Mathematics and numerics for data assimilation and state estimation – Lecture 19

Summer semester 2020

Overview

Summary lecture 18

- Stochastic processes, filtrations and Wiener processes.

- Plan for today: Itô integrals, theory and numerical integration of stochastic differential equations (SDE)

$$V_t = V_0 + \int_0^t b(V_s)ds + \int_0^t \sigma(V_s)dW_s$$

Overview

Construction of stochastic integrals

Seeking to make sense of the SDE

$$V_t = V_0 + \int_0^t b(V_s)ds + \int_0^t \sigma(V_s)dW_s$$

we need to define the stochastic integral.
Riemann-Stieltjes approach: let $|\Delta|$ denote the largest timestep in a mesh over $[0, t]$ and

$$\int_0^t \sigma(V_s)dW_s = \lim_{|\Delta|\to 0} \sum_k \sigma(V_{t_k^*})(W_{t_{k+1}} - W_{t_k})$$

for some $t_k^* \in [t_k, t_{k+1}]$.

Problem: these integrals are well-defined provided $\sigma(V_t)$ is continuous (which is reasonable to assume) and W_t has bounded total variation – which almost surely is not the case for the Wiener process.

Implication: different choices of t_k^* may lead to different integral values (both pathwise and in expectation).

Example

Consider the integral $\int_0^t W_s dW_s$, and three different choices for integration point:

$$t_k^* = \begin{cases} \text{left: } t_k \quad \text{giving} \quad I^L = \sum_k W_{t_k}(W_{t_{k+1}} - W_{t_k}) \\ \text{right: } t_{k+1} \quad \text{giving} \quad I^R = \sum_k W_{t_{k+1}}(W_{t_{k+1}} - W_{t_k}) \\ \text{middle: } t_{k+1/2} \quad \text{giving} \quad I^M = \sum_k W_{t_{k+1/2}}(W_{t_{k+1}} - W_{t_k}) \end{cases}$$

And

$$\mathbb{E}\left[I^L \right] = \sum_k \mathbb{E}\left[W_{t_k}(W_{t_{k+1}} - W_{t_k}) \right] \overset{W_{t_k} \perp (W_{t_{k+1}} - W_{t_k})}{=} \sum_k \mathbb{E}\left[W_{t_k} \right] \mathbb{E}\left[W_{t_{k+1}} - W_{t_k} \right] = 0,$$

while

$$\begin{aligned} \mathbb{E}\left[I^R \right] &= \sum_k \mathbb{E}\left[W_{t_{k+1}}(W_{t_{k+1}} - W_{t_k}) \right] \\ &= \sum_k \mathbb{E}\left[\Big((W_{t_{k+1}} - W_{t_k}) + W_{t_k} \Big)(W_{t_{k+1}} - W_{t_k}) \right] \\ &= \sum_k \mathbb{E}\left[(W_{t_{k+1}} - W_{t_k})^2 \right] + I^L = \sum_k (t_{k+1} - t_k) = t \end{aligned}$$

and $\quad \mathbb{E}\left[I^M \right] = t/2.$

Overview

Itô integral

Given a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, with $\mathcal{F}_t = \mathcal{F}_t^W$, the Itô integral is defined by

$$\int_0^t \sigma(V_s)dW_s := \lim_{|\Delta|\to 0} \sum_k \sigma(V_{t_k})(W_{t_{k+1}} - W_{t_k})$$

where Δ denotes a mesh/subdivision of $[0, t]$ and one assumes that both V_t and W_t are $\mathcal{F}_t$-adapted.
It remains to describe what we mean by "$=$" in the above definition.

Integrals of simple and $\mathcal{F}_t$-adapted functions

Given a mesh $\{\tau_k\}_{k=0}^n$ over an interval $[S, T]$, we consider simple functions of the form

$$\phi_n(\omega, t) := \sum_{j=1}^{n-1} e_j(\omega) \mathbb{1}_{[\tau_j, \tau_{j+1})}(t)$$

with e_j being $\mathcal{F}_{\tau_j}$-measurable. This makes also ϕ_n $\mathcal{F}_t$-measurable.

