5 Stochastic differential equations

The stochastic differential equation (SDE) we shall discuss has the form

$$dX(t) = \alpha(X(t), t)dt + \beta(t, X(t))dB(t), \quad X(0) = x_0, \tag{28}$$

where the coefficients $\alpha : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}, \beta : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ are two given functions, x_0 is a given initial value, and B is a Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A solution to the SDE is a random process X which satisfies (28) for t assuming values in a given interval $[0, T)$ or $[0, \infty)$. The form (28) is a shorthand notation for the equation

$$X(t) = x_0 + \int_0^t \alpha(u, X(u))du + \int_0^t \beta(u, X(u))dB(u),$$

which involves the Ito integral.

We stress that the source of randomness of the solution if the BM B, but not the coefficients. A consequence of this is the Markov property:

Theorem 5.1. *The solution to* (28) *is a Markov process.*

While the proof is technical, the intuitive content of the theorem must be clear. Given $X(t) = x$, the increment of X over a small time interval is determined by $\alpha(t, x), \beta(t, x)$ and the Brownian increment which does not depend on the history prior to time t.

5.1 Examples of SDE's

Like for ordinary DE's, solution to SDE's is rarely available as a closed analytic expression via α, β and B. We shall consider some special cases, where a solution can be found explicitly by simple manipulations with the Ito formula.

Example 5.2. Central for the Black-Scholes theory is the SDE

$$dX(t) = \mu X(t)dt + \sigma X(t)dB(t), \quad X_0 = x_0, \tag{29}$$

with $x_0 > 0$. Although we know that the solution is a geometric BM, we will employ this instance to introduce a new technique.

Let us try to find solution to (29) in the form $X(t) = f(t, B(t))$ with some to be determined fuinction f. Applying Ito's formula

$$df(t, B(t)) = \left(f_t(t, B(t)) + \frac{1}{2} f_{xx}(t, B(t)) \right) dt + f_x(t, B(t))dB(t).$$

On the other hand, substituting $X(t) = f(t, B(t))$ in (29) we obtain

$$dX(t) = \mu f(t, B(t))dt + \sigma f(t, B(t))dB(t),$$

whence

$$\left(f_t(t, B(t)) + \frac{1}{2} f_{xx}(t, B(t)) \right) dt + f_x(t, B(t))dB(t) = \mu f(t, B(t)dt + \sigma f(t, B(t))dB(t),$$

or in simplified notation

$$\left(f_t + \frac{1}{2} f_{xx}\right) dt + f_x dB = \mu f dt + \sigma f dB.$$

To find f we need to match the 'dt' and 'dB' coefficients:

$$\mu f = f_t + \frac{1}{2} f_{xx}, \qquad \sigma f = f_x.$$

Keeping t as parameter, the second equation $f_x/f = \sigma$ is an ODE in the variable x. Since $f_x/f = \partial \log f / \partial x$, the general solution is

$$f(t, x) = e^{\sigma x + g(t)},$$

where $g(t)$ is an arbitrary 'constant' of integration depending on t. Substituting this in the first equation $\mu f = f_t + \frac{1}{2} f_{xx}$ and cancelling common factors results in $g'(t) = \mu - \frac{1}{2}\sigma^2$, therefore $g(t) = (\mu - \frac{1}{2}\sigma^2)t + c$. To meet $X(0) = x_0$ we should take the integration constant $c = \log x_0$. It follows that the solution is the gBM

$$X(t) = x_0 \exp\left((\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)\right),$$

as anticipated.

