Example: the Ornstein-Uhlenbeck process

$$dX_t = -\beta (X_t - \alpha) dt + \sigma dW_t$$

where $\beta > 0$, $\alpha \in \mathbb{R}$, $\sigma > 0$ and $X_0 = x_0$.

Solution:

$$X_t = \alpha + (x_0 - \alpha) e^{-\beta t} + \sigma \int_0^t e^{-\beta (t-s)} dW_s$$

Note that this is a sum of deterministic terms and an integral of a deterministic function with respect to a Wiener process with normally distributed increments. The distribution is thus normal.

The conditional expectation is

$$E[X_t|X_0 = x_0] = E\left[\alpha + (x_0 - \alpha) e^{-\beta t} + \sigma \int_0^t e^{-\beta (t-s)} dW_s\right]$$

The conditional variance is

$$\text{Var}[X_t|X_0 = x_0] = E\left[(\sigma \int_0^t e^{-\beta (t-s)} dW_s)^2\right]$$

Use Ito’s isometry to obtain

$$\text{Var}[X_t|X_0 = x_0] = \sigma^2 E\left[\int_0^t e^{-2\beta (t-s)} ds\right] = \frac{\sigma^2}{2\beta} (1 - e^{-2\beta t})$$

Thus $(X_t|X_0 = x_0) \sim N(\alpha + (x_0 - \alpha) e^{-\beta t}, \frac{\sigma^2}{2\beta} (1 - e^{-2\beta t}))$.

Asymptotically $X_t \sim N(\alpha, \frac{\sigma^2}{2\beta})$ (or always if $X_0 \sim N(\alpha, \frac{\sigma^2}{2\beta})$).

Parameter interpretation in the OU-process

Example: population growth model

Also called the geometric Brownian motion

$$dN_t = a N_t dt + \sigma N_t dW_t$$

The Itô solution:

$$N_t = N_0 \exp \left\{ \left( a - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right\}$$

The Stratonovich solution:

$$N_t = N_0 \exp \{ at + \sigma W_t \}$$

\[ \beta = 0.01, \sigma = 1 \] \[ \beta = 0.1, \sigma = 1 \] \[ \beta = 0.01, \sigma = 0.5 \]

$\beta$ : how "strongly" the system reacts to perturbations (the "decay-rate" or "growth-rate")

$\sigma^2$ : the variation or the size of the noise.

$\alpha$ : the asymptotic mean
Qualitative behavior of the Itô solution

\[ N_t = N_0 \exp \left\{ (a - \frac{1}{2} \sigma^2) t + \sigma W_t \right\} \]

- If \( a > \frac{1}{2} \sigma^2 \) then \( N_t \to \infty \) when \( t \to \infty \), a.s.
- If \( a < \frac{1}{2} \sigma^2 \) then \( N_t \to 0 \) when \( t \to \infty \), a.s.
- If \( a = \frac{1}{2} \sigma^2 \) then \( N_t \) will fluctuate between arbitrary large and arbitrary small values as \( t \to \infty \), a.s.

Whereas for the Stratonovich solution we have

\[ N_t = N_0 \exp \{ at + \sigma W_t \} \]

- If \( a > 0 \) then \( N_t \to \infty \) when \( t \to \infty \), a.s.
- If \( a < 0 \) then \( N_t \to 0 \) when \( t \to \infty \), a.s.

... just like in the deterministic case.

Apparently it makes a huge difference which interpretation we choose.

Note though:
If \( W_t \) is independent of \( N_t \) we would expect that

\[ E[N_t] = E[N_0] e^{at} \]

i.e. the same as when there is no noise in \( a_t \). Let us check:

Let

\[ Y_t = e^{\sigma W_t} \]

and apply Ito’s formula

\[ dY_t = \frac{1}{2} \sigma^2 e^{\sigma W_t} dt + \sigma e^{\sigma W_t} dW_t \]

i.e.

\[ Y_t = Y_0 + \frac{1}{2} \sigma^2 \int_0^t e^{\sigma W_s} ds + \sigma \int_0^t e^{\sigma W_s} dW_s \]

