Lecture 2: Itô Calculus and Stochastic Differential Equations

Simo Särkkä
Aalto University
November 4, 2014
During the last lecture we treated SDEs as white-noise driven differential equations of the form

\[ \frac{dx}{dt} = f(x, t) + L(x, t) w(t), \]

For linear equations the approach worked ok. But there is something strange going on:
- The use of chain rule of calculus led to wrong results.
- With non-linear differential equations we were completely lost.
- Picard-Lindelöf theorem did not work at all.

The source of all the problems is the everywhere discontinuous white noise \( w(t) \).

So how should we really formulate SDEs?
Integrating the differential equation from $t_0$ to $t$ gives:

$$x(t) - x(t_0) = \int_{t_0}^{t} f(x(t), t) \, dt + \int_{t_0}^{t} L(x(t), t) \, w(t) \, dt.$$ 

- The first integral is just a normal Riemann/Lebesgue integral.
- The second integral is the problematic one due to the white noise.
- This integral cannot be defined as Riemann, Stieltjes or Lebesgue integral as we shall see next.
In the Riemannian sense the integral would be defined as

$$\int_{t_0}^{t} \mathbf{L}(\mathbf{x}(t), t) \mathbf{w}(t) \, dt = \lim_{n \to \infty} \sum_{k} \mathbf{L}(\mathbf{x}(t^*_k), t^*_k) \mathbf{w}(t^*_k) (t_{k+1} - t_k),$$

where $t_0 < t_1 < \ldots < t_n = t$ and $t^*_k \in [t_k, t_{k+1}]$.

- **Upper and lower sums** are defined as the selections of $t^*_k$ such that the integrand $\mathbf{L}(\mathbf{x}(t^*_k), t^*_k) \mathbf{w}(t^*_k)$ has its maximum and minimum values, respectively.

- The Riemann integral exists if the upper and lower sums converge to the same value.

- Because white noise is **discontinuous everywhere**, the Riemann integral **does not exist**.
Stieltjes integral is more general than the Riemann integral.

In particular, it allows for discontinuous integrands.

We can interpret the increment $w(t) \, dt$ as increment of another process $\beta(t)$ such that

$$\int_{t_0}^{t} L(x(t), t) \, w(t) \, dt = \int_{t_0}^{t} L(x(t), t) \, d\beta(t).$$

It turns out that a suitable process for this purpose is the Brownian motion —
Brownian motion

1. **Gaussian increments:**

   \[ \Delta \beta_k \sim N(0, Q \Delta t_k), \]

   where \( \Delta \beta_k = \beta(t_{k+1}) - \beta(t_k) \) and \( \Delta t_k = t_{k+1} - t_k \).

2. **Non-overlapping increments are independent.**

- \( Q \) is the **diffusion matrix** of the Brownian motion.
- Brownian motion \( t \mapsto \beta(t) \) has **discontinuous derivative** everywhere.
- **White noise** can be considered as the formal derivative of Brownian motion \( w(t) = d\beta(t)/dt \).
Stieltjes integral is defined as a limit of the form

\[
\int_{t_0}^{t} L(x(t), t) \, d\beta = \lim_{n \to \infty} \sum_{k} L(x(t^*_k), t^*_k) [\beta(t_{k+1}) - \beta(t_k)],
\]

where \( t_0 < t_1 < \ldots < t_n \) and \( t^*_k \in [t_k, t_{k+1}] \).

The limit \( t^*_k \) should be independent of the position on the interval \( t^*_k \in [t_k, t_{k+1}] \).

But for integration with respect to Brownian motion this is not the case.

Thus, Stieltjes integral definition does not work either.
Attempt 3: Lebesgue integral

- In Lebesgue integral we could interpret $\beta(t)$ to define a “stochastic measure” via $\beta((u, v)) = \beta(u) - \beta(v)$.
- Essentially, this will also lead to the definition

$$\int_{t_0}^{t} L(x(t), t) \, d\beta = \lim_{n \to \infty} \sum_k L(x(t_k^*), t_k^*) \left[ \beta(t_{k+1}) - \beta(t_k) \right],$$

where $t_0 < t_1 < \ldots < t_n$ and $t_k^* \in [t_k, t_{k+1}]$.
- Again, the limit should be independent of the choice $t_k^* \in [t_k, t_{k+1}]$.
- Also our “measure” is not really a sensible measure at all.
- $\Rightarrow$ Lebesgue integral does not work either.
The solution to the problem is the Itô stochastic integral.