Basics of Gaussian processes We call process $(Y(t), t \geq 0)$ *Gaussian* if the joint distribution of $Y(t_1), \ldots, Y(t_k)$ is a multivariate normal distribution for any choice of k and $t_1 < \cdots < t_k$. We know that a multivariate normal distribution (recall Section 1) is characterised by the mean vector and the covariance matrix. Likewise, the finite-dimensional distributions of a Gaussian process are uniquely determined by the mean function $m(t) = \mathbb{E}Y(t)$ and the covariance function $r(s,t) = \mathrm{Cov}(Y(s), Y(t))$. In particular, the Brownian motion can be characterised as the Gaussian process with $m(t) = 0, r(s,t) = s \wedge t$ (where $s, t \geq 0$) [3]. The processes with $m(t) = 0$ for all t are called centered.

Let ψ be a nonrandom function. The stochastic integral $Y(t) = \int_0^t \psi(u) dB(u)$, considered as a random function of the upper limit $t \geq 0$, is a Gaussian process. To see this, choose a partition $0 = t_0 < t_1 < \cdots < t_k$. The increment $Y(t_{i+1}) - Y(t_i) = \int_{t_i}^{t_{i+1}} \psi(u) dB(u)$ is determined by $(B(t) - B(t_i),\ t \in [t_i, t_{i+1}])$; this is obvious if ψ is a piecewise-constant nonrandom function, and follows by an approximation of ψ in general. By the independence of increments of the BM, the 'pieces' $(B(t) - B(t_i),\ t \in [t_i, t_{i+1}])$ are idependent for distinct i, hence the increments $Y(t_{i+1}) - Y(t_i)$ are also independent. Again by the definition of the Ito integral, the increments $Y(t_{i+1}) - Y(t_i)$ have normal distribution (this is first shown for piecewise-constant ψ, then extended to the general ψ).

[3] The independence of increments and their normal distribution follow easily, but the continuity of paths needs further comment. We say that two processes with the same finite-dimensional distributions are *versions* of the same process. For the Gaussian process with $m(t) = 0, r(s,t) = s \wedge t$ there exists a version with continuous paths; this continuous version is the standard Brownian motion.

Together with the independence, the latter implies that the increments $Y(t_{i+1}) - Y(t_i)$, $i < k$, are jointly normal, but then also $Y(t_1), \dots, Y(t_k)$ are jointly normal, as linear combinations of the increments. Since $Y(t) = \int_0^t \psi(u)dB(u)$ is a martingale, the process is centered. To compute the covariance function we assume first that $s < t$, then

$$\begin{aligned}
\mathrm{Cov}(Y(s), Y(t)) = \mathrm{Cov}\left(\int_0^s \psi(u)dB(u), \int_0^t \psi(u)dB(u)\right) &= \\
\mathbb{E}\left(\int_0^s \psi(u)dB(u) \int_0^t \psi(u)dB(u)\right) &= \\
\mathbb{E}\left(\int_0^s \psi(u)dB(u)\left[\int_0^s \psi(u)dB(u) + \int_s^t \psi(u)dB(u)\right]\right) &= \\
\mathrm{Var}\left(\int_0^s \psi(u)dB(u)\right) = \int_0^s \psi^2(u)du&,
\end{aligned}$$

where we used the Ito isometry and the relation called 'the orthogonality of increments':

$$\mathbb{E}\left(\int_0^s \psi(u)dB(u) \int_s^t \psi(u)dB(u)\right) = \mathbb{E}(Y(s)(Y(t)-Y(s)) = \mathbb{E}(Y(s))\mathbb{E}(Y(t)-Y(s)) = 0.$$

For arbitrary $s, t \geq 0$ the covariance function is

$$\mathrm{Cov}(Y(s), Y(t)) = \mathrm{Var}(Y(s \wedge t)) = \int_0^{s \wedge t} \psi^2(u)du.$$

Finally, note that if $(X(t), t \geq 0)$ is a Gaussian process, then also $(g(t)X(t), t \geq 0)$ is Gaussian for arbitrary nonrandom function g (left as an exercise).

