6 Stochastic Differential equations (1 lecture)

This section is mainly based on lecture notes of Prof. Steve Kou. Standard references for this topic include Chapter 5 of Oksendal or Chapter 5 of "Brownian Motion and Stochastic Calculus" 2nd edition, by Karatzas and Shreve.

Many models in finance rely on the stochastic integral equation

$$X_t = X_0 + \int_0^t b(s, X_s)dt + \int_0^t \sigma(s, X_s)dW_s, \tag{6.1}$$

or equivalently, in a form of the stochastic differential equation

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

In this chapter, we study two issues: (a) whether the solution exists and, if so, when it is unique; (b) how to solve the equation, at least for some interesting special cases.

First, we have to assign a proper meaning to the solution(s) of a stochastic differential equation (see (5.18))

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

or written component-wise as

$$dX_t^{(i)} = b_i(t, X_t)dt + \sum_{j=1}^m \sigma_{ij}(t, X_t)dW_t^{(j)}, \qquad i = 1, \cdots, n,$$

where $W_t = (W_t^{(1)}, \cdots, W_t^{(m)})$ is a standard m-dimensional Brownian motion (by definition, the components are required to be independent).

There are at least two ways to assign proper meanings to the above stochastic differential equations. First, we may want to find a solution X_t when W_t is given. This is called a strong solution. The second interpretation, which is called a weak solution, allows one to construct X_t and W_t simultaneously. See Questions 4 and 6 of Homework V, and if you want to read more, see Section 5.3 of Oksendal and Section 5.3 of "Brownian Motion and Stochastic Calculus" (2nd edition) by Karatzas and Shreve).

In finance, most of the time we are interested in strong solutions. For example, when we study the replication of an option payoff with the underlying asset following a stochastic differential equation, we talk about strong solutions.

6.1 Lipschitz and Linear Growth Conditions

To study the existence and uniqueness of the strong solution of a stochastic differential equations, we need to discuss some regularity conditions on the drift and volatility coefficients $b(t, x)$ and $\sigma(t, x)$.

The coefficients $b(t,x)$ and $\sigma(t,x)$ are said to satisfy the Lipschitz condition if for every $0 \le t < \infty$ and every x and y,

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le C|x-y|, \tag{6.2}$$

and are said to satisfy the linear growth condition if for every $0 \le t < \infty$ and every x

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \le D\left(1 + |x|^2\right),$$

where C and D are constants.

An important result in stochastic differential equations is that if the coefficients $b(t,x)$ and $\sigma(t,x)$ satisfy the Lipschitz condition and the linear growth condition, then the strong solution of (6.1) exists and is unique, in the sense that if $\tilde{X}_t$ is another solution then

$$\mathbb{P}\left(X_t = \tilde{X}_t, \text{ for all } 0 \le t < \infty\right) = 1.$$

The proof[51] can be found in Oksendal Theorem 5.2.1.

Example 6.1 *Consider the stochastic differential equation*

$$dX_t = 3X_t^{1/3}dt + 3X_t^{2/3}dW_t, \qquad X_0 = 0,$$

where W_t is a standard Brownian motion with $W_0 = 0$. By Itô formula, we can see that a solution is given by $X_t = W_t^3$, because

$$dX_t = dW_t^3 = 3W_t^2 dW_t + \frac{1}{2}6W_t dt = 3X_t^{1/3}dt + 3X_t^{2/3}dW_t,$$

with $X_0 = W_0^3 = 0$. However, that Lipschitz condition fails. Indeed, the Lipschitz condition reduces to

$$|3x^{1/3} - 3y^{1/3}| + |3x^{2/3} - 3y^{2/3}| < C|x-y|;$$

however, this is violated at $x = 0$ because

$$\lim_{y \to x} \frac{|x^\gamma - y^\gamma|}{|x-y|} = \left|\frac{dx^\gamma}{dx}\right| = |\gamma x^{\gamma-1}| = \left|\frac{\gamma}{x^{1-\gamma}}\right| \to \infty \quad \text{when } x \to 0, \quad \text{if } 0 < \gamma < 1, \tag{6.3}$$

and we cannot find a finite constant $C < \infty$. In fact, we can find another solution by noting that $X_t \equiv 0$ is also a solution.

Some examples in finance may not satisfy the Lipschitz conditions. For example, consider the following stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t^\gamma dW_t, \quad 0 < \gamma < 1. \tag{6.4}$$

[51] which is essentially the same as the existence and uniqueness proof for ordinary differential equation, i.e., you construct a Picard sequence and show that it is a Cauchy sequence and hence converge. Then you show that the limit satisfy (6.1).

The motivation of introducing this process is to model the situation when the volatility tends to decrease when the stock price X_t increases: $\frac{dX_t}{X_t} = \mu dt + \sigma \frac{1}{X_t^{1-\gamma}} dW_t$.

For this stochastic differential equation, the Lipschitz condition reduces to

$$|\mu x - \mu y| + |\sigma x^\gamma - \sigma y^\gamma| \le C|x - y|,$$

which implies that

$$\frac{|x^\gamma - y^\gamma|}{|x - y|} \le \frac{C}{\sigma} < \infty.$$

However, this is violated at $x = 0$ when $0 < \gamma < 1$ because of (6.3).