Thus

\[ E[Y_t] = E[Y_0] + \frac{1}{2} \sigma^2 \int_0^t E[e^{\sigma W_s}] ds + \sigma E \left[ \int_0^t e^{\sigma W_s} dW_s \right] \]

We obtain the differential equation for \( E[Y_t] \):

\[ \frac{d}{dt} E[Y_t] = \frac{1}{2} \sigma^2 E[Y_t] \quad ; \quad E[Y_0] = 1 \]

so that

\[ E[Y_t] = E[e^{\sigma W_t}] = e^{\sigma^2 t/2} \]
Finally
\[ E[N_t] = E \left[ N_0 \exp \left\{ (a - \frac{1}{2} \sigma^2) t + \sigma W_t \right\} \right] \]
\[ = E[N_0] \exp \left\{ (a - \frac{1}{2} \sigma^2) t \right\} E \left[ \exp \{ \sigma W_t \} \right] \]
\[ = E[N_0] \exp \left\{ (a - \frac{1}{2} \sigma^2) t \right\} \exp \left\{ \frac{1}{2} \sigma^2 t \right\} \]
\[ = E[N_0] e^{at} \]

exactly as we expected! However, for the Stratonovich solution, the same calculations give
\[ E[N_t] = E[N_0] e^{(\tilde{a} + \sigma^2/2)t} \]

where \( \tilde{a} \) is seen to be a different parameter from \( a \).

**Examples from ODEs**

The equation
\[ \frac{dx_t}{dt} = x_t^2 , \quad x_0 = 1 \]
does not satisfy the linear growth condition. It has the unique solution
\[ x_t = \frac{1}{1-t} ; \quad 0 \leq t < 1 \]
but no global solution (defined for all \( t \)).

The linear growth condition ensures that the solution \( X_t \) does not explode, i.e. \( |X_t| \) does not tend to \( \infty \) in finite time.

**An existence and uniqueness result**

Linear growth and local Lipschitz conditions:
For each \( N \in \mathbb{N} \) there exists a constant \( K_N \) such that
\[ |b(x, t)| + |\sigma(x, t)| \leq K_N (1 + |x|) \]
and
\[ |b(x, t) - b(y, t)| + |\sigma(x, t) - \sigma(y, t)| \leq K_N (x - y) \]
for all \( t \in [0, \mathbb{N}] \) and for all \( x \), where \( |\sigma|^2 = \text{tr} \sigma \sigma^T \).

Then
\[ dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = U, \quad U \perp \{ W_t \}_{t \geq 0} \]
has a unique \( t \)-continuous solution \( X_t \).

**Examples from ODEs**

The equation
\[ \frac{dx_t}{dt} = 3x_t^{2/3} , \quad x_0 = 0 \]
does not satisfy the Lipschitz condition at \( x = 0 \). It has more than one solution:
\[ x_t = \begin{cases} 0 & \text{for } t \leq a \\ (t - a)^3 & \text{for } t > a \end{cases} \]
for any \( a > 0 \).

The Lipschitz condition ensures that a solution \( X_t \) is unique: If \( X_t^{(1)} \) and \( X_t^{(2)} \) are two \( t \)-continuous processes satisfying the conditions then
\[ X_t^{(1)} = X_t^{(2)} \quad \text{for all } t \leq T, \text{ a.s.} \]
The solution $X_t$ where drift and diffusion coefficients fulfill the growth and Lipschitz conditions is a strong solution:

- the version of $W_t$ is given in advance
- The solution $X_t$ is $\mathcal{F}_t^U$-adapted

$\mathcal{F}_t^U$ is the filtration generated by the initial $U$ and $W_s, s \leq t$.

If only $b(\cdot)$ and $\sigma(\cdot)$ are given, and we ask for a pair of processes $(\tilde{X}_t, \tilde{W}_t)$ then the solution is called a weak solution.

Strong uniqueness means pathwise uniqueness, weak uniqueness means that any two solutions are identical in law, i.e. have the same finite dimensional distributions.

A strong solution is also a weak solution.

There are SDEs with no strong solution, but still a unique weak solution.

