Stochastic Calculus, Courant Institute, Fall 2020
http://www.math.nyu.edu/faculty/goodman/teaching/StochCalc2020/index.html

Week 4

Jonathan Goodman, November, 2020
Tentative

1 SDE models, diffusions

This is the most important of the six classes. It describes how stochastic differential equations, SDEs, are used to create models of random processes in continuous time with continuous paths. The class starts with some terminology and the philosophy related to SDE models. This and the next section are all definitions and theory. The applications come in Sections 3 and 5, and in the exercises. There is less motivation here because the motivation is similar to Week 2 (for Ite's lemma) and Week 3 (for backward equations).

A random process X_t is an *Ito process* if it may be represented as an indefinite integral in the form

$$X_t = \int_0^t a_s\, ds + \int_0^t b_s dW_s \ . \tag{1}$$

The coefficients a_s and b_s may be random, but it is understood that that b_s is non-anticipating. It is not necessary that a_t and b_t be functions of W_t or X_t. For example, we could have

$$X_t = t^2 + \int_0^t b_s\, dW_s \ , \quad b_s = \int_0^s u\, dW_u \ .$$

The integrands a_s and b_s on the right side of the Ito process formula (1) determine the *infinitesimal mean* and *infinitesimal variance* of X_t. Suppose $dt > 0$ is an infinitesimal but non-zero increment of time (a more mathematical version is coming) and $dX = X_{t+dt} - X_t$ is the corresponding increment of X. The infinitesimal mean is a_t means

$$a_t\, dt = \mathrm{E}\big[\, dX_t \mid W_{[0,t]} \,\big] \ . \tag{2}$$

The conditional expectation on the right is the conditional expectation given that you know everything relevant that happened up to time t. In this case, the only thing you might or might not know is the value of the Brownian motion path. We discuss this issue more below.

It might be more familiar to math-trained people to take a small but not infinitesimal increment $\Delta t > 0$. The corresponding increment of X is $\Delta X = X_{t+\Delta t} - X_t$. The infinitesimal mean formula is

$$a_t\, \Delta t = \mathrm{E}\big[\, \Delta X \mid W_{[0,t]} \,\big] + o(\Delta t) \ . \tag{3}$$

I think of the informal version (2) as a shorthand way to write this. I believe (personal belief, others disagree) that more formal statements like (3) do not make people who use them more likely to reason correctly. I see plenty of finance and economics papers written in terms of the fanciest mathematical formalism that make elementary reasoning mistakes that would be less likely using simpler and more intuitive reasoning such as (2).

The infinitesimal variance is

$$b_t^2 \, dt = \text{var}\big(\, dX \mid W_{[0,t]} \, \big) \ . \tag{4}$$

This is the same as the expected square of the increment

$$b_t^2 \, dt = \text{E}\Big[\, (dX)^2 \mid W_{[0,t]} \, \Big] \ . \tag{5}$$

This is because

$$\begin{aligned} \text{var}(\, dX \mid \cdot \,) &= \text{E}\Big[\, (dX)^2 \mid \cdot \, \Big] - (\text{E}[\, dX \mid \cdot \,])^2 \\ &= \text{E}\Big[\, (dX)^2 \mid \cdot \, \Big] - a_t^2 \, dt^2 \ . \end{aligned}$$

In the language of Week 2, the dt^2 term on the right is *tiny* and can be ignored. In the Δt language, (4) would be

$$b_t^2 \Delta t = \text{var}\big(\, \Delta X \mid W_{[0,t]} \, \big) + o(\Delta t) \ . \tag{6}$$

This is equivalent to an expected square formula

$$b_t^2 \Delta t = \text{E}\Big[\, (\Delta X)^2 \mid W_{[0,t]} \, \Big] + o(\Delta t) \ . \tag{7}$$

The derivation is almost the same. If you're not used to "big Oh" and "little oh" reasoning, you can use the less formal version with differentials given above, or you can look it up in Wikipedia.