As another example, consider the Cox-Ingersoll-Ross (1985) (CIR) interest rate model:

$$dr_t = -\alpha(r_t - \beta)dt + \sigma\sqrt{r_t}dW_t, \tag{6.5}$$

with $\alpha > 0$, where r_t represents the spot interest rate. In this case the Lipschitz condition is

$$| - \alpha(x - \beta) + \alpha(y - \beta)| + |\sigma\sqrt{x} - \sigma\sqrt{y}| \le C|x - y|,$$

which is again violated at $x = 0$ for the same reason. One can prove that if $r_0 > 0$, the solution remains non-negative, and r_t can reach zero only if $\sigma^2 > 2\alpha\beta$. If $\sigma^2 \le 2\alpha\beta$, r_t will never be zero. The solution $r(t)$, given $r(u)$ for some $u < t$, is up to a scale factor, a non-central chi-square distribution. A proof can be found in "Interest Rate Models: An Introduction" (2004) by Andrew Cairns or Section 3.4 of "Monte Carlo Methods in Financial Engineering" by Paul Glasserman. See Question 8 in Homework V and Question 8 in Homework VI for related exercises.

To rigorously analyze these stochastic differential equations violating the Lipschitz condition or linear growth condition, one may have to study the theory of diffusion processes with proper boundary conditions on an interval (the interval may be of finite or infinite length). An excellent introduction to the subject is Karlin and Taylor (1988, A Second Course in Stochastic Processes, Academic Press, Vol. 2), in which there is a long chapter devoted to diffusion processes and boundary conditions; see also Section 5 of "Brownian Motion and Stochastic Calculus" 2nd edition by Karatzas and Shreve.

Remark: Before we move to the application, we like to list a few useful facts: Since W_s is a continuous function of s, so is $f(W_s)$ and therefore, by the fundamental theorem of calculus,

$$\frac{d}{dt}\int_0^t f(W_s)ds = f(W_t).$$

However, we cannot take $\frac{d}{dt}$ on $\int_0^t f(W_s)dW_s$ (e.g. $\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}$ which is not differentiable). But we still have

$$d\left(\int_0^t f(W_s)dW_s\right) = f(W_t)dW_t. \tag{6.6}$$

Why? Because if we call $Y_t = \int_0^t f(W_s)dW_s$, then $Y_0 = 0$ and (6.6) simply says that $dY_t = f(W_t)dW_t$, which is of course true since by defintion, $dY_t = f(W_t)dW_t$ means

$$Y_t - Y_0 = \int_0^t f(W_s)dW_s$$

which is true by the definition of Y_t.

Following Section 6.3 of Shreve II, we will state the Markov property of (6.1) which extends that of (4.33) [52]. Let $0 \le t \le T$ be given, and let $f(y)$ be any function[53]. Denote by

$$g(t,x) = \mathbb{E}^{t,x}[f(X(T))] \tag{6.7}$$

the expectation of $f(X(T))$, where $X(T)$ is the solution of (6.1) with initial condition $X(t) = x$. Note that there is nothing random about $g(t,x)$; it is an ordinary function of the two dummy variables t and x.

Theorem 6.1 *Let $X(u)$, $u \ge 0$, be a solution of (6.1) with initial condition given at time 0. Then, for $0 \le t \le T$,*

$$\mathbb{E}[f(X(T))|\mathcal{F}(t)] = g(t, X(t)) \overset{(6.7)}{=} \mathbb{E}^{t,X(t)}[f(X(T))] \tag{6.8}$$

By definition[54], the above theorem immediately implies the following result:

Corollary 6.1 *Solutions to (6.1) are Markov processes.*

According to Shreve II, the details of the proof of Theorem 6.1 are quite technical and are not given in Shreve II [55], but the intuitive content is clear. Suppose the process $X(u)$ beings at time zero, being generated by (6.1), and one watches it up to time t. Suppose now one is asked, based on this information, to compute the conditional expectation of $f(X(T))$, where $T > t$. Then one should pretend that the process is starting at time t at its current position, generate the solution to the stochastic differential equation corresponding to this initial condition, and compute the expected value of $f(X(T))$ generated in this way. In other words, replace $X(t)$ by a dummy x in order to hold it constant, compute $g(t,x) = \mathbb{E}^{t,x}[f(X(T))]$, and after computing this function, put the random variable $X(t)$ back in place of the dummy x.

[52]In the current context, the $g(x)$ in (4.36) should be called $g(s,x)$ with g defined by (6.7).

[53]Technically, we should required f to be so called Boreal-measurable. All functions you will meet in application satisfies this requirement.

[54]The discrete Markov process is defined in Question 7 of Homework III. The continuous Markov process is defined to be the $X(t)$ that satisfies (4.33) : for any $0 \le s \le t \le T$ and for any f, there is a g so that $\mathbb{E}[f(X(t))|\mathcal{F}(s)] = g(X(s))$.

[55]Note that a proof is indeed given in the solution of Question 5 of Homework VI.

6.2 Some Immediate Applications in Finance

First we use stochastic differential equations to analyze stock prices. Consider an one-dimensional stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \qquad X_0 = x_0.$$

with $\mu, x_0 \in (-\infty, \infty)$ and $\sigma > 0$ being constants. This equation can be used to model the stock price with mean return μ, and volatility σ.

The Lipschitz and linear growth conditions are satisfied. By Itô formula,

$$\begin{aligned} d\log(X_t) =& \frac{1}{X_t} dX_t + \frac{1}{2}\left(-\frac{1}{X_t^2}\right)(dX_t)^2 = \frac{1}{X_t} dX_t + \frac{1}{2}\left(-\frac{1}{X_t^2}\right)\sigma^2 X_t^2 dt \\ =& \frac{1}{X_t}(\mu X_t dt + \sigma X_t dW_t) - \frac{1}{2}\sigma^2 dt \\ =& \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma W_t. \end{aligned}$$

By definition, the above "differential equation" is another way to rewrite

$$\begin{aligned} \log X_t =& \log X_0 + \int_0^t \left(\mu - \frac{1}{2}\sigma^2\right) ds + \int_0^t \sigma dW_s \\ =& \log X_0 + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t. \end{aligned}$$