Example 5.3. *Ornstein-Uhlenbeck process* is defined as the solution to the SDE

$$dX(t) = -\alpha X(t)dt + \sigma dB(t), \quad X_0 = x_0, \tag{30}$$

where α, σ are positive constants. Without the diffusion term the equation is the ODE $dX(t) = -\alpha X(t)dt$ with solution $x_0 e^{-\alpha t}$ converging to 0 as $t \to \infty$. Thus it is natural to expect that the distribution of $X(t)$ will converge to some limit as $t \to \infty$.

The solution of SDE cannot be found in the form $X(t) = f(t, B(t))$. We leave checking this as an exercise.

Let us try to find a solution in the form

$$X(t) = a(t)\left(x_0 + \int_0^t b(u)dB(u)\right), \tag{31}$$

where $a(t), b(t)$ two smooth functions, $a(0) = 1$. The RHS of (31) is a Gaussian process, because the functions $a(t), b(t)$ are nonrandom (see the above remarks on the Gaussian processes).

The product rule applies in the classical form, because $dtdB(t) = 0$, and it yields

$$dX(t) = a'(t)\left(x_0 + \int_0^t b(u)dB(u)\right)dt + a(t)b(t)dB(t),$$

which in the view of (31) can be written as

$$dX(t) = \frac{a'(t)}{a(t)} X(t)dt + a(t)b(t)dB(t).$$

Matching the coefficients with (30) results in

$$\frac{a'(t)}{a(t)} = -\alpha, \quad a(t)b(t) = \sigma.$$

Recalling $a(0) = 1$ we solve these as

$$a(t) = e^{-\alpha t}, \quad b(t) = \sigma e^{\alpha t}.$$

Whence

$$X(t) = e^{-\alpha t}\left(x_0 + \sigma \int_0^t e^{\alpha u} dB(u)\right) = x_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-u)} dB(u).$$

The distribution of $X(t)$ is normal, with

$$\mathbb{E}X(t) = x_0 e^{-\alpha t} \to 0 \quad \text{as } t \to \infty,$$
$$\text{Var}X(t) = \sigma^2 \int_0^t e^{-2\alpha(t-u} du = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}) \to \frac{\sigma^2}{2\alpha} \quad \text{as } t \to \infty.$$

From the limits of the mean and the variance, the limit distribution of $X(t)$ is normal, that is $X(t) \xrightarrow{d} \mathcal{N}(0, \frac{\sigma^2}{2\alpha})$, as expected.

Example 5.4. The *Brownian bridge* (BB) on $[0, 1]$ is the Brownian motion 'forced to visit 0 at time 1'. Literally this means conditioning the BM $(B(t), t \in [0, 1])$ on the event $\{B(1) = 0\}$, but care is necessary because the conditioning is on the event of probability $\mathbb{P}(B(1) = 0) = 0$. A way to deal with the difficulty is to introduce the BB as a limit of the processes obtained by conditioning the BB on the event $\{|B(1)| < \epsilon\}$, while sending $\epsilon \to 0$.

A classical direct construction defines the BB as $B^\circ(t) = B(t) - tB(1)$, $t \in [0, 1]$. It is not hard to see that $B^\circ(t)$ is a Gaussian process with $\mathbb{E}B^\circ(t) = 0$ and $\text{Cov}(B^\circ(s), B^\circ(t)) = s(1-t)$ for $0 \le s \le t \le 1$. One disadvantage of this definition is that B° is not adapted to the natural filtration of the BM. For instance, at time $t = 1/2$ the observer of BM cannot calculate $B^\circ(1/2)$, since $B(1)$ is yet unknown.