$$\begin{aligned} \text{var}(\, \Delta X \mid \cdot \,) &= \text{E}\Big[\, (\Delta X)^2 \mid \cdot \, \Big] - (\text{E}[\, dX \mid \cdot \,])^2 + o(\Delta t) \\ &= \text{E}\Big[\, (\Delta X)^2 \mid \cdot \, \Big] - \big(a_t^2 \, \Delta t + o(\Delta t)\big)^2 + o(\Delta t) \\ \text{var}(\, \Delta X \mid \cdot \,) &= \text{E}\Big[\, (\Delta X)^2 \mid \cdot \, \Big] + o(\Delta t) \ . \end{aligned} \tag{8}$$

The infinitesimal mean and infinitesimal variance formulas come from the Ito process representation and properties of integrals and continuous functions. For example,

$$\text{E}\big[\, \Delta X \mid W_{[0,t]} \, \big] = \ \text{E}\left[\int_t^{t+\Delta t} a_s \, ds \mid W_{[0,t]} \right]$$

The expectation of the Ito integral part is zero. If a_s is a continuous function of s, then

$$\int_t^{t+\Delta t} a_s \, ds = a_t \Delta t + o(\Delta t) \ .$$

For the infinitesimal variance, which is the same as the infinitesimal square, let Y_t be the Brownian motion integral

$$Y_t = \int_0^t b_s \, dW_s \; .$$

The increment of this is

$$\Delta Y = \int_t^{t+\Delta t} b_s \, dW_s$$

The Ito isometry formula from Week 2 gives

$$\mathrm{E}\left[\left(\int_t^{t+\Delta t} b_s \, dW_s \right)^2 \mid \cdot \right] = \int_t^{t+\Delta t} \mathrm{E}\left[b_s^2 \mid \cdot \right] \, ds \; .$$

The conditioning is the Brownian motion path up to time t. This means that in the conditional expectation b_t is known and $b_s \approx b_t$ if $s \approx t$ (because b_s is a continuous function of s). Therefore

$$\int_t^{t+\Delta t} \mathrm{E}\left[b_s^2 \mid \cdot \right] \, ds = b_t^2 \Delta t + o(\Delta t) \; .$$

You can check, as we checked (8), that the ds integral in (1) changes this by a "tiny" amount, which means $o(\Delta t)$. This shows that the Ito process (1) has infinitesimal mean (3) and infinitesimal variance (6) if a_t and b_t are continuous functions of t. What I call "infinitesimal variance" is more commonly called *quadratic variation*. The infinitesimal mean is *drift*.

This reasoning is used to write integral expressions for stochastic processes that satisfy specified drift and quadratic variation conditions. Suppose you have a stochastic process X_t and some reasoning suggests that the drift is a_t and the quadratic variation is μ_t. You pick any square root $b_t^2 = \mu_t$. Then the integral (1) has the desired properties. From this we learn that an Ito process is completely determined by its infinitesimal mean and variance. You might of this as Gaussian-like. Gaussian random variables are determined by their mean and variance. But Ito processes do not have to be Gaussian.

An Ito process is a *diffusion* if it is also a *Markov process*. The *Markov property* is that the distribution of the future, conditional on the past, is the same as the distribution of the future conditional on the present. This means that the distribution of the increments (see Week 1 for the definition of increment processes) in the future of t depends on X_t alone, not on values of X_s for $s < t$. Many probability classes discuss the Markov property when talking about Markov chains. The Markov property for Ito processes is less technical. We saw that an Ito process is determined by its infinitesimal mean and variance. Therefore, an Ito Markov process, a diffusion process, is determined by the the infinitesimal mean and variance at time t as functions of X_t. If X_t is such a process, there are functions $a(x,t)$ and $b(x,t)$ so that

$$dX_t = a(X_t,t) \, dt + b(X_t,t) \, dW_t \; . \tag{9}$$

This is a *stochastic differential equation* or *SDE*.