Hence

$$X_t = X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}.$$

In Question 1 of Homework III, we have computed $\mathbb{E}[e^{\theta W_t}] = e^{\frac{1}{2}\theta^2 t}$. Another interesting way to compute the mean is to use the Itô formula. It is clear that for any constant θ

$$d(e^{\theta W_t}) = \theta e^{\theta W_t} dW_t + \frac{1}{2}\theta^2 e^{\theta W_t} dt.$$

Therefore,

$$e^{\theta W_t} = 1 + \int_0^t \theta e^{\theta W_s} dW_s + \frac{1}{2}\theta^2 \int_0^t e^{\theta W_s} ds,$$

and taking expectation (recall $\mathbb{E}$[Itô integral] $= 0$),

$$\mathbb{E}(e^{\theta W_t}) = 1 + \frac{1}{2}\theta^2 \int_0^t \mathbb{E}(e^{\theta W_s}) ds$$

Taking derivative of the deterministic function $\mathbb{E}(e^{\theta W_t})$ with respect to t, we have

$$\frac{d}{dt}\mathbb{E}(e^{\theta W_t}) = \frac{1}{2}\theta^2 \mathbb{E}(e^{\theta W_t}), \qquad \mathbb{E}(e^{\theta W_0}) = 1,$$

which is equivalent to a simple differential equation

$$\frac{dY_t}{dt} = \frac{1}{2}\theta^2 Y_t, \qquad Y_0 = 1,$$

whose solution is

$$\mathbb{E}(e^{\theta W_t}) = e^{\frac{1}{2}\theta^2 t}$$

By the same method (see Question 1 of Homework VI), we can get

$$\mathrm{Var}(e^{\theta W_t}) = e^{2\theta^2 t} - e^{\theta^2 t}. \tag{6.9}$$

(Recall that in Question 1 of Homework III, we have obtained by direct calculation that of $X \sim N(\mu, \sigma^2)$, then $\mathbb{E}X = e^{\mu + \frac{1}{2}\sigma^2}$ and $\mathrm{Var}(e^X) = e^{2\mu}\left(e^{2\sigma^2} - e^{\sigma^2}\right)$.)

Remark: If $X_t = X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$. It is easy to see that the median of X_t is

$$m(X_t) = X_0 e^{(\mu - \frac{1}{2}\sigma^2)t}$$

One puzzle appears. If $0 < \mu < \frac{1}{2}\sigma^2$, then we know that as $T \to \infty$,

$$\frac{\log(X_T/X_0)}{T} = \left(\mu - \frac{1}{2}\sigma^2\right) + \sigma \frac{W_T}{T} \to \mu - \frac{1}{2}\sigma^2 < 0$$

almost surely, and

$$\frac{\log(m(X_T)/X_0)}{T} = \mu - \frac{1}{2}\sigma^2 < 0.$$

Therefore, we have

$$\log(X_T) \to -\infty, \qquad \log(m(X_T)) \to -\infty$$

almost surely as $T \to \infty$ if $0 < \mu < \frac{1}{2}\sigma^2$. But we also know that

$$\mathbb{E}[X_T] = X_0 e^{\mu T} \to \infty.$$

What really happens is that the distribution of X_T is highly skewed to the right (i.e. with small probability X_T can be very large, leading to large values in terms of the expectation), and big volatility (hence big risk) may drive most of investors to bankrupt even before the expected big fortune arrives.

Another interesting feature revealed here is the connection between the stochastic differential equations (SDE) and partial differential equations (PDE). It appears that the mean of a solution of SDE may be a solution of PDE. There is a general theorem, Feynman-Kac theorem, addressing this point, see Question 5 of Homework VI. See also Question 7 of Homework VI.

6.3 Ornstein-Uhlenbeck Process

Besides the geometric Brownian motion, another widely used process in finance is the Ornstein-Uhlenbeck process, which is the solution of the SDE

$$dX_t = -\alpha(X_t - C)dt + \sigma dW_t, \tag{6.10}$$

where $\alpha>0$, $\sigma > 0$, and $C \in \mathbb{R}$, all being constant and X_0 may be a random variable. It has been used in a famous paper by Vasicek (1977) to model interest rate. The most important feature of (6.10) is mean reversion: the drift term in the SDE is negative if X_t is above C, and is positive if X_t is below C: It is indeed this specialty attracts Vasicek to use the process, since the interest rate appears to be mean reverting if the interest rate is too high, it will drop, and vice versa.

Like in Question 5 of Homework V, we can verify by Itô formula that the unique solution of (6.10) is

$$X_t = C(1 - e^{-\alpha t}) + X_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s. \tag{6.11}$$

For the mean and variance, we have

$$\mathbb{E}[X_t] = C(1 - e^{-\alpha t}) + \mathbb{E}[X_0]e^{-\alpha t},$$

and

$$\begin{aligned}\text{Var}[X_t] =& \text{Var}[X_0]e^{-2\alpha t} + \sigma^2 \int_0^t e^{-2\alpha(t-s)} ds \\ =& \text{Var}[X_0]e^{-2\alpha t} + \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}).\end{aligned}$$

As a remark, if Y_s is a deterministic process, then $\int_0^t Y_s dW_s$ is a Gaussian process [56], and many explicit calculations are possible for Gaussian processes (just like many explicit calculations are possible for multivariate normal distributions).

But if Y_s is itself a stochastic process, $\int_0^t Y_s dW_s$ may no longer be a Gaussian process. E.g. $\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}$.

6.4 Integration of Geometric Brownian Motion

Recall that an Asian option has payoff $(\frac{1}{m}\sum_{j=1}^m S(t_j) - K)^+$ for some fixed set of dates $0 = t_0 < t_1 < \cdots < t_m = T$. Hence it can be interesting to consider the arithmetic average of geometric Brownian motion

$$\frac{1}{t} A_t^{(v)} \stackrel{\text{def}}{=} \frac{1}{t} \int_0^t e^{2(vs + W_s)} ds.$$

[56] Z_t is a Gaussian process means that $(Z_{t_1}, Z_{t_2}, \cdots, Z_{t_k})$ are multivariate normal distribution for any positive integer k and any sequence $t_1, t_2, \cdots, t_k$.

