hence, using the previous example we obtain

$$\int_0^t \Big(\int_0^s W_u dW_u) dW_s = \frac{1}{6}W_t^3 - \frac{1}{2}tW_t = \frac{1}{3!}t^{\frac{3}{2}} h_3\left(\frac{W_t}{\sqrt{t}}\right),$$

where $h_3(x) = x^3 - 3x$ is the third order Hermite polynomial. The general case is left as an exercise. □

Table of Contents

SDE: Wellposed-ness

We fist establish the classical well-posedness result for the stochastic differential equation.

$$d\boldsymbol{X}_t = \boldsymbol{b}(\boldsymbol{X}_t, t)dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t) \cdot d\boldsymbol{W}_t,$$

Theorem

Let $\boldsymbol{X} \in \mathbb{R}^n, \boldsymbol{W} \in \mathbb{R}^m$. Suppose the coefficients $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{\sigma} \in \mathbb{R}^{n\times m}$ satisfy global Lipschitz and linear growth conditions as

$$|\boldsymbol{b}(\boldsymbol{x}, t) - \boldsymbol{b}(\boldsymbol{y}, t)| + |\boldsymbol{\sigma}(\boldsymbol{x}, t) - \boldsymbol{\sigma}(\boldsymbol{y}, t)| \leq K|\boldsymbol{x} - \boldsymbol{y}|,$$

$$|\boldsymbol{b}(\boldsymbol{x}, t)|^2 + |\boldsymbol{\sigma}(\boldsymbol{x}, t)|^2 \leq K(1 + |\boldsymbol{x}|^2)$$

for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, t \in [0, T]$, where K is a positive constant and $|\cdot|$ means the Frobenius norm. Assume the initial value $\boldsymbol{X}_0 = \xi$ is a random variable which is independent of $\mathcal{F}_\infty^{\boldsymbol{W}}$ and satisfies $\mathbb{E}|\xi|^2 < \infty$. Then SDE has a unique t-continuous solution $\boldsymbol{X}_t \in \mathcal{V}[0, T]$.

Diffusion process

- The SDEs driven by Wiener processes is the typical Markov process which is also called the *diffusion processes* in stochastic analysis.

Diffusion process

- The SDEs driven by Wiener processes is the typical Markov process which is also called the *diffusion processes* in stochastic analysis.
- The diffusion process is defined for a Markov process $\{\boldsymbol{X}_t\}$ with continuous trajectory and its transition density $p(\boldsymbol{x}, t|\boldsymbol{y}, s)$ $(t \geq s)$ satisfies the following conditions for any $\epsilon > 0$:

$$\lim_{t \to s} \frac{1}{t-s} \int_{|\boldsymbol{x}-\boldsymbol{y}|<\epsilon} (\boldsymbol{x}-\boldsymbol{y}) p(\boldsymbol{x}, t|\boldsymbol{y}, s) d\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{y}, s) + O(\epsilon),$$

$$\lim_{t \to s} \frac{1}{t-s} \int_{|\boldsymbol{x}-\boldsymbol{y}|<\epsilon} (\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^T p(\boldsymbol{x}, t|\boldsymbol{y}, s) d\boldsymbol{x} = \boldsymbol{a}(\boldsymbol{y}, s) + O(\epsilon).$$

  $\boldsymbol{b}(\boldsymbol{y}, s)$ is called the drift of the considered diffusion process and $\boldsymbol{a}(\boldsymbol{y}, s)$ is called the diffusion matrix at time s and position $\boldsymbol{y}$.

Ornstein-Uhlenbeck process

Example (Ornstein-Uhlenbeck process)

$$dX_t = -\gamma X_t dt + \sigma dW_t.$$

The Ornstein-Uhlenbeck process (OU process) has fundamental importance in statistical physics since it serves as the simplest model for many complex diffusion dynamics.

Ornstein-Uhlenbeck process

Example (Ornstein-Uhlenbeck process)

$$dX_t = -\gamma X_t dt + \sigma dW_t.$$

The Ornstein-Uhlenbeck process (OU process) has fundamental importance in statistical physics since it serves as the simplest model for many complex diffusion dynamics.