There are *weak* and *strong* ways to understand an Ito process. The *weak* way to that X_t is a stochastic process determined by its infinitesimal mean and variance as (2) and (4). This is natural from a modeling point of view. In the *strong* interpretation you think of X_t as a function of the Brownian motion path $W_{[0,t]}$ through the integral representation (1). This is convenient for computing and analysis. The weak interpretation The theory of Ito processes says that if X_t has continuous sample paths and has infinitesimal mean and variance

2 Ito's lemma for diffusion processes

Let X_t be an Ito process specified in the strong way (1), which we write as

$$dX_t = a_t dt + b_t dW_t \; . \tag{10}$$

For a general Ito process, as opposed to a diffusion process, the coefficients a_t and b_t are any random but non-anticipating functions. Suppose $f(x,t)$ is some function, the corresponding "Ito's lemma" is

$$df(X_t,t) = \partial_x f(X_t,t)\, dX_t + \left[\partial_t f(X_t,t) + \frac{1}{2} b_t^2 \, \partial_x^2 f(X_t,t) \right] dt \; . \tag{11}$$

The "Ito rule" behind this is $(dX_t)^2 = b_t^2 \, dt$. We can write this in a different way using the Ito formula (10) for dX. The formula becomes (for me, anyway) more clear when we leave out the arguments X_t, t everywhere, so $f(X_t,t) = f$ and $\partial_t f(X_t,t) = \partial_t f$, etc.

$$df = b_t \partial_x f \, dW_t + \left[\partial_t f + a_t \partial_x f + \frac{1}{2} b_t^2 \, \partial_x^2 f \right] dt \; . \tag{12}$$

This formula has the consequence that $Y_t = f(X_t,t)$ is a martingale if and only if the coefficient of dt is zero, which is

$$f(X_t,t) \text{ is a martingale } \iff \partial_t f + a_t \partial_x f + \frac{1}{2} b_t^2 \, \partial_x^2 = 0 \; . \tag{13}$$

This is the main step in deriving backward equations for general diffusion processes.

The derivation follows the one from Week 2. You expand in a Taylor series. As in Week 2, we write f for $f(X_t,t)$, etc.

$$\begin{aligned} df(X_t,t) &= f(X_t + tdX, t+dt) - f(X_t, t_) \\ &= \partial_x f dX + \frac{1}{2} \partial_x^2 f \, (dX)^2 + \partial_t f dt + \text{ tiny terms} \; . \end{aligned}$$

The "tiny terms" are things like $\frac{1}{2} \partial_t f (dt)^2$ and $\frac{1}{6} \partial_x^3 f (dX)^3$. The more subtle step is

$$(dX)^2 = b_t^2 dt + \text{ tiny } \; .$$

Part of being an Ito process is the square increment expectation:

$$\mathrm{E}\left[\,(dX)^2 \mid W_{[0,t]}\right] = b_t^2 dt \ .$$

This means that

$$\mathrm{E}\left[\,(dX)^2 - b_t^2 dt \mid W_{[0,t]}\right] = 0 \ .$$

We showed in Week 2 that this is tiny.

A more formal explanation of Ito's lemma would go like the one in Week 2. Look back at that one to follow this. Here, I will be "sketchy", to give you an idea what's going on without explaining it completely. I hope this helps you understand and use the Ito formulas (11) or (12). The formula (12) is understood as shorthand for

$$\begin{aligned} f(X_{T_2}, T_2) - f(X_{T_1}, T_1) = &\int_{T_1}^{T_2} \partial_t f(X_s, s)\, dW_s \\ &+ \int_{T_1}^{T_2} \left[\partial_t f(X_s, s) + a_s \partial_x f(X_s, s) + \frac{1}{2} b_s^2\, \partial_x^2 f(X_s, s) \right] ds \ . \end{aligned}$$

The discrete times corresponding to a small Δt are $t_k = k\Delta t$. The corresponding increments are

$$\begin{aligned} \Delta X_k &= X_{t_{k+1}} - X_{t_k} \\ \Delta f_k &= f(X_{t_{k+1}}, t_{k+1}) - f(X_{t_k}, t_k) \ . \end{aligned}$$

We write f_k for $f(X_{t_k}, t_k)$, and similarly for derivatives.

$$\begin{aligned} f(X_{T_2}, T_2) - f(X_{T_1}, T_1) &\approx \sum_{T_1 \le t_k < T_2} \Delta f_k \\ &= \sum_{T_1 \le t_k < T_2} \partial_x f_k \Delta X_k + \frac{1}{2} \partial_x^2 f_k (\Delta X_k)^2 + \partial_t f_k \Delta t + \cdots \ . \end{aligned}$$