Remark: Note that the above conditions are sufficient conditions, not necessary conditions.

**Sufficient condition for the existence of a unique weak solution**

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = U$$

- $a(x) = \sigma(x)\sigma(x)^T$
- $a(x)$ continuous
- $a(x)$ strictly positive definite for all $x$
- There exists a constant $K$ such that
  $$|a_{ij}(x)| \leq K(1 + |x|^2)$$
  $$|b_i(x)| \leq K(1 + |x|)$$

for all $i, j = 1, \cdots, d$ and $x$.

The solution is a strong Markov process.

**Transition densities:**

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

$$y \mapsto p(t, x, y)$$

Conditional density of $X_t$ given $X_0 = x$; also conditional density of $X_{t+s}$ given $X_s = x$.

Data: $X_{t_1}, \cdots, X_{t_n}, \quad t_1 < \cdots < t_n$.

Likelihood function:

$$L(\theta) = \prod_{i=1}^{n} p(t_i - t_{i-1}, X_{t_{i-1}}, X_{t_i}; \theta)$$
• Chapman-Kolmogorov equation:
  \[ p(t + s, x, y) = \int p(t, z, y)p(s, x, z)dz \]

• Kolmogorov’s backward equation:
  \[ \frac{1}{2} \sigma^2(x) \frac{\partial^2 p}{\partial x^2} + b(x) \frac{\partial p}{\partial x} = \frac{\partial p}{\partial t} \]
  with the initial condition
  \[ p(t, x, y) \to \delta_x \text{ as } t \to 0. \]
  \( \delta_x \) is the Dirac measure at \( x \).

• Kolmogorov’s forward equation:
  \[ \frac{1}{2} \frac{\partial^2}{\partial y^2} \left[ \sigma(y)^2 p \right] - \frac{\partial}{\partial y} \left[ b(y)p \right] = \frac{\partial p}{\partial t} \]
  (Fokker-Planck equation)

Examples:

• Ornstein-Uhlenbeck
  \[ dX_t = -\beta(X_t - \alpha)dt + \sigma dW_t \]
  Remember that
  \( (X_t | X_0 = x_0) \sim N(\alpha + (x - \alpha)e^{-\beta t}, \sigma^2(1 - e^{-2\beta t})/2\beta) \).

• Cox-Ingersoll-Ross
  \[ dX_t = \theta_X^{-1}(X_t) dt + \sigma_X dW_t \]
  \( \theta > 0, \alpha > 0, \sigma > 0. \)

\[ p(t, x, y) = \beta \left( \frac{y}{x} \right)^{\frac{1}{2}\nu} \exp \left[ \frac{\beta}{\nu} \left( x + y \right) \right] \]
\[ \times \exp \left[ \frac{-\beta(x + y)}{\exp(\beta t) - 1} \right] \left( \begin{array}{c} \frac{\beta \sqrt{xy}}{\sinh(\beta t)} \\ \sinh(\beta t) \end{array} \right), \]
where \( \beta = 2\theta \sigma^{-2} \) and \( \nu = \beta \alpha - 1. \)

\( I_{\nu} \) is a modified Bessel function with index \( \nu \).

The transition density is a non-central \( \chi^2 \)-distribution.

Examples:

• Radial Ornstein-Uhlenbeck
  \[ dX_t = \theta X_t^{-1} \left( X_t \right) dt + dW_t, \]
  \( \theta > 0. \)

\[ p(t, x, y) = \frac{(y/x)^{\theta \nu} \sqrt{xy} \exp(-y^2 + (\theta + \frac{1}{2})t)}{\sinh(t)} \]
\[ \times \exp \left[ \frac{-\left( x^2 + y^2 \right)}{\exp(2t) - 1} \right] I_{\theta - \frac{1}{2}} \left( \frac{xy}{\sinh(t)} \right), \]
\( I_{\nu} \) is a modified Bessel function with index \( \nu \).
Taylor expansions

Review of deterministic expansions:
Consider
\[ \frac{d}{dt} x_t = a(x_t) \]
with initial value \( x_{t_0} \) for \( t \in [t_0, T] \), and \( a(\cdot) \) is sufficiently smooth.
We can write
\[ x_t = x_{t_0} + \int_{t_0}^{T} a(x_s) ds \]