In general, it is difficult to compute $A_t^{(v)}$ directly, as $A_t^{(v)}$ is path dependent; for example, the arithmetic average up to today is correlated with the arithmetic average up to yesterday.

However, it is possible to link $A_t^{(v)}$ to a simpler process, $Y_t^{(v)}$

$$\frac{1}{t}Y_t^{(v)} = \frac{1}{t}\int_0^t e^{2(v(t-s)+W_t-W_s)}ds.$$

Since the two processes

$$\{W_s, 0 \le s \le t\}, \qquad \{W_t - W_{t-s}, 0 \le s \le t\}$$

have the same probabilistic law, we have for every $t \ge 0$, $A_t^{(v)}$ equal in distribution to $Y_t^{(v)}$, because

$$\begin{aligned} A_t^{(v)} &= \int_0^t e^{2(vs+W_s)}ds = \int_0^t e^{2(vs+W_t-W_{t-s})}ds \\ &= \int_0^t e^{2(v(t-u)+W_t-W_u)}du = Y_t^{(v)} \end{aligned}$$

Now, apply Itô formula for the product (see Question 9 of Homework V), we have

$$\begin{aligned} dY_t^{(v)} =& d\left(e^{2vt+2W_t}\int_0^t e^{-2vs-2W_s}ds\right) \\ =& d\left(e^{2vt+2W_t}\right)\int_0^t e^{-2vs-2W_s}ds + e^{2vt+2W_t}d\left(\int_0^t e^{-2vs-2W_s}ds\right) \\ &+ d\left(e^{2vt+2W_t}\right)d\left(\int_0^t e^{-2vs-2W_s}ds\right). \end{aligned}$$

Since

$$\begin{aligned} d\left(e^{2vt+2W_t}\right) =& e^{2vt+2W_t}(2vdt + 2dW_t) + \frac{1}{2}e^{2vt+2W_t}4dt \\ =& e^{2vt+2W_t}\left((2v+2)dt + 2dW_t\right), \\ d\left(\int_0^t e^{-2vs-2W_s}ds\right) &= e^{-2vt-2W_t}dt, \end{aligned}$$

we have

$$\begin{aligned} dY_t^{(v)} =& e^{2vt+2W_t}\left((2v+2)dt + 2dW_t\right)\int_0^t e^{-2vs-2W_s}ds \\ &+ e^{2vt+2W_t}e^{-2vt-2W_t}dt \\ =& Y_t^{(v)}((2v+2)dt + 2dW_t) + dt \\ =& (2(v+1)Y_t^{(v)} + 1)dt + 2Y_t^{(v)}dW_t. \end{aligned}$$

As we will demonstrate next, we can solve the above equation explicitly. This result is taken from Dufresne (1989, Insurance: Mathematics and Economics). It has been used by Linetsky (2004) to price continuous Asian options, and is extended to Levy processes by Carmona, Petit and Yor (1997).

6.5 Explicit Strong Solutions for One-dimensional Linear Equations

Explicit solutions for SDE are rarely available, among which are the solution for the one-dimensional ($n = 1$) stochastic differential equation

$$dX_t = [b(t)X_t + a(t)]dt + \sum_{j=1}^{m}[\sigma_j(t)X_t + \eta_j(t)]dW_t^{(j)}, \tag{6.12}$$

where $W_t = (W_t^{(1)}, \cdots, W_t^{(m)})$ is an m-dimensional standard Brownian motion, and the coefficients b, a, σ_j, η_j all uniformly bounded and measurable processes.

Theorem 6.2 *The unique strong solution of (6.12) is*

$$X_t = Z_t\left\{X_0 + \int_0^t \frac{1}{Z_s}\left[a(s) - \sum_{j=1}^{m}\sigma_j(s)\eta_j(s)\right]ds + \sum_{j=1}^{m}\int_0^t \frac{\eta_j(s)}{Z_s}dW_s^{(j)}\right\}, \tag{6.13}$$

where

$$Z_t = \exp\left\{\int_0^t b(s)ds + \sum_{j=1}^{m}\int_0^t \sigma_j(s)dW_s^{(j)} - \frac{1}{2}\sum_{j=1}^{m}\int_0^t \sigma_j^2(s)ds\right\}.$$

Proof: We want to verify that the X_t given in (6.13) satisfies (6.12). For notation simplicity, we put $X_t = Z_t\xi_t$, where $\xi_0 = X_0$ and

$$d\xi_t = \frac{1}{Z_t}\left[a(s) - \sum_{j=1}^{m}\sigma_j(t)\eta_j(t)\right]dt + \sum_{j=1}^{m}\frac{\eta_j(t)}{Z_t}dW_t^{(j)}.$$

By Itô formula,

$$\begin{aligned} dZ_t =& Z_t\left\{b(t)dt + \sum_{j=1}^{m}\sigma_j(t)dW_t^{(j)} - \frac{1}{2}\sum_{j=1}^{m}\sigma_j^2(t)dt\right\} + \frac{1}{2}Z_t\sum_{j=1}^{m}\sigma_j^2(t)dt \\ =& Z_t\left\{b(t)dt + \sum_{j=1}^{m}\sigma_j(t)dW_t^{(j)}\right\}. \end{aligned}$$