We use (1) to replace ΔX_k with $b_{t_k} \Delta W_k + a_{t_k} \Delta t$. This is because

$$\Delta X_k = \int_{t_k}^{t_{k+1}} b_s dW_s + \int_{t_k}^{t_{k+1}} a_s dt \ .$$

We need to know what happens when we replace b_s with b_{t_k}, since

$$\int_{t_k}^{t_{k+1}} b_{t_k} dW_s = b_{t_k} \Delta W_k \ .$$

The difference is

$$\int_{t_k}^{t_{k+1}} (\, b_s - b_{t_k}) \, dW_s \ .$$

This "should be small", because b_s is close to b_{t_k}.

The weak interpretation of Ito's lemma is justified in a possibly simpler way using the weak understanding of the Ito process. The weak understanding of Ito's lemma in the dW form (12) is

$$\mathrm{E}\big[\,\Delta f \mid W_{[0,t]}\,\big] = \left[\partial_t f + a_t \partial_x f + \frac{1}{2} b_s^2\, \partial_x^2 f\right] \Delta t + o(\Delta t) \tag{14}$$

$$\mathrm{E}\big[\,(\Delta f)^2 \mid W_{[0,t]}\,\big] = b_t^2\, \Delta t + o(\Delta t)\ . \tag{15}$$

Both of these may be seen using Taylor expansions. It's easy to get these terms and to guess that other terms are "tiny" in the sense of being $o(\Delta t)$. It's not so easy to prove it (as far as I know, maybe someone can correct me). For example, there are terms involving $(\Delta X)^3$ and $(\Delta X)^4$. If X_t were Brownian motion, the cubic term would be zero and the quartic term would be $3\Delta t^2$. When X_t is a more complicated Ito process, the appropriate sized are harder to prove.

3 Ornstein Uhlenbeck

The *Ornstein Uhlenbeck* process, or *OU* process, satisfies the SDE

$$dX_t = -aX_t dt + \sigma dW_t\ . \tag{16}$$

We imagine that the *friction coefficient* a is positive, but many facts about the OU process are true for any a. Ornstein and Uhlenbick (Einstein first, but he had too many things named after himn) used this as a model of the velocity of small particle in water. The term $-aX_t dt$ represents friction from the water trying to make the particle move slower. The term σdW_t represents random impacts of water molecules on the particle, which make it move. The particle slows down quickly when a and keeps its velocity longer when a is small. The noise coefficient σ determines the strength of random forces that take X_t from its equilibrium position $X = 0$.

The OU process is used to model physical or financial quantities that have an equilibrium position or value when the equilibrium is disturbed by random noise. If the equilibrium position is $x_* \neq 0$, the SDE (16) may be modified to

$$dX_t = -a(X - x_*)dt + \sigma dW\ . \tag{17}$$

In finance, this might apply to the short term interest rate, r_t. In the model, there is a "natural" rate r_* that is disturbed by random "shocks". But r_t will tend, on average, to revert to r_* over time.

The OU SDE (16) may be solved using the *integrating factor* method you learned in your class on differential equations (and possibly forgot). We give the calculations in the language of stochastic calculus and Ito's lemma. We write the equation as

$$dX_t + aX_t dt = \sigma dW_t\ .$$

Then we multiply by the *integrating factor*, e^{at}. On the left, you have

$$e^{at} dX_t + e^{at} aX_t dt\ .$$

In differential equations class, this is more likely written using derivatives, and the calculation would be

$$e^{at}\frac{dX}{dt} + e^{at}aX = \frac{d}{dt}\left(e^{at}X\right) \ .$$

We can do the corresponding calculation for the OU process using Ito's lemma. We want to calculate $d(e^{at}X_t)$. Here are the calculations in the language of Ito's lemma in the form (11)

$$\begin{aligned} f(x,t) &= e^{at}x \\ \partial_x f &= e^{at} \\ \partial_x^2 f &= 0 \\ \partial_t f &= ae^{at}x \\ d\left[\,f(X_t,t)\right] = d\left[\,e^{at}X_t\right] &= e^{at}dX_t + ae^{at}X_t dt \ . \end{aligned}$$