The Itô integral is given by

$$\int_S^T \phi_n(t, \omega) dW_t := \lim_{|\Delta| \to 0} \sum_k \phi_n(t_k, \omega)(W_{t_{k+1}} - W_{t_k}) = \sum_{j=0}^{n-1} e_j(\omega)(W_{\tau_{j+1}} - W_{\tau_j})$$

Motivation: Summing over a finer mesh $\Delta \supset \{\tau_k\}_{k=0}^n$ leads to telescoping sums of Wiener increments over each $\tau - \textit{interval}$: if $[t_{k_1}, t_{k_2}) = [\tau_j, \tau_{j+1})$, then $\phi_n(\cdot, \omega)|_{[\tau_j, \tau_{j+1})} = \phi_n(\tau_j, \omega)$ and

$$\sum_{k=k_1}^{k_2-1} \phi_n(t_k, \omega)(W_{t_{k+1}} - W_{t_k}) = \phi_n(\tau_j, \omega) \sum_{k=k_1}^{k_2-1} (W_{t_{k+1}} - W_{t_k}) = e_j(\omega)(W_{\tau_{j+1}} - W_{\tau_j})$$

Properties of simple-function stochastic integrals

Since $e_j(\omega)$ is $\mathcal{F}_{\tau_j}$-measurable, it turns out that

$$e_j \perp \Delta W_k := W_{\tau_{k+1}} - W_{\tau_k} \quad \text{for any } k \geq j,$$

(since $\mathcal{F}_{\tau_j} \perp \sigma(\{W_s - W_{\tau_j}\}_{s \geq \tau_j})$).

Property 1: The Itô integral has mean zero:

$$\mathbb{E}\left[\int_S^T \phi_n(t,\cdot)dW_t\right] = \sum_{j=0}^{n-1} \mathbb{E}\left[\, e_j(\cdot)\Delta W_j\right] = \sum_{j=0}^{n-1} \mathbb{E}\left[\, e_j(\cdot)\right] \mathbb{E}\left[\, \Delta W_j\right] = 0$$

Property 2: Itô isometry:

$$\mathbb{E}\left[\left(\int_S^T \phi_n(t,\cdot)dW_t\right)^2\right] = \mathbb{E}\left[\int_S^T \phi_n^2(t,\cdot)dt\right]$$

Indepencence of σ−algebras vs rv [cf. Durrett]

Given two rv on $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ and $Y : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ defined on the same probability space, we recall that

$$X \perp Y \iff \mathbb{P}(X^{-1}(B_1) \cap Y^{-1}(B_2)) = \mathbb{P}(X^{-1}(B_1))\mathbb{P}(Y^{-1}(B_2)) \quad \forall B_1, B_2 \in \mathcal{B}.$$

The independence condition is equivalent to

$$\mathbb{P}(C_1 \cap C_2) = \mathbb{P}(C_1)\mathbb{P}(C_2) \qquad \forall C_1 \in \sigma(X) \text{ and } C_2 \in \sigma(Y),$$

since any $C_1 \in \sigma(X)$ can be written $C_1 = X^{-1}(B_1)$ for some $B_1 \in \mathcal{B}$ and any $C_2 \in \sigma(Y)$, $C_2 = Y^{-1}(B_2)$ for some $B_2 \in \mathcal{B}$.

Equivalence $\perp$ of rv and $\perp$ of σ-algebras: $X \perp Y \iff \sigma(X) \perp \sigma(Y)$.

This naturally extends to point evaluations etc of stochastic processes. E.g.,

$$e_j \perp \Delta W_j \iff \sigma(e_j) \perp \sigma(\Delta W_j)$$

And this holds since since $\sigma(e_j) \subset \mathcal{F}_{\tau_j} \perp \sigma(\{W_s - W_{\tau_j})\}_{s \geq \tau_j}) \supset \sigma(\Delta W_j)$.