Hence, again by Itô formula (Question 9 of Homework V),

$$\begin{aligned} dX_t =& Z_t d\xi_t + \xi_t dZ_t + d\xi_t dZ_t \\ =& \left[a(t) - \sum_{j=1}^{m}\sigma_j(t)\eta_j(t)\right]dt + \sum_{j=1}^{m}\eta_j(t)dW_t^{(j)} \\ &+ \xi_t Z_t\left\{b(t)dt + \sum_{j=1}^{m}\sigma_j(t)dW_t^{(j)}\right\} + \sum_{j=1}^{m}\sigma_j(t)\eta_j(t)dt \\ =& [a(t) + X_t b(t)]dt + \sum_{j=1}^{m}[X_t\sigma_j(t) + \eta_j(t)]dW_t^{(j)}. \end{aligned}$$

The uniqueness and existence follows from the Lipschitz and linear growth conditions. □

Example 6.2 *With $a(t) = 0$, $\eta_j(t) = 0$, immediately from (6.13), we see that*

$$X_t = Z_t X_0, \quad Z_t = \exp\left\{\int_0^t b(s)ds + \sum_{j=1}^m \int_0^t \sigma_j(s)dW_s^{(j)} - \frac{1}{2}\sum_{j=1}^m \int_0^t \sigma_j^2(s)ds\right\}.$$

Thus, the strong solution of the equation

$$dX_t = b(t)X_t dt + \sum_{j=1}^m \sigma_j(t)X_t dW_t^{(j)}$$

is

$$X_t = X_0 \exp\left\{\int_0^t \left(b(s) - \frac{1}{2}\sum_{j=1}^m \sigma_j^2(s)\right) ds + \sum_{j=1}^m \int_0^t \sigma_j(s)dW_s^{(j)}\right\}.$$

This result generalizes the geometric Brownian motion, which has constant coefficient drift and volatility.

Example 6.3 *If we let $m = 1$, $a(t) = \alpha C$, $b(t) = -\alpha$, $\eta_i(t) = \sigma$, and $\sigma_i(t) = 0$, then we arrive at the SDE of Ornstein-Uhlenbeck process. Indeed, in this case*

$$Z_t = \exp\left\{\int_0^t b(s)ds\right\} = e^{-\alpha t}$$

and we have

$$\begin{aligned}
X_t =& Z_t\left\{X_0 + \alpha C\int_0^t \frac{1}{Z_s}ds + \sigma\int_0^t \frac{1}{Z_s}dW_s\right\} \\
=& e^{-\alpha t}\left\{X_0 + \alpha C\int_0^t e^{\alpha s}ds + \sigma\int_0^t e^{\alpha s}dW_s\right\} \\
=& X_0 e^{-\alpha t} + \alpha C e^{-\alpha t}\frac{e^{\alpha t} - 1}{\alpha} + \sigma e^{-\alpha t}\int_0^t e^{\alpha s}dW_s \\
=& X_0 e^{-\alpha t} + C(1 - e^{-\alpha t}) + \sigma\int_0^t e^{-\alpha(t-s)}dW_s,
\end{aligned}$$

which is consistent with our previous result.

6.6 Homework VI

(Only submit solutions to Questions 6, 8.)

1. Use Itô formula, like how we calculate $\mathbb{E}(e^{\theta W_t})$ in the notes, to compute $\text{Var}(e^{\theta W_t})$. [Hint: Let $Y_t = e^{\theta W_t} - e^{\frac{1}{2}\theta^2 t}$. Consider dY_t^2. The solution of $\frac{dZ_t}{dt} = aZ_t + f(t)$ with a constant a, which can be rewritten as $\frac{d}{dt}\left(e^{-at}Z_t\right) = e^{-at}f(t)$, is $Z_t = Z_0 e^{at} + \int_0^t e^{a(t-s)} f(s)ds$.]

   Solution: $\text{Var}(e^{\theta W_t}) = \mathbb{E}\left(Y_t^2\right)$. $Y_0 = 0$.

$$dY_t = \theta e^{\theta W_t} dW_t + \frac{1}{2}\theta^2 e^{\theta W_t} dt - \frac{1}{2}\theta^2 e^{\frac{1}{2}\theta^2 t} dt.$$

$$\begin{aligned} dY_t^2 &= 2Y_t dY_t + (dY_t)^2 \\ =&2Y_t\left(\theta e^{\theta W_t} dW_t + \frac{1}{2}\theta^2 e^{\theta W_t} dt - \frac{1}{2}\theta^2 e^{\frac{1}{2}\theta^2 t} dt\right) + \theta^2 e^{2\theta W_t} dt \\ =&2Y_t\theta e^{\theta W_t} dW_t + \theta^2 Y_t^2 dt + \theta^2 e^{2\theta W_t} dt \end{aligned}$$

$$Y_t^2 = Y_0^2 + \int_0^t 2Y_s \theta e^{\theta W_s} dW_s + \int_0^t \left(\theta^2 Y_s^2 + \theta^2 e^{2\theta W_s}\right) ds$$

   Hence

$$\begin{aligned} \mathbb{E}Y_t^2 &= \int_0^t \left(\theta^2 \mathbb{E}(Y_s^2) + \theta^2 \mathbb{E}e^{2\theta W_s}\right) ds \\ &= \int_0^t \theta^2 \mathbb{E}(Y_s^2) + \theta^2 e^{2\theta^2 s} ds. \end{aligned}$$

   Let $Z_t = \mathbb{E}Y_t^2$. Then

$$\frac{dZ_t}{dt} = \theta^2 Z_t + \theta^2 e^{2\theta^2 t}, \qquad Z_0 = 0.$$

   By the hint, the solution is $Z_t = \int_0^t e^{\theta^2(t-s)}\theta^2 e^{2\theta^2 s} ds = e^{2\theta^2 t} - e^{\theta^2 t}$. This is $\text{Var}(e^{\theta W_t})$.