Therefore, the SDE (16) is equivalent to

$$d\left[\,e^{at}X_t\right] = e^{at}dW_t \ .$$

The next step in the integrating factor method is to integrate the perfect differential (in differential equations, the perfect derivative), to get

$$e^{aT}X_t - X_0 = \int_0^T e^{at}dW_t \ .$$

Finally, you multiply by e^{-aT} to arrive at a formula for X_T

$$X_T = e^{-aT}X_0 + \int_0^T e^{-a(T-t)}\, dW_t \ . \tag{18}$$

This solution formula reveals most of the important facts about the OU process. One is that the OU process "forgets" its initial state, X_0. As $T \to \infty$, the influence of the initial state, which is $e^{-aT}X_0$ disappears exponentially. Another is that the distribution of X_T converges to a Gaussian as $T \to \infty$. In fact, the whole right side of (18) is Gaussian, with mean $e^{-aT}X_0$. An Ito integral with respect to Brownian motion where the integrand is just a function of t (i.e, is not random), is Gaussian. This is the Gaussian nature of Brownian motion. You can "see" it using the approximation

$$\int_0^T c_t dW_t \approx \sum_{t_k<T} c_{t_k}\Delta W_k \ .$$

The right side is a sum of independent Gaussians ΔW_k with weights c_{t_k} that are not random. Therefore, the right side is mean zero Gaussian for any $\Delta t > 0$. If the variance has a limit, then the distribution has a limit, which is the Gaussian with that variance. In this case, the variance is given by the Ito isometry formula

$$\mathrm{E}\left[\left(\int_0^T e^{-a(T-t)}dW_t\right)^2\right] = \int_0^T \mathrm{E}\left[\left(e^{-a(T-t)}\right)^2\right] dt \ .$$

The quantity in square braces $[\cdots]$ on the right is not random, so the integral is

$$\mathrm{E}\left[\left(\int_0^T e^{-a(T-t)} dW_t\right)^2\right] = \int_0^T e^{-2a(T-t)}\, dt = \frac{1}{2a}\left(1 - e^{-2aT}\right) \; .$$

In particular

$$\lim_{T\to\infty} \mathrm{var}(X_t) = \frac{1}{2a} \; . \tag{19}$$

The OU model (17) is an example of an *equilibrium model* because there is an equilibrium distribution, an equilibrium PDF, for X_t. The PDF of X_t converges to the equilibrium distribution as $t \to \infty$. The process X_t itself does not stop moving, but for large t it moves "within" the equilibrium distribution. Brownian motion is not an equilibrium model because the variance of W_t goes to infinity as $t \to \infty$. The OU model can maintain an equilibrium because of the *mean reversion* term $-a(X - x_*)dt$. The mean of X_t reverts to x_*, but X_t itself (as was just said) fluctuates about x_* with a variance of about $\frac{1}{2a}$ when t is large.

Equilibrium models may be appropriate for some quantities in financial markets, such as interest rates. They are not may not be appropriate for prices of traded assets (stocks). If the stock price were mean reverting, you would be able to profit by buying whenever $X_t < x_*$ and selling whenever $X_t > x_*$. That said, a Nobel Memorial Prize in economics was given to people who showed empirically that actual stock prices are slightly mean reverting.

4 Backward equations for diffusions

Let X_t satisfy an SDE (9). If $V(x)$ is a payout function, the corresponding value function is

$$f(x,t) = \mathrm{E}[\, V(X_T) \mid X_t = x] \; . \tag{20}$$

The value function satisfies the backward equation

$$\partial_t f + a(x)\partial_x f + \frac{1}{2} b^2 \partial_x^2 f = 0 \; . \tag{21}$$

This PDE has *final conditions* $f(x,T) = V(x)$. The final condition and the PDF determine $f(x,t)$ for $t \leq T$. This equation may be justified using one of the justifications from Week 3. If f satisfies this, then (13) implies that $f(X_t,t)$ is a martingale. That means that

$$\mathrm{E}[\, f(X_T,t) \mid X_t = x] = f(x,t) \; .$$

The final condition $f(x,T) = V(x)$ shows that this f is precisely the value function (20).