Proof of Itô isometry:

$$\begin{aligned}
\mathbb{E}\left[\left(\int_S^T \phi_n(t,\cdot)dW_t\right)^2\right] &= \mathbb{E}\left[\sum_{j,k} e_j e_k \Delta W_j \Delta W_k\right] \\
&= \sum_j \mathbb{E}\left[e_j^2 \Delta W_j^2\right] + 2\sum_{j<k} \mathbb{E}\left[\sum_{j,k} e_j e_k \Delta W_j \Delta W_k\right] \\
&= \sum_j \mathbb{E}\left[e_j^2\right] \mathbb{E}\left[\Delta W_j^2\right] + 2\sum_{j<k} \mathbb{E}\left[e_j e_k \Delta W_j\right] \mathbb{E}\left[\Delta W_k\right] \\
&= \sum_j \mathbb{E}\left[e_j^2\right] (\tau_{j+1} - \tau_j) \\
&= \mathbb{E}\left[\int_S^T \phi_n^2(t,\cdot)dt\right]
\end{aligned}$$

Where we used that $e_j \perp \Delta W_j$ and that for $k > j$, $e_j e_k \Delta W_j \perp \Delta W_k$ (since $\mathcal{F}_{\tau_k} \perp \sigma(\{W_s - W_{\tau_k}\}_{s \geq \tau_k})$.

We next extend the Itô integral to more general integrands:

Definition 1

Let $\mathcal{V}[S, T]$ be the class of functions $f(t, \omega) \in \mathbb{R}$ that satisfying

- $f : [S, T] \times \Omega \to \mathbb{R}$ is $\mathcal{B} \times cF$-measurable (i.e., $f^{-1}(B) \in \mathcal{B} \times \mathcal{F}$ for any $B \in \mathbb{R}$)
- f is $\mathcal{F}_t$-adapted, (i.e., $f(t, \cdot)$ is $\mathcal{F}_t$-measurable for each $t \in [S, T]$)
- $f \in L^2(\Omega; L^2[S, T])$ meaning $\mathbb{E}^\omega\left[\int_S^T f^2(t, \omega) dt\right] < \infty$.

[ELV-E 7] For any $f \in \mathcal{V}[S, T]$ there exists a sequence of simple fcns $\{\phi_n\} \subset \mathcal{V}[S, T]$ such that

$$\|f - \phi_n\|^2_{L^2(\Omega; L^2[S,T])} = \mathbb{E}\left[\int_S^T (\phi_n(t, \cdot) - f(t, \cdot))^2 \, dt\right] \to 0 \quad \text{as } n \to \infty.$$

This implies that $\{\phi_n\}$ is Cauchy in the Banach space $L^2(\Omega; L^2[S, T])$.

Definition of Itô integral

We define
$$\int_S^T f(t,\omega)dW_t \overset{L^2(\Omega)}{:=} \lim_{n\to\infty} \int_S^T \phi_n(t,\omega)dW_t$$

This limit exists, since by Itô isometry,

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_S^T \phi_n(t,\cdot)dW_t - \int_S^T \phi_m(t,\cdot)dW_t\right)^2\right] \\
&\quad = \mathbb{E}\left[\left(\int_S^T \phi_n(t,\cdot) - \phi_m(t,\cdot)dW_t\right)^2\right] \\
&\quad = \mathbb{E}\left[\int_S^T (\phi_n(t,\cdot) - \phi_m(t,\cdot))^2 dt\right] \\
&\quad = \|\phi_n - \phi_m\|^2_{L^2(\Omega;L^2[S,T])} \to 0 \quad \text{as } m,n\to\infty.
\end{aligned}$$

Properties of the Itô integral

For $f, g \in \mathcal{V}[S, T]$ and $u \in [S, T]$, the following integral properties extend from simple-function setting:

- Mean zero: $\mathbb{E}\left[\int_S^T f dW_t\right] = 0$,
- Itô isometry: $\mathbb{E}\left[\left(\int_S^T f dW_t\right)^2\right] = \mathbb{E}\left[\int_S^T f^2 dt\right]$,
- partition of integral: $\int_S^T f dW_t \overset{a.s.}{=} \int_S^u f dW_t + \int_u^T f dW_t$,
- for any scalar $c \in \mathbb{R}$, $\int_S^T f + cg dW_t \overset{a.s.}{=} \int_S^T f dW_t + c \int_S^T g dW_t$,
- $\int_S^T f dW_t$ is $\mathcal{F}_T$-measurable.