Solution. The SDE is equivalent to

$$dX_t + \gamma X_t dt = \sigma dW_t.$$

By applying Ito's formula to $e^{\gamma t} X_t$, we get

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

Integrating from 0 to t we have

$$e^{\gamma t} X_t - X_0 = \int_0^t (\gamma e^{\gamma s} X_s ds + e^{\gamma s} dX_s) = \int_0^t \sigma e^{\gamma s} dW_s.$$

Ornstein-Uhlenbeck process

Thus the solution is

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

Ornstein-Uhlenbeck process

Thus the solution is

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

Define $Q_t := \int_0^t e^{-\gamma(t-s)} dW_s$, then it is easy to show that Q_t is a Gaussion process with

$$\mathbb{E}Q_t = 0, \qquad \mathbb{E}Q_t^2 = \int_0^t \mathbb{E}e^{-2\gamma(t-s)} ds = \frac{1}{2\gamma}(1 - e^{-2\gamma t}).$$

Therefore, X_t is also a Gaussian process if X_0 is Gaussian, and the limit behavior of X_t is

$$X_t \xrightarrow{d} N\left(0, \frac{\sigma^2}{2\gamma}\right), \qquad (t \to +\infty).$$

This equation is called the SDE with additive noise since the coefficient of dW_t term is just a constant.

Geometric Brownian motion

Example (Geometric Brownian motion)

$$dN_t = rN_t dt + \alpha N_t dW_t, \quad r, \alpha > 0.$$

This model has strong background in mathematical finance, in which N_t *represents the asset price,* r *is the interest rate and* α *is called the volatility.*

Geometric Brownian motion

Example (Geometric Brownian motion)

$$dN_t = rN_t dt + \alpha N_t dW_t, \quad r, \alpha > 0.$$

This model has strong background in mathematical finance, in which N_t represents the asset price, r is the interest rate and α is called the volatility.

Solution. Divide N_t to both sides we have $dN_t/N_t = rdt + \alpha dW_t$. In deterministic calculus $1/N_t dN_t = d(\log N_t)$, so we apply Ito's formula to $\log N_t$, then

$$\begin{aligned} d(\log N_t) &= \frac{1}{N_t} dN_t - \frac{1}{2N_t^2}(dN_t)^2 \\ &= \frac{1}{N_t} dN_t - \frac{1}{2N_t^2}\alpha^2 N_t^2 dt \\ &= \frac{1}{N_t} dN_t - \frac{\alpha^2}{2} dt. \end{aligned}$$

Geometric Brownian motion

Substitute the equation of dN_t we get

$$d(\log N_t) = (r - \frac{\alpha^2}{2})dt + \alpha dW_t.$$

Geometric Brownian motion

Substitute the equation of dN_t we get

$$d(\log N_t) = (r - \frac{\alpha^2}{2})dt + \alpha dW_t.$$

Integrate from 0 to t to both sides

$$\log N_t - \log N_0 = (r - \frac{\alpha^2}{2})t + \alpha W_t,$$

$$N_t = N_0 \exp\left\{(r - \frac{\alpha^2}{2})t + \alpha W_t\right\}.$$

This equation is called the SDE with multiplicative noise since the coefficient of dW_t term depends on N_t.

Langevin equation

Example (Langevin equation)

Mathematically a mesoscopic particle obeys the following well-known Langevin equation by Newton's Second Law

$$\begin{cases} d\boldsymbol{X}_t &= \boldsymbol{V}_t dt, \\ md\boldsymbol{V}_t &= \Big(-\gamma \boldsymbol{V}_t - \nabla V(\boldsymbol{X}_t)\Big)dt + \sqrt{2\sigma}d\boldsymbol{W}_t, \end{cases}$$

where γ is frictional coefficient, $V(\boldsymbol{X})$ is external potential, $\boldsymbol{W}_t$ is standard Wiener process, and σ is the strength of fluctuating force.