The idea is to fix the choice to $t^*_k = t_k$, and define the integral as

$$\int_{t_0}^{t} L(x(t), t) \, d\beta(t) = \lim_{n \to \infty} \sum_{k} L(x(t_k), t_k) [\beta(t_{k+1}) - \beta(t_k)].$$

This Itô stochastic integral turns out to be a sensible definition of the integral.

However, the resulting integral does not obey the computational rules of ordinary calculus.

Instead of ordinary calculus we have Itô calculus.
Consider the white noise driven ODE

$$\frac{dx}{dt} = f(x, t) + L(x, t) \, w(t).$$

This is actually defined as the Itô integral equation

$$x(t) - x(t_0) = \int_{t_0}^{t} f(x(t), t) \, dt + \int_{t_0}^{t} L(x(t), t) \, d\beta(t),$$

which should be true for arbitrary $t_0$ and $t$.

Settings the limits to $t$ and $t + dt$, where $dt$ is “small”, we get

$$dx = f(x, t) \, dt + L(x, t) \, d\beta.$$ 

This is the canonical form of an Itô SDE.
Let’s formally divide by $dt$, which gives

$$\frac{dx}{dt} = f(x, t) + L(x, t) \frac{d\beta}{dt}.$$ 

Thus we can interpret $d\beta/dt$ as white noise $w$.

Note that we cannot define more general equations

$$\frac{dx(t)}{dt} = f(x(t), w(t), t),$$

because we cannot re-interpret this as an Itô integral equation.

White noise should not be thought as an entity as such, but it only exists as the formal derivative of Brownian motion.
Consider the stochastic integral

\[ \int_0^t \beta(t) \, d\beta(t) \]

where \( \beta(t) \) is a standard Brownian motion (\( Q = 1 \)).

Based on the ordinary calculus we would expect the result \( \beta^2(t)/2 \)—but it is wrong.

If we select a partition \( 0 = t_0 < t_1 < \ldots < t_n = t \), we get

\[ \int_0^t \beta(t) \, d\beta(t) = \lim \sum_k \beta(t_k)[\beta(t_{k+1}) - \beta(t_k)] \]

\[ = \lim \sum_k \left[ -\frac{1}{2}(\beta(t_{k+1}) - \beta(t_k))^2 + \frac{1}{2}(\beta^2(t_{k+1}) - \beta^2(t_k)) \right] \]
We have
\[
\lim \sum_k \frac{1}{2} \left( \beta(t_{k+1}) - \beta(t_k) \right)^2 \to -\frac{1}{2} t
\]
and
\[
\lim \sum_k \frac{1}{2} \left( \beta^2(t_{k+1}) - \beta^2(t_k) \right) \to \frac{1}{2} \beta^2(t).
\]
Thus we get the (slightly) unexpected result
\[
\int_0^t \beta(t) \, d\beta(t) = -\frac{1}{2} t + \frac{1}{2} \beta^2(t).
\]
This is unexpected only if we believe in the chain rule:
\[
\frac{d}{dt} \left[ \frac{1}{2} x^2(t) \right] = \frac{dx}{dt} x.
\]
But it is not true for a (Itô) stochastic process \( x(t) \)!
Itô formula

Assume that \( x(t) \) is an Itô process, and consider arbitrary (scalar) function \( \phi(x(t), t) \) of the process. Then the Itô differential of \( \phi \), that is, the Itô SDE for \( \phi \) is given as

\[
d\phi = \frac{\partial \phi}{\partial t} \, dt + \sum_i \frac{\partial \phi}{\partial x_i} \, dx_i + \frac{1}{2} \sum_{ij} \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) \, dx_i \, dx_j
\]

\[
= \frac{\partial \phi}{\partial t} \, dt + (\nabla \phi)^T \, dx + \frac{1}{2} \, \text{tr} \left\{ \left( \nabla \nabla^T \phi \right) \, dx \, dx^T \right\},
\]

provided that the required partial derivatives exists, where the mixed differentials are combined according to the rules

\[
dx \, dt = 0
\]

\[
dt \, d\beta = 0
\]

\[
d\beta \, d\beta^T = Q \, dt.
\]
Consider the Taylor series expansion:

$$\phi(x + dx, t + dt) = \phi(x, t) + \frac{\partial \phi(x, t)}{\partial t} dt + \sum_i \frac{\partial \phi(x, t)}{\partial x_i} dx_i$$

$$+ \frac{1}{2} \sum_{ij} \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) dx_i dx_j + \ldots$$

To the first order in $dt$ and second order in $dx$ we have

$$d\phi = \phi(x + dx, t + dt) - \phi(x, t)$$

$$\approx \frac{\partial \phi(x, t)}{\partial t} dt + \sum_i \frac{\partial \phi(x, t)}{\partial x_i} dx_i + \frac{1}{2} \sum_{ij} \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) dx_i dx_j.$$ 

In deterministic case we could ignore the second order and higher order terms, because $dx \ d^T x$ would already be of the order $dt^2$.