Overview

Definition 2 (1-D Itô process)

Given a Wiener process W_t defined on $(\Omega, \mathcal{F}, \mathbb{P})$, an Itô process over $[0, T]$ is defined by

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

where $\sigma \in \mathcal{V}[0, T]$ and $b : \Omega \times [0, T] \to \mathbb{R}$ is $\mathcal{F}_t$-adapted and $\int_0^T |b(t,\omega)|dt < \infty$ for a.a. ω. Or, equivalently,

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

Question: For an Itô process X_t and $f \in C^2(\mathbb{R})$, what is the "Itô chain rule" for computing $df(X_t) =?$,

The classic chain rule yields: $$df(X_t) = f'(X_t)dX_t + \underbrace{\frac{1}{2}f''(X_t)dX_t^2 + \dots}_{h.o.t.}$$

but since X_t has less regularity than in classic settings, it turns out that some "classic h.o.t." needs to be reclassified as leading order.

Quadratic variation of the Wiener process

The quadtratic variation of W_t over $[0, t]$ is defined as

$$[W, W]_t := \lim_{|\Delta|\downarrow 0} \sum_k (W_{t_{k+1}} - W_{t_k})^2$$

It can be shown that for any $t \geq 0$,

$$[W, W]_t \overset{L^2(\Omega)}{=} t \quad \text{meaning} \quad \mathbb{E}\left[([W, W]_t - t)^2 \right] = 0.$$

We employ this property to motivate the following Itô integration:

$$\int_0^t W_s dW_s \approx \sum_j W_{t_j}(W_{t_{j+1}} - W_{t_j}) = \dots$$

$$= \frac{W_t^2}{2} - \frac{1}{2}\sum_j (W_{t_{j+1}} - W_{t_j})^2 \to \frac{W_t^2}{2} - \frac{t}{2}.$$

This corresponds to the differential equation

$$W_t dW_t = \frac{dW_t^2}{2} - \frac{dt}{2} \quad \text{or equivalently} \quad dW_t^2 = 2W_t dW_t + dt$$

Note that this is different from the classic chain rule: $dW_t^2 = 2W_t dW_t$.

Theorem 3 (ELV-E 7.6)

Assume $f \in \mathcal{V}[0,T]$ is bounded and continuous for $t \in [0,T]$ for almost all ω. Then, in probability,

$$\lim_{|\Delta|\downarrow 0} \sum_j f(t_j^*, \omega)(W_{t_{j+1}} - W_{t_j})^2 = \int_0^T f(s,\omega)ds$$

for any choice $t_j^* \in [t_j, t_{j+1}]$

This motivates formally writing $(dW_t)^2 = dt$, and by introducing the additional formal h.o.t. rules

$$(dt)^2 = 0, \quad \text{and} \quad dtdW = dWdt = 0$$

we derive for the Itô process

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

and $f \in C^2(\mathbb{R})$, the **1D Itô's formula**:

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 = \left(f'(X_t)b + \frac{1}{2}f''(X_t)\sigma^2\right)dt + f'(X_t)dW \tag{1}$$

Application of Itô's formula

To evaluate

$$X_t = \int_0^t W_s dW_s$$

consider the detour of introducing $f(x) = x^2/2$ and noting that

$$X_t = \int_0^t f'(W_s) dW_s.$$

Next, apply Itô's formula to $Y_t = f(W_t)$:

$$dY_t = f'(W_t)dW_t + \frac{1}{2}f''(W_t)(dW_t)^2 = W_t dW_t + \frac{dt}{2}.$$

Integrating both sides yields,

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

Itô integrals in higher dimensions

Multidimensional Itô integrals of the form

$$\int_0^T \sigma(t, \omega) dW_t$$

where

- each component of $\sigma : [0, T] \times \Omega \to \mathbb{R}^{d \times n}$ belongs to the function space $\mathcal{V}[0, T]$ and
- the components of $W_t : \Omega \times [0, T] \to \mathbb{R}^n$ are independent Wiener processes.

See [ELV-E 7.2] for more details on this and Itô's formula in higher dimensions.