Langevin equation

Example (Langevin equation)

Mathematically a mesoscopic particle obeys the following well-known Langevin equation by Newton's Second Law

$$\begin{cases} d\boldsymbol{X}_t &= \boldsymbol{V}_t dt, \\ md\boldsymbol{V}_t &= \Big(-\gamma \boldsymbol{V}_t - \nabla V(\boldsymbol{X}_t) \Big) dt + \sqrt{2\sigma} d\boldsymbol{W}_t, \end{cases}$$

where γ is frictional coefficient, $V(\boldsymbol{X})$ is external potential, $\boldsymbol{W}_t$ is standard Wiener process, and σ is the strength of fluctuating force.

In the case that the external force is zero, we have

$$md\boldsymbol{V}_t = -\gamma \boldsymbol{V}_t dt + \sqrt{2\sigma} d\boldsymbol{W}_t.$$

This is exactly an Ornstein-Uhlenbeck process for $\boldsymbol{V}_t$.

Fluctuation-Dissipation Relation

- In the limit $t \to \infty$, we have

$$\langle \frac{1}{2} m \boldsymbol{V}^2 \rangle = \frac{3\sigma}{2\gamma}.$$

Fluctuation-Dissipation Relation

- In the limit $t \to \infty$, we have

$$\langle \frac{1}{2} m \boldsymbol{V}^2 \rangle = \frac{3\sigma}{2\gamma}.$$

- From equilibrium thermodynamics, the average kinetic energy must obey the rule

$$\langle \frac{1}{2} m \boldsymbol{V}^2 \rangle = \frac{3k_B T}{2}.$$

Fluctuation-Dissipation Relation

- In the limit $t \to \infty$, we have

$$\langle \frac{1}{2} m \boldsymbol{V}^2 \rangle = \frac{3\sigma}{2\gamma}.$$

- From equilibrium thermodynamics, the average kinetic energy must obey the rule

$$\langle \frac{1}{2} m \boldsymbol{V}^2 \rangle = \frac{3k_B T}{2}.$$

- Thus we obtain the well-known fluctuation-dissipation relation:

$$\sigma = k_B T \gamma.$$

Fluctuation-Dissipation Relation

- In the limit $t \to \infty$, we have

$$\langle \frac{1}{2} m \boldsymbol{V}^2 \rangle = \frac{3\sigma}{2\gamma}.$$

- From equilibrium thermodynamics, the average kinetic energy must obey the rule

$$\langle \frac{1}{2} m \boldsymbol{V}^2 \rangle = \frac{3k_B T}{2}.$$

- Thus we obtain the well-known fluctuation-dissipation relation:

$$\sigma = k_B T \gamma.$$

- It can be proved that in this case the diffusion coefficient

$$D := \lim_{t \to \infty} \frac{\langle (\boldsymbol{X}_t - \boldsymbol{X}_0)^2 \rangle}{6t} = \frac{k_B T}{\gamma}$$

which is called Einstein's relation.

Brownian dynamics]

Example (Brownian dynamics)

In the high γ case, the velocity $\boldsymbol{V}_t$ will always stay at an equilibrium Gaussian distribution, which means formally we can take $d\boldsymbol{V}_t = 0$. Then the Langevin equation is approximated by

$$d\boldsymbol{X}_t = -\frac{1}{\gamma}\nabla V(\boldsymbol{X}_t)dt + \sqrt{\frac{2k_B T}{\gamma}}d\boldsymbol{W}_t,$$

which is called Brownian dynamics or Smoluchowski approximation.

Table of Contents

Stratonovich integral: Definition

Definition

The Stratonovich (or Fisk-Stratonovich) integral is defined as the limit of the following approximation

$$\int_0^T f(t,\omega) \circ dW_t \approx \sum_j \frac{f(t_j) + f(t_{j+1})}{2} (W_{t_{j+1}} - W_{t_j}).$$

Stratonovich integral: Definition

Definition

The Stratonovich (or Fisk-Stratonovich) integral is defined as the limit of the following approximation

$$\int_0^T f(t,\omega) \circ dW_t \approx \sum_j \frac{f(t_j) + f(t_{j+1})}{2}(W_{t_{j+1}} - W_{t_j}).$$

Remark.

- We use the special notation $\circ$ for stochastic integral to distinguish the Itô and Stratonovich integrals.
- Following the similar way as in the definition for the Ito integral, we can also establish a consistent stochastic calculus based on the Stratonovich integral.