In the stochastic case we know that $dx \ d^T x$ is potentially of the order $dt$, because $d\beta \ d^T \beta$ is of the same order.
If we apply the Itô formula to $\phi(x) = \frac{1}{2}x^2(t)$, with $x(t) = \beta(t)$, where $\beta(t)$ is a standard Brownian motion, we get

$$d\phi = \beta \, d\beta + \frac{1}{2} \, d\beta^2$$

$$= \beta \, d\beta + \frac{1}{2} \, dt,$$

as expected.
Itô differential of \( \sin(\omega x) \)

Assume that \( x(t) \) is the solution to the scalar SDE:

\[
\frac{dx}{dt} = f(x) \, dt + d\beta,
\]

where \( \beta(t) \) is a Brownian motion with diffusion constant \( q \) and \( \omega > 0 \). The Itô differential of \( \sin(\omega x(t)) \) is then

\[
\frac{d[\sin(x)]}{dt} = \omega \cos(\omega x) \, dx - \frac{1}{2} \omega^2 \sin(\omega x) \, dx^2 \\
= \omega \cos(\omega x) [f(x) \, dt + d\beta] - \frac{1}{2} \omega^2 \sin(\omega x) [f(x) \, dt + d\beta]^2 \\
= \omega \cos(\omega x) [f(x) \, dt + d\beta] - \frac{1}{2} \omega^2 \sin(\omega x) \, q \, dt.
\]
Let’s consider the **linear multidimensional time-varying SDE**

\[ dx = F(t) x \, dt + u(t) \, dt + L(t) \, d\beta \]

Let’s define a (deterministic) **transition matrix** \( \Psi(t, t_0) \) via the properties

\[
\frac{\partial \Psi(\tau, t)}{\partial \tau} = F(\tau) \, \Psi(\tau, t) \\
\frac{\partial \Psi(\tau, t)}{\partial t} = -\Psi(\tau, t) \, F(t) \\
\Psi(\tau, t) = \Psi(\tau, s) \, \Psi(s, t) \\
\Psi(t, \tau) = \Psi^{-1}(\tau, t) \\
\Psi(t, t) = I.
\]
Multiplying the above SDE with the integrating factor $\Psi(t_0, t)$ and rearranging gives

$$\Psi(t_0, t) \, dx - \Psi(t_0, t) F(t) \, x \, dt = \Psi(t_0, t) \, u(t) \, dt + \Psi(t_0, t) \, L(t) \, d\beta.$$ 

Itô formula gives

$$d[\Psi(t_0, t) \, x] = -\Psi(t, t_0) \, C(t) \, x \, dt + \Psi(t, t_0) \, dx.$$ 

Thus the SDE can be rewritten as

$$d[\Psi(t_0, t) \, x] = \Psi(t_0, t) \, u(t) \, dt + \Psi(t_0, t) \, L(t) \, d\beta.$$ 

where the differential is a Itô differential.
Integration (in Itô sense) from $t_0$ to $t$ gives

$$\Psi(t_0, t) x(t) - \Psi(t_0, t_0) x(t_0) = \int_{t_0}^{t} \Psi(t_0, \tau) u(\tau) \, d\tau + \int_{t_0}^{t} \Psi(t_0, \tau) L(\tau) \, d\beta(\tau).$$

Rearranging gives the full solution

$$x(t) = \Psi(t, t_0) x(t_0) + \int_{t_0}^{t} \Psi(t, \tau) u(\tau) \, d\tau + \int_{t_0}^{t} \Psi(t, \tau) L(\tau) \, d\beta(\tau).$$
Let’s consider LTI SDE

\[ d\mathbf{x} = \mathbf{F} \mathbf{x} \, dt + \mathbf{L} \, d\beta. \]

The transition matrix now reduces to the matrix exponential:

\[ \Psi(t, t_0) = \exp(\mathbf{F}(t - t_0)) \]

\[ = \mathbf{I} + \mathbf{F}(t - t_0) + \frac{\mathbf{F}^2(t - t_0)^2}{2!} + \frac{\mathbf{F}^3(t - t_0)^3}{3!} + \ldots \]