We aim at an alternative construction of the BB as an Ito process that drifts towards 0. When the position at time $t < 1$ is x it is most natural to drift at rate $-x/(1-t)$, where $1 - t$ is the remaining time. On this way we arrive at the SDE

$$dX(t) = -\frac{X(t)}{1-t} dt + dB(t). \tag{32}$$

We can solve the SDE in the form (31) like in the previous example, we find

$$\frac{a'(t)}{a(t)} = -\frac{1}{1-t}, \quad a(t)b(t) = 1.$$

Solving for a, b

$$a(t) = 1 - t, \quad b(t) = \frac{1}{1-t},$$

we find the solution

$$X(t) = (1-t)\int_0^t \frac{1}{1-u} dB(u), \quad t \in [0,1].$$

The integrand is nonrandom, hence X is a centered Gaussian process. The covariance function is, for $s < t$

$$\mathbb{E}[X(s)X(t)] = \mathbb{E}\left[(1-t)(1-s)\int_0^s \frac{1}{1-u} dB(u) \cdot \int_0^t \frac{1}{1-u} dB(u)\right] =$$

$$(1-s)(1-t)\text{Var}\int_0^s \frac{1}{1-u} dB(u) =$$

$$(1-s)(1-t)\int_0^s \frac{1}{(1-u)^2} du = (1-s)(1-t)s/(1-s) = s(1-t).$$

Comparing with the mean and covariance function of B°, we see that X is a BB. In constrast to $B^\circ(t) = B(t) - tB(1)$, however, the process X is adapted to the natural filtration of the BM B.

SDE with coefficients depending linearly on x. The last three examples are special cases of the SDE

$$dX(t) = [p_1(t) + p_2(t)X(t)]dt + [q_1(t) + q_2(t)X(t)]dB(t).$$

It can be shown that the solution is $X(t) = Y(t)Z(t)$, where

$$Z(t) = \exp\left(\int q_2(u)dB(u) + \int_0^t (p_2(u) - \frac{1}{2}q_2^2(u))du\right)$$

is a generalised gBM and

$$Y(t) = x_0 + \int_0^t \frac{p_1(u) - q_1(u)q_2(u)}{Z(u)} du + \int_0^t \frac{q_1(u)}{Z(u)} dB(u).$$

This follows by expanding the differentials dY, dZ and using the product rule for $d(XY)$.

5.2 Existence of solutions

Although explicit solution of SDE is rarely possible, existence and uniqueness of the solution hold under fairly general assumptions on the coefficients.

Theorem 5.5. *Suppose the coefficients of SDE* (28) *for* $(t \in [0,T])$ *satisfy the Lipschitz condition*

$$|\alpha(t,x) - \alpha(t,y)| + |\beta(t,x) - \beta(t,y)| \le K|x-y|$$

and the growth condition

$$|\alpha(t,x)| + |\beta(t,x)| \le K(1+|x|)^2,$$

with some constant $K > 0$*. Then* (28) *has a unique solution* $X(t), t \in [0,T]$*, which is*

(i) *continuous,*

(ii) *adapted to the natural filtration of the BM,*

(iii) *uniformly bounded in the mean-square sense*

$$\sup_{t\in[0,T]} \mathbb{E}(X^2(t)) < \infty.$$

(The proof, which will not be given here, resembles the Picard method of iterations from the theory of ODE.)

To appreciate the nature of the conditions on the coefficients, let us look at some phenomena with ODE's. Consider the ODE $x' = x^2$. The growth condition for $\alpha(t) = t^2$ does not hold. The solution $x(t) = (1-t)^{-1}$ 'explodes' at $t = 1$, so there is no continuous solution on $[0, T]$ for $T > 1$.

For the ODE $x' = \sqrt{x}$ the Lipschitz condition does not hold, because the square root function has unbounded derivative near 0. One solution is $x(t) = 0$ and another solution is $x(t) = \frac{1}{4}t^2$. Thus the uniqueness fails.

Example 5.6. A very interesting SDE is the *Tanaka equation*

$$dX(t) = \mathrm{sgn}X(t)\, dB(t), \tag{33}$$

where sgn is the sign function ($\mathrm{sgn}\, x = 1$ for $x \geq 0$, $\mathrm{sgn}\, x = -1$ for $x < 0$). The Lipschitz condition does not hold, because sgn is discontinuous at $x = 0$, so Theorem 5.5 does not apply. We will show that *there is no solution X adapted to the natural Brownian filtration* $\mathcal{F}^B = (\mathcal{F}_t^B,\ t \geq 0)$.