   Another Solution (which essentially changes the computation in the notes for $\mathbb{E}e^{\theta W_t}$ to $\mathbb{E}e^{2\theta W_t}$): $\text{Var}(e^{\theta W_t}) = \mathbb{E}[e^{2\theta W_t}] - [\mathbb{E}e^{\theta W_t}]^2$. Let $Y_t = e^{2\theta W_t}$. Then

$$dY_t = e^{2\theta W_t} 2\theta dW_t + \frac{1}{2} e^{2\theta W_t} 4\theta^2 dt.$$

$$Y_t = Y_0 + \int_0^t e^{2\theta W_s} 2\theta dW_s + \int_0^t 2\theta^2 e^{2\theta W_s} ds.$$

Hence

$$\begin{aligned}\mathbb{E}Y_t =& \mathbb{E}Y_0 + \int_0^t \left(2\theta^2 \mathbb{E}e^{2\theta W_s}\right) ds \\ =& \mathbb{E}Y_0 + \int_0^t \left(2\theta^2 \mathbb{E}Y_s\right) ds.\end{aligned}$$

Let $Z_t = \mathbb{E}Y_t$. Then

$$\frac{dZ_t}{dt} = 2\theta^2 Z_t, \qquad Z_0 = 1.$$

The solution is $Z_t = e^{2\theta^2 t}$. Similarly, $\mathbb{E}e^{\theta W_t} = e^{\frac{1}{2}\theta^2 t}$. So $\mathrm{Var}(e^{\theta W_t}) = \mathbb{E}[e^{2\theta W_t}] - [\mathbb{E}e^{\theta W_t}]^2 = e^{2\theta^2 t} - (e^{\frac{1}{2}\theta^2 t})^2 = e^{2\theta^2 t} - e^{\theta^2 t}$.

2. (Brownian bridge. Section 4.7 of Shreve II) Given real numbers α, β and given $T > 0$, consider the one-dimensional equation for $0 \le t \le T$:

$$dX_t = \frac{\beta - X_t}{T - t} dt + dW_t, \qquad X_0 = \alpha. \tag{6.14}$$

Identity the $b, a, m, \sigma_j, \eta_j$ in (6.12) and then use (6.13) to write down the solution.

Solution: $b(t) = -1/(T-t)$, $a(t) = \beta/(T-t)$, $m = 1$, $\sigma_1 = 0$, $\eta_1 = 1$. By (6.13), $Z_t = \exp\left(\int_0^t -1/(T-s)ds\right) = \exp\left(\int_0^t d\log(T-s)\right) = \exp(\log(T-t) - \log T) = (T-t)/T$.

$$\begin{aligned}X_t &= \frac{T-t}{T}\left(\alpha + \int_0^t \frac{T}{T-s}\left(\frac{\beta}{T-s}\right) ds + \frac{T}{T-s} dW_s\right) \\ &= \alpha\left(1 - \frac{t}{T}\right) + \beta\frac{t}{T} + (T-t)\int_0^t \frac{dW_s}{T-s}. \end{aligned} \tag{6.15}$$

Since we are here, we want to say a few more words about this X_t. If we set $Z_t = (T-t)\int_0^t \frac{dW_s}{T-s}$, we find that

$$\mathbb{E}Z_t = 0, \qquad \mathrm{Var}[Z_t] = (T-t)^2 \int_0^t \frac{1}{(T-s)^2} ds = \frac{t(T-t)}{T}.$$

As $t \uparrow T$, $\mathrm{Var}[Z_t] \to 0$, which means that its randomness disappears and Z_t converges to its mean, which is 0, as $t \uparrow T$. This observation suggests that it makes sense to define

$$Y_t = \begin{cases} (T-t)\int_0^t \frac{1}{T-s} dW_s & 0 \le t < T, \\ 0 & t = T. \end{cases}$$

Then Y_t is a continuous process. (See also Lemma 6.9 of Section 5.6 of Karatzas & Shreve.) Since $X_0 = \alpha$ and $X_T = \beta$ are prescribed values, X_t is called Brownian bridge.

But because there is a drift term in (6.14), X_t is not a martingale and hence is not a Brownian motion. Another way to construct Brownian bridge from α to β is

$$B_t = \alpha\left(1 - \frac{t}{T}\right) + \beta\frac{t}{T} + \left(W_t - \frac{t}{T}W_T\right), \qquad 0 \le t \le T. \tag{6.16}$$

Please note that $\{X_t\}$ is a Gaussian process, which means that for any $t_1 < t_2 \cdots < t_n$, $(X_{t_1}, \cdots, X_{t_n})$ are jointly normally distributed. So is $\{B_t\}$. It is obvious that when $\alpha = 0 = \beta$, $0 \le s \le t \le T$, $\mathbb{E}[X_t] = 0 = \mathbb{E}[B_t]$. We can also verify

$$\begin{aligned}\mathbb{E}[X_tX_s] &= (T-t)(T-s)\left(\mathbb{E}\left[\int_0^s \frac{1}{(T-u)^2}du\right] + \cancel{\mathbb{E}\left[\left(\int_0^s \frac{1}{(T-u)}dW_u\right)\left(\int_s^t \frac{1}{(T-u)}dW_u\right)\right]}\right)\\ &= s - \frac{st}{T},\end{aligned}$$

$$\begin{aligned}\mathbb{E}[B_tB_s] &= \mathbb{E}\left[W_tW_s\right] - \frac{s}{T}\mathbb{E}\left[W_tW_T\right] - \frac{t}{T}\mathbb{E}\left[W_sW_s\right] + \frac{ts}{T^2}\mathbb{E}\left[W_TW_T\right]\\ &= s - \frac{st}{T}.\end{aligned}$$

Hence $\mathbb{E}[X_tX_s] = \min(s,t) - \frac{st}{T} = \mathbb{E}[B_tB_s]$. As a result, the Gaussian processes $\{X_t\}$ and $B_t\}$ have the same distribution. However, X_t is $\mathcal{F}_t$-adpated, B_t is not.