Connection between Ito and Stratonovich SDEs

Proposition

It turns out that If X_t satisfies the SDE

$$dX_t = b(X_t, t)dt + \sigma(X_t, t) \circ dW_t$$

in the Stratonovich sense, then X_t satisfies the modified Itô SDE

$$dX_t = \left(b(X_t, t) + \frac{1}{2}\partial_x \sigma\sigma(X_t, t)\right)dt + \sigma(X_t, t)dW_t.$$

Connection between Ito and Stratonovich SDEs

Proposition

It turns out that If X_t satisfies the SDE

$$dX_t = b(X_t, t)dt + \sigma(X_t, t) \circ dW_t$$

in the Stratonovich sense, then X_t satisfies the modified Itô SDE

$$dX_t = \Big(b(X_t, t) + \frac{1}{2}\partial_x \sigma\sigma(X_t, t)\Big)dt + \sigma(X_t, t)dW_t.$$

Proof.

- To understand this, we assume the solution X_t of the Stratonovich SDE satisfies

$$dX_t = \alpha(X_t, t)dt + \beta(X_t, t)dW_t.$$

Connection between Ito and Stratonovich SDEs

- By the definition of the Stratonovich integral

$$\int_0^t \sigma(X_s, s) \circ dW_s$$
$$\approx \sum_j \frac{1}{2}(\sigma(X_{t_j}, t_j) + \sigma(X_{t_{j+1}}, t_{j+1}))(W_{t_{j+1}} - W_{t_j}).$$

Connection between Ito and Stratonovich SDEs

- By the definition of the Stratonovich integral

$$\int_0^t \sigma(X_s, s) \circ dW_s$$
$$\approx \sum_j \frac{1}{2}(\sigma(X_{t_j}, t_j) + \sigma(X_{t_{j+1}}, t_{j+1}))(W_{t_{j+1}} - W_{t_j}).$$

- Additionally, we have

$$X_{t_{j+1}} = X_{t_j} + \alpha(X_{t_j}, t_j)\Delta t_j + \beta(X_{t_j}, t_j)\Delta W_{t_j} + h.o.t.,$$

$$\sum_j \sigma(X_{t_{j+1}}, t_{j+1})\Delta W_{t_j}$$
$$= \sum_j \Big(\sigma(X_{t_j}, t_j)\Delta W_{t_j} + \partial_t \sigma(X_{t_j}, t_j)\Delta t_j \Delta W_{t_j}$$
$$+ \partial_x \sigma \alpha(X_{t_j}, t_j)\Delta t_j \Delta W_{t_j} + \partial_x \sigma \beta(X_{t_j}, t_j)\Delta W_{t_j}^2 + h.o.t.\Big)$$
$$\to \int_0^t \sigma(X_s, s) dW_s + \int_0^t \partial_x \sigma \beta(X_s, s) ds$$

Connection between Ito and Stratonovich SDEs

- Summarizing the above results we obtain that X_t satisfies

$$dX_t = \left(b(X_t, t) + \frac{1}{2}\partial_x \sigma \beta(X_t, t)\right)dt + \sigma(X_t, t)dW_t.$$

Connection between Ito and Stratonovich SDEs

- Summarizing the above results we obtain that X_t satisfies

$$dX_t = \Big(b(X_t,t) + \frac{1}{2}\partial_x \sigma \beta(X_t,t)\Big)dt + \sigma(X_t,t)dW_t.$$

- To make the two SDEs consistent, we take

$$\beta(x,t) = \sigma(x,t), \quad \alpha(x,t) = b(x,t) + \frac{1}{2}\partial_x \sigma\sigma(x,t).$$

$\square$

Connection between Ito and Stratonovich SDEs

- Summarizing the above results we obtain that X_t satisfies

$$dX_t = \Big(b(X_t, t) + \frac{1}{2}\partial_x \sigma \beta(X_t, t)\Big)dt + \sigma(X_t, t)dW_t.$$

- To make the two SDEs consistent, we take

$$\beta(x, t) = \sigma(x, t), \quad \alpha(x, t) = b(x, t) + \frac{1}{2}\partial_x \sigma\sigma(x, t).$$