Notice first that the process $Y(t) = \int_0^t Z(u)dB(u)$ is a BM, provided $|Z(t)| = 1$ for all t where Z is an adapted process (possibly, adapted to some filtration for the BM different from $\mathcal{F}^B$). This follows by Lévy's theorem, since Y is a continuous martingale, with quadratic variation

$$\langle Y\rangle(t) = \mathbb{E}\int_0^t (Z(u))^2 du = t.$$

Suppose X is a solution to (33) adapted to $\mathcal{F}^B$. Choosing $Z(t) = \mathrm{sgn}X(t)$ we have $|Z(t)| = 1$ and by the above X is a BM. On the other hand, multiplying both parts of (33) by $\mathrm{sgn}\, X(t)$ we get

$$dB(t) = \mathrm{sgn}X(t)dX(t),$$

or, equivalently,

$$B(t) = \int_0^t \mathrm{sgn}X(u)dX(u)$$

(where both X and B are BM's).

For a smooth function $\int_0^t \mathrm{sgn}f(u)df(u) = |f(u)|$ (check this for $f(u) = \sin u$). For the BM there is an additional term:

$$\int_0^t \mathrm{sgn}X(u)dX(u) = |X(t)| - L(t),$$

where $L(t)$ is the *local time at* 0 for the BM $(X(t), t \in [0,T])$[4]. The local time $(L(t), t \geq 0)$ is a process adapted to the filtration $\mathcal{F}^{|X|} = (\mathcal{F}_t^{|X|},\ t \geq 0)$ associated with the absolute value process $|X(t)|$. But then also $B(t) = \int_0^t \mathrm{sgn}X(u)dX(u) = |X(t)| - L(t)$ is adapted to $\mathcal{F}^{|X|}$. Now, we have X adapted to $\mathcal{F}^B$ (as solution to (33)), and B adapted to $\mathcal{F}^{|X|}$. Let $\mathcal{F}^X$ be the natural filtration of BM X. We have

$$\mathcal{F}_t^X \subset \mathcal{F}_t^B \subset \mathcal{F}_t^{|X|}.$$

But this is a contradiction, because $\mathcal{F}_t^{|X|}$ is strictly smaller than $\mathcal{F}_t^X$. Plainly, the inclusion $\mathcal{F}_t^X \subset \mathcal{F}_t^{|X|}$ implies that we could determine the sign $\mathrm{sgn}X(t)$ from $(|X(u)|, u \leq t)$, which is impossible because the sign of BM at any given time t is independent of the absolute value $(|X(u)|, u \leq t)$[5].

The contradiction shows that (33) has no solution adapted to $\mathcal{F}^B$, as in Theorem 5.5.

A solution adapted to $\mathcal{F}^B$ is called *strong*. Strong solution presumes that the BM B is given and $X(t)$ is 'computable' from $(B(u), u \leq t)$. The nonexistence of strong solution for Tanaka's equation suggests to relax the adaptedness condition, hence to widen the concept of solution to SDE. For given coefficients $\alpha(t,x), \beta(t,x)$ a *weak* solution to (28) is a pair of processes $(X(t), \widetilde{B}(t))$ which satisfy the SDE, with $\widetilde{B}$ being a BM. Weak solution does not require that X be adapted to the natural filtration of $\widetilde{B}$. Of course, each strong solution is a weak solution.

It is easy to construct a weak solution to Tanaka's equation. To that end, let X be *any* BM, and consider another BM $\widetilde{B}$ defined as the stochastic integral

$$\widetilde{B}(t) = \int_0^t \mathrm{sgn}X(u)dX(u).$$

Then $d\widetilde{B}(t) = \mathrm{sgn}X(t)dX(t)$, whence $dX(t) = \mathrm{sgn}X(t)d\widetilde{B}(t)$ so X is a weak solution to Tanaka's equation.