3. Use the following formula (which can be derived using Itô isometry together with the identity $\mathbb{E}(XY) = \frac{1}{2}\left(\mathbb{E}(X+Y)^2 - \mathbb{E}X^2 - \mathbb{E}Y^2\right)$)

$$\mathbb{E}\left(\int_0^t Y_u^{(1)}dW_u \cdot \int_0^t Y_u^{(2)}dW_u\right) = \mathbb{E}\left(\int_0^t Y_u^{(1)}Y_u^{(2)}dt\right). \tag{6.17}$$

to show that for the Ornstein-Uhlenbeck process the covariance is given by

$$\mathrm{Cov}(X_s, X_t) = \mathbb{E}(X_s - \mathbb{E}X_s, X_t - \mathbb{E}X_t) = \left(\mathrm{Var}[X_0] + \frac{\sigma^2}{2\alpha}(e^{2\alpha(\min(t,s))} - 1)\right)e^{-\alpha(t+s)}.$$

[Hint: Assume $s < t$. You can break $\int_0^t f(u)dW_u$ into the sum of $\int_0^s f(u)dW_u$ and $\int_s^t f(u)dW_u$. Note that we have learned in Theorem 4.1 that $\int_0^s f(u)dW_u$ is $\mathcal{F}_s$-measurable while $\int_s^t f(u)dW_u$ only depends on increments of the Brownian motion that is independent of $\mathcal{F}_s$. Hence the two parts are independent and are both mean zero.]

Solution: $dX_t = -\alpha(X_t - C)dt + \sigma dW_t$ or $X_t = C(1-e^{-\alpha t}) + X_0e^{-\alpha t} + \sigma\int_0^t e^{-\alpha(t-\tau)}dW_\tau$. $X_t - \mathbb{E}X_t = (X_0 - \mathbb{E}X_0)e^{-\alpha t} + \sigma\int_0^t e^{-\alpha(t-\tau)}dW_\tau$. Hence

$$\begin{aligned}\mathrm{Cov}(X_s, X_t) =& \mathbb{E}\Big[\left((X_0 - \mathbb{E}X_0)e^{-\alpha t} + \sigma\int_0^t e^{-\alpha(t-\tau)}dW_\tau\right)\cdot\left((X_0 - \mathbb{E}X_0)e^{-\alpha s} + \sigma\int_0^s e^{-\alpha(s-\tau)}dW_\tau\right)\Big]\\ =& \mathbb{E}(X_0 - \mathbb{E}X_0)^2e^{-\alpha(t+s)} + \mathbb{E}\int_0^{\min(t,s)}\sigma^2e^{-\alpha(t+s)+2\alpha\tau}d\tau\\ =& e^{-\alpha(t+s)}\left(\mathrm{Var}[X_0] + \frac{\sigma^2}{2\alpha}(e^{2\alpha(\min(t,s))} - 1)\right).\end{aligned}$$

4. Verify by Itô formula that the unique solution of (6.10) is (6.11).

   Solution:
$$X_t = C(1 - e^{-\alpha t}) + X_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s.$$
$$e^{\alpha t} X_t = C(e^{\alpha t} - 1) + X_0 + \sigma \int_0^t e^{\alpha s} dW_s.$$
$$e^{\alpha t} dX_t + \alpha e^{\alpha t} X_t dt \overset{Q9,HW-IV}{=} d(e^{\alpha t} X_t) = C\alpha e^{\alpha t} dt + \sigma e^{\alpha t} dW_t.$$
   Hence
$$dX_t = -\alpha X_t dt + C\alpha dt + \sigma dW_t.$$
   Note that if you let $Y_t = X_t - C$, the Y_t would satisfy the equation in Question 5 of Homework V where we have shown how to "derive" the solution instead of "verify" that the given equation satisfies the equation.

5. (Feynman-Kac formula) Given functions b, σ, and f, consider the stochastic differential equation
$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$
   and the deterministic partial differential equation
$$\partial_t g(t, x) + b(t, x)\partial_x g(t, x) + \frac{1}{2}\sigma^2(t, x)\partial_x^2 g(t, x) = 0 \tag{6.18}$$
   with terminal condition
$$g(T, x) = f(x) \qquad \text{for all } x. \tag{6.19}$$
   Show that $g(t, X(t))$ is a martingale, and the solution of (6.18) can be represented as the expectation of $f(X(T))$ under the condition $X(t) = x$:
$$g(t, x) = \mathbb{E}^{t,x}[f(X(T))]. \tag{6.20}$$
   Proof: By Itô formula (5.9)
$$\begin{aligned} dg(t, X_t) &= \partial_t g dt + \partial_x g dX + \frac{1}{2}\partial_x^2 g(t, x)(dX)^2 \\ &= \partial_t g dt + \partial_x g(b dt + \sigma dW_t) + \frac{1}{2}\partial_x^2 g(t, x)\sigma^2 dt \\ &= \left(\partial_t g + b\partial_x g + \frac{1}{2}\sigma^2 \partial_x^2 g\right) dt + \sigma \partial_x g dW_t \\ &\overset{(6.18)}{=} \sigma \partial_x g dW_t. \end{aligned} \tag{6.21}$$

Hence $g(t, X(t))$ is a martingale since Itô integrals are martingales. Moreover, integrating $dg(t, X_t) = \sigma \partial_x g dW_t$ from t to T, we get

$$g(T, X_T) - g(t, X_t) = \int_t^T \sigma(s, X_s) \partial_x g(s, X_s) dW_s.$$

Hence

$$\mathbb{E}^{t,x}[g(T, X(T))] = \mathbb{E}^{t,x}[g(t, X(t) + \int_t^T \sigma(s, X_s)\partial_x g(s, X_s) dW_s] \stackrel{\text{iii) of Thm 4.1}}{=} \mathbb{E}^{t,x}[g(t, X(t)].$$

But $\mathbb{E}^{t,x}[g(T, X(T))] \stackrel{(6.19)}{=} \mathbb{E}^{t,x}[f(X(T))]$ while $\mathbb{E}^{t,x}[g(t, X(t)] = g(t, x)$ as $\mathbb{E}^{t,x}$ requires $X(t) = x$. This proves (6.20).

