Lecture 2: Itô Calculus and Stochastic Differential Equations

Simo Särkkä

Aalto University

November 4, 2014

Contents

SDEs as white noise driven differential equations

- During the last lecture we treated SDEs as white-noise driven differential equations of the form

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, t) + \mathbf{L}(\mathbf{x}, t)\,\mathbf{w}(t),$$

- 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 $\mathbf{w}(t)$.
- So how should we really formulate SDEs?

Equivalent integral equation

- Integrating the differential equation from t_0 to t gives:

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

- 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.

Attempt 1: Riemann integral

- In the Riemannian sense the integral would be defined as

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

  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.

Attempt 2: Stieltjes integral

- Stieltjes integral is more general than the Riemann integral.
- In particular, it allows for discontinuous integrands.
- We can interpret the increment $\mathbf{w}(t)\,\mathrm{d}t$ as increment of another process $\beta(t)$ such that

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

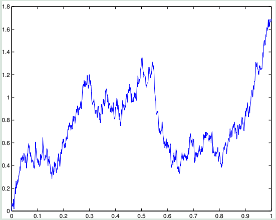
- It turns out that a suitable process for this purpose is the Brownian motion —

Brownian motion

1. Gaussian increments:

$$\Delta\beta_k \sim \mathrm{N}(0, \mathbf{Q}\,\Delta t_k),$$

   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.



- $\mathbf{Q}$ 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 $\mathbf{w}(t) = \mathrm{d}\beta(t)/\mathrm{d}t$.

Attempt 2: Stieltjes integral (cont.)

- Stieltjes integral is defined as a limit of the form

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

  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

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

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

Attempt 4: Itô integral

- 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

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

- 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.

Itô stochastic differential equations

- Consider the white noise driven ODE

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, t) + \mathbf{L}(\mathbf{x}, t)\,\mathbf{w}(t).$$

- This is actually defined as the Itô integral equation

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

  which should be true for arbitrary t_0 and t.
- Settings the limits to t and $t + \mathrm{d}t$, where $\mathrm{d}t$ is "small", we get

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\;\mathrm{d}t + \mathbf{L}(\mathbf{x}, t)\;\mathrm{d}\beta.$$

- This is the canonical form of an Itô SDE.

Connection with white noise driven ODEs

- Let's formally divide by $\mathrm{d}t$, which gives

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, t) + \mathbf{L}(\mathbf{x}, t)\,\frac{\mathrm{d}\boldsymbol{\beta}}{\mathrm{d}t}.$$

- Thus we can interpret $\mathrm{d}\boldsymbol{\beta}/\mathrm{d}t$ as white noise $\mathbf{w}$.
- Note that we cannot define more general equations

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}(t), \mathbf{w}(t), t),$$

  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.

Stochastic integral of Brownian motion

- Consider the stochastic integral
$$\int_0^t \beta(t)\,\mathrm{d}\beta(t)$$
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
$$\begin{aligned}
\int_0^t \beta(t)\,\mathrm{d}\beta(t) &= \lim \sum_k \beta(t_k)[\beta(t_{k+1}) - \beta(t_k)] \\
&= \lim \sum_k \left[-\frac{1}{2}(\beta(t_{k+1}) - \beta(t_k))^2 \right. \\
&\qquad \left. + \frac{1}{2}(\beta^2(t_{k+1}) - \beta^2(t_k)) \right]
\end{aligned}$$

Stochastic integral of Brownian motion (cont.)

- We have

$$\lim \sum_k -\frac{1}{2}(\beta(t_{k+1}) - \beta(t_k))^2 \longrightarrow -\frac{1}{2}t$$

and

$$\lim \sum_k \frac{1}{2}(\beta^2(t_{k+1}) - \beta^2(t_k)) \longrightarrow \frac{1}{2}\beta^2(t).$$

- Thus we get the (slightly) unexpected result

$$\int_0^t \beta(t)\, \mathrm{d}\beta(t) = -\frac{1}{2}t + \frac{1}{2}\beta^2(t).$$

- This is unexpected only if we believe in the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2}x^2(t)\right] = \frac{\mathrm{d}x}{\mathrm{d}t}\, x.$$

- But it is not true for a (Itô) stochastic process $x(t)$!