5.3 Feynman-Kac connection

By valuation of the European call option we encountered a partial DE (18) for the price of the option considered as a function of the current stock value and time. A similar connection, variation of the Feynman-Kac theorem, appears in the more general context of SDE's.

Theorem 5.7. *Consider the SDE*

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dB(t),$$

and let $h : \mathbb{R} \to \mathbb{R}$ be a function. For fixed r, define the function

$$g(t,x) = \mathbb{E}_{t,x}[e^{-r(T-t)}h(X(T))], \quad t \in [0,T]$$

[4]The local time of BM X is the limit $L(t) = \lim_{\epsilon \to 0} \epsilon^{-1} \int_0^t 1(|X(u)| < \epsilon/2)du$. Thus L is adapted to $\mathcal{F}^{|X|}$.

[5]$\mathbb{P}(\mathrm{sgn}B(t) = \pm 1) = \frac{1}{2}$.

where the expectation $\mathbb{E}_{t,x}$ refers to the solution of SDE with $X(t) = x$ (it is assumed that the expectation is finite for all t and x). Then $g(t, x)$ satisfies the PDE

$$g_t(t, x) + \alpha(t, x)g_x(t, x) + \frac{1}{2}\beta^2(t, x)g_{xx}(t, x) = rg(t, x) \tag{34}$$

and the terminal condition

$$g(T, x) = h(x).$$

Proof. By the Markov property of the solution to the SDE, the definition of g can be written as $e^{-rt}g(t, X(t)) = \mathbb{E}[e^{-rT}h(X(T))|\mathcal{F}_t]$, where $(\mathcal{F}_t)$ is a background filtration. By the 'tower property' this defines a martingale, whence the SDE follows by equating the 'dt' coefficient of the Ito differential to 0:

$$d(e^{-rt}g(t, X(t)) = e^{-rt}(-rg + g_t + \alpha g_x + \frac{1}{2}\beta^2 g_{xx})dt + e^{-rt}\beta g_x dB.$$

□

5.4 Systems of SDE's

A k-dimensional BM is a vector-valued process $\vec{B}(t) = (B_1(t), \ldots, B_k(t))^*$ ($*$ stays for transposition), where B_i's are independent standard BM. In the vector notation, a system of k SDE's has the form

$$\vec{X}(t) = \vec{\alpha}(t, \vec{X}(t))dt + \beta(t, \vec{X}(t))d\vec{B}(t),$$

where $\vec{\alpha}(t, \vec{X}(t)) = (\vec{\alpha}_1(t, \vec{X}(t)), \cdots, \vec{\alpha}_k(t, \vec{X}(t)))^*$ and $\beta(t, \vec{X}(t)) = (\beta_{ij})_{i,j=1,\ldots,k}$.

Example 5.8. The complex-valued process $e^{iB(t)}$ can be called the BM on the circle. The real and imaginary parts are, respectively, $X(t) = \cos B(t), Y(t) = \cos B(t)$. Calculating the differentials,

$$\begin{aligned} dX(t) &= -\sin B(t)dB(t) - \frac{1}{2}\cos B(t)dt = -Y(t)dB(t) - \frac{1}{2}X(t)dt \\ dY(t) &= \cos B(t)dB(t) - \frac{1}{2}\sin B(t)dt = X(t)dB(t) - \frac{1}{2}Y(t)dt, \end{aligned}$$

which in the vector notation becomes

$$d\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} dt + \begin{bmatrix} -Y(t) & 0 \\ X(t) & 0 \end{bmatrix}\begin{bmatrix} dB_1(t) \\ dB_2(t) \end{bmatrix}$$

The second BM B_2 does not affect the solution.