As for Brownian motion, there are other backward equations for other functions of the process. Consider a continuous payout random variable

$$Y = \int_0^T R(X_s) ds \; .$$

We form a value function at time t by considering only the payouts in the future of t:

$$Y_t = \int_t^T R(X_s)\, ds \; . \tag{22}$$

The value function for this is

$$f(x,t) = \mathrm{E}\left[\int_t^T R(X_s)\, ds \mid X_t = x \right] \; . \tag{23}$$

We find the PDE that f satisfies by looking for a martingale. We find a value function by looking for a martingale related to this. So consider the process

$$Z_t = f(X_t,t) - \int_t^T R(X_s,s)\, ds = f(X_t,t) - Y_t \; .$$

This satisfies $Z_T = 0$, because both (22) and (23) give zero. Ask what PDE f should satisfy so that Z_t is a martingale. In a small increment of time, the integral Y_t decreases by $R(X_t)\, dt$:

$$dY_t = -R(X_t)\, dt \; .$$

With Ito's lemma, we get

$$dZ_t = QdW_t + \left[\partial_t f(X_t,t) + a(X_t)\partial_x f(X_t,t) + \frac{1}{2} b(X_t)^2 \partial_x^2 f(X_t,t) + R(X_t) \right] dt \; .$$

There is a formula for Q, but I didn't write it to emphasize that it doesn't matter. Z_t is a martingale if

$$\partial_t f(X_t,t) + a(X_t)\partial_x f(X_t,t) + \frac{1}{2} b(X_t)^2 \partial_x^2 f(X_t,t) + R(X_t) = 0 \; .$$

Therefore, we should solve

$$\partial_t f(x,t) + a(x)\partial_x f x,t) + \frac{1}{2} b(x)^2 \partial_x^2 f(X_t,t) + R(x) = 0 \; . \tag{24}$$

We should use final condition $f(x,T) = 0$, which expresses the fact that the running payout pays zero in zero time.

5 Geometric Brownian motion

A *geometric Brownian motion* is a diffusion process that satisfies the SDE

$$dS_t = \mu S_t dt + \sigma S_t dW_t \; . \tag{25}$$

The parameters μ and σ are the *expected rate of return* (or just *expected return* and *volatility* respectively. This seems to be a natural model of the price of a traded asset (a stock). The price goes up or down by an amount proportional to S_t. This means that the probability of $S = 100 \to S = 102$ is the same as the probability $300 \to 309$. Both are 2% increases.

This equation may be solved using basic differential equation ideas supplemented with Ito's lemma. The differential equations method is *separation of variables*, which means putting S and derivatives of S on one side and then integrating both sides. The first step of this program is

$$\frac{1}{S_t} dS_t = \mu dt + \sigma dW_t \ . \tag{26}$$

The indefinite integral of the right side is

$$\int_0^T \mu dt + \int_0^T \sigma dW_t = \mu T + \sigma W_T \ .$$

In ordinary calculus, you would recognize $\frac{1}{s}\frac{ds}{dt} = \frac{d}{dt}\log(s)$. We do the corresponding calculation for the diffusion process using Ito's lemma with

$$\begin{aligned} f(s) &= \log(s) \\ \partial_s f(s) &= \frac{1}{s} \\ \partial_s^2 f(s) &= -\frac{1}{s^2} \\ \partial_t f(s) &= 0 \\ (dS_t)^2 &= \sigma^2 S_t^2 dt \ . \end{aligned}$$

Ito's lemma (11) then gives

$$d\log(S_t) = \frac{1}{S_t} dS_t - \frac{1}{2}\,\sigma^2 S_t^2\,\frac{1}{S_t^2}\,dt \ .$$

We re-write this for our calculation in the form

$$\frac{1}{S_t} dS_t = d\log(S_t) + \frac{\sigma^2}{2} dt \ .$$

This puts (26) into the form

$$d\log(S_t) = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dW_t \ .$$

We integrate both sides from $t = 0$ to $t = T$ to find

$$\log(S_T) - \log(S_0) = \left(\mu - \frac{\sigma^2}{2}\right) T + \sigma W_T \ .$$

With a little more algebra, this is the solution formula

$$S_T = S_0 e^{\sigma W_T + \left(\mu - \frac{\sigma^2}{2}\right) T} \ . \tag{27}$$

Exercise 2 looks at this solution from different points of view.