Itô formula

Itô formula

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

$$
\begin{aligned}
\mathrm{d}\phi &= \frac{\partial \phi}{\partial t}\,\mathrm{d}t + \sum_i \frac{\partial \phi}{\partial x_i}\,\mathrm{d}x_i + \frac{1}{2}\sum_{ij}\left(\frac{\partial^2 \phi}{\partial x_i \partial x_j}\right)\mathrm{d}x_i\,\mathrm{d}x_j \\
&= \frac{\partial \phi}{\partial t}\,\mathrm{d}t + (\nabla \phi)^{\mathsf{T}}\,\mathrm{d}\mathbf{x} + \frac{1}{2}\operatorname{tr}\left\{\left(\nabla\nabla^{\mathsf{T}}\phi\right)\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{x}^{\mathsf{T}}\right\},
\end{aligned}
$$

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

$$
\begin{aligned}
\mathrm{d}\mathbf{x}\,\mathrm{d}t &= 0 \\
\mathrm{d}t\,\mathrm{d}\boldsymbol{\beta} &= 0 \\
\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\boldsymbol{\beta}^{\mathsf{T}} &= \mathbf{Q}\,\mathrm{d}t.
\end{aligned}
$$

Itô formula: derivation

- Consider the Taylor series expansion:

$$\phi(\mathbf{x} + \mathrm{d}\mathbf{x}, t + \mathrm{d}t) = \phi(\mathbf{x}, t) + \frac{\partial \phi(\mathbf{x}, t)}{\partial t}\, \mathrm{d}t + \sum_i \frac{\partial \phi(\mathbf{x}, t)}{\partial x_i}\, \mathrm{d}x_i + \frac{1}{2} \sum_{ij} \left(\frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) \mathrm{d}x_j\, \mathrm{d}x_j + \ldots$$

- To the first order in $\mathrm{d}t$ and second order in $\mathrm{d}\mathbf{x}$ we have

$$\begin{aligned} \mathrm{d}\phi &= \phi(\mathbf{x} + \mathrm{d}\mathbf{x}, t + \mathrm{d}t) - \phi(\mathbf{x}, t) \\ &\approx \frac{\partial \phi(\mathbf{x}, t)}{\partial t}\, \mathrm{d}t + \sum_i \frac{\partial \phi(x, t)}{\partial x_i}\, \mathrm{d}x_i + \frac{1}{2} \sum_{ij} \left(\frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) \mathrm{d}x_i\, \mathrm{d}x_j. \end{aligned}$$

- In deterministic case we could ignore the second order and higher order terms, because $\mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{x}^\mathsf{T}$ would already be of the order $\mathrm{d}t^2$.
- In the stochastic case we know that $\mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{x}^\mathsf{T}$ is potentially of the order $\mathrm{d}t$, because $\mathrm{d}\boldsymbol{\beta}\, \mathrm{d}\boldsymbol{\beta}^\mathsf{T}$ is of the same order.

Itô formula: example 1

Itô differential of $\beta^2(t)/2$

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

$$\begin{aligned} \mathrm{d}\phi &= \beta \, \mathrm{d}\beta + \frac{1}{2} \, \mathrm{d}\beta^2 \\ &= \beta \, \mathrm{d}\beta + \frac{1}{2} \, \mathrm{d}t, \end{aligned}$$

as expected.

Itô formula: example 2

Itô differential of $\sin(\omega\, x)$

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

$$\mathrm{d}x = f(x)\,\mathrm{d}t + \mathrm{d}\beta,$$

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

$$\begin{aligned}
\mathrm{d}[\sin(x)] &= \omega\,\cos(\omega\, x)\,\mathrm{d}x - \frac{1}{2}\omega^2\,\sin(\omega\, x)\,\mathrm{d}x^2 \\
&= \omega\,\cos(\omega\, x)\,[f(x)\,\mathrm{d}t + \mathrm{d}\beta] - \frac{1}{2}\omega^2\,\sin(\omega\, x)\,[f(x)\,\mathrm{d}t + \mathrm{d}\beta]^2 \\
&= \omega\,\cos(\omega\, x)\,[f(x)\,\mathrm{d}t + \mathrm{d}\beta] - \frac{1}{2}\omega^2\,\sin(\omega\, x)\,q\,\mathrm{d}t.
\end{aligned}$$