Overview

Existence and uniqueness of Itô SDE

Theorem 4 (ELV-E 7.14)

For the Itô SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad \text{for} \quad t \in [0, T], \quad X_t|_{t=0} = X_0$$

with coefficients $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times n}$ *and* W *an* $n-$*dimensional Wiener process, assume that for some* $K > 0$ *that*

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq K|x - y|$$
$$|b(x)|^2 + |\sigma(x)|^2 \leq K(1 + |x|^2)$$

for all $x, y \in \mathbb{R}^d$ *and that* $X_0 \in L^2(\Omega)$ *is independent from the history of the Wiener paths:* $\sigma(X_0) \perp \mathcal{F}_T^W$. *Then there exists a unique solution* $X \in L^2(\Omega; L^2[0, T])$ *satisfying* $X \in \mathcal{V}[0, T]$ *for each component.*

Remark: Unless X_0 is deterministic, the filtration must be augmented $\mathcal{F}_t = \mathcal{F}_t^W \vee \sigma(X_0) = \sigma(X_0, \{W_s\}_{s\leq t})$.

Proof ideas:

Existence: can be derived through a Picard iteration argument:

$$X_t^{(k+1)} = X_0 + \int_0^t b(X_s^{(k)})\, ds + \int_0^t \sigma(X_s^{(k)}) dW_s$$

and $X_t^{(0)} := X_0$.

Uniqueness in $L^2(\Omega; L^2[0,T])$: Given a pair of solutions $X, \hat{X}$, Itô isometry and the regularity of the coefficients yield

$$\begin{aligned}\mathbb{E}\left[|X_t - \hat{X}_t|^2\right] &\leq 2\mathbb{E}\left[\left(\int_0^t b(X_s) - b(\hat{X}_s) ds\right)^2\right] \\ &\quad + 2\mathbb{E}\left[\int_0^t (\sigma(X_s) - \sigma(\hat{X}_s))^2 ds\right] \\ &\leq 2K^2(1+t)\int_0^t \mathbb{E}\left[|X_s - \hat{X}_s|^2\right] ds\end{aligned}$$

By Grönwall's inequality, $X_t \overset{a.s.}{=} \hat{X}_t$ for all $t \in [0,T] \cap \mathbb{Q}$. Result follows by the (a.s.) continuity of solutions.

Example: Geometric Brownian Motion

$$dN_t = rN_t dt + \alpha N_t dW_t, \qquad N_t|_{t=0} = N_0$$

N_t the non-negative price of an asset, $r, \alpha >$ interest rate and volatility. Assuming $N_t > 0$ (once $N_t = 0$, it will remain 0-valued),

$$\frac{dN_t}{N_t} = rdt + \alpha dW_t,$$

Applying Ito's formula to $Y_t = \log(N_t)$ yields

$$\begin{aligned} d\log(N_t) &= \frac{1}{N_t} dN_t - \frac{1}{2N_t^2}(dN_t)^2 \\ &= \frac{rN_t dt + \alpha N_t dW_t}{N_t} - \frac{N_t^2 \alpha^2 dt}{2N_t^2} \\ &= (r - \alpha^2/2)dt + \alpha dW_t \end{aligned}$$

and thus

$$N_t = N_0 e^{(r-\alpha^2/2)t + \alpha W_t}.$$

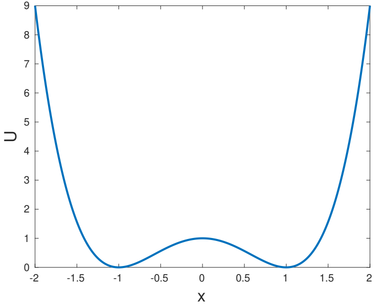
Langevin equation

$$dX_t = V_t dt$$
$$m dV_t = (-\gamma V_t - U'(X_t))dt + \sigma dW_t$$

Particle-velocity system (X, V) in a force field potential $U : \mathbb{R} \to \mathbb{R}$.
Friction coefficient γ, σ - magnitude of noise force

This is a "stochastic version" the newtonian dynamics

$$\dot{x} = v$$
$$m\dot{v} = -U'(x)$$



Potentials with local minima lead to pseudo-stable states for X_t.

Summary

- Have introduced stochastic integrals and differential equations.
- SDE extend the previously studied dynamics $\Psi(V_j) + \xi_j$ in many ways:
  1. the dynamics may now be nonlinear in both the drift and the diffusion coefficient,
  2. the noise enters in a more general way (not only as additive noise) through the diffusion coefficient,
  3. the dynamics is now continuous . . . so one may generalize observation frequency as well.
- Next time: Filtering problems with SDE dynamics.