The solution formula (27) allows you to find the PDF of S_T. The distribution is called *log-normal* because $\log(S_T)$ is normal. Exercise 5 asks you to calculate and plot this density. Other than plotting, I don't know much use for it. If I want to calculate expectations of functions of S_T, I use the Gaussian distribution of W_T instead. We will see this in Week 5, when we derive the *Black Scholes formula*. The formula (27) shows that $S_T > 0$. You will see in Exercise 5 that computations using the Euler Maruyama method (29) might fail to give positive results if the time step is too large.

The formula (27) has the striking feature that the usual growth rate of S_T for large T is slower than $e^{\mu T}$. This is strikingly clear when $\mu = 0$. Then the process (25) is a martingale, so its expected value does not change with time. However, for large T, W_T is on the order of $\sqrt{T}$ so $e^{\sigma W_T}$ is the smaller part and $e^{-\frac{1}{2}\sigma^2 T}$ is the larger part. This GBM converges to zero in distribution. You will see this in the plots of Exercise 5. The expectation

$$\mathrm{E}[S_T] = S_0$$

is achieved by having rare paths much larger than S_0 while typical paths are much smaller. This is an example of a strongly *skewed* distribution with a PDF that is not symmetric around the mean.

6 Computational methods

This section discusses two computational problems. One is generating sample paths for a diffusion process from the SDE. The other is finite difference methods for solving the backward equation. Most modeling projects involve first formulating a stochastic model such as an SDE and then doing computer work of some kind to explore the behavior of the model. The material here should seem natural, given similar methods for simpler problems we have already done. These methods do not produce the exact solution, neither for sample paths nor for backward equations. They have parameters Δt or Δx. As these parameters go the zero, the computed solution converges to the actual (model) solution. The trick in practice is to choose Δt or Δx small enough to get the accuracy you need. The computer time increases as Δt and Δx decrease, so you don't want to take these parameters smaller than necessary.

Consider the SDE (9). Choose a Δt and approximation times $t_k = k\Delta t$. Denote the values of the approximate sample path by

$$X_k^{\Delta t} \approx X_{t_k} \ .$$

An optimistic approximation to the SDE for non-zero Δt would be

$$X_{k+1}^{\Delta t} = a(X_k^{\Delta t})\, \Delta t + b(X_k^{\Delta t})\, \Delta W_k \ . \tag{28}$$

Everything here is known, except possibly ΔW_k. We take this to be the increment of Brownian motion over the time increment Δt. These increments (as we have seen since Week 1) are Gaussian with mean zero and variance Δt. You can ask the Gaussian random number generator to give you random variables with that distribution (see the code `StockSim.py`), or you can ask for $Z_k \sim \mathcal{N}(0,1)$ and take $\Delta W_k = \sqrt{\Delta t} Z_k$. In this case, the computer program would implement the formula

$$X_{k+1}^{\Delta t} = a(X_k^{\Delta t})\, \Delta t + b(X_k^{\Delta t})\, \sqrt{\Delta t}\, Z_k \ , \quad Z_k \sim \mathcal{N}(0,1) \ . \tag{29}$$

This is the *Euler Maruyama* method, which is how diffusion processes are usually simulated.

This may seem odd, particularly if you have experience with numerical methods for ordinary or partial differential equations. For those problems there are families of sophisticated and extremely accurate methods, including Runge Kutta methods, finite element methods, and so on. There are whole graduate courses devoted to such methods. The simplest method, which is Euler's method, is explained in the first class. The rest of the course explains better methods. Yet, for SDE, there do not seem to be methods that are much better than the simple Euler Maruyama method (28).

7 Exercises

1. Let X_t be an OU process with a deterministic starting point $X_0 = x_0$. Let $u(x,t)$ be the PDF for X_t.

   (a) Use the solution formula (18) to show that $u = \mathcal{N}(\mu_t, v_t)$ and find formulas for the mean μ_t and variance v_t.

   (b) Find the solution formula like (18) for the process 17 that reverts to x_*.