Solutions of linear SDEs

- Let's consider the linear multidimensional time-varying SDE

$$\mathrm{d}\mathbf{x} = \mathbf{F}(t)\,\mathbf{x}\;\mathrm{d}t + \mathbf{u}(t)\;\mathrm{d}t + \mathbf{L}(t)\;\mathrm{d}\boldsymbol{\beta}$$

- Let's define a (deterministic) transition matrix $\boldsymbol{\Psi}(t, t_0)$ via the properties

$$\begin{aligned}
\partial\boldsymbol{\Psi}(\tau, t)/\partial\tau &= \mathbf{F}(\tau)\,\boldsymbol{\Psi}(\tau, t) \\
\partial\boldsymbol{\Psi}(\tau, t)/\partial t &= -\boldsymbol{\Psi}(\tau, t)\,\mathbf{F}(t) \\
\boldsymbol{\Psi}(\tau, t) &= \boldsymbol{\Psi}(\tau, s)\,\boldsymbol{\Psi}(s, t) \\
\boldsymbol{\Psi}(t, \tau) &= \boldsymbol{\Psi}^{-1}(\tau, t) \\
\boldsymbol{\Psi}(t, t) &= \mathbf{I}.
\end{aligned}$$

Solutions of linear SDEs (cont.)

- Multiplying the above SDE with the integrating factor $\mathbf{\Psi}(t_0, t)$ and rearranging gives

$$\mathbf{\Psi}(t_0, t)\, \mathrm{d}\mathbf{x} - \mathbf{\Psi}(t_0, t)\, \mathbf{F}(t)\, \mathbf{x}\, \mathrm{d}t = \mathbf{\Psi}(t_0, t)\, \mathbf{u}(t)\, \mathrm{d}t + \mathbf{\Psi}(t_0, t)\, \mathbf{L}(t)\, \mathrm{d}\beta.$$

- Itô formula gives

$$\mathrm{d}[\mathbf{\Psi}(t_0, t)\, \mathbf{x}] = -\mathbf{\Psi}(t, t_0)\, C(t)\, \mathbf{x}\, \mathrm{d}t + \mathbf{\Psi}(t, t_0)\, \mathrm{d}\mathbf{x}.$$

- Thus the SDE can be rewritten as

$$\mathrm{d}[\mathbf{\Psi}(t_0, t)\, \mathbf{x}] = \mathbf{\Psi}(t_0, t)\, \mathbf{u}(t)\, \mathrm{d}t + \mathbf{\Psi}(t_0, t)\, \mathbf{L}(t)\, \mathrm{d}\beta.$$

where the differential is a Itô differential.

Solutions of linear SDEs (cont.)

- Integration (in Itô sense) from t_0 to t gives

$$\begin{aligned}
&\boldsymbol{\Psi}(t_0, t)\,\mathbf{x}(t) - \boldsymbol{\Psi}(t_0, t_0)\,\mathbf{x}(t_0) \\
&\qquad = \int_{t_0}^{t} \boldsymbol{\Psi}(t_0, \tau)\,\mathbf{u}(\tau)\;\mathrm{d}\tau + \int_{t_0}^{t} \boldsymbol{\Psi}(t_0, \tau)\,\mathbf{L}(\tau)\;\mathrm{d}\boldsymbol{\beta}(\tau).
\end{aligned}$$

- Rearranging gives the full solution

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

Solutions of linear LTI SDEs

- Let's consider LTI SDE

$$\mathrm{d}\mathbf{x} = \mathbf{F}\,\mathbf{x}\;\mathrm{d}t + \mathbf{L}\;\mathrm{d}\boldsymbol{\beta}.$$