The above calculation contains a proof of Theorem 6.1: Given any f, using the b and σ in $dX_t = bdt + \sigma dW_t$, we can form partial differential equations (6.18)+(6.19) and solve for g. Then $g(t, X(t))$ is a martingale by the calculation (6.21). Hence

$$\mathbb{E}[f(X(T))|\mathcal{F}(t)] \stackrel{(6.19)}{=} \mathbb{E}[g(T, X(T))|\mathcal{F}(t)] \stackrel{\text{martingale}}{=} g(t, X(t)) \tag{6.22}$$

The above equation says that given all information up to t, to calculate the expectation of $f(X(T))$, only the value of $X(t)$ matters. Hence $g(t, X(t)) = \mathbb{E}^{t,X(t)}[f(X(T))]$ and the g obtained from solving (6.18)+(6.19) is the same as the g defined in (6.7). This proves (6.8).

6. (Black-Scholes-Merton equation) Given constant r, functions b, σ, and f, consider the stochastic differential equation

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

and the deterministic partial differential equation

$$\partial_t g(t, x) + b(t, x)\partial_x g(t, x) + \frac{1}{2}\sigma^2(t, x)\partial_x^2 g(t, x) = rg(t, x) \tag{6.23}$$

with terminal condition

$$g(T, x) = f(x) \qquad \text{for all } x. \tag{6.24}$$

Fix $T > 0$, show that $e^{-rt}g(t, X(t))$ is a martingale, and the solution $g(t, x)$ of (6.23)+(6.24) can be represented as

$$g(t, x) = \mathbb{E}^{t,x}[e^{-r(T-t)}f(X(T))] \tag{6.25}$$

where $\mathbb{E}^{t,x}$ means the expectation under the condition that $X(t) = x$.

7. Consider

$$dX_t = [a(t)X_t + b(t)]dt + c(t)dW_t$$

with X_0 = constant. Let $\mathbb{E}[X_t] = m(t)$ and $\text{Var}[X_t] = v(t)$. Show that

$$\frac{dm(t)}{dt} = a(t)m(t) + b(t),$$
$$\frac{dv(t)}{dt} = 2a(t)v(t) + c^2(t).$$

Solution: The differential form is equivalent to the integral form

$$X_t - X_0 = \int_0^t [a(s)X_s + b(s)]ds + \int_0^t c(s)dW_s. \tag{6.26}$$

Hence by taking $\mathbb{E}$ on both sides, we obtain

$$m(t) - m(0) = \int_0^t [a(s)m(s) + b(s)]ds \tag{6.27}$$

which is equivalent to (or by taking $\frac{d}{dt}$)

$$\frac{dm(t)}{dt} = a(t)m(t) + b(t). \tag{6.28}$$

By Itô formula

$$d(X_t^2) = 2X_t dX_t + \frac{1}{2}2(dX_t)^2 = 2X_t[a(t)X_t + b(t)]dt + 2X_t c(t)dW_t + c^2(t)dt$$

After writing into integral form and taking expectation, we get

$$\mathbb{E}[X_t^2] - \mathbb{E}[X_0^2] = 2\int_0^t a(s)\mathbb{E}[X_s^2]ds + 2\int_0^t b(s)m(s)ds + \int_0^t c^2(s)ds. \tag{6.29}$$

But from (6.28) we know

$$\frac{d}{dt}m^2(t) = 2m(t)\frac{d}{dt}m(t) = 2a(t)m^2(t) + 2b(t)m(t)$$

or

$$m^2(t) - m^2(0) = 2\int_0^t a(s)m^2(t) + 2\int_0^t b(s)m(s)ds.$$

Subtracting the above equation from (6.29) and taking derivative with respect to t, we obtain

$$\frac{dv(t)}{dt} = 2a(t)v(t) + c^2(t)$$

with $v(t) = \mathbb{E}[X_t^2] - m^2(t)$.

8. Consider the CIR model

$$dY_t = (\gamma - \beta Y_t)dt + \sigma\sqrt{Y_t}dW_t, \qquad Y_0 = a, \tag{6.30}$$

where γ, β, σ, a are positive constants.

a) Write down equation (6.30) in integral form.

b) Let $u(t) = \mathbb{E}[Y_t]$. Use the integral form of (6.30) to show that $u(t)$ satisfies the differential equation

$$\frac{d}{dt}u(t) = \gamma - \beta u(t), \qquad u(0) = a.$$

c) By Itô formula, show that

$$dY_t^2 = Y_t(2\gamma + \sigma^2 - 2\beta Y_t)dt + 2\sigma\left(\sqrt{Y_t}\right)^3 dW_t. \tag{6.31}$$

d) Let $v(t) = \mathbb{E}[Y_t^2]$. Use the integral form of (6.31) to find a differential equation that is satisfied by $v(t)$. This equation can contain $u(t)$ and you do not need to solve for $u(t)$ from b).

e) Use b) and d) to determine the constant C so that

$$\mathrm{Var}[Y_t] = \mathbb{E}[Y_t^2] - (\mathbb{E}[Y_t])^2 \to C \qquad \text{when} \quad t \to \infty. \tag{6.32}$$

[Hint: You do not have to solve the differential equations in b) and d). Suppose you know $u(t) \to U$ and $v(t) \to V$ for some constants U and V when $t \to \infty$, what can you say about U and V?]