   (c) Show that $\Pr(X_t < 0) > 0$ for any x_0 and $\sigma > 0$ and $t > 0$ and x_*.

   (d) Suppose X_t satisfies the SDE (9) and $u(x,t)$ is the PDF of X_t. The *forward equation* is

   $$\partial_t u = -\partial_x \left(a(x) u(x,t) \right) + \frac{1}{2} \partial_x^2 \left(b(x)^2 u(x,t) \right) \ .$$

   Check by explicit calculation that the solution formula from part (a) satisfies the forward equation for (16).

   (e) Show that if $X_t \sim \mathcal{N}(0, \frac{1}{2a})$, then X_T has the same distribution for any $T > t$. *Hint*, show that this satisfies the forward equation.

   (f) Use your formulas for the mean and variance of X_t to find the value function for an OU process (16) mean reverting to $x_* = 0$. Show that your formula satisfies the backward equation and appropriate final condition.

2. The geometric Brownian motion SDE 25 may be solved using the *log variable transformation*. There are several equivalent ways to derive the transformation.

(a) Set $X_t = \log(S_t)$. Use Ito's lemma to find the SDE that X_t satisfies. Show that $X_t = a + bt + cW_t$ is a solution. The process X_t is *Brownian motion with drift*. This is the derivation given above, explained slightly differently.

(b) Write the backward equation for S_t. Consider the log change of variables $f(s,t) = g(\log(s), t)$. Show that this g satisfies

$$\partial_t g + \alpha \partial_x g + \beta \partial_x^2 g = 0 \ .$$

Find the relation between α and β here to a, b, and c from part (a). What diffusion process has this PDF as its backward equation?

3. Assume that the higher moments of the diffusion process are of the size they would be for Brownian motion, which is

$$\mathrm{E}\Big[\, |\Delta X|^3 \Big] \leq C \Delta t^{\frac{3}{2}}$$
$$\mathrm{E}\Big[\, |\Delta X|^4 \Big] \leq C \Delta t^2$$

Suppose $f(x,t)$ has partial derivatives up to order 4 in both variables. Define $\Delta f = f(X_{t+\Delta t}, t + \Delta t) - f(x,t)$ Show that

$$\mathrm{E}[\, \Delta f \mid X_t = x] = \left[\partial_t f(x,t) + a(x)\partial_x f(x,t) + \frac{1}{2} b(x)^2 \partial_x^2 f(x,t) \right] \Delta t + o(\Delta t) \ .$$

Show that

$$\mathrm{E}\big[\, (\Delta f)^2 \mid X_t = x \big] = \left[b(x) \partial_x f(x,t) \right]^2 \Delta t + o(\Delta t) \ .$$

This is a *weak* version of Ito's lemma.

4. Consider the Ornstein Uhlenbeck process (16) and running payout

$$\int_t^T X_s^2 \, ds \ .$$

Evaluate the value function (23) explicitly using variance formulas for the OU process. Verify that this function satisfies the backward equation (24).

5. The code `StockSim.py` does the Euler Maruyama method to compute a geometric Brownian motion governed by the SDE (26).

(a) Run with a larger T and Δt and see that it is common to produce negative approximate prices. For this, you do not need so many paths and the plots will not look good. This is OK because the results are not good either.

(b) Find a formula for $u(s,t)$, which is the PDF of S_T. You can use the solution formula (27) for this. Add the exact PDF to the plot and see what Δt you need to get a good match.

(c) The problem gets harder for larger T. Do a calculation with larger T to see that the PDF of S_T is not at all symmetric.

(d) A *volatility surface* model makes σ a function of s. A volatility *skew* adds a slope, which is $\sigma(s) = \sigma_0 + \sigma_1(s - S_0)$. A volatility *smile* adds a positive quadratic term. Experiment with volatility skew, both positive and negative and see how this impacts the PDF. Add two curves to the plot, one with positive and one with negative skew. Choose the slopes s_1 so that the PDF is noticeably different but not completely different. At this point, your plot will have four curves. Make sure the vol surface curve legend labels have the corresponding skew values. Volatility skew and smile are used to explain the observed fact that market option prices for put options that are unlikely to be "in the money" are much higher than the Black Scholes theory says they should be.