- The transition matrix now reduces to the matrix exponential:

$$\begin{aligned}\boldsymbol{\Psi}(t, t_0) &= \exp\left(\mathbf{F}\,(t - t_0)\right) \\ &= \mathbf{I} + \mathbf{F}\,(t - t_0) + \frac{\mathbf{F}^2\,(t - t_0)^2}{2!} + \frac{\mathbf{F}^3\,(t - t_0)^3}{3!} + \ldots\end{aligned}$$

- The solution simplifies to

$$\mathbf{x}(t) = \exp\left(\mathbf{F}\,(t - t_0)\right)\,\mathbf{x}(t_0) + \int_{t_0}^{t} \exp\left(\mathbf{F}\,(t - \tau)\right)\,\mathbf{L}\;\mathrm{d}\boldsymbol{\beta}(\tau).$$

- Corresponds to replacing $\mathbf{w}(\tau)\;\mathrm{d}\tau$ with $\mathrm{d}\boldsymbol{\beta}(\tau)$ in the heuristic solution.

Solutions of linear LTI SDEs

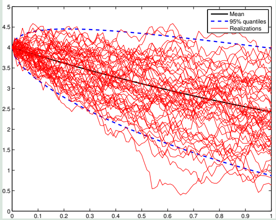
Solution of Ornstein–Uhlenbeck equation

The complete solution to the scalar SDE

$$\mathrm{d}x = -\lambda\, x \; \mathrm{d}t + \mathrm{d}\beta, \qquad x(0) = x_0,$$

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

$$x(t) = \exp(-\lambda\, t)\, x_0 + \int_0^t \exp(-\lambda\,(t-\tau)) \; \mathrm{d}\beta(\tau).$$



Non-linear SDEs

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

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\,\mathrm{d}t + \mathbf{L}(\mathbf{x}, t)\,\mathrm{d}\boldsymbol{\beta}.$$

- 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:

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

  where $\Delta\beta_k \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}\,\Delta t)$.
- The method might now look more natural, because $\Delta\beta_k$ is just a finite increment of Brownian motion.

Existence and uniqueness of solutions

- The existence and uniqueness conditions for SDE solutions can be proved via stochastic Picard iteration:

$$
\begin{aligned}
\varphi_0(t) &= \mathbf{x}_0 \\
\varphi_{n+1}(t) &= \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(\varphi_n(\tau), \tau)\, \mathrm{d}\tau + \int_{t_0}^{t} \mathbf{L}(\varphi_n(\tau), \tau)\, \mathrm{d}\beta(\tau).
\end{aligned}
$$

- The iteration converges and thus the SDE has unique strong solution provided that the following are met:
  - Functions $\mathbf{f}$ and $\mathbf{L}$ grow at most linearly in $\mathbf{x}$.
  - Functions $\mathbf{f}$ and $\mathbf{L}$ are Lipschitz continuous in $\mathbf{x}$.
- A strong solution means a solution $\mathbf{x}$ for a given β — strong uniqueness implies that the whole path is unique.
- We can also have a weak solution which is some pair $(\tilde{\mathbf{x}}, \tilde{\beta})$ which solves the SDE.
- Weak uniqueness means that the distribution is unique.

Stratonovich calculus

- The symmetrized stochastic integral or the Stratonovich integral can be defined as follows:

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

  where $t_k^* = (t_k + t_k)/2$ 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.

Stratonovich calculus (cont.)

Conversion of Stratonovich SDE into Itô SDE

The following SDE in Stratonovich sense

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\, \mathrm{d}t + \mathbf{L}(\mathbf{x}, t) \circ \mathrm{d}\boldsymbol{\beta},$$

is equivalent to the following SDE in Itô sense

$$\mathrm{d}\mathbf{x} = \tilde{\mathbf{f}}(\mathbf{x}, t)\, \mathrm{d}t + \mathbf{L}(\mathbf{x}, t)\, \mathrm{d}\boldsymbol{\beta},$$

where

$$\tilde{f}_i(\mathbf{x}, t) = f_i(\mathbf{x}, t) + \frac{1}{2} \sum_{jk} \frac{\partial L_{ij}(\mathbf{x})}{\partial x_k} L_{kj}(\mathbf{x}).$$

Summary

- 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 defition, 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.