□

In the high dimensions, one can derive similarly

$$d\boldsymbol{X}_t = \Big(\boldsymbol{b}(\boldsymbol{X}_t, t) + \frac{1}{2}\nabla_x \boldsymbol{\sigma} : \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\Big)dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t) \cdot d\boldsymbol{W}_t$$

where $(\nabla_x \boldsymbol{\sigma} : \boldsymbol{\sigma})_i := \sum_{jk} \partial_k \sigma_{ij}\sigma_{kj}$, if $\boldsymbol{X}$ satisfies

$$d\boldsymbol{X}_t = \boldsymbol{b}(\boldsymbol{X}_t, t)dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t) \circ d\boldsymbol{W}_t.$$

Stratonovich integral: Property

Properties of the Stratonovich integral

- The Stratonovich integral satisfies the Newton-Leibnitz chain rule

$$df(X_t) = f'(X_t) \circ dX_t = f'(X_t) b(X_t, t) dt + f'(X_t) \sigma(X_t, t) \circ dW_t$$

Stratonovich integral: Property

Properties of the Stratonovich integral

- The Stratonovich integral satisfies the Newton-Leibnitz chain rule

$$df(X_t) = f'(X_t) \circ dX_t = f'(X_t) b(X_t, t) dt + f'(X_t) \sigma(X_t, t) \circ dW_t$$

- The corresponding multi-dimensional form is

$$\begin{aligned} df(\boldsymbol{X}_t) =& \nabla f(\boldsymbol{X}_t) \circ d\boldsymbol{X}_t \\ =& \nabla f(\boldsymbol{X}_t) \cdot \boldsymbol{b}(\boldsymbol{X}_t, t) dt + \nabla f(\boldsymbol{X}_t) \cdot \boldsymbol{\sigma}(\boldsymbol{X}_t, t) \circ d\boldsymbol{W}_t. \end{aligned}$$

Stratonovich integral: Property

Properties of the Stratonovich integral

- The Stratonovich integral satisfies the Newton-Leibnitz chain rule

$$df(X_t) = f'(X_t) \circ dX_t = f'(X_t)b(X_t, t)dt + f'(X_t)\sigma(X_t, t) \circ dW_t$$

- The corresponding multi-dimensional form is

$$\begin{aligned} df(\boldsymbol{X}_t) =& \nabla f(\boldsymbol{X}_t) \circ d\boldsymbol{X}_t \\ =& \nabla f(\boldsymbol{X}_t) \cdot \boldsymbol{b}(\boldsymbol{X}_t, t)dt + \nabla f(\boldsymbol{X}_t) \cdot \boldsymbol{\sigma}(\boldsymbol{X}_t, t) \circ d\boldsymbol{W}_t. \end{aligned}$$

- The Itô isometry and mean zero property no longer hold for the Stratonovich integral.

Wong-Zakai type theorem

- For each fixed realization ω of W_t, we want to solve X_t by treating $W.(\omega)$ like a deterministic forcing term.
- But the ODE can not be solved in the classical case because of the rough property of the path of the Brownian motion.
- Note that the C^1 functions on $[0,T]$ are dense in $C[0,T]$, we regularize the Brownian motion path from the following way

$$W^m \to W \text{ in } L^\infty[0,T] \text{ norm } \text{ as } m \to \infty,$$

  where $W^m \in C^1[0,T]$, the differential equation

$$dX_t^m = b(X_t^m, t)dt + \sigma(X_t^m, t)dW_t^m$$

  can be solved in the classical sense.
- Then it can be proved that

$$X^m \to X \text{ in } L^\infty[0,T], \quad m \to \infty, \ a.s.$$

The limit X_t is precisely the Stratonovich solution of the SDE.

Wong-Zakai type theorem

Therefore, Stratonovich interpretation is useful in physics.

- In realistic situations, the noise term $\dot{W}$ is usually not "white" but a smoothed colored noise since the idealistic white noise must be supplied with infinite energy from external environment.
- This smoothed colored noise exactly corresponds to some regularization of the white noise, which falls into the regime in the Wong-Zakai type smoothing argument.