The solution simplifies to

\[ \mathbf{x}(t) = \exp(\mathbf{F}(t - t_0)) \mathbf{x}(t_0) + \int_{t_0}^{t} \exp(\mathbf{F}(t - \tau)) \mathbf{L} \, d\beta(\tau). \]

Corresponds to replacing \( \mathbf{w}(\tau) \, d\tau \) with \( d\beta(\tau) \) in the heuristic solution.
Solutions of linear LTI SDEs

Solution of Ornstein–Uhlenbeck equation

The complete solution to the scalar SDE

$$\text{d}x = -\lambda x \, \text{d}t + \text{d}\beta, \quad x(0) = x_0,$$

where $\lambda > 0$ is a given constant and $\beta(t)$ is a Brownian motion is

$$x(t) = \exp(-\lambda t) x_0 + \int_0^t \exp(-\lambda (t - \tau)) \, \text{d}\beta(\tau).$$
Non-linear SDEs

- There is no general solution method for non-linear SDEs

\[ d\mathbf{x} = f(\mathbf{x}, t) \, dt + L(\mathbf{x}, t) \, d\beta. \]

- Sometimes we can use transformation/other methods from deterministic setting and replace chain rule with Itô formula.

- However, we can still use the Euler–Maruyama method presented last time:

\[ \hat{\mathbf{x}}(t_{k+1}) = \hat{\mathbf{x}}(t_k) + f(\hat{\mathbf{x}}(t_k), t_k) \Delta t + L(\hat{\mathbf{x}}(t_k), t_k) \Delta \beta_k, \]

where \( \Delta \beta_k \sim N(0, Q \Delta t) \).

- The method might now look more natural, because \( \Delta \beta_k \) is just a finite increment of Brownian motion.
The existence and uniqueness conditions for SDE solutions can be proved via stochastic Picard iteration:

\[ \varphi_0(t) = x_0 \]

\[ \varphi_{n+1}(t) = x_0 + \int_{t_0}^{t} f(\varphi_n(\tau), \tau) \, d\tau + \int_{t_0}^{t} L(\varphi_n(\tau), \tau) \, d\beta(\tau). \]

The iteration converges and thus the SDE has unique strong solution provided that the following are met:
- Functions \( f \) and \( L \) grow at most linearly in \( x \).
- Functions \( f \) and \( L \) are Lipschitz continuous in \( x \).

A strong solution means a solution \( x \) for a given \( \beta \) — strong uniqueness implies that the whole path is unique.

We can also have a weak solution which is some pair \((\tilde{x}, \tilde{\beta})\) which solves the SDE.

Weak uniqueness means that the distribution is unique.
The symmetrized stochastic integral or the Stratonovich integral can be defined as follows:

\[
\int_{t_0}^{t} L(x(t), t) \circ d\beta(t) = \lim_{n \to \infty} \sum_{k} L(x(t^*_k), t^*_k) [\beta(t_{k+1}) - \beta(t_k)],
\]

where \( t^*_k = (t_k + t_{k+1})/2 \) is the midpoint.

- Recall that in Itô integral we had the starting point \( t^*_k = t_k \).
- Now the Itô formula reduces to the rule from ordinary calculus.
- Stratonovich integral is not a martingale which makes its theoretical analysis harder.
- Smooth approximations to white noise converge to the Stratonovich integral.
Conversion of Stratonovich SDE into Itô SDE

The following SDE in **Stratonovich sense**

\[ \mathrm{d}x = f(x, t) \, \mathrm{d}t + L(x, t) \circ \mathrm{d}\beta, \]

is equivalent to the following SDE in **Itô sense**

\[ \mathrm{d}x = \tilde{f}(x, t) \, \mathrm{d}t + L(x, t) \, \mathrm{d}\beta, \]

where

\[ \tilde{f}_i(x, t) = f_i(x, t) + \frac{1}{2} \sum_{jk} \frac{\partial L_{ij}(x)}{\partial x_k} L_{kj}(x). \]
White noise formulation of SDEs had some problems with chain rule, non-linearities and solution existence.

We can reduce the problem into existence of integral of a stochastic process.

The integral cannot be defined as Riemann, Stieltjes or Lebesgue integral.

It can be defined as an Itô stochastic integral.

Given the definition, we can define Itô stochastic differential equations.

In Itô stochastic calculus, the chain rule is replaced with Itô formula.

For linear SDEs we can obtain a general solution.

Existence and uniqueness can be derived analogously to the deterministic case.

Stratonovich calculus is an alternative stochastic calculus.