Let \( f : \mathbb{R} \to \mathbb{R} \) be a continuously differentiable function. By the chain rule
\[ \frac{d}{dt} f(x_t) = a(x_t) f'(x_t) \]
Define the operator
\[ Lf = a f' \]
where ' denotes differentiation with respect to \( x \).
Express the above equation for \( f(x) \) in integral form
\[ f(x_t) = f(x_{t_0}) + \int_{t_0}^{t} Lf(x_s) ds \]
Note that if \( f(x) = x \) then \( Lf = a, L^2 f = La \) and
\[ x_t = x_{t_0} + \int_{t_0}^{t} a(x_s) ds \]

Applying again to the function \( f = La \) to obtain
\[ x_t = x_{t_0} + a(x_{t_0}) \int_{t_0}^{t} ds + \int_{t_0}^{t} \int_{t_0}^{s} La(x_z) dz ds \]
\[ = x_{t_0} + a(x_{t_0}) \int_{t_0}^{t} ds + \int_{t_0}^{t} \int_{t_0}^{s} a(x_z) dz ds \]
\[ = x_{t_0} + a(x_{t_0})(t - t_0) + R_1 \]
which is the simplest non-trivial expansion for \( x_t \).
For a general $r + 1$ times continuously differentiable function $f$ we obtain the classical Taylor formula in integral form

$$f(x_t) = f(x_{t_0}) + \sum_{l=1}^{r} \frac{(t-t_0)^l}{l!} L^l f(x_{t_0}) + \int_{t_0}^{t} \cdots \int_{t_0}^{s_r} L^{r+1} f(x_{s_1}) ds_1 \cdots ds_{r+1}$$

For $f$ twice continuously differentiable, Ito’s formula yields

$$f(X_t) = f(X_{t_0}) + \int_{t_0}^{t} \left( b(X_s)f'(X_s) + \frac{1}{2} \sigma^2(X_s)f''(X_s) \right) ds + \int_{t_0}^{t} \sigma(X_s) dW_s$$

Note that for $f(x) = x$ we have $L^0 f = b$ and $L^1 f = \sigma$, and the original equation for $X_t$ is obtained

$$X_t = X_{t_0} + \int_{t_0}^{t} b(X_s) ds + \int_{t_0}^{t} \sigma(X_s) dW_s$$

The Ito-Taylor expansion

Iterated application of Ito’s formula!

Consider

$$X_t = X_{t_0} + \int_{t_0}^{t} b(X_s) ds + \int_{t_0}^{t} \sigma(X_s) dW_s$$

We introduce the operators

$$L^0 f = b f' + \frac{1}{2} \sigma^2 f''$$
$$L^1 f = \sigma f'$$

Like in the deterministic expansions, we apply Ito’s formula to the functions $f = b$ and $f = \sigma$ and obtain

$$X_t = X_{t_0} + \int_{t_0}^{t} \left( b(X_{t_0}) + \int_{t_0}^{s} L^0 b(X_z) dz + \int_{t_0}^{s} L^1 b'(X_z) dW_z \right) ds$$
$$+ \int_{t_0}^{t} \left( \sigma(X_{t_0}) + \int_{t_0}^{s} L^0 \sigma(X_z) dz + \int_{t_0}^{s} L^1 \sigma'(X_z) dW_z \right) dW_s$$

$$= X_{t_0} + b(X_{t_0}) \int_{t_0}^{t} ds + \sigma(X_{t_0}) \int_{t_0}^{t} dW_s + R$$
$$= X_{t_0} + b(X_{t_0})(t - t_0) + \sigma(X_{t_0})(W_t - W_{t_0}) + R$$

This is the simplest non-trivial Ito-Taylor expansion of $X_t$ involving single integrals with respect to both time and the Wiener process. The remainder contains multiple integrals with respect to both.
In the previous expansion we had
\[
R = \int_{t_0}^{t} \int_{t_0}^{s} L^0 b(X_z) dz \, ds + \int_{t_0}^{t} \int_{t_0}^{s} L^1 b(X_z) dW_z \, ds \\
+ \int_{t_0}^{t} \int_{t_0}^{s} L^0 \sigma(X_z) dz \, dW_s + \int_{t_0}^{t} \int_{t_0}^{s} L^1 \sigma(X_z) dW_z \, dW_s
\]
Note that \(dz \, ds, dW_z \, ds\) and \(dz \, dW_s\) “scales like 0”, whereas \(dW_z \, dW_s\) scales like \(dt\), comparable to the terms in the simplest expansion with two single integrals.
We therefore continue the expansion by applying the Ito formula to \(f = L^1 \sigma\).

The next Ito-Taylor expansion becomes
\[
X_t = X_{t_0} + b(X_{t_0}) \int_{t_0}^{t} ds + \sigma(X_{t_0}) \int_{t_0}^{t} dW_s + L^1 \sigma(X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW_z \, dW_s + \tilde{R}
\]
with remainder
\[
\tilde{R} = \int_{t_0}^{t} \int_{t_0}^{s} L^0 b(X_z) dz \, ds + \int_{t_0}^{t} \int_{t_0}^{s} L^1 b(X_z) dW_z \, ds \\
+ \int_{t_0}^{t} \int_{t_0}^{s} L^0 \sigma(X_z) dz \, dW_s + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{\hat{s}} L^1 \sigma(X_u) du \, dW_z \, dW_s \\
+ \int_{t_0}^{t} \int_{t_0}^{s} L^1 \sigma(X_u) dW_u \, dW_z \, dW_s
\]

Consider the Itô stochastic differential equation
\[
dX_t = b(X_t) \, dt + \sigma(X_t) \, dW_t
\]
and a time discretization
\[
0 = t_0 < t_1 < \cdots < t_j < \cdots < t_N = T
\]
Put
\[
\Delta_j = t_{j+1} - t_j \\
\Delta W_j = W_{t_{j+1}} - W_{t_j}
\]
Then
\[
\Delta W_j \sim N(0, \Delta_j)
\]

**Numeric solutions**

When no explicit solution is available we can approximate different characteristics of the process by simulation. (Realizations, moments, qualitative behavior etc). We use the approximations from the Ito-Taylor expansions.

- Different schemes (Euler, Milstein, higher order schemes...)
- Rate of convergence (Weak and strong)
The Euler-Maruyama scheme

We approximate the process $X_t$ given by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t; \quad X(0) = x_0$$

at the discrete time-points $t_j, 1 \leq j \leq N$ by

$$Y_{t_{j+1}} = Y_{t_j} + b(Y_{t_j})\Delta_j + \sigma(Y_{t_j})\Delta W_j; \quad Y_{t_0} = x_0$$

where $\Delta W_j = \sqrt{\Delta_j} \cdot Z_j$, with $Z_j \sim N(0,1)$ for all $j$.

The Milstein scheme

We can even do better!

We approximate $X_t$ by

$$Y_{t_{j+1}} = Y_{t_j} + b(Y_{t_j})\Delta_j + \sigma(Y_{t_j})\Delta W_j + \frac{1}{2}\sigma(Y_{t_j})\sigma'(Y_{t_j})\{(\Delta W_j)^2 - \Delta_j\}$$

(now Milstein...)

where the prime $'$ denotes the derivative.
The Milstein scheme

The Milstein scheme converges in the strong sense with order 1:

\[ E(|X_T - Y_{t_N}|) \leq K\delta \]

We could regard the Milstein scheme as the proper generalization of the deterministic Euler-scheme.

If \( b(X_t) \) does not depend on \( X_t \) the Euler-Maruyama and the Milstein scheme coincide.

Multi-dimensional diffusions:

Euler scheme: Similar.

Milstein scheme: Involves multiple Wiener integrals.

\[ \int_{n\delta}^{(n+1)\delta} \int_{n\delta}^{s} dW_u^{(1)} dW_s^{(2)} \]

Simulation schemes are based on stochastic Ito-Taylor expansions that are formally obtained by iterated use of Ito’s formula.

Kloeden and